

Tiziano Vecchiato

Valutare l'impatto sociale con metriche adeguate

La sfida degli esiti non può essere evitata. La differenza tecnica tra esito e impatto è sostanziale: gli esiti sono beneficio diretto per i destinatari mentre l'impatto è beneficio esteso alla comunità. La valutazione di impatto sociale è chiamata ad entrare nel merito dei potenziali a disposizione per meglio identificare quanto i servizi di welfare riescono a redistribuire bene comune. L'outcome e l'impatto sociale sono infatti questioni di elevato interesse sociale, non confinabile nella sfera della valutazione privata, visto che le soluzioni dei problemi umani sono «bene comune» da condividere su più vasta scala.

Il problema

Nel dibattito sulla valutazione di impatto vengono considerate soprattutto condizioni per ottimizzare i fattori produttivi e migliorare le performance economiche e sociali, ma quasi sempre si tratta di misurare di *output* e non di *outcome*. «Fare molto» per «produrre poco» non può bastare e le verifiche di esito e di impatto non possono fare sconti. L'enfasi sulla qualità di processo ha attirato per molti anni le attenzioni istituzionali, con certificazioni e accreditamenti. Ingenti quantità di tempo e risorse hanno mascherato il *deficit* di investimento

proprio su quello che qui ci interessa: come valutare gli esiti e l'impatto sociale per quantificare i benefici e misurare il valore sociale reso possibile (Vecchiato T, 2014a; 2014b).

Il prestazionismo è molto diffuso. Domina le prassi attuali di welfare, mortificando l'esercizio delle responsabilità professionali e sociali. Invece di trasformare capacità e risorse si accontenta di

fare quello che viene chiesto, senza domandarsi se serve, con quanti benefici a vantaggio di chi. La filiera del *logic model* «*input, output, activity, outcome, impact*» viene così trasformata in procedura da seguire con risultati contabilizzati, senza riconoscere il rendimento e la rigenerazione delle risorse che le verifiche di impatto dovrebbero evidenziare.

La valutazione di esito prepara condizioni affidabili per la valutazione di impatto sociale. In gioco non è solo la misura dei risultati, ma l'esercizio delle responsabilità necessarie per conseguire bene ulteriore a vantaggio dei destinatari diretti e indiretti.

Esiti e impatto

La sfida degli esiti non può essere evitata. La differenza tecnica tra esito e impatto è sostanziale: gli esiti sono beneficio diretto

AUTORE

■ Tiziano Vecchiato, direttore Fondazione «Emanuela Zancan», Padova.

per i destinatari mentre l'impatto è beneficio esteso alla comunità. La differenza tra *outcome* «intenzionali e non intenzionali» è primo criterio per discriminare ciò che si raggiunge perché cercato e ciò che si ottiene per effetto indotto o, nei casi migliori, in «concorso al risultato». Tecnicamente si distingue tra *outcome* di tipo alfa, cioè misura di esito diretto delle azioni realizzate e *outcome* di tipo beta, che identifica l'esito aggiuntivo reso possibile dalla personalizzazione delle risposte. Oltre a questo e a certe condizioni si possono misurare indici di tipo gamma, quando il valore è ulteriore, non grazie alle azioni messe in campo ma per il concorso dei destinatari all'esito, a vantaggio di quanti fanno parte del loro spazio di vita e/o della «collettività» (Fondazione Emanuela Zancan 2012; 2013; 2014; 2015).

Si tratta di una progressione generativa, frutto congiunto di esito e di impatto. Insieme mettono a valore l'esercizio delle responsabilità «per e con», così da conseguire il massimo possibile di beneficio personale e sociale.

Tecnicamente è estensione delle potenzialità del *logic model*, se si accetta di misurare gli esiti «ulteriori», quelli eccedenti, oltre le proprie attese, a beneficio della comunità. Sono esiti non sempre prefigurati ma misurabili (Fernandez E. e altri, 2015; Vecchiato T., 2013). Sono riconoscibili nell'incontro generativo di responsabilità, a vantaggio proprio e di altri, nell'incontro tra diritti e doveri, produttivi di dividendo sociale. La valutazione di impatto sociale è chiamata ad entrare nel merito di questi valori, per meglio identificare quanto i servizi di welfare riescono a realizzare e a redistribuire. Avviene più facilmente, con un cambio di paradigma, se da costo diventano investimento produttivo di valore economico e sociale (Bezze M. e Vecchiato T., 2012; Maffucci A.N. e altri, 2011).

Indicazioni dal dibattito europeo

Le sfide per il *Social Impact Investment* sono sintetizzate dal Geces (2014) in 6 punti. Le misure di impatto sociale non possono essere qualcosa di indefinito e non direttamente quantificabile secondo parametri economici. Significa che, diversamente dal passato, non si può pensare a soluzioni univoche, visto che gli *outcome* da misurare hanno, come è naturale che sia, contenuti diversi e, anche per questo, richiedono metriche sensibili alle diverse specificità da misurare. La possibilità di accettare la non univocità delle metriche non è scontata. Il timore è di rinunciare a livelli di affidabilità basati su indici standardizzati.

Una prima condizione per operare con metriche diverse viene anzitutto dalla loro padronanza, associata all'utilizzo di una sintassi in grado di gestire varie metodiche, sapendo che i loro risultati possono essere riportati a unità di misura affidabili. Metriche diverse sono necessarie per valutare gli esiti in modo trasparente e adeguato nei differenti contesti in cui si ottengono. Saranno tanto più affidabili, come vedremo, quanto più saranno misure dirette e non approssimate (*proxy*) delle dimensioni da valutare.

Spesso i tassi di approssimazione vengono ridotti con tecniche di coinvolgimento degli aventi interesse. Ma il cointeressamento e la corresponsabilizzazione nella valutazione sono e rimangono dei *proxy*, non trasformabili in indici di esito diretto. Tecnicamente non dovrebbero cioè diventare valori di esito e impatto dell'investimento (Investing for good, 2015; The Good analyst, 2012), ma compensare il quadro quando non è abbastanza a fuoco.

Misure di interesse sociale

Le premesse appena esposte sono «requisiti sintattici» necessari per gestire misure di interesse sociale. Non rispettarli significa

accettare un inevitabile indebolimento dei risultati di quella che dovrebbe essere *Social Impact Investment Evaluation*. Non a caso al punto 6 delle raccomandazioni Geces si evidenzia la necessità di metriche adeguate, cioè: (a) capaci di interfacciarsi con le specificità dei contenuti e dei contesti; (b) isomorfe, così da rendere possibile il confronto tra risultati, senza che le metriche utilizzate diventino un ostacolo, visto che sono un'opportunità per rendere più affidabili le comparazioni tra risultati ottenuti con diverse azioni².

I servizi sanitari e sociali sono un terreno di investimento privilegiato nel sociale. Le ragioni sono riconducibili ai fondamenti della socialità solidale. Il primo è la raccolta fiscale. Fa convergere grandi quantità di risorse, però non gestite con logiche di «investimento» ma di «trasferimento». Le perdite sono considerevoli, a causa dell'inefficiente sistema di trasporto delle risorse dal cittadino che conferisce fino al finanziamento dei servizi necessari per la comunità. La rete di redistribuzione ha infatti molte falle e distruzioni. All'offerta pubblica si aggiunge l'offerta privata, profit e non profit, che si aspetta adeguata remunerazione economica e/o sociale.

I risultati sono ad esempio calcolati con bilanci di salute, in modo epidemiologico, ammettendo che i fattori intervenienti sono numerosi e che gli esiti di salute sono descrivibili soltanto con indici generali e approssimativi. Vengono da una scatola nera poco decifrabile con risultati non trasparenti basati su tecniche di certificazione «di processo».

Negli ultimi 10 anni altre valutazioni si sono concentrate su risultati gestionali. Hanno trovato nei Lea (Livelli essenziali di assistenza) e nei costi standard due questioni su cui impegnarsi. Si è così capito che i problemi di metrica non sono compatibili con scelte politiche avulse dalle evidenze disponibili, senza valutazioni preventive di impatto. Nella definizione dei Lea è prevalso il pregresso giuridico (includendo nei Lea quanto previsto da norme indipenden-

ti da criteri di appropriatezza e di evidenza), mettendo così a nudo la mancanza di significato univoco sul modo di intendere il termine «essenziale». Ne hanno risentito i diritti fondamentali delle persone più deboli.

Problemi di trasparenza e affidabilità

Analoghe criticità emergono dalla disputa sui costi standard. Cosa significa «standard»? Non possono essere costi medi, visto che il costo di una prestazione appropriata non può essere mediocre. Per standard si può intendere: *gold* standard, medio standard, basso standard. Sempre standard è. I bisogni meritevoli di diventare diritti fondamentali delle persone cosa meritano? Si potrebbe rispondere che dipende dalla composizione dei fattori produttivi necessari per garantire risposte efficaci. Non basta cioè che le risposte siano qualitative, appropriate, «processate» in modi standardizzati. Il salto di sostanza è dato dal «se servono», se aiutano veramente, se sono appunto «efficaci». Il diritto umano ad avere aiuto effettivo non dovrebbe cioè essere trasformato in diritto ad ottenere prestazioni. Non è la stessa cosa e la verifica di esito lo sa. Dovrebbe saperlo anche la valutazione di impatto, se è ad essa strettamente connessa. Ma normalmente questo non avviene e chi valuta non sempre ha prima misurato gli esiti, mentre chi fa valutazione di impatto non necessariamente ha prima misurato gli indici di efficacia a cui ricondurla.

Si è così potuto lucrare sulla cedevolezza dei metodi, potendo contare su verifiche inadeguate. Se come vedremo, si adottassero modalità più lineari, dirette, le variabili rappresenterebbero in modi più affidabili le condizioni di partenza e di arrivo, riducendo i rischi di discrezionalità e di manipolazione.

Ogni sistema di rilevazione deve poter valorizzare i passaggi dal generale allo specifico come avviene in Sisclass (Bezze M.

e altri, 2005), con regole di composizione in grado di classificare i contenuti a diversi livelli (Ezell M. e altri, 2011). Ma un problema è come utilizzare dati di fonte diversa con criteri di equivalenza e in modo interoperabile. In alcune ricerche è stata utilizzata una combinazione di descrittori degli interventi/servizi. Gli attributi utilizzati per classificare i contenuti di offerta sono ad esempio stati: l'attività (il contenuto principale o fondamentale), le modalità (come viene realizzato quel contenuto), le finalità (perché viene realizzato il contenuto), il target (i destinatari dell'attività). Dalla loro combinazione si è visto come fosse possibile identificare l'unità di offerta (servizio) e i contenuti di offerta erogati, tenendo conto che nei settori sanitario e sociale le definizioni non sempre rispecchiano la differenza tra attività erogata, nomi, contenuti ad essa attribuiti e sistemi di remunerazione.

Oltre le soluzioni che conosciamo

Le logiche moderne in particolare la logica delle risorse computazionali (Girard J.-Y., 1987), si chiedono come le combinazioni dei fattori si possono caratterizzare in termini di *input* e *output* in diversi «contesti di condizionalità», quando i risultati possono ad esempio essere interpretati «in sequenza» o «come conseguenza» di specifiche risorse.

Ci possono aiutare due regole. La prima si basa sulla condizionalità classica e può essere espressa così $[M \rightarrow E]$. Si legge: «se *Measurement* allora *Evaluation*» o anche «se puoi valutare (*verum dicere*) prima devi aver verificato (*verum facere*)». La seconda è $[OE \rightarrow IE]$, cioè «se *Outcome Evaluation* allora *Impact Evaluation*» o anche «puoi fare valutazione di impatto se e solo se hai fatto valutazione di *outcome*». In entrambi i casi il condizionale materiale ($\rightarrow$) ci dice che le operazioni del «verificare» e del «valutare» e le operazioni di «*Outcome Evaluation*» e «*Impact Evaluation*» sono in relazione. Significa

che non è possibile valutare se prima non si è verificato. Significa che si devono evitare le misure *praxy* tutte le volte che i valori osservati possono corrispondere a metriche osservabili. Significa ancora che è meglio evitare l'intermediazione interpretativa degli *stakeholder*, quando si possono utilizzare variabili osservabili direttamente. È meglio infatti utilizzare misure lineari per poter gestire le funzioni del «verificare» e del «valutare» in modi più affidabili. Ai sensi della regola numero 1, non è corretto «invertire la sequenza» cioè raccogliere giudizi a cui attribuire un punteggio in modo da «valutare» per poi «verificare» o peggio ancora «valutare senza verificare». La regola numero 2 entra nel merito del rapporto tra valutazione di esito e valutazione di impatto: non posso fare «valutazione di impatto» se prima non ho fatto «valutazione di esito». È una regola esigente, perché chiede di non confondere l'esito con l'impatto. Impone di raccogliere misure di *outcome* diretto per i destinatari e, solo dopo, misurare i benefici estesi alla comunità di riferimento. È quello che la Fondazione Zancan fa con PersonaLAB (*Personalised Environment for Research on Services, Outcomes and Need Assessment*, www.personalab.org).

Strade praticabili?

Cosa succederebbe se applicassimo le due regole allo SROI (2012) o ad altre soluzioni utilizzate da enti finanziari? Sarebbe uno stress test impegnativo, visto che propone soluzioni «per semplificare» le fonti di informazione. Lo fanno con prescrizioni finalizzate ad associare valori «di osservazione» e «di giudizio», così da poterli confrontare con tecniche di *benchmark*. Si è pensato, in questo modo, di mettere un argine alla discrezionalità, accettando costi e rischi non indifferenti. Se analizziamo certe procedure di misurazione vediamo infatti quanto siano «*time consuming*» con tecniche di «*reductio ad unum*» che nascono da valutazioni contingenti formulate dai soggetti interessati.

Sappiamo che la tecnica di riduzione da giudizio a valore economico equivalente è un tipo di semplificazione accettata, ma il risultato è che avremo comunque misure *proxy*, gestite in modo «analogico» e anche per questo non direttamente espressive dei fattori osservati e misurati. Di conseguenza ne risentiranno le misure di *outcome* e quelle di impatto, a causa dell'insufficiente focalizzazione sugli «oggetti» da misurare e valutare. Sono sostituiti con indici formulati da «soggetti» implicati nell'investimento, tecnicamente «soggettivi». Il livello di affidabilità dei risultati dipenderà da quanto e come la metodologia è stata seguita. Sotto questa luce è una evoluzione metodologica del «*logic mode*», di cui assume le componenti grammaticali, compensando la carenza di sintassi con prescrizioni procedurali e metodologiche. Sono necessarie per gestire i processi di attribuzione. Anche per questo l'affidabilità dei risultati dei metodi quali ad esempio lo SROI è attribuibile al metodo e non abbastanza al merito.

Conclusioni

Le considerazioni proposte chiedono di misurare di volta in volta i problemi con metriche adeguate alla natura delle questioni considerate, per poi standardizzare con indici di equivalenza. Possono essere utilizzati per confrontare, su più vasta scala, la capacità di *outcome* e di impatto degli investimenti. È possibile farlo anche con problemi complessi, descrivibili con variabili multiassiali. È un modo per evitare la *black box*, in cui collocare variabili complesse, per associare ad esse «valori di giudizio». Il doppio rischio a cui si va incontro è l'eccesso di soggettività e le dinamiche di *group-think* (Manz C.C. e Neck C.P., 1995), cioè il pensiero che con dinamiche di consenso si riducono le differenze. Queste criticità hanno ostacolato avanzamenti sostanziali della conoscenza e nel nostro caso la possibilità di meglio valutare l'esito e l'impatto degli investimenti nell'area dei servizi umani.

Per questo le domande di Geces sono altrettante sfide per un migliore esercizio delle responsabilità pubbliche e private. L'*outcome* e l'impatto sociale sono infatti questioni di elevato interesse sociale, non confinabile nella sfera della valutazione privata visto che le soluzioni dei problemi umani sono «bene comune» da condividere su più vasta scala.

Note

- 1 Relazione tenuta da Tiziano Vecchiato il 10 ottobre 2015 nella giornata conclusiva della XV edizione delle Giornate di Bertinoro «Generare impatto sociale».
- 2 L1. Impact measurement should be an effort to describe synthetically what social impact an organization (or its funders) wishes to achieve and/or has achieved. However:
 - a. social impact should not be limited to «soft» behavior of people or interactions among them, but should include economic value
 - b. social impact can flow not only from changes made in the lives of people outside the organization, but also from its inner workings (participation etc.)
 - c. involvement of stakeholders for which measurement is done produces value per se.
2. There is no single best solution in impact measurement. Therefore, at no level of government (local, national, supranational e.g. EU, G8 or UN) there should be an attempt to prescribe method and metrics to be used to assess outcomes or impact. This is explicitly to preserve diversity in social sector modes of intervention and to avoid distorting perception of value and allotment of funds according to specific quantitative indicators.
3. However, it is legitimate and desirable to request minimum standards in the process of social impact measurement and the reporting of results. Some debate is still necessary as to the precise content of such requirements and who should impose them and check compliance.
4. There should be an agreed vocabulary of measurement. We should expect that most good measurement be done on outcomes, with some effort to clarify im-

pact (outcome due to the social intervention).

5. Third-party involvement in measurement is a plus. The extent of it depends on what is being measured and is maximal when qualitative answers are sought.

6. It is desirable to get more specific on metrics than the European approach did, as a form of guidance rather than regulation of the social sector. Doing this while respecting diversity is perhaps possible by focusing on specific sectors.

RIFERIMENTI BIBLIOGRAFICI

- Bezze M., Faenzi G., Lippi A., Paganelli L., Pompei A. e Vecchiato T. (2005), *La classificazione dei servizi e degli interventi sociali*, in «Studi Zancan», 2, pp. 91-131.
- Bezze M. e Vecchiato T. (2012), *La lotta alla povertà con un welfare generativo*, in «Studi Zancan», 6, pp. 11-30.
- Ezell M., Spath R., Zeira A., Canali C., Fernandez E., Thoburn, J. e Vecchiato, T. (2011), *An International Classification System for Child Welfare Programs*, in «Children and Youth Services Review», 33, pp. 1847-1854.
- Fernandez E., Zeira A., Vecchiato T. e Canali C. (a cura di) (2015), *Theoretical and Empirical Insights into Child and Family Poverty. Cross National Perspectives*, Springer, New York.
- Fondazione Emanuela Zancan (2012), *Vincere la povertà con un welfare generativo. La lotta alla povertà. Rapporto 2012*, Il Mulino, Bologna.
- Fondazione Emanuela Zancan (2013), *Rigenerare capacità e risorse. La lotta alla povertà. Rapporto 2013*, Il Mulino, Bologna.
- Fondazione Emanuela Zancan (2014), *Welfare generativo. Responsabilizzare, rendere, rigenerare. La lotta alla povertà. Rapporto 2014*, Il Mulino, Bologna.
- Fondazione Emanuela Zancan (2015), *Cittadinanza generativa. La lotta alla povertà. Rapporto 2015*, Il Mulino, Bologna.
- Girard J.-Y. (1987), *Linear Logic*, in «Theoretical Computer Science», 50, pp. 1-102.
- Investing for good (2015), *Dictionary of indicators*, in <http://www.investingforgood.co.uk/#!dictionary-of-indicators/cfcg>.
- Maluccio A.N., Canali C., Vecchiato T., Lightburn A., Aldgate J. e Rose W. (a cura di) (2011), *Improving Outcomes for Children and Families: Finding and Using International Evidence*, Jessica Kingsley, London and Philadelphia.
- Manz C.C. e Neck C.P. (1995), *Teamthink: Beyond the Groupthink Syndrome in Self-Managing Work Teams*, in «Journal of Managerial Psychology», 10, 1, pp. 7-15.
- The Good analyst (2012), *Methodology for impact analysis and assessment (MLAA) rating*, in «Scope Impact Report» Nov. 2012, www.goodanalyst.com.
- The SROI Network Accounting for Value, Human Foundation (2012), *Guida al ritorno sull'investimento sociale*, Human Foundation, Roma.
- Vecchiato T. (2013), *Obiettivi ed esiti misurabili*, in «Studi Zancan», 2, pp. 5-16.
- Vecchiato T. (2014a), *Verso nuovi diritti e doveri sociali: la sfida del welfare generativo*, in «Diritto e Società», 1, pp. 153-161.
- Vecchiato T. (2014b), *IM indicators and best practice in H&SC (Health and Social Care)*, in *La finanza che include: gli investimenti ad impatto sociale per una nuova economia. Rapporto Italiano della Social Impact Investment Task Force istituita in ambito G8, Presidenza del Consiglio dei Ministri*, Roma.



SUMMARY

The challenge of outcomes can not be avoided. The technical difference between outcome and impact is substantial: outcome represents the direct benefit for beneficiaries, whereas impact is the benefit extending towards the entire community. The evaluation of social impact is required to investigate the available potential so as to better identify how much welfare services manage to redistribute the common good. Outcome and social impact are issues of relevant social interest, which can not be confined to the sphere of private evaluation, since solutions to human problems are «common good» to be shared on a greater scale.