

# UNA CRITICA DELLA VALUTAZIONE REALISTA<sup>1</sup>

Alberto Martini

2009

Questo capitolo è dedicato ad un approccio che si pone esplicitamente in antitesi alla logica controfattuale. Si tratta della *realist* (o *realistic*) *evaluation*, termine che traduciamo con “valutazione realista”. Di essa sono proponenti ed esponenti principali due studiosi inglesi, Ray Pawson e Nick Tilley, coautori di una monografia-manifesto pubblicata nel 1997, *Realistic Evaluation*, e di numerosi saggi e articoli negli anni successivi [Pawson 2002; 2006; Pawson e Tilley 1997; 2004].

Al lettore che voglia approfondire questo approccio consigliamo di fare riferimento ai testi originali, in quanto la trattazione di questo capitolo non rende giustizia della complessità analitica e dello spessore epistemologico delle argomentazioni dei due autori (né del loro inconfondibile stile espositivo, ricco di metafore e riferimenti arguti). Le prossime pagine si limitano ad una esposizione commentata della critica realista al metodo sperimentale e alla logica controfattuale.

I due autori sono animati da una forte carica antagonista nei confronti di altri approcci, ma in particolare del metodo sperimentale. Nella monografia del 1997 affermano, senza mezzi termini:

*For us, the experimental paradigm constitutes a heroic failure* [Pawson e Tilley 1997, 8]

Emblematicamente, il 1997 è anche l'anno in cui il governo federale messicano lancia la valutazione sperimentale di *PROGRESA*, cioè il più vasto studio randomizzato di un intervento di politica sociale mai condotto al di fuori degli Stati Uniti. Nel 1997 si insedia nel Regno Unito il governo Blair, che apre le porte ad un massiccio uso di *pilot studies* per sperimentare le proprie politiche sociali. Il decennio successivo al 1997 vede il diffondersi senza precedenti dell'utilizzo del metodo sperimentale, anche al di fuori degli Stati Uniti, a cui era stato tradizionalmente confinato. Per nulla intimidito, anzi forse stimolato da questo successo, Pawson nel 2006 ne ribadisce la “futilità”:

*... using randomized controlled trials as the lead investigatory strategy. ... I will return to the futility of this approach in the next chapter* [Pawson 2006, 18].

Perché tanta avversione? La prima critica agli “sperimentalisti” è quella di prestare *ottusamente* attenzione solo al “se” una politica “funziona”, invece di capirne il “perché”:

---

<sup>1</sup> Capitolo eliminato per ragioni di spazio da “Valutare il successo delle politiche pubbliche” Il Mulino, Bologna, 2009

*Experimentalists have pursued too single-mindedly the question of whether a program works, at the expense of knowing why it works [Pawson e Tilley 1997, XV].*

L'avversione riguarda tutta la logica contro fattuale:

*The realist synthesis sees no opportunity to ponder counterfactual questions [Pawson 2006, 180].*

Ben altre, secondo i due autori, sono le domande che devono porsi i valutatori:

*Realist evaluation asks not, 'What works?' or 'Does this program work?' but asks instead, 'What works for whom in what circumstances and in what respects, and how?' [Pawson e Tilley 2004, 1].*

Quindi la valutazione realista non si chiede "Quale soluzione funziona?" e neppure "Questo intervento funziona?", ma si chiede invece "Quale soluzione funziona per chi, in quali circostanze, in che senso, e come?". Tutto condivisibile, con una cautela: *porsi* domande più sfidanti e ambiziose è abbastanza facile, ma non è garanzia sufficiente di *trovare* risposte plausibili.

Con una leggera punta di snobismo, la valutazione realista si assegna obiettivi cognitivi decisamente superiori a quelli perseguiti da altri, segnatamente dagli economisti, che si accontentano della "semplice" quantificazione degli effetti di una politica:

*Realist evaluation is chary about serving up those stable net effects conclusions that are understandably beloved by economic modellers anxious to help wrest greatest utility from our precious and finite resources [Pawson e Tilley 2004, 22]*

Traducendo liberamente: "la valutazione realista è riluttante a produrre quelle stime di effetti netti predilette dagli economisti, così ansiosi di aiutare ad ottenere il massimo dalle nostre preziose e scarse risorse".

Queste citazioni, rappresentative dello stile caustico dei due autori, sottolineano l'*antagonismo* tra due obiettivi cognitivi diversi, la quantificazione degli effetti e l'individuazione dei meccanismi che possono spiegare gli effetti. Che siano obiettivi diversi è fuori di dubbio: che siano antagonisti, e quindi che perseguire il primo sia di *ostacolo* a perseguire il secondo, è tesi che non condividiamo.

## 1. EFFETTI DI CAUSE E CAUSE DI EFFETTI

---

Paul Holland, nel suo articolo del 1986 sull'inferenza causale, introduce la distinzione tra *effetti di cause* e *cause di effetti*. Sono i primi l'oggetto dell'inferenza causale: *data una causa potenziale* (nel nostro caso un trattamento, una politica pubblica) *come si stabilisce se essa ha prodotto effetti?* (che essi siano desiderati o meno, attesi o meno non fa una differenza sostanziale: quello che conta è determinare se da una certa azione sono scaturite conseguenze ad essa imputabili in senso causale).

Il secondo termine della distinzione, *cause di effetti*, sottintende una problematica conoscitiva affatto diversa: *dato un certo effetto, realizzato o atteso, quali ne sono le possibili cause, intese nel senso di meccanismi causali?*

I metodi illustrati nei capitoli precedenti sono concepiti esclusivamente per dare una risposta al primo problema conoscitivo, la quantificazione degli effetti, con lo scopo di

contribuire a stabilire se una politica pubblica ha avuto successo o meno. Riteniamo questo un obiettivo fondamentale per la valutazione delle politiche, se non altro alla luce dell'esigenza di rendere conto dell'utilizzo delle risorse pubbliche. Stabilire segno e dimensione degli effetti di una politica non implica però negare l'importanza del secondo problema conoscitivo, *capire perché* una certa politica funziona oppure non funziona. Questo è senza dubbio l'obiettivo ultimo della valutazione, a cui però si può arrivare gradualmente, con un processo cumulativo, partendo dalle risposte date al primo quesito.

La nostra tesi è che sia sensato perseguire una *sequenza logica di interrogativi* riguardo alla valutazione degli effetti di una singola politica pubblica. Prima è indispensabile affrontare e rispondere all'interrogativo di base: possiamo *attribuire* alla politica attuata i cambiamenti osservati nel fenomeno di interesse? Questo è già di per sé un obiettivo conoscitivo formidabile, come abbiamo visto nei capitoli precedenti, e riteniamo vada perseguito per primo in ordine logico. Se non si sa se un effetto c'è oppure no, dubitiamo sia utile (o possibile) cercarne spiegazioni; non crediamo sia sensato interrogarsi sul *perché* di qualcosa di cui non si conosce neppure *il segno*.

### 1.1 La variabilità degli effetti

Una volta stabilito che un certo intervento "funziona" in media, restano senza dubbio molte questioni aperte. Anzi, è la risposta stessa al primo quesito a mettere meglio a fuoco le domande successive. Una riguarda il "per chi e in quali contesti" l'intervento funziona. Questo è un problema che è ben presente a chi si occupa di analisi controfattuale, e su questo i valutatori realisti non dicono nulla di sostanzialmente nuovo, al di là dell'enfasi polemica con cui pongono le questione.

Chiunque abbia stimato l'effetto di una politica si è posto il problema di come l'effetto varia a seconda delle caratteristiche dei beneficiari e del particolare contesto (geografico, organizzativo) in cui l'intervento di trova ad operare. L'ostacolo alla produzione di questa conoscenza dettagliata non è l'insipienza degli sperimentalisti, come sostengono i realisti, ma è molto più banale (e difficile da superare): per poter stimare effetti disaggregati occorrono molte osservazioni, cioè campioni molto grandi, di modo da avere una base empirica sufficiente a cogliere con la necessaria precisione gli effetti per i sottogruppi. Occorre inoltre che le caratteristiche dei contesti non siano troppo correlate con le caratteristiche dell'intervento, in modo da poterne isolare l'effetto [Bloom, Hill e Riccio 2005].

Quindi nulla di particolarmente originale nell'affermare l'importanza di capire "cosa funziona per chi e in quale contesto". Il problema dell'eterogeneità degli effetti è anzi al centro dell'elaborazione metodologica dell'analisi controfattuale<sup>2</sup>.

### 1.2 Una tesi radicale

Il vero "stacco" tra analisi controfattuale e valutazione realista riguarda la comprensione dei meccanismi che generano gli effetti, un obiettivo enormemente più sfidante dei precedenti. Obiettivo che a nostro parere andrebbe perseguito in modo selettivo e mirato, poggiandosi sull'evidenza via via accumulata da analisi più semplici. Questo non trova

---

<sup>2</sup> Si veda ad esempio Heckman, Smith e Clements [1997] e Bloom, Hill e Riccio [2005].

d'accordo i proponenti della valutazione realista, che sostengono una tesi molto radicale: *la totale inutilità di ogni analisi che si limiti a fare attribuzioni di causalità senza partire dai meccanismi che spiegano gli effetti.*

Le motivazioni apparenti di questa critica radicale e distruttiva sono molteplici: in parte si tratta di argomentazioni di natura epistemologica, riguardanti il concetto stesso di causalità; in parte si tratta del modo di affrontare il tema della complessità della realtà sociale, a cui i due autori portano una sensibilità schiettamente sociologica; in parte si tratta di un giudizio contingente che i due autori danno, nella monografia del 1997, di un filone particolare di studi sperimentali condotti nel campo della prevenzione del crimine e della riabilitazione dei detenuti. Si tratta di un campione molto parziale e poco rappresentativo della valutazione controfattuale, in quanto ignora quasi completamente l'esperienza statunitense. In generale, i due autori forniscono, consapevolmente o meno, una versione un po' caricaturale della sperimentazione sociale e soprattutto sembrano ignorare i suoi più significativi sviluppi metodologici. Ad esempio, la critica secondo cui l'esperimento sia una *black box* di cui non si capisce il funzionamento è ben presente alla comunità che pratica la sperimentazione sociale. Valga per tutte la seguente citazione di Howard Bloom, uno dei principali metodologi della randomizzazione.

*This limitation of the experimental paradigm is often referred to as the 'black box' problem. Simply put, the problem is that experiments are good at documenting the linkages, or the lack thereof, between an intervention (the input to the black box) and the outcomes (the output of the black box), but they provide little or no direct information about why the intervention did or did not affect the outcome [Bloom et al. 2005, 77].*

Tralasciamo quindi la critica un po' gratuita e disinformata al metodo sperimentale ed esaminiamo invece i primi due elementi cardine di differenziazione della valutazione realista rispetto alla logica controfattuale: l'enfasi sulla complessità dei programmi sociali e la diversa idea di causalità.

## 2. COMPLESSITÀ E CAUSALITÀ

---

Al cuore della logica controfattuale sta il tentativo di semplificare un problema conoscitivo complesso, quello dell'attribuzione causale. Parte importante di questo tentativo di semplificazione sta nell'idea di considerare la politica pubblica come "trattamento" rivolto a determinati "soggetti". Nessuno pensa realmente che le cose siano così semplici, si tratta solamente di un tentativo di "modellare" la realtà per mettere in evidenza ciò che è essenziale.

### 2.1 L'elogio della complessità

Questa strategia viene rifiutata in blocco dalla valutazione realista, che argomenta esattamente l'opposto:

*In describing social interventions I have striven to convince the reader that they are not 'treatments' aimed at 'subjects'. Rather, they are complex systems thrust amidst complex systems [Pawson 2006, 168].*

E in modo più caustico:

*By now it should be clear that realist evaluation has little use for a 'find-the-intervention-X-that-cures-problem-Y' notion of programme building [Pawson e Tilley 2004, 8].*

La sostanza del discorso è la seguente: i programmi sono frutto di complesse e sofisticate interazioni sociali; la complessità va esaltata come opportunità, come elemento fondante; la complessità richiede una particolare sensibilità per essere decifrata; solo i sociologi hanno questa sensibilità.

*Realists regard programmes as rather sophisticated social interactions set amidst a complex social reality.*

*Realism carries a profoundly sociological view on social change.*

*Realist evaluation has a distinctive account of the nature of programmes and how they work, of what is involved in explaining and understanding programmes, of the research methods that are needed to understand the workings of programmes, and of the proper products of evaluation research [Pawson e Tilley 2004, passim].*

## 2.2 I programmi come teorie

Dal discorso sulla complessità dei programmi sociali, che a tratti rischia la tautologia ("complesso è ciò che è complesso"), emerge un'argomentazione centrale all'approccio realista, quello dei "programmi come teorie da testare", che lo avvicina al filone della *theory-based evaluation* [Chen 1990; Weiss 1997]:

*The cornerstone of the realist project is a distinctive viewpoint on how intervention bring about change. It is only by understanding and probing its apparatus of change that one can evaluate a programme. According to realist evaluation programmes are 'theories', they are 'embedded', they are 'active', and they are parts of 'open systems' [Pawson e Tilley 2004, 2].*

*Interventions are theories. This is the most fundamental realist claim about interventions. A more conventional perspective sees interventions in more tangible terms such as collection of resources, equipment and personnel but, for the realist, such resources are theories incarnate [Pawson 2006, 26].*

Messa a confronto con tanta passione, la logica controfattuale emerge in tutta la sua ottusa freddezza. Quindi quella che nell'approccio controfattuale è la "strategia di identificazione", che elimina i fattori di confondimento per isolare l'effetto della variabile-trattamento T sulla variabile-risultato Y, per i realisti diventa sinonimo di perdita di informazione e quindi della capacità di spiegare la realtà:

*Experimental evaluation is also  $X \Rightarrow Y$  fixated and reduces the business of explanation to drawing causal inferences between variables... the goal is to establish a relationships between programs (X) and outcome (Y). Any other process, action, idea, disposition has to be treated as a potential 'confounding variable' (Z) and experimental controls are put in place in an attempt to rule out their action [Pawson e Tilley 1997, 69].*

Nell'ambito dell'inferenza causale, i fattori di confondimento sono responsabili di attribuzioni spurie, per questo è importante eliminarne l'influenza. Tuttavia è semplicemente falso che qualsiasi altra informazione, "process, action, idea, disposition", sia

un fattore di confondimento. Anzi, ogni strategia di identificazione trae la sua plausibilità da una conoscenza quanto più approfondita del processo di selezione. La critica realista su questo punto diventa caricaturale. Ma veniamo ad un punto cruciale:

*The key point is that this way of perceiving causation, whether it is achieved by statistical control (as in multivariate analysis techniques) or by use of control groups (as in experimental techniques), overlooks a vital explanatory ingredient. To repeat, that ingredient is not a variable but a mechanism. Social interventions only and always work through the action of mechanisms, through a process of weaving resources and reasoning together. Without this being the first item on the research agenda, all subsequent work on program outcomes will remain a mystery [Pawson e Tilley 1997, 69].*

Secondo i due autori, isolando il contributo netto della variabile-trattamento sulla variabile-risultato, si perde di vista un ingrediente fondamentale, il “meccanismo”. Se i meccanismi non sono il primo elemento dell’agenda di ricerca, tutto il lavoro di analisi degli effetti è inutile. L’idea che ci possano essere due fasi, di cui la seconda è un approfondimento della prima, non sfiora i valutatori realisti: o si parte dai meccanismi, o si perde tempo. Ogni tentativo di “semplice” attribuzione causale è inutile:

*The caution of realist evaluation applies, inevitably, in respect of the issue of ‘attribution’. In its classic guise, this refers to the problem of how and how safely we may infer that a programme is responsible for outcomes observed. Realism rejects this particular formulation arguing that that programmes are active, and thus it is the operation of particular mechanisms acting in context that brings about change [Pawson e Tilley 2004, 15].*

### **2.3 Due idee di causalità**

Cruciale è il riferimento fatto in precedenza alla diversa “way of perceiving causation” che differenzerebbe i due approcci. I realisti riprendono una distinzione classica nella filosofia della scienza, tra due modi fondamentalmente diversi di concepire la causalità:

- la *successionist logic*, che parte dal presupposto che una “causa” non può essere osservata, ma la relazione di causa-effetto può essere “inferita” sulla base dell’osservazione di regolarità empiriche o della loro produzione mediante “manipolazione”;
- la *generative logic*, che al contrario stabilisce relazioni di causa-effetto solo individuando i meccanismi che generano gli effetti.

Queste due concezioni di causalità hanno radici filosofiche molto antiche, la prima da alcuni viene fatta risalire ad Eraclito (ma è principalmente associata a Hume), la seconda ad Aristotele. Sono il prodotto di secoli di dibattito epistemologico, e ci lascia perplessi scegliere tra di loro come se fosse la scelta della squadra del cuore, o l’uno o l’altro. La ricerca bio-medica è un esempio dell’uso congiunto delle due logiche, una basata sulla scoperta di meccanismi e l’altra sull’osservazione di regolarità empiriche. Lo sviluppo di nuovi farmaci si basa tanto sull’identificazione dei meccanismi biologici che spiegano perché una certa molecola produce certe alterazioni, quanto sul ricorso alla sperimentazione clinica, che si basa non su meccanismi ma su inferenza derivante da regolarità empiriche. E queste regolarità empiriche sono interpretabili come “causali”

proprio perché sono stati eliminati i fattori di confondimento mediante l'utilizzo della randomizzazione.

Diremo di più: è proprio la "complessità" dei processi sociali a indurre talvolta all'uso della sperimentazione clinica per verificare se un certo farmaco funziona. Se gli individui fossero perfettamente programmabili e prevedibili nelle loro reazioni, basterebbe aver individuato i principi attivi di un farmaco in laboratorio per predire l'efficacia del farmaco una volta applicato su larga scala. Ma l'applicazione su larga scala porta con sé elementi di complessità e incertezza: non tutti i pazienti prendono il farmaco come prescritto; non tutti vivono nel contesto ideale; alcuni assumono altre sostanze o cibi che possono interferire con il farmaco che si sta sperimentando. Per questo i nuovi farmaci vengono testati non solo in laboratorio, ma mediante studi randomizzati su popolazioni reali. Proprio per incorporare la complessità e non credere solamente ai meccanismi biologici.

I realisti traggono da tutto ciò la lezione opposta: l'uso della randomizzazione per la valutazione degli interventi pubblici è sorpassata, mentre è la valutazione realista ad aver ereditato il potenziale esplicativo delle scienze naturali. Nelle parole di Tilley:

*Realists have a different account of causality. The realist understands causality in terms of underlying causal mechanisms generating regularities. Natural science is replete with the investigation of mechanisms explaining observed patterns. Realists understand experiments in the natural sciences as the creation of conditions in which mechanisms will be activated. In laboratories scientists create artificial conditions in which those causal mechanisms which they conjecture to exist will be activated. (...) Realistic evaluation is simply an application of this insight to the examination of social programmes. Its concern is with understanding causal mechanisms and the conditions under which they are activated to produce specific outcomes. Pawson and I thus believe that the kind of evaluation that we are advocating is strictly consonant with the ways in which science is done. Natural scientists rarely do experiments in the ways in which experimental evaluation is conventionally considered [Tilley 2000, 10].*

### **3. MECCANISMI, CONTESTI, RISULTATI: LA METODOLOGIA REALISTA IN BREVE**

---

Vediamo ora i punti salienti della metodologia che i realisti propongono in alternativa alla aborrita analisi contro fattuale.

Si tratta di un processo iterativo che comincia con l'elaborazione di una "teoria" dell'intervento, fondata sulla triade "meccanismi - contesto - risultati attesi", da cui si derivano ipotesi specifiche, del tipo "quell'azione può raggiungere quel risultato per quei soggetti in quel contesto".

*Realist evaluation stresses four key linked concepts for explaining and understanding programmes: 'mechanism', 'context', 'outcome pattern', and 'context-mechanism-outcome pattern configuration' [Pawson e Tilley 2004, 5].*

I meccanismi sono "ciò che produce gli effetti":

*Mechanisms describe what it is about programmes and interventions that bring about any effects. Mechanisms thus explicate the logic of an intervention; they trace the destiny of a*

*programme theory, they pinpoint the ways in which the resources on offer may permeate into the reasoning of the subjects [Pawson e Tilley 2004, 6].*

Anche il contesto assume un ruolo particolare di supporto o di ostacolo all'agire dei meccanismi:

*For realism, it is axiomatic that certain contexts will be supportive to the programme theory and some will not. And this gives realist evaluation the crucial task of sorting the one from the other [Pawson e Tilley 2004, 6].*

I risultati sono le conseguenze volute e non volute di meccanismi che operano nei diversi contesti:

*Outcome-patterns comprise the intended and unintended consequences of programmes, resulting from the activation of different mechanisms in different contexts [Pawson e Tilley 2004, 7].*

Fin qui è difficile essere in disaccordo. Il problema è la mole enorme di ipotesi che scaturiscono dai molteplici meccanismi che operano in contesti diversi e generano risultati di ogni tipo. Come si domina l'esplosione del numero di tali combinazioni? Senza ovviamente ridurre tutto alla vituperata "relazione tra variabili".

*Outcome patterns can take many forms and programmes should be tested against a range of output and outcome measures. Much is to be learned by monitoring programmes across a range of such measures. We may find an influence at point A or in respect of characteristic B. But no change may be discernible at time C or in relation to property D. Then again, we may find a quite unexpected movement at E and an unwanted outcome at F. Deciphering the reasons for such a variegated pattern can give us vital clues to the workings of programmes [Pawson e Tilley 2004, 7].*

Che è come dire: tutto può scaturire dalle infinite configurazioni possibili di meccanismi-contesti-risultati. Come se ne viene a capo? Come si garantisce che tutto ciò ci dia effettivamente "vital clues to the workings of programmes" e non ci lasci con una gran confusione?

*In general, there are an infinite number of explanations for why, when and how a programme works but there are only so many ways in which a programme might be improved. Realist evaluators need not wait to figure out the totality of explanations but should concentrate, therefore, on those programme ideas and variations, which are 'on the table' [Pawson e Tilley 2004, 16].*

Delle infinite possibili spiegazioni del perché un intervento funziona occorre concentrarsi su quelle che sono "sul tavolo". Ma questo rappresenta (una rara ammissione!) un limite dell'approccio realista.

*And this conundrum spells out the ineluctable limitation of realist evaluation. Programmes modify and transmute and are constantly being tried out in fresh circumstances, so there are always more and new potential explanations of their efficacy [Pawson e Tilley 2004, 16].*

Ci sia consentita la battuta: "chi di complessità ferisce, di complessità perisce". Alla fine tutto si risolve con il semplice buon senso del "fa quello che puoi" nel distinguere tra le spiegazioni alternative.

*At this point, realist evaluation resorts to the pragmatism principle, which says simply - go as far as you can in sorting and sifting the rival explanations [Pawson e Tilley 2004, 16].*

E scopriamo che il problema dell'attribuzione causale, che sta al centro della logica contro fattuale, non è del tutto risolto:

*All well and good, but one still has the problem of attribution. From the preceding examples it is clear that a programme may operate through many different mechanisms. So, how can we be sure that a particular mechanism or set of mechanisms is in fact responsible for bringing about change in a given context? [Pawson e Tilley 2004, 15].*

I meccanismi sono importanti, ma un dato intervento può operare attraverso vari meccanismi. Come possiamo stabilire che un dato meccanismo è responsabile per avere prodotto un certo cambiamento? Il lettore di questo manuale può pensare a qualche strategia di identificazione. Il realista invece si accontenta che il meccanismo "dia un senso" al risultato osservato:

*Thus, attribution is dealt with when we accept that action of a mechanism makes sense of the particular outcome pattern observed. Now, **sense-making** is not the platform on which other traditions of evaluation are built but it is the **cornerstone of the realist approach** [Pawson e Tilley 2004, 16].*

Si tratta di paradigmi culturali chiaramente diversi: quello controfattuale, basato su un rigido impianto di metodi quantitativi, e quello realista, in cui il *sense-making* è la pietra angolare dell'intero edificio concettuale.

### **3.1 Le fasi del disegno realista**

Le descrizioni più dettagliate delle singole fasi del disegno realista lasciano un'impressione di entusiastica dedizione, tenuta assieme da un filo logico molto tenue. Il primo passo è sempre la formulazione di una teoria del programma, a cui tutto può (giustamente) contribuire, ma una buona dose di immaginazione (sociologica?) è cruciale.

*Realist evaluation normally begins by eliciting and formalising the programme theories to be tested. There can be various sources of these including documents, programme architects, practitioners, previous evaluation studies and social science literature. Hence documentary analysis, interviews and library searches may all be involved. Interviews with practitioners are deemed especially important as discussions of apparent programme successes and failures can lead to fine-grained hypotheses about what works for whom and in what circumstances and respects. This stage is the launching pad for realist evaluation and is, in many ways, its most distinctive phase. What is involved is bringing the imagination to bear in 'thinking though' how a programme works [Pawson e Tilley 2004, 10].*

Ma l'immaginazione non basta, occorrono i dati per testare le ipotesi. I dati possono essere di ogni tipo e tratti da ogni possibile fonte:

*The next stage is to collect data that will allow interrogation of these embryonic hypotheses. Programme documents and/or practitioners will have suggested that a particular resource is vital, and that a particular way of interpreting it is the key to success. The initial hypotheses will also often cover the type of client who is better placed to succeed and some institutional locations where this is most likely to happen. Data gathering has the task of trying to match information to these various leads. [Pawson e Tilley 2004, 10].*

Il valutatore deve avventarsi su ogni tipo di dati: dati amministrativi, interviste di ogni tipo, confronti pre-post (solamente?), *focus group*, e così via.

*The evaluator has, quite literally, to scavenge for the best data to test out the theories. Existing administrative records might be put to use, stakeholders of all type might be interviewed and shadowed, dedicated before-and-after measures might be designed and put in place, focus groups might be assembled to unearth reasons for crucial choices, and so on [Pawson e Tilley 2004, 10].*

La terza fase comporta testare le ipotesi mediante i dati raccolti. Questo può essere fatto in ogni modo possibile:

*The third stage is to subject a whole package of CMOC hypotheses to systematic test, using data sets assembled as above. The purpose of the analysis is to see if the model will explain the complex footprint of outcomes left by the programme. There is no single analytic method suitable for this purpose and the design of data analysis is a matter of the subtlety of the proposed theories and the availability of data [Pawson e Tilley 2004, 10].*

L'unica tattica consigliata è fare confronti tra sottogruppi, identificando "vincitori" e "perdenti" tra i beneficiari e i "pro" e i "contro" nell'erogazione dei servizi.

*The primary tactic is thus to interrogate these hypotheses by making sub-group comparisons. Overall, the explanatory theory is investigated by devising and testing out multiple comparisons identifying winners and losers amongst subjects and pros and cons in programme delivery [Pawson e Tilley 2004, 10].*

La domanda che si porrebbe Pawson stesso se non fosse parte in causa è la seguente: "Riusciranno i nostri eroi a "dare senso" a tutte le possibili variazioni osservate?"

*Will the emerging theory explain implementation variations, impact variations, socio-demographic sub-group variations, temporal outcome variations, personal attribute outcome variations, biological make-up outcome variations and so on? [Pawson e Tilley 2004, 10].*

Siamo allo stadio finale, l'interpretazione dell'analisi. Le teorie di come il programma funziona sono confermate o confutate dall'analisi precedente? In qualche caso sì, in qualche caso no, conclusione non inaspettata.

*The final stage is the assessment and interpretation of the analysis. Have the theories about how the programme worked been supported or refuted by the proceeding analysis? Judgement on this score is invariably mixed, with some output and outcome variations being clear and intelligible, whilst others remain quite puzzling [Pawson e Tilley 2004, 10].*

#### 4. UNA PROSPETTIVA AFFASCINANTE, MA QUANTO REALISTICA?

---

La valutazione realista rischia di finire schiacciata sotto il peso della sua stessa ambizione intellettuale. Sembrano riconoscerlo gli stessi proponenti, quando affermano:

*We have also to acknowledge important shortcomings in realist evaluation. It is intellectually enormously challenging [Pawson e Tilley 2004, 22].*

Infatti:

*Realist evaluation requires sustained thinking and imagination to work through programme theory, to define expected outcome patterns, and to figure out exactly what footprints or data signatures to look for and where to find them to test or arbitrate between theories. None of this is easy. It requires advanced theoretical understanding, abilities to design research and techniques for data analysis [Pawson e Tilley 2004, 22].*

La nostra “valutazione” conclusiva della valutazione realista è un misto di sincero apprezzamento e di altrettanto sincero scetticismo. Apprezzamento per i formidabili obiettivi cognitivi che questo approccio si pone, per l’abilità argomentativa e la profonda cultura dei due autori, e per aver attirato l’attenzione su una problematica che tendeva ad essere trascurata, il “perché” le politiche funzionano o non funzionano.

Restiamo invece sinceramente scettici sulla possibilità di ottenere, con i metodi proposti dalla valutazione realista, risultati così convincenti da essere in grado di soppiantare completamente la logica contro fattuale. Non basta farsi domande ambiziose, bisogna fornire risposte plausibili. L’approccio controfattuale pone una semplice domanda ma schiera un apparato poderoso di metodi per trovare risposte credibili. La valutazione realista fa l’opposto: parte con una salva di domande ambiziose, ma poi si riduce ad una serie di raccomandazioni di buon senso: formula una teoria, trova i dati, testa le teoria, affina la teoria.

Restiamo scettici sulla fattibilità di scoprire tutti i meccanismi causali che legano politiche e fenomeni sociali. Lunga è ancora la strada di una tale comprensione: nel frattempo, l’approccio controfattuale rimane a nostro parere la scelta migliore per rispondere alla domanda “*what works*”. Una domanda che è ben lungi dal trovare una risposta nella gran parte dei casi: ci pare francamente prematuro metterla da parte e con essa il formidabile apparato metodologico costruito con il contributo di una vasta comunità internazionale. Ci parrebbe più sensato un atteggiamento più cauto e modesto da parte dei valutatori realisti, che si ponga in un’ottica complementare e non antagonista.

Anche a voler aderire senza riserve a questo paradigma, va riconosciuto lo scarso realismo di una sua generalizzazione al complesso delle politiche pubbliche. Ma su questo i due autori invece non hanno dubbi e affermano con grande immodestia:

*The realist approach can reach every corner of evaluation methodology from design to data collection to analysis to cumulation. ... We believe that realistic evaluation can reach every corner of the programming and policy field [Pawson e Tilley 1997, XVII].*

Concludiamo queste conclusioni con una citazione di due autori americani, anch’essi critici dell’eccessiva fede nel metodo sperimentale:

*This is not to say that valid conclusions can never be drawn from experimental studies of ‘what works’, but rather that those who do such research need to cultivate **skepticism and humility**—a realistic assessment of the difficulties and limits that are inherent in their approach [Erickson e Gutierrez 2002, 24].*

Questa è un’affermazione che ci sentiamo di sottoscrivere appieno.

## BIBLIOGRAFIA

---

- Bloom H.S., Richburg-Hayes L., and Black A.R. [2005], *Using Covariates to Improve Precision*, New York, NY, MDRC Working Papers on Research Methodology, 2005.
- Holland P. [1986], *Statistics and Causal Inference*, in «Journal of the American Statistical Association», 1986, n. 81, pp. 945-960.
- Chen H.T. [1990], *Theory-driven Evaluations*, Beverly Hills, CA, Sage Publications, 1990.
- Pawson R. [2002], *Theorizing the Interview*, in Fielding N. et al (eds.), *Interviewing*, Beverly Hills, CA, Sage, 2002.
- Pawson R. [2006], *Evidence-based Policy: a Realistic Perspective*, London, UK, Sage, 2006.
- Pawson R., Tilley N. [1997], *Realistic Evaluation*, London, Sage Publications, 1997.
- Pawson R., Tilley N. [2004], *Realistic Evaluation*, in Matthieson S. (ed) *Encyclopaedia of Evaluation*, Beverly Hills, CA, Sage, 2004.
- Tilley N., [2000], *Experimentation and criminal justice policies in the UK*, in «Crime and Delinquency», Vol. 46, N. 2, 2000.
- Weiss C., [1997], *Evaluation*, Upper Saddle River, NJ, Prentice Hall Inc., 1997.