

Dimensioni della valutazione: criteri per la costruzione di uno strumento

Premessa

Il tema assegnatomi è piuttosto complesso e ampio, e il tempo concesso alla comunicazione orale piuttosto breve. Mi scuso per le inevitabili semplificazioni che dovrò fare; anche se in questo testo scritto avrò maggiore possibilità di spiegarmi, limiterò comunque la scrittura facendo riferimento a testi dove le materie affrontate sono dibattute con maggior approfondimento.

Alcuni equivoci sulla tecnica valutativa e i suoi strumenti

Come scrive Nicoletta Stame, guardando la recente legislazione “l’Italia sembrerebbe il paese di Bengodi della valutazione, salvo poi subito accorgersi che tali leggi [Bassanini, 286/99, ecc.] non sono applicate”². La valutazione è penetrata nelle abitudini e nel lessico amministrativo italiano, e un modulo valutativo non manca in qualunque formulario per progetto a bando, ma sulle conseguenze concrete sul piano della qualità valutativa è lecito avere qualche dubbio³.

A fronte di un interesse spesso superficiale o di facciata delle Amministrazioni pubbliche, si contrappone una competenza valutativa sempre più diffusa, ma non particolarmente brillante. In Italia, malgrado la maggiore aspettativa che avevamo una decina di anni fa, la qualità valutativa complessiva (teorica e tecnica) non si è elevata quanto si sarebbe auspicato⁴. La valutazione si limita spesso all’assemblaggio di alcune tecniche di ricerca, piuttosto semplici e solitamente interpretate in modo piuttosto piatto e banale (ma semmai nobilitate da analisi statistiche sofisticatissime). Ricerche valutative basate su questionari improvvisati, su pochi focus group (o presunti tali⁵), su manciate di oscuri indicatori, sono prodotte in numero che mi appare assolutamente superiore a quelle più articolate e giustificate metodologicamente⁶. Le ragioni sono molteplici, e mi limito qui a un breve sommario:

1. Una dimenticanza a monte: la teoria.

C’è ormai molta produzione di valutazione, in Europa e in Italia, che mostra una corsa verso il tecnicismo: si applica l’equazione “valutazione = tecniche valutative”, ignorando la necessità di una “teoria della tecnica”, che giustifichi ogni scelta operativa che – dato un contesto e un mandato – conduca verso l’espressione del giudizio valutativo.

E’ piuttosto facile avere una controprova empirica di questa affermazione: è sufficiente interrogarsi – leggendo un rapporto valutativo – sulla ragione delle scelte delle diverse tecniche e procedure; quasi di regola manca – nel testo – ogni giustificazione metodologica seria, e si salta a piè pari dalle descrizioni preliminari di contesto alla presentazione dei risultati, ottenuti

¹ Claudio Bezzi, valutatore, Icroodus Srl. bezzi@valutazione.it.

² Nicoletta Stame, *La valutazione nella P.A.*, in “Valutazione 2000. Esperienze e riflessioni”, a cura di Mauro Palumbo, Franco Angeli, Milano 2000; su questo punto si veda anche il saggio di Giuseppe Mele, *La valutazione e la riforma della Pubblica Amministrazione. Gli spazi e le prospettive applicative nell’analisi dell’azione amministrativa*, in Mauro Palumbo (a cura di), *Valutazione 2000. Esperienze e riflessioni*, cit., ancora più severo nel giudizio.

³ Una riflessione completa e condivisibile in pieno in Mita Marra, *Alcuni aspetti e problemi della valutazione nella P.A.*, “Rassegna Italiana di Valutazione”, a. VIII, n. 29, 2004.

⁴ Una riflessione in merito in Claudio Bezzi, *Dieci anni di RIV*, “Rassegna Italiana di Valutazione”, n. 33, 2005.

⁵ Che i focus group – tecnica peraltro sovrastimata in Italia – siano spesso solo “presunti”, è stato da me discusso in *Il Focus group: consideriamone i limiti*, “Sociologia e Ricerca Sociale” n. 76/77, 2005.

⁶ Cfr. Claudio Bezzi, *Cosa valutiamo, esattamente?*, “Rassegna Italiana di Valutazione”, n. 27, 2004.

in un qualche modo appena accennato. In effetti, negli ultimi anni, appaiono a volte delle cornici giustificative che si rifanno alle teorie più in voga e dibattute al momento (in questi anni in particolare la *Realistic Evaluation* e la *Theory-based Evaluation*), ma spesso questi riferimenti teorici sembrano più una concessione al lettore erudito che non una necessità concreta con ricadute empiriche specifiche, e dopo tali riferimenti si illustrano pratiche di ricerca sostanzialmente scollegate, quando non in contraddizione, con quanto prima declamato. Non collegare l'approccio empirico a una teoria ha delle ricadute abbastanza serie in tutto il lavoro valutativo; significa infatti che si ritengono *neutrali* le tecniche, e *indifferenti*, e non invece sistematicamente soggette a limiti specifici che sono accettabili solo se controllati. Tale "controllo" ha a che fare solo con una conoscenza in merito al tipo di informazione che si rileva in quel modo, e alla sua pertinenza entro certi obiettivi cognitivi sostenibili alla luce di finalità più generali, che hanno appunto a che fare con la teoria. Esempio immaginario per capirsi: se utilizzo una determinata matrice multicriteriale costruita a partire da determinati criteri, forse la logica che scaturisce è pertinente con un approccio "realista" (à la Pawson) capace di rivelarmi qualcosa di pertinente sui meccanismi; non so se quello che ho detto – giusto per esemplificare – sia vero o no, ma se non sono in grado di stabilire questo genere di connessioni, allora *perché avrei scelto di costruire una matrice multicriteriale sulla base di quei determinati criteri, anziché fare i soliti tre focus group?*

2. Un equivoco a valle: la metodologia.

Consequentemente si equivocano "metodo", "metodologia", "tecniche" e "strumenti", e i limiti intrinseci in ogni approccio e ogni strumento di ricerca sono ignorati o sottovalutati. Sembra spesso mancare un disegno generale della ricerca valutativa e una cornice giustificativa delle scelte operative. Sembra sfuggire – in una cospicua maggioranza di casi – che le tecniche sono l'ultimo anello di una catena che parte molto più a monte con la definizione di un mandato valutativo, la costruzione di specifiche domande valutative, e diverse altre fasi e considerazioni che solo in seguito possono contribuire alla scelta di determinate tecniche e procedure piuttosto di altre⁷.

Ovviamente questa è una conseguenza del punto precedente, e ci serve per non lasciare galleggiare la teoria in un vuoto empirico. Fra la teoria e la rilevazione e analisi di informazioni, c'è invece questa necessità di creare dei nessi, e ciò costituisce il complesso metodologico. Per illustrare meglio questa continuità propongo una sintesi estrema nella figura sottostante.

Livello epistemologico-metodologico	↔	Livello metodologico-operativo	↔	Livello delle definizioni operative
Disegno₁		Disegno₂		Tecniche
<ul style="list-style-type: none"> Riflessione sulla conoscibilità dell'oggetto di valutazione Precondizioni necessarie per l'avvio della valutazione Conseguenze operative ed etiche, oltre che cognitive, della valutazione Valutazione della valutabilità, metavalutazione, etc. 		<ul style="list-style-type: none"> Traduzione del disegno₁ in un determinato percorso adatto al caso specifico (opportunità, coerenza, plausibilità...) Definizione specifica delle precondizioni e delle conseguenze, e analisi delle ricadute tecniche nella valutazione specifica 		<ul style="list-style-type: none"> Scelta delle tecniche che soddisfano le precondizioni del disegno₂ (validità, affidabilità, ...) nel caso specifico, alla luce anche dei vincoli strutturali della ricerca Definizioni operative relative alla raccolta e analisi delle informazioni

Indipendentemente dal lessico adottato in questa tabella⁸ si vede abbastanza chiaramente qual è il percorso che qui viene giudicato ottimale: una lettura teorica (dove la "teoria", come si vede, è legata all'oggetto della valutazione), cui segue una concezione metodologica generale, entro la quale stabilire gli specifici percorsi empirici.

⁷ Questo è il tema centrale del mio *Il disegno della ricerca valutativa*, Nuova edizione rivista e aggiornata, FrancoAngeli, Milano 2003 (prima ed. 2001).

⁸ Ripresa da Claudio Bezzi, *Il disegno della ricerca valutativa*, cit.

3. Il risultato: un potenziale operazionismo sterile.

Se non si pone attenzione a tali problemi, il risultato è un esercizio tecnico sterile, definibile operazionismo, che produce risultati di maniera, o poco pertinenti, o poco validi, o comunque scollegati dalle reali necessità indicate nel mandato⁹.

Questo ambito problematico ha una grande rilevanza perché investe sia la qualità della valutazione (la sua validità, la sua autorevolezza, la sua credibilità) sia il suo concreto utilizzo (che deve partire dalla sua comprensione da parte dei potenziali utilizzatori per arrivare sino alla credibilità e potenziale condivisione dei giudizi valutativi finali), temi evidentemente a cuore ai valutatori, ma che i valutatori devono sostenere con una maggiore competenza teorico-metodologica.

Una nota finale: si suggerisce, da alcune parti, che la valutazione non sia e non debba essere un fatto metodologico, e che la valutazione fondata sulla validità delle sue tecniche sia una concezione antiquata, ormai superata da impostazioni e filosofie nuove (quarte generazioni, valutazioni finalizzate all'*empowerment*, e via discorrendo). Il pluralismo è un dovere valutativo, e c'è spazio anche per tali approcci non metodologici o addirittura anti-metodologici¹⁰, ma l'opinione di chi scrive è che una valutazione senza riflessione metodologica abbia il valore di una *chat*, che la valutazione finalizzata all'*empowerment* sia una meritoria azione politica che non si debba chiamare valutazione, e che la "valutazione democratica"¹¹ avrà senso quando sarà apparsa una valutazione fascista o in qualche modo almeno tirannica.

L'esplosione degli approcci partecipati, *bottom-up*, qualitativi, nati da una giusta presa di distanza dall'eccessivo formalismo tecnico, non deve portare alla liquidazione dell'assunto di base che la valutazione è soprattutto una forma di ricerca sociale applicata, e che deve pertanto confrontarsi con i problemi anche epistemologici e metodologici della ricerca sociale.

Elementi teorico-metodologici per la costruzione di uno strumento valutativo

Dalla denuncia dell'operazionismo sterile, alla riflessione su un "come si fa" più adeguato, il passo è lungo e complicato. E bisogna ricordare che molti sono gli approcci e le filosofie di fondo che il pluralismo valutativo accoglie; vale a dire: quello che segue riflette la filosofia valutativa dell'Autore, non certo dell'intera comunità dei valutatori.

Ogni strumento è innanzitutto un concetto inserito in un "testo" metodologico più ampio. Cosa significa ciò? Occorre vedere ogni strumento non già come serie preordinata di procedure tecniche, ma come operazioni cognitive, integrate, finalizzate a produrre conoscenza. Se decido di costruire un questionario, per esempio, lo faccio perché *il formato delle informazioni* che il questionario mi produce è conforme sia al mandato valutativo, al contesto, alle diverse necessità organizzative e di budget, etc., sia alla mia capacità operativa e di analisi, sia alla più generale organizzazione e analisi dei dati che ne farò per esprimere un giudizio valutativo conforme al mandato. E nell'ambito di quel questionario compirò numerose scelte tecniche (p.es. in relazione alla strutturazione degli *item*, al lessico utilizzato per formulare le domande, al campione, etc.), ognuna delle quali è parte dello

⁹ Abbiamo precisi riscontri empirici, se solo vogliamo riconoscerli; si veda per esempio quanto scrive, criticamente, Roberto Righetti, *Trascurare l'evaluando*, "Rassegna Italiana di Valutazione", n. 27, 2004.

¹⁰ Anti-metodologico è p.es. quello noto di Guba e Lincoln, *Fourth Generation Evaluation*, Sage, Newbury Park, CA., 1989, che ha fatto più proseliti che non valutazioni (per una critica radicale si veda almeno Saville Kushner, *The Limits of Constructivism in Evaluation*, "Evaluation", vol. 2, n. 2, 1996); per quanto riguarda la nota *empowerment evaluation* vale la pena di leggere la stroncatura che ne fanno Michael Quinn Patton e Michael Scriven nella recensione all'ultimo volume di Fettermann (con replica dell'interessato): *Toward Distinguishing Empowerment Evaluation and Placing It in a Larger Context: Take Two*, "American Journal of Evaluation", Vol. 26, n. 3, pp. 408-417, 2005.

¹¹ Per tutti si veda Floc'hlay Beatrice e Plottu Eric, *Democratic Evaluation. From Empowerment Evaluation to Public Decision-Making*, "Evaluation", vol. 4, n. 3, July 1998, parzialmente tradotto come *La valutazione democratica delle decisioni pubbliche*, "Queste istituzioni", a. XXVIII, n. 124, 2001.

stesso tipo di ragionamento che, in sintesi, si può formulare col tradizionale “perché?”; perché per esprimere un giudizio valutativo pertinente, valido, etc. etc., ritengo più utile un questionario del tale tipo, somministrato a un tale campione, analizzato poi in qual modo, e così via? Se non si sa rispondere a tali domande, probabilmente si è fatta una cattiva ricerca valutativa; cattiva perché non argomentabile, forse casuale, basata su quel che il valutatore credeva di saper fare (e non già su quello che credeva utile).

Si è accennato al questionario a mo’ di esempio, naturalmente; stesso discorso per i focus group, per le interviste, per ogni tecnica econometrica, per approcci sperimentali o presunti tali, e così via enumerando.

Il legame fra le operazioni tecniche e le informazioni che si raccolgono è di tipo semiotico, come quello fra indicatori e concetto, tipico della ricerca sociale. Il discorso ci porterebbe troppo lontano, ma ha a che fare, sostanzialmente, con la possibilità/capacità di *argomentare* ogni scelta che viene compiuta¹². “Ogni scelta” significa certamente anche quelle tecniche, quelle che diamo per scontate perché le abbiamo imparate così, così sono scritte nei manuali, così tutte le usano.

Queste false certezze metodologiche affliggono in particolare la valutazione più formalizzata, quella fondata su approcci macroeconomici, sull’analisi costi-benefici, sull’approccio sperimentale, o comunque su quello che gode di un’ampio apparato statistico. Purtroppo a me pare che dietro apparenti formalismi tecnici si nascondono, spesso, scelte poco motivate, errori o difetti o leggerezze sapientemente (o inconsapevolmente?) occultati, banalità dei risultati. Non sempre, ovviamente, ma spesso. E per essere ecumenico dico subito che anche dietro approcci “qualitativi” vedo spesso – in realtà – superficialità teorico-metodologica, inconsistenza empirica, banalità dei risultati e dei giudizi valutativi.

Le procedure che legano lo strumento o la tecnica al mandato (a monte) e specialmente al risultato (a valle) non riguardano solo la validità statistica dei dati (che altro non è che un particolare tipo di legame semiotico) ma anche altri aspetti, dalla *face validity* indicata da Patton¹³, alle necessità strategiche del contesto, fino alla fondamentale attenzione a quell’apprendimento organizzativo che appare sempre più come la principale vocazione implicita della valutazione.

Volendo indicare gli elementi cruciali nella costruzione di una procedura valutativa, di uno strumento valutativo, mi sento di indicarne almeno due: il contesto (con una connotazione antropologica specifica, relativa alla cultura locale, sia essa una cultura “sociale” – di popolazione – o professionale, o organizzativa, etc.) e la necessità della più ampia partecipazione degli *stakeholder* all’impresa valutativa.

Relativamente al contesto occorre ricordare che in un significato sociologicamente rilevante è tornato in auge da qualche anno con la *realistic evaluation* di Pawson¹⁴, approccio ormai ritualmente citato dai più. Nella *realistic evaluation* il contesto è – scusandomi per la banalizzazione – il luogo dove agiscono quegli importanti “meccanismi” che animano ogni Programma, e la cui scoperta e analisi sono il cuore – per Pawson – del mandato valutativo. Anche se non viene detto in modo esplicito sembrano i meccanismi l’elemento essenziale della coppia, ma forse dobbiamo soffermarci con maggiore attenzione proprio su quei contesti che sono, in realtà, substrato sociale, complessità culturale, ambiente linguistico che performa ogni nostra valutazione. Si tratta insomma di avere più attenzione agli aspetti culturali, ad indossare con più convinzione occhiali antropologici, per valorizzare gli elementi *culturali* che finiscono per *determinare* la

¹² Personalmente mi sembra illuminante l’equivalenza fra valutazione e argomentazione. Si veda Ville Valovirta, *Evaluation Utilization as Argumentation*, “Evaluation”, vol 8, n.1, January 2002.

¹³ Nel celebre *Utilization-focused Evaluation*, Sage, Beverly Hills, CA, 3rd ed., 1998.

¹⁴ In particolare Ray Pawson, Nick Tilley, *Realistic Evaluation*, Sage, London, 1997.

possibilità di individuare i meccanismi, che – di più – stabiliscono *cosa siano* i meccanismi per quella comunità, *come siano* interpretati, *come realmente agiscano*¹⁵.

In secondo luogo la partecipazione. Tale partecipazione non ha alcuna finalità “democratica” o di *empowerment*, ma costituisce la necessaria risposta *metodologica* alla coniugazione fra contesto, apprendimento organizzativo, e analisi concettuale finalizzata al giudizio valutativo. La partecipazione diventa una *necessità* metodologica quando si concepisce la valutazione come un procedimento argomentativo (non meramente come sequenza di operazioni tecniche), l’evaluando come una parte del contesto, e il contesto come eminentemente linguistico e culturale. A quel punto il valutatore *ha bisogno* di ri-co-costruire i modelli di evaluando che sono nella testa degli *stakeholder*, sia per realizzare una valutazione valida *per loro* (la *face validity* di Patton), sia per stimolare quell’apprendimento organizzativo che altrimenti non si darebbe.

Elementi tecnici finali per la costruzione di strumenti valutativi

Con una eccessiva sinteticità, volendo indicare alcuni elementi tecnici finali che – in relazione agli elementi teorico-metodologico sopra accennati – possano portare verso una corretta ed efficace costruzione di “strumenti” valutativi, mi sento di segnalare i seguenti:

1. Evitare l’adesione acritica a qualunque forma di *fast thought*, la scelta di soluzioni basate sul “si è sempre fatto così!”, la pedante riproposizione di idee manualistiche e scolastiche, la convinta adesione alle mode, specie le più tecnicistiche, specie quelle basate su procedure di una qualche supposta “qualità”. Porsi il dubbio intellettuale e avere un atteggiamento critico costante come abito precipuo del valutatore. Anche negli elementi più tecnici.
2. Non pensare mai che le tecniche risolvano i problemi cognitivi: le tecniche sono protesi, e le dobbiamo utilizzare solo in quanto estensioni della nostra capacità argomentativa; e sono per giunta infide, ognuna con i suoi bravi difetti e limiti che occorre conoscere, salvo produrre errori sistematici (*bias*) che inficiano i giudizi valutativi finali.
3. Occorre saper scegliere le procedure operative, le soluzioni tecniche, gli strumenti concettuali, più appropriati in ogni singolo caso, e quindi occorre conoscere un ventaglio sufficientemente ampio di tecniche, e occorre avere competenze metodologiche abbastanza solide, tanto da poter modificare e adattare quelle procedure operative, e saper utilizzare tecniche miste, integrate fra loro¹⁶.
4. Attenzione ad alcuni elementi molto comuni, quasi inevitabili, ma sovente oggetto di errori clamorosi; prima di tutti gli “indicatori”: gli indicatori non sono un *set* di numeri, che ci sembrano belli, scelti fra i pochi che abbiamo a disposizione. Gli indicatori sono complesse costruzioni concettuali che *poi* possono anche assumere un formato numerico, ma a monte di tutto c’è – ancora una volta – un ragionamento, e non la scelta più o meno casuale di basi numeriche¹⁷.

¹⁵ Ho trattato compiutamente di tutto questo dal punto di vista teorico e metodologico nel mio *Evaluation Pragmatics*, “Evaluation”, in corso di stampa, 2005.

¹⁶ Le tecniche miste (o “triangolazione”) è un esempio interessante di una nuova frontiera metodologica in valutazione che fa sempre più capolino anche in Italia senza una chiara consapevolezza di cosa implichi. Per una rassegna problematica si veda Jennifer C. Greene and Valerie J. Caracelli, eds. (1997), “Advanced in Mixed-Method Evaluation: The Challenger and Benefits of Integrating Diverse Paradigms”, *New Directions for Evaluation*, 74; in italiano si può leggere Paolo Parra Spiani, *Triangolazione e processi valutativi*, RIV 24/2001, poi ampliato dall’Autore nel suo *Triangolazione e privato sociale. Strategie per la ricerca valutativa*, Bonanno editore, Acireale-Roma 2004.

¹⁷ Gli indicatori *non sono* (o non sono necessariamente) dei numeri, bensì elementi informativi facilmente operativizzabili (misurabili) che *indicano*, con un rapporto incerto, semantico (o secondo alcuni Autori, probabilistico), un concetto non altrimenti indagabile; quindi, per esempio, il rossore sul volto è – fra l’altro – un indicatore di timidezza, la professione è – fra l’altro – un indicatore della capacità di spesa di un individuo, come il consumo energetico nazionale è un indicatore del livello medio dei consumi e quindi della ricchezza media nazionale. Una bibliografia minima sugli indicatori (e sulle correlate questioni dei concetti, dimensioni, indici, etc., che brevemente vedremo in questo paragrafo) prenderebbe numerose pagine; mi limito a segnalare Paul F. Lazarsfeld, *Dai concetti agli indici empirici*, in “L’analisi empirica nelle scienze sociali. I - Dai concetti agli indici empirici”, a cura di R. Boudon e

5. Ogni valutazione si situa in un contesto specifico. Il concetto di *contesto* non fa affatto riferimento a una dimensione geografica circoscritta, ma ad un ambito culturale e sociale: una comunità di pratiche, un'organizzazione, un insieme identificabile di attori sociali accomunati da identità di obiettivi. C'è quindi indubbiamente un contesto specifico nel singolo servizio sociale territoriale, ma c'è il contesto delle *politiche* sociali nazionali (che raccoglie tutti gli attori interessati direttamente dalla loro definizione, gestione, uso, conseguenze); c'è il contesto "umbro" come c'è il contesto delle politiche del lavoro europee; e così via. Ebbene la valutazione non può essere generico, standardizzato, amorfo assemblaggio di procedure valide ovunque e comunque, ma deve essere "ritagliata"¹⁸ su ogni specifico contesto; ma "contesto" – come già ricordato – viene così ad assumere un significato socio-antropologico: la cultura professionale che il quell'organizzazione si esprime; l'agenda delle priorità che quegli *stakeholder* individuano; le logiche generali, gli obiettivi e le finalità che in quelle politiche si affermano, nel linguaggio loro proprio. Una valutazione attenta a questa fondamentale dimensione non può che adottare una metodologia specifica, flessibile, capace di rispondere a quelle determinate domande valutative¹⁹.
6. Ecco perché non può esserci valutazione efficace senza coinvolgimento degli *stakeholder*; se guardiamo al contesto, e intendiamo costruire domande valutative a partire dalle sue specificità, dobbiamo avere un dialogo non formale con gli attori rilevanti, e costruire strumenti con una chiara *face validity*, perché *loro* devono capire quello facciamo come valutatori, altrimenti non saranno in grado di compenetrarsi con la valutazione e non sapranno indirizzare il valutatore verso dimensioni concrete, strumenti efficaci, giudizi valutativi utilizzabili.

P.F. Lazarsfeld, Il Mulino, Bologna 1969; Alberto Marradi, *Concetti e metodi per la ricerca sociale*, La Giuntina, Firenze 1987; Alberto Marradi, "Referenti, pensiero e linguaggio: una questione rilevante per gli indicatori", *Sociologia e Ricerca Sociale*, n. 43, 1994; Luca Ricolfi, "Sul rapporto di indicazione: l'interpretazione semantica e l'interpretazione sintattica", *Sociologia e Ricerca Sociale*, n. 39, 1992; Leonardo Cannavò, *Teoria e pratica degli indicatori nella ricerca sociale. I – Teorie e problemi della misurazione sociale*, Ed. LED, Milano 1999.

Per una riflessione specifica sugli indicatori in valutazione si veda invece Mauro Palumbo, *Gli indicatori valutativi*, "Rassegna Italiana di Valutazione", a. VII, n. 27, 2003, pp. 107-129.

¹⁸ Come ci ricordano esplicitamente Peter Rossi, Howard E. Freeman e Mark W. Lipsey, *Evaluation. A systematic approach*, 6th ed., Sage, Thousand Oaks, CA., 1999.

¹⁹ Cfr. Claudio Bezzi, *Evaluation Pragmatics*, cit.