

Cogest- Ciclo di conferenze su "Conti pubblici e controlli: prospettive e problemi"

7 novembre 2002, Aula Magna della Scuola Superiore della P.A.

"La valutazione delle politiche pubbliche: chi l'ha vista?"

ALBERTO MARTINI, associato di Statistica Economica Università del Piemonte Orientale,
Direttore Progetto Valutazione, Torino

discussant: **NICOLETTA STAME**, ordinario Univesità "La Sapienza" di Roma

CAIANIELLO (testo ridotto)

Buonasera, benvenuti tutti. E grazie di essere venuti, in verità grazie b dobbiamo dire tutti invece alla professoressa Stame e al professor Martini, persone che non è necessario presentare, in questa materia della valutazione sono tra i massimi esperti in Italia. La pubblica amministrazione, come ben sappiamo, a differenza delle aziende private, non ha il metro dei prezzi del mercato per poter dare un valore al proprio prodotto e stabilire se la propria attività sia soddisfacente o meno, e deve perciò avere altri strumenti per lanciarsi in quello che anch'esso è in qualche modo un mercato, quello del giudizio degli elettori sui governanti e anche quello della competitività con gli altri Paesi. Strumenti fra i quali primeggiano le tecniche valutative affinché l'azione amministrativa si svolga con successo su entrambi i versanti, come gli illustri relatori vorranno spiegarci. Do quindi volentieri la parola al professor Martini.

Professor MARTINI

Grazie Presidente. Oggi io e la professoressa Stame giochiamo un ruolo complementare. Conoscete la storia del bicchiere mezzo vuoto e mezzo pieno? Io sono il bicchiere mezzo vuoto e la prof.ssa Stame per definizione il bicchiere mezzo pieno, quindi io inizierò parlandovi della metà vuota del bicchiere.

Comincio con una citazione, che mi piace moltissimo, di un ex magistrato della Corte dei Conti, Sergio Ristuccia. *"Di valutazione e di cultura della valutazione si parla molto. Spesso con quella amena leggerezza davanti alla quale non si capisce se tutto sia già chiaro e definito ovvero ci si riferisca ad una strada lunga, un po' difficile...Di questa leggerezza i dibattiti nella pubblica amministrazione offrono i più variegati esempi (Queste Istituzioni, 1996)"*

Qualche anno fa scriveva queste parole per me profetiche sulla valutazione. Avrebbe dovuto scriverle oggi invece che nel 1996. Di valutazione, di cultura della valutazione si parla molto, si parla molto e si parla, si parla tanto. Sei anni dopo queste parole di Ristuccia, se ne parla soprattutto.

Infatti da qui viene il titolo del mio intervento, La valutazione *Chi l'ha vista?* Tutti ne hanno sentito parlare ma poi vista non si capisce bene chi l'abbia vista. Tutti ne hanno sentito parlare, tutti anzi la invocano, ma mai dopo averla vista all'opera, soprattutto se utilizzata bene.

Ma, continua Ristuccia: spesso con quella amena leggerezza (ed è bellissimo questo tocco) davanti alla quale non si capisce se tutto sia già chiaro e definito ovvero ci si riferisca ad una strada lunga e un po' difficile.... Di questa leggerezza i dibattiti nella pubblica amministrazione offrono i più variegati esempi – e qui decidete voi se siete d'accordo oppure no. Questa amena leggerezza, ahimè, permane, generando quella che, a mio parere, è un pochino una bolla. Non speculativa, ma una bolla valutativa. E tutte le bolle prima o poi scoppiano.

Il professor Caianiello notava nella sua introduzione il fatto che nel '99 riempiva questa sala, oggi invece siete in pochi. Oggi, forse la gente comincia a stufarsi di sentire soltanto parlare di valutazione, e io sono molto preoccupato di questo. Sono preoccupato per il fatto che la bolla possa scoppiare perché di mestiere mi occupo di valutazione.

Nello stesso periodo in cui Ristuccia scriveva queste parole profetiche, due studiosi inglesi, molto diversi da Ristuccia, due sociologi, Ray Pawson e Nick Tilley, scrivevano nell'introduzione a un loro libro, forse il più famoso: *The term evaluation carries so much baggage that one is in danger of dealing not so much with methodology than with incantation*” (Ray Pawson and Nick Tilley, *Realistic Evaluation*, 1997, pag. 2)

Detto in inglese fa tutto un altro effetto, ma anche in italiano non perde la sua forza: *il termine valutazione è carico di un tale fardello di significati che uno ha l'impressione di avere a che fare più con un incantesimo che con una metodologia*. E ahimè è così, lo so che è brutto dire queste cose, però qui abbiamo a che fare più con un incantesimo. Quello per cui io mi batto, per cui mi batterò oggi nel mio discorso è la speranza che questo incantesimo diventi più una metodologia. Smetta cioè di essere un incantesimo che va bene in tutte le situazioni, senza che però nessuno sappia esattamente che cosa sia, una cosa che si invoca così, con questo tocco di magia.

Che cosa alimenta questo incantesimo? Perché da 4-5 anni a questa parte, da 2 o 3 quasi in modo ossessivo, si parla di valutazione da tutte le parti? Cosa è che ha determinato questa bolla valutativa? Alla base vi è una ragione nobilissima: il desiderio diffuso, non solo in Italia ma in tutto il mondo di migliorare le prestazioni della pubblica amministrazione. Detta nel gergo valutativo, a voi molto familiare è la necessità di migliorarne l'efficienza e l'efficacia. Questi due termini significano cose molto concrete, ma ahimè – e su questo tornerò dopo – sono stati mitizzati diventando un qualcosa che automaticamente una volta misurato origina dei miglioramenti. Questa è un' l'essenza della valutazione come bacchetta magica: l'efficienza, l'efficacia e l'idea che misurandole in qualche modo tutto migliori.

Quindi dietro questa bolla, questo grande innamoramento per la valutazione c'è il desiderio di migliorare la pubblica amministrazione, che ha tante forme e che si esplica in tanti modi. La

specificità della valutazione rispetto al decentramento, alla concertazione, alla privatizzazione e alla semplificazione, tutti gli altri "one" che abbiamo inventato per migliorare la prestazione della pubblica amministrazione, è questo ingrediente aggiuntivo, che la rende molto affascinante: la fiducia, spesso ingenua nel metodo scientifico. L'idea è che con i dati, con l'analisi, con gli indicatori in qualche modo si riesca a mettere ordine in un processo decisionale altrimenti irrazionale. E' molto radicata la presunzione che i dati (altra parola magica) e i loro derivati, le tabelle, i numeri, gli indicatori, consentano decisioni migliori.

Il convergere di queste due suggestioni, la fiducia nel metodo scientifico e il desiderio di migliorare la pubblica amministrazione, ha contribuito a un rapidissimo processo di istituzionalizzazione della valutazione. L'esempio migliore è il decreto legislativo 286 che curiosamente voi chiamate "decreto sui controlli". Il titolo del decreto tuttavia non ha la parola "controllo". E' molto retorico e ridondante: "meccanismi e strumenti di monitoraggio e valutazione dei costi, dei rendimenti e dei risultati dell'attività svolta dalle amministrazioni pubbliche". Se guardate il comma b) dell'articolo 1 con occhi diversi, ci trovate tutta l'ingenuità di questo approccio. "Le pubbliche amministrazioni si dotano di strumenti adeguati a" "Si dotano", ok, non si sa bene come ma si dotano, "di strumenti adeguati a verificare l'efficacia e l'efficienza (ma qui c'è anche l'economicità) dell'azione amministrativa, al fine di ottimizzare il rapporto tra costi e risultati". Cioè passa in questo modo l'idea che misurando qualcosa si abbiano effetti reali, addirittura l'ottimizzazione del rapporto tra costi e risultati. Misurare è la stessa cosa che migliorare. Ma sappiamo tutti che non è così.

Ma il pezzo migliore di retorica istituzionale sulla valutazione deve ancora arrivare. E' bellissimo, questo forse non lo avete mai visto. Decreto ministeriale del 2000, istituzioni delle lauree specialistiche. Tra le 102 classi di laurea specialistica è stata istituita la classe delle lauree specialistiche in metodi per l'analisi valutativa dei sistemi complessi. E' interessante la descrizione di cosa devono saper fare questi laureati. C'è tutto il carico di aspettative che si gettano addosso alla valutazione, non c'è nulla di sbagliato in quello che vi leggerò, però notate quanta enfasi e quante aspettative si costruiscono dietro questa parola di cui poi nessuno conosce realmente i contenuti.

I laureati nei corsi di laurea specialistica della classe devono: avere le competenze necessarie per intervenire nel processo di decisione, di monitoraggio e di valutazione di politiche e programmi, offrendo al decisore o al controllore elementi di giudizio ex ante e ex post sugli effetti previsti e reali dell'azione, riducendo la complessità nell'ambito decisionale, migliorando l'efficienza e l'efficacia - e qui c'è anche la qualità dell'attività, anche dal punto di vista organizzativo - ottimizzando le risorse disponibili e favorendo i processi di trasparenza. E qui suona l'inno nazionale!

Questi laureati dovrebbero essere in grado di fare tutte queste cose! A questo punto è lecito chiedersi: ma se bastava così poco perché non lo abbiamo fatto prima? Se bastava imparare due o tre cose misteriose, ed ottimizzare così le risorse, migliorare l'efficacia dell'azione pubblica, favorire

la trasparenza, ridurre la complessità. A me questa sembra, parafrasando il titolo di un film con Tom Cruise, “missione impossibile”.

Davvero questa magica valutazione fa tutte queste cose? Prima che facevate?! Perché non lo avete fatta prima allora?! Quello che manca soprattutto, e che io tendo a richiamare, è una riflessione sui meccanismi reali che possono tradurre la valutazione in miglioramento. Guardiamoci negli occhi e cominciamo a pensare, a riflettere su quali meccanismi producono tutte queste ottime cose. Non scarichiamoci la coscienza dicendo "valutando succedono queste cose magnifiche". Perché alla base di tutto sta fondamentalmente un grandioso atto di fede, che in questi anni abbiamo fatto e che ha alimentato questa bolla - che però prima o poi scoppia.

Ma io mi chiedo, la valutazione migliora davvero l'azione pubblica? Succede sempre e comunque anche quando lo sforzo intellettuale e le risorse impiegate in questa valutazione, o controllo di gestione, o quello che volete, sono minime? Perché spesso le risorse intellettuali e materiali che ci si mette dietro sono veramente minime. Non ci sono mai contraddizioni, tensioni, effetti perversi? E, soprattutto, basta misurare qualcosa perché questo migliori?

Io vi offro adesso due esempi per vivacizzare un po' il discorso. Il primo banalissimo, però secondo me significativo, legato alla mia esperienza di docente universitario: la legge 370 del 99 impone ai nuclei di valutazione di ciascuna università di verificare l'opinione degli studenti sulla qualità della didattica dei docenti. Ma attraverso quali meccanismi tale valutazione può migliorare la didattica? Ma quale sistema di incentivi, di segnali porta i docenti a migliorare la didattica? Nella mia Università abbiamo pensato fondamentalmente a un meccanismo di *shaming*, cioè contiamo sul fatto di far vergognare i docenti. Cosa facciamo? prendiamo i risultati di questi questionari, li restituiamo al singolo docente, mostrando i voti che ha preso rispetto ai singoli criteri (cura, preparazione nelle lezioni, capacità di motivare l'interesse della materia, ecc.) in confronto ai voti presi dai colleghi. L'idea è che il singolo professore si vergogni e sia spinto a migliorarsi se si vede sotto la media dei colleghi. Potrebbe non funzionare, ma almeno un “meccanismo” abbiamo provato ad inventarlo, non ci siamo limitati a misurare. Questo ci sembrava un modo per utilizzare la valutazione, non solo per invocarla.

Due mesi fa il Comitato nazionale per la valutazione del sistema universitario mi manda una lettera in cui mi si dice che il Comitato propone un questionario minimo che i nuclei di valutazioni delle singole Università è opportuno adottino al fine di garantire una omogenea rilevazione su scala nazionale. Proprio per assicurare la comparabilità dei dati è necessario che tale questionario sia adottato ecc. ecc. Questo è un *altro uso* della valutazione. Ciò che sta dietro l'imposizione di un questionario unico per tutti è un'altra idea, quella di rilevare dati sulla soddisfazione degli studenti in modo omogeneo e poi fare un qualche tipo di confronto alias graduatoria tra gli atenei.

Bene, cosa succede? Che il meccanismo di *shaming* che ci siamo inventati per far funzionare la valutazione non funziona più. Se gli studenti sono critici con i propri docenti (e funziona lo *shaming*) il loro Ateneo ne risente nel confronto con gli altri. Questo crea un conflitto di interessi.

Qual è la morale? Che la valutazione non è raccolta di dati, è utilizzo di dati per migliorare. Apparentemente la rilevazione dell'opinione degli studenti sulla qualità della didattica dei docenti è sempre la stessa cosa ma in realtà, se ragioniamo sui meccanismi che possono tradurre questa valutazione in miglioramento dell'azione pubblica, sono due meccanismi diversi, anzi sono due meccanismi in questo caso incompatibili, per cui o fai uno o fai l'altro. O fai lo *shaming* interno o fai il confronto tra Università. Non puoi fare tutte e due le cose! La legge non dice che bisogna usarli in un modo o nell'altro, se però ci mettiamo a riflettere un attimo sull'utilizzo vediamo che emergono delle incompatibilità.

Un esempio meno banale – e qui sparo molto più in alto – è il seguente. Attraverso quali meccanismi la valutazione può migliorare le politiche del lavoro? Costruendo qualche indicatore opportunamente etichettato come di efficacia oppure di impatto? Bene, vi faccio vedere quella che secondo me è una mostruosità valutativa. Questi sono indicatori di impatto utilizzati per valutare (altro uso del termine) il fondo sociale europeo, e sono contenuti nel quadro comunitario di sostegno che è un documento sottoscritto dal governo nazionale e dall'Unione Europea, quindi una cosa abbastanza impegnativa. Guardate all'indicatore “quota di occupati alle dipendenze con rapporto di lavoro temporaneo” che viene definito come numero di occupati con lavoro temporaneo, quindi precari, sul numero di occupati alle dipendenze, quindi i precari diviso il totale degli occupati alle dipendenze. Secondo gli estensori di questo documento questo rapporto/indicatore *misura il successo delle politiche di flessibilizzazione del lavoro*. In poche parole se ci sono tanti precari siamo contenti perché questo semplice rapporto misura il successo delle politiche di flessibilizzazione del lavoro! Questo documento però l'ha sottoscritto il governo italiano con Bruxelles! Di chi sia colpa non lo so, io mi limito a rilevare quali sono i rischi se ci si fa prendere dall'andazzo “è valutazione, quindi bastano un po' di indicatori”.

Il successivo indicatore (efficacia dei centri per l'impiego) se utilizzato realmente può addirittura produrre effetti perversi. I centri per l'impiego sono quelli che hanno sostituito gli uffici di collocamento, per semplicità pensate a rinnovati uffici di collocamento che dovrebbero offrire servizi di sostegno all'inserimento lavorativo. L'efficacia dei centri per l'impiego in questo documento è definita come numero degli utenti diviso il numero delle persone in cerca di lavoro. Quindi uno presume che se questo rapporto cresce, il centro per l'impiego è più efficace. Che significa? Io francamente non lo so.... Comunque l'efficacia è una cosa buona: essere efficaci vuol dire essere più bravi degli altri. Quindi il centro è più efficace se il numero degli utenti sul totale delle persone in cerca di occupazione è più alto.

Come si fa allora a diventare più efficaci? Si offrono servizi sempre meno intensivi, si contano come utenti quelli che sono soltanto di passaggio al Centro per l'Impiego, si convocano i disoccupati per un colloquio di un quarto d'ora, che non serve assolutamente a nulla..... Però se fai un colloquio di due ore, un bilancio di competenza ecc. che serve a qualcosa probabilmente, lo devi fare su un numero di utenti molto più piccolo, e allora diventi inefficace. Quindi se utilizzato per giudicare la performance dei singoli centri per l'impiego questo indicatore porterebbe a risultati assurdi. Per diventare efficaci dobbiamo offrire a un numero più grande possibile di persone

servizi sempre meno intensivi e probabilmente meno utili, invece che concentrarsi sui disoccupati di lunga durata e le fasce deboli.

Io sono convinto che il diffondersi di una concezione superficiale e magniloquente allo stesso tempo della valutazione vada contrastato - questa è la mia missione - e vada invece contrapposto un approccio parsimonioso nei concetti. Soprattutto la valutazione deve essere minimalista negli obiettivi. Non possiamo pensare che la valutazione, qualunque cosa essa sia, possa realizzare tutte queste cose stupende. Dobbiamo pensare a cosa realmente può fare e attraverso quali meccanismi di miglioramento una attività di tipo cognitivo (perché di questo si tratta fondamentalmente), possa migliorare l'azione pubblica. Prima di intraprendere una qualsiasi valutazione, prima di invocarla, è necessario - secondo me - chiarire quali sono le aspirazioni al miglioramento. Questa cosa che invochiamo continuamente a che aspirazione corrisponde?

Io ho provato a catalogare queste aspirazioni, ma vi devo avvertire che non vanno prese alla lettera. Le varie aspirazioni tendono senza dubbio a combinarsi, a sovrapporsi nella pratica, ma questa non è una buona ragione per dire che sono tutte la stessa cosa, e non sono esaustive. Vi faccio vedere come si può concettualizzare la valutazione in modo molto concreto, cercando di chiarire quali sono i meccanismi possibili di miglioramento che questa cosa misteriosa può innescare. Cinque aspirazioni per la valutazione, prima vi dico i nomi.

l'aspirazione meritocratica

l'aspirazione razionalizzatrice

l'aspirazione resocontista

l'aspirazione conoscitiva

l'aspirazione partecipativa

La valutazione vuole fare tante cose: io ne ho individuate cinque. Tuttavia vi voglio convincere del fatto che per fare queste sono necessari strumenti, tempi e meccanismi di utilizzo profondamente diversi. E' inutile che continuiamo a dire che esiste un'unica valutazione: esistono invece tante attività cognitive che rispondono a esigenze diverse.

L'aspirazione meritocratica è la prima e forse più primitiva funzione della valutazione – identificare i migliori per premiarli, che possono essere progetti presentati per il finanziamento, possono essere dirigenti da valutare, possono essere approcci alternativi ad un problema. Comunque il meccanismo fondamentale, il modo con cui la valutazione vorrebbe contribuire in questo caso specifico è quello di consentire di discernere chi è meritevole e chi non è meritevole. Ma meritevole rispetto a cosa? Rispetto ad una serie di criteri pre-identificati che possono essere diversamente pesati e che avranno bisogno di qualcuno che esprima dei giudizi. Il tutto è molto ben rappresentabile con una tabella, dove sulle righe ci sono le cose tra cui dobbiamo scegliere e sulle colonne i criteri e in ogni cella il giudizio che qualche esperto dà rispetto a ciascun criterio su ciascuna azione alternativa. Il prodotto finale di questa forma di valutazione è identificare il migliore da selezionare o il migliore da premiare. Questa è una delle accezioni di valutazione, ma se voi girate per gli uffici di un'amministrazione regionale e gridate ad alta voce "valutazione", quelli che vi sentono al 90% probabilmente pensano a questa cosa. Per loro valutazione è selezione

di progetti, comunitari, regionali. Le regioni in particolare fanno tantissima attività di smistamento di fondi e di assegnazione di risorse sulla base di procedure competitive.

Una seconda aspirazione, che è diversa nei suoi meccanismi cognitivi e nell'utilizzo, è l'aspirazione a razionalizzare. Possiamo chiamarla l'aspirazione del controllo, del controllore, interno o esterno che sia (il controllo manageriale o il controllo direzionale). L'essenza di questa valutazione, non è premiare i migliore ma identificare le cose che non vanno. Questo vuole fare la Corte dei Conti nella sua attività di controllo sulla gestione, non vuole premiare i migliori, non vuole generare nuova conoscenza, ma essenzialmente vuole o vorrebbe identificare anomalie nel funzionamento di organizzazioni pubbliche. A che fine? Al fine di correggerle, dove i meccanismi di correzione saranno diversi a seconda che i controlli siano interni o siano esterni. L'essenza qui è identificare cose che non vanno. Quindi l'aspirazione razionalizzatrice si esplica nella realizzazione di strumenti di controllo, interno e esterno, del funzionamento di singole organizzazioni. "Ho identificato qualcosa che non va e te lo segnalo e quindi ti incentivo a correggere il problema."

Quindi giudicare il funzionamento di singole organizzazioni, di cui di fatto poi concretamente si tende a giudicare costi e qualità. Potete parlare di efficienza, efficacia, economicità, ma giudicare il funzionamento di un'organizzazione pubblica, per contribuire alla sua razionalizzazione, significa dare dei giudizi comparativi facendo dei confronti con degli standard che possono richiedere l'osservazione di altre organizzazioni. Alla fine, al di là di tutta la retorica, questo è l'oggetto del contendere, quando vogliamo razionalizzare vogliamo mandare dei segnali che correggano anomalie rispetto a costi eccessivi o a qualità scadente.

Questo è diverso da segnalare il merito, qui anzi si vuole identificare quello che non va, ma non da un punto di vista di una graduatoria, quindi non è più lo schema a matrice di prima, non è più quello il processo cognitivo. Questo è il confronto con standard, non il premiare o dare punteggi rispetto a certi criteri di giudizio.

Butto lì una provocazione: si dà spesso per scontato, e c'è anche scritto nel 286, che la valutazione dei dirigenti deve tenere conto delle risultanze del controllo di gestione. A tutti sembrerà giusto. Secondo me è profondamente sbagliato perché non tiene conto dei conflitti latenti tra motivazioni diverse. Un meccanismo come il controllo di gestione per avere un senso, a mio parere, deve identificare le cose che non vanno, e deve essere uno strumento della dirigenza. Come si può pensare che la stessa dirigenza nel momento in cui identifica anomalie utilizza poi quella roba per venire giudicata sulla propria performance? Volere sempre mettere le cose assieme significa annacquare e talvolta anche produrre degli effetti perversi. Quindi il controllo di gestione risponde fondamentalmente, se è una strategia di valutazione, a finalità di razionalizzazione e deve identificare anomalie, non premiare i migliori. Le due cose secondo me vanno tenute separate.

La terza aspirazione, ho usato questo termine orribile "resocontista" volontariamente, anche per sottolineare il fatto che il termine *accountability* che adesso sta andando di moda tantissimo, non ha una traduzione in italiano. *Accountability* è l'esigenza di rendere e chiedere conto dei risultati

ottenuti. *To be accountable* non è fare il contabile, ma rendere conto della propria attività. Qui la frase simbolo è: occorre chiedere e rendere conto dei risultati dell'azione pubblica, che è diverso da individuare le anomalie. Qui l'aspirazione resocontista vuole rispondere a criteri e esigenze di *accountability* e si esplica nella messa in opera di sistemi di descrizione e misurazione dell'attività al fine di fine ultimo di legittimarla.

Quindi quando facciamo rendicontazione - però se io dico rendicontazione voi pensate a quella contabile per quello dico resocontista – facciamo un tipo particolare di valutazione. Tutta la valutazione dei fondi strutturali dell'Unione Europea è fondamentalmente ispirata a questa logica “io ti do i soldi, tu misurami e descrivimi cosa hai fatto di questi soldi, quale attività hai messo in opera”. Il che non è l'atto di identificare anomalie, perché tipicamente si fa alla fine, si fa quando non si può più avere realmente del feed-back, non è una preoccupazione di tipo manageriale quella che sta dietro questa funzione valutativa.

Quindi siamo arrivati ad analizzare tre aspirazioni: meritocrazia, razionalizzazione, rendicontazione o *accountability*. La quarta è quella che in fondo ha ispirato persone come Nicoletta Stame e il sottoscritto a occuparci di valutazione, ed è quella che nel mondo anglosassone, in particolare negli Stati Uniti, sta dietro al termine *evaluation*. *Evaluation* in italiano non si traduce con valutazione, o anzi valutazione non si traduce con *evaluation*. La lingua italiana è molto povera e traduce con valutazione uno spettro di termini che in inglese è molto più ampio. *Evaluation*, nella pratica statunitense ma anglosassone in generale, significa fondamentalmente apprendere lezioni dall'esperienza. Quindi sostanzialmente è l'analisi critica dell'attività svolta per capire se funziona o se non funziona. Il che non significa razionalizzarla e neppure rendere conto, ma è fare ricerca valutativa, e quindi *evaluation research*.

Quindi la frase simbolo qui è: occorre apprendere lezioni utili dallo svolgimento di un'azione pubblica. Non è rendere conto ma è capire se ciò che si fa funziona o non funziona, per generalizzare queste lezioni all'attività futura. Quindi l'aspirazione conoscitiva si esplica soprattutto nell'identificare i meccanismi che portano all'azione che si sta valutando, ad influenzare il problema su cui si interviene e nel capire gli ostacoli che impediscono a questi meccanismi di agire. Dal punto di vista cognitivo questa attività valutativa è molto più ambiziosa, proprio perché più elevati e complessi sono i suoi obiettivi.

Se voglio fare *evaluation research* non posso farla anno per anno, costringerla nella camicia di forza della rendicontazione annuale. Anche alla Corte dei Conti, se mai volete fare *evaluation research*, scordatevi l'anno solare. L'anno solare è un periodo di riferimento che va bene per attività di tipo rendicontativo: ogni anno ti do dei soldi e ogni anno voglio sapere cosa hai fatto, ma i problemi pubblici su cui l'azione pubblica interviene non hanno una scadenza annuale: il problema della disoccupazione, la tossicodipendenza, la povertà, oppure le carenze strutturali di capitalizzazione delle imprese artigiane e tutto quello che volete, non hanno cadenza annuale. Quindi l'attività di valutazione e l'aspirazione di tipo conoscitivo ha orizzonti tematici e temporali diversi. Identificare meccanismi con cui l'azione pubblica produce risultati è cosa diversa dal

rendicontare i risultati ottenuti, sono tutte e due attività nobili ma non possiamo metterle sullo stesso piano.

Infine, c'è anche un'altra aspirazione che sta diventando sempre più importante, sotto l'etichetta valutazione c'è anche l'aspirazione a far partecipare, il cosiddetto *empowerment*. Gli stakeholders di un'azione pubblica sono chiamati a far presente i loro punti di vista. Questa è la più confusa delle aspirazioni forse, ma anche la più nuova. Gli stessi sistemi di *customer satisfaction* non corrispondono semplicemente a un'esigenza di razionalizzazione, ma servono fondamentalmente a far esprimere giudizi a tutti coloro che hanno un interesse legittimo a quell'azione pubblica. E anche questa rientra sotto il grande cappello della valutazione.

La lista delle aspirazioni potrebbe allungarsi, non è detto che siano solo 5. Cinque le ho messe io perché stavano bene nella pagina. Non è la stessa cosa - voglio convincervi - identificare i migliori, evidenziare anomalie, rendere conto dei risultati, apprendere lezioni utili e far partecipare: sono attività molto diverse tra di loro. E queste profonde diversità vengono troppo spesso ignorate, etichettando tutto come valutazione.

Ma quel che è peggio è che si è andati alla ricerca di un minimo comune denominatore per tutte queste aspirazioni, di un metodo pass partout, che sia valido per tutti gli usi, accattivante, facile da spiegare, e che poi richieda anche pochi soldi e poche idee. E questo minimo comune denominatore è proprio povero ed è troppo appiattito sulla pretesa di misurare tutto e di ridurre tutto ad indicatori di efficienza ed efficacia. Tutti quei tipi di attività e di aspirazioni non possono essere soddisfatti da indicatori di efficienza ed efficacia.

Efficienza cosa significa? Parliamoci chiaro, significa minimizzazione dei costi. Poi potete elucubrare ma significa questo: far le cose spendendo meno. Efficiente è quell'azione pubblica, progetto, politica, organizzazione, servizio, quello che volete, che produce una prestazione al costo minimo (a parità di qualità del prodotto-servizio, se non c'è la parità della qualità è facile produrre a costo minimo). Ora dare un giudizio di efficienza è una cosa maledettamente difficile, non basta utilizzare, come spesso si suggerisce, ci si illude, un indicatore che più o meno abbia la forma di costo per unità di prestazione. Quello è semplicemente un descrittore. Io ho un'organizzazione che fornisce una certa prestazione, so quanto spende per quella prestazione, quindi ho un sistema di contabilità analitica. Tutte cose ottime.

Alla fine ricavo un indicatore di costo per unità di prestazione, e lo etichetto come indicatore di efficienza. Ma non mi indica l'efficienza, perché un giudizio di efficienza su una organizzazione richiede ben altro. Richiede innanzitutto di trovare uno standard. Non mi basta sapere quale è il costo per unità di prestazione, per dare un giudizio devo fare un confronto, per fare un confronto devo avere uno standard, che rappresenti il livello ottimale, cioè minimo del costo per quella prestazione con quelle caratteristiche. Ma questo è enormemente più difficile rispetto a descrivere e misurare il costo per unità di prestazione, richiede confronto con altre unità, ma richiede anche di utilizzare un metro per garantire che confronti di costi unitari siano fatti a parità di qualità delle

prestazioni, e questo è tre ordine di grandezza più difficile. Quello che spesso si chiama indicatore di efficienza, non mi indica l'efficienza, mi indica quale è il costo per quella organizzazione.

Ho molto più da dire sull'efficacia, che spesso viene presentata in un binomio inscindibile con l'efficienza. In realtà il binomio è scindibilissimo, perché la rilevanza dell'efficienza e la rilevanza dell'efficacia sono molto diverse. I giudizi di efficienza sono molto rilevanti, e molto difficili da dare, come ho argomentato prima, ma sono rilevanti all'interno di organizzazioni che producono servizi ragionevolmente standardizzati.

I giudizi di efficacia sono invece invocati in situazioni enormemente più eterogenee tali però da fargli perdere un significato compiuto. Io ho tentato questo esercizio di far vedere cinque usi molto diversi - oggi il numero magico è sempre cinque - 5 accezioni in cui il termine efficace è rilevante, per farvi vedere come gli strumenti per dare questo giudizio di efficacia devono necessariamente essere diversi. Se vi chiedessi cosa è l'efficacia tutti direste quello che c'è scritto sullo schermo, perché a questo siamo abituati: l'efficacia è la capacità di raggiungere gli obiettivi. Ci sono anche altre definizioni però questa mette tutti d'accordo nella sala. Ma che significa obiettivi? Obiettivo è un termine foriero di ambiguità e di confusione.

Efficacia può significare capacità di produrre effetti desiderati sui destinatari di una politica, questo è il termine più pregnante di efficacia che probabilmente ci sarebbe nel vocabolario, efficacia vuol dire avere effetti, produrre in senso causale gli effetti, cioè le modificazioni che erano l'obiettivo di questo intervento pubblico. Ma questa è una possibile definizione di efficacia, perché spesso se ne incontra un'altra, altrettanto nobile, ma molto diversa dal punto di vista della sua misurazione, che è la capacità di raggiungere e coinvolgere i destinatari di una politica. Questo spesso viene definito come efficacia.

In un ambito più organizzativo, l'efficacia è spesso citata come capacità di servire tutti gli utenti che domandano un servizio, . L'assenza di razionamento è uguale all'efficacia, se li serviamo tutti il servizio è efficace. Il grado con cui ci invece si presentano delle code è il grado in cui il servizio è inefficace. Questa nozione nobilissima è molto diversa dalla precedente. Spesso efficacia viene definita come capacità di soddisfare le aspettative degli utenti che ricevono il servizio, la *customer satisfaction*.

Più spesso ancora in ambito gestionale efficacia è la capacità di rispettare gli impegni programmati. I comuni si danno degli obiettivi che non sono obiettivi di trasformazione, ma sono cose da fare, impegni che un dirigente si prende e il grado con cui gli impegni sono stati rispettati viene spesso definito efficacia.

Tutto perfettamente legittimo, e l'unico denominatore comune a queste accezioni è la capacità di raggiungere gli obiettivi. Ma in che senso raggiungere gli obiettivi? A seconda del contesto cambia il significato del termine obiettivo, quindi obiettivo può voler dire produrre un effetto - questo era l'obiettivo della politica - raggiungere destinatari, servire tutti gli utenti, soddisfare gli utenti e

rispettare gli impegni. Si potrebbe abolire il termine efficacia e sostituirlo con: impatto, grado di copertura, assenza di razionamento, grado di soddisfazione degli utenti e realizzazione degli impegni programmati. Queste sono le cose di cui realmente parliamo.

Invece continua l'alone di mistero che circonda la valutazione, l'uso di termini pass partout che danno questo falso senso di sicurezza. La sollecitazione che io porto è quella di cominciare a ragionare in modo più critico su queste cose. Quando parliamo di efficacia che cosa intendiamo realmente? Quali sono i meccanismi che poi portano a dei miglioramenti dell'efficacia? Bisogna uscire dal generico ed essere più specifici. Non è pensabile valutare l'efficacia in ambiti così diversi utilizzando indicatori di efficacia.

Sarebbe troppo bello che cose così diverse potessero essere misurate e apprezzate usando lo stesso strumento. Per valutare – io direi verificare – il rispetto degli impegni programmati - semplici indicatori di cose fatte possono essere sufficienti, dovevi fare dieci cose, ne hai fatte cinque, la tua efficacia è del 50%. Io non parlerei di efficacia in questo caso perché rispettare impegni e tabelle di marcia è fondamentale per una buona gestione, siamo assolutamente tutti d'accordo su questo, ma non significa fare cose utili, fare cose gradite e fare cose efficaci. Significa rispettare gli impegni, quindi chiamiamolo rispetto degli impegni. Invece all'estremo opposto per valutare la capacità di un intervento di ottenere gli effetti desiderati, è impensabile usare semplici indicatori, - avete visto in quel caso delle politiche del lavoro che assurdità si tendono a dire quando si appiattisce l'analisi all'uso di semplici descrittori - ed è invece necessaria una analisi di impatto che utilizzi gli strumenti analitici adeguati al caso. Ad ogni esigenza conoscitiva va accostato lo strumento adeguato.

Morale della favola: secondo me non ci porta da nessuna parte questa tendenza ad invocare la valutazione come se fosse una bacchetta magica, che costa poco e produce magicamente risultati in tutte le situazioni, come dicevo prima, questo porta solo alla disillusione. Prima o poi le bolle scoppiano tutte, come ben sa chi ha investito in borsa negli ultimi anni, e scoppierà prima o poi anche la bolla valutativa se non diamo un bel colpo di remi in termini di concretezza e di rigore.

Io ho quattro regole di buona valutazione da proporre.

Se vi chiedono di valutare chiedetegli cosa vogliono sapere, perché valutare non significa nulla, bisogna sapere che cosa si vuole ottenere con la valutazione. Diversamente si continua a fare aria fritta. Chiedetegli il perché vogliono sapere questa cosa, e soprattutto chiedetegli quale meccanismo di miglioramento la valutazione può innescare. Interrogiamoci sul perché questa cosa è utile non perché lo dice il decreto 286. Scusate il mio atteggiamento poco riverente nei confronti della legislazione, ma il fatto che lo dica il decreto 286 a me non convince neanche un po'. Voglio capire a che serve, non mi basta il fatto che me lo dica una legge.

Se vi parlano di dati ricordate che i dati non sono mai dati, ma sono presi. Quindi è un alibi dire "non ci sono i dati". Non è vero che non ci sono i dati. Se vuoi fare una valutazione, se vuoi

confrontare, vuoi capire, vuoi rendicontare, vuoi individuare anomalie, ti andrai, col tempo dovuto e le risorse dovute, raccogliere i dati che servono. I dati uno se li va a prendere, non è che non ci sono i dati, perché quella domanda a cui tu vuoi rispondere è probabilmente la prima volta che viene posta, e quindi nessuno ha pensato a raccogliere questi dati, quindi non sono “dati”.

Quando vi chiedono di valutare gli effetti di qualche cosa, chiedetegli se si rendono conto di quanto è difficile valutare effetti. Io sono stupefatto dalla facilità con cui questo termine viene usato. Tutte le scienze sociali, la sociologia e l'economia hanno da sempre al centro questo grosso rovello, come si individuano legami di causa-effetto tra fenomeni. Nella valutazione invece magicamente la valutazione degli effetti è diventata una sorta di barzelletta. No, un attimo, valutare effetti vuol dire individuare legami di causa-effetto tra uno stimolo che è rappresentato dall'azione pubblica e un cambiamento osservato, ma è maledettamente difficile, non basta prendere 2-3 numeretti. Diversamente avete visto cosa accade: avete visto in che trappole concettuali sono finiti quelli che hanno scritto il QCS - il rapporto tra gli occupati precari e gli occupati totali misura il successo delle politiche alla flessibilizzazione, è un pochino più complesso. Quindi quando vi dicono valutare gli effetti chiedetegli "ma vi rendete conto di quanto è difficile?"

Non dimenticate che la principale tecnica di valutazione, alla fine, è il buon senso, non vi fate incantare da statistici ed econometrici, da gente che propone effetti speciali, perché nella gran parte delle situazioni l'ingrediente principale non sono i dati ma il buon senso e soprattutto il primo punto: chiedersi perché si vuole valutare e perché dovrebbe migliorare le cose.

CAIANIELLO (testo ridotto)

Se mi è consentito, solo 30 secondi sul discorso del professor Martini (malgrado la grande soggezione in cui mi ha messo): Vorrei dire anzitutto che chi parte tardi come nel caso italiano ha il vantaggio di vedere gli errori fatti da quelli che l'hanno preceduto per evitarli. Una seconda considerazione è questa: è bene mettere in guardia, come ha fatto il professor Martini, contro ingenui ed euforici ottimismo; purché – e credo che siamo tutti d'accordo – ciò non induca scetticismo, occorrendo vedere in quale momento e contesto storico stiamo parlando. Ora, nel nostro paese, un pericolo grosso, secondo me, è quello costituito dalla resistenze di apparati e burocrazie tradizionali, ed è quindi altrettanto necessario chiarire che le indubbie difficoltà di attuazione delle nuove metodologie non debbono offrire alibi e pretesti a quelle resistenze.

La parola alla Prof. Stame.

Professoressa STAME

Conosco il professor Martini da un po' di tempo e posso dirvi che in precedenti occasioni non ha mai riempito tanto il bicchiere come questa volta, smentendo le aspettative. Oggi Martini ha detto moltissime cose su come la valutazione potrebbe essere fatta bene e in modo utile; altre volte l'ho sentito parlare soprattutto di quanto le cose andassero male. E credo che non sia casuale, perché in

realtà egli ha riflettuto sulla situazione che si è creata per il modo in cui la valutazione è arrivata da noi (ossia per imposizione dei Fondi Strutturali e delle riforme amministrative), e si è posto delle domande relative a cosa si può fare per contrastare certe tendenze negative che hanno accompagnato questo ingresso della valutazione.

Io vorrei fare solo un breve commento. Nel suo intervento la parola che è ricorsa di più è stata "meccanismi", una parola chiave nel dibattito internazionale sulla valutazione in questo momento, dominato dagli approcci alla "valutazione basata sulla teoria", che ruotano intorno all'idea di meccanismo causale (Carol Weiss) e di meccanismo e contesto (Ray Pawson). Martini ha usato due accezioni di meccanismo:

- a) il meccanismo attraverso cui opera la valutazione (ed è stato l'esempio degli insegnanti e dello "shaming");
- b) il meccanismo causale che all'interno di un programma fa sì che una situazione possa cambiare nel modo desiderato.

Il secondo è il tema su cui si è sempre ragionato di più. La valutazione è un tipo di ricerca al servizio dell'amministrazione pubblica per cercare di capire come le proposte legislative e i cambiamenti amministrativi influiscano su una realtà sociale che è molto diversa da una zona all'altra. Secondo la valutazione realistica (Pawson e Tilley), per ottenere determinati risultati auspicati (outcome) occorre analizzare la situazione nella quale ci si trova (contesto) e capire in che modo esso reagisca agli strumenti della politica (meccanismi).

Recentemente al congresso europeo della valutazione che si è tenuto a Siviglia, Ray Pawson parlando dei meccanismi che sono alla base delle politiche e dei programmi si è rifatto al titolo di un libro famoso che si chiama "*Carrots, sticks and sermons*" (carote, bastoni e prediche) e ha detto che tutto sommato i meccanismi principali dei programmi si possono ridurre a questi tre, perché la maggior parte dei programmi mirano o a indurre a fare qualche cosa, o a reprimere qualche cosa che si vuole evitare, oppure a cercare di convincere. Ma se i meccanismi sono abbastanza semplici, cosa è che cambia? cambiano i contesti, cambiano le situazioni nelle quali si opera; e, si badi bene, il contesto non cambia semplicemente perché una città del nord è diversa da una del sud, una grande da una piccola oppure - come si dice abitualmente - sono tutte diverse. I contesti variano perché ci sono delle realtà che reagiscono diversamente al cambiamento, alcune sono portate al cambiamento perché aspirano a inserirsi all'interno di una realtà, altre sono meno portate perché non vogliono cambiare e oppongono strategie di resistenza. E in ciascuno di questi diversi contesti farà differenza se si usa un bastone o una carota, o se ci si limita a una predica.

Questo modo di pensare ai programmi tiene conto, tra l'altro, della difficoltà a chiarire gli obiettivi di cui parlava Martini. Nell'ambito della valutazione dei programmi molti dicono – e secondo me ciò è alla base di tutte le difficoltà cui accennava Martini - "so cosa sono gli obiettivi, ora provo a misurare se li ho raggiunti o meno". Ma non basta volere un cambiamento in una certa direzione, bisogna sapere che quel cambiamento non potrà ottenersi in modo conforme nelle diverse realtà; quindi, se voglio un certo cambiamento devo sapere che in una situazione si potrà ottenere

attraverso un meccanismo e in un'altra situazione attraverso uno diverso; solo articolando meccanismo e contesto posso capire se l'obiettivo di un cambiamento desiderato può essere raggiunto nella realtà.

Torniamo ora al primo modo in cui Martini ha usato il meccanismo. Qui ha detto cose molto interessanti e nuove, a cui vorrei collegarmi. Mi è sembrata molto utile la distinzione tra i diversi meccanismi attraverso cui opera la valutazione: premiare ciò che ha funzionato (e che corrisponde a ciò a cui si aspira) o correggere le anomalie facendo vergognare chi è rimasto indietro.

Io penso che, presi insieme, questi due modi di pensare al meccanismo si riferiscano anche a diversi approcci alla valutazione, che a mio parere si distinguono a seconda del criterio in base al quale si dà un giudizio sull'azione oggetto di analisi, e che sono più facilmente riconducibili alle situazioni da valutare. Mi riferisco alla differenza tra dare un giudizio in base ad un criterio di qualità (uno standard) o in base ad un obiettivo: il criterio di qualità indica qualcosa a cui ci si può avvicinare, in una scala ordinale, mentre il raggiungimento di un obiettivo indica che si è ottenuto un cambiamento, qualcosa che prima non c'era (non importa se poi questo cambiamento possa essere espresso in una forma quantitativa, di aumento o diminuzione di qualche valore). Questi due criteri si riferiscono a situazioni diverse: lo standard si riferisce alla erogazione di una prestazione ordinaria, in un servizio, mentre l'obiettivo si riferisce ad un programma straordinario. Solitamente però si confonde tra le due cose, e lo si fa anche nella recente legislazione quando si dice: i dirigenti vengono valutati sul fatto se hanno ottenuto gli obiettivi. Ma è molto diverso dire se i dirigenti hanno perseguito il miglioramento di un servizio o se hanno ottenuto dei risultati di cambiamento corrispondenti agli obiettivi di un programma. Ad esempio, se valuto il servizio fornito dal sistema della istruzione superiore, devo vedere se una università funziona meglio, perché ad esempio è migliore il rapporto docenti-studenti; se valuto un servizio ospedaliero, devo vedere se il servizio è migliore perché ci sono meno code, ecc.. Invece in un programma vogliamo fare delle cose nuove che prima non si facevano, introdurre un nuovo modo di insegnare, un nuovo curriculum, un nuovo tipo di servizio; la parola "programma" va insieme a "innovativo", i programmi europei premiano l'innovatività. E quindi per valutarlo devo capire cosa è stato fatto e cosa si è ottenuto.

Nel primo caso c'è tutta la problematica di come si definiscono gli standard: non potrei essere più d'accordo di quello che ha detto Martini in proposito. Nel secondo caso bisogna ragionare sugli obiettivi e su come si riarticolano, e soprattutto bisogna tenere in conto tutte le forze interne al programma ed esterne che concorrono al risultato (i famosi problemi di attribuzione) e smetterla con l'aspirazione alla generalizzabilità, secondo cui basterebbe avere trovato che un risultato è stato raggiunto da una parte per poterlo realizzare dappertutto.

Questi sono i problemi che ci si deve porre avvicinandosi alla valutazione, in modo che essa diventi una abitudine, al di là del fatto se è un obbligo. Io sono d'accordo con Martini nella sua polemica contro l'istituzionalizzazione, ma non perché penso che non bisogna creare nuclei di valutazione (non facciamo altro che insegnare, creare competenze valutative, dire qui ci vorrebbe

un nucleo di valutazione), ma perché la cosa importante che devono saper fare le persone incaricate del ruolo della valutazione è esattamente rispondere alle domande della “valutazione modesta” contenute nell'ultima slide, ossia “cosa vogliamo sapere”, “perché la vogliamo sapere”, “quale meccanismo di miglioramento la valutazione può innescare”, “se ci si rende conto di quanto è difficile valutare gli effetti” ecc.

Questo modo di pensare è molto distante dal modello di valutazione che ci viene presentato attualmente nelle amministrazioni pubbliche in Italia. Penso in particolare al proliferare di linee guida, alle idee del “mainstreaming”, ecc., che partono tutte dalla convinzione che una volta identificato che qualche cosa vada bene poi dopo si possa riprodurre da tutte le parti. Il concetto più ricorrente è quello di *best practice*: è uno dei modi come si fa *benchmarking*: si cerca di identificare l'esempio migliore per poi eventualmente premiarlo, e additarlo come esempio. Tuttavia, la *best practice* viene spesso individuata come modello di un'azione che si vuole imporre a tutti quanti, e in base al quale si giudica chi è migliore e chi peggiore. Ma poiché nella realtà le situazioni (i contesti) sono molto diversi, una pratica considerata buona (non: migliore) da una parte può solo costituire un buon esempio da cui trarre insegnamenti da adattare in una situazione diversa, niente più. C'è una grande differenza tra proporre una *best practice* che sia qualcosa che debba essere generalizzata a tutti quanti, e consentire alle persone che sono situate in diversi ambienti di imitare, prendere l'esempio positivo che hanno visto da qualche altra parte.

Ciò rimanda a una concezione diversa di amministrazione pubblica: o imporre dall'alto un modello che sia uguale per tutti, o dare a ciascuno nella sua situazione la possibilità di capire cosa funziona, cosa si vuol fare nel proprio ambiente e applicarlo. Qui entrerebbero anche certe aspirazioni alla democrazia, all'empowerment, ma bisogna fare attenzione a come si usa il concetto di partnership, perché mi pare che se ne faccia un certo abuso, come ad esempio nelle politiche sociali. Se voi leggete la legge 328/00 (sui servizi sociali), la 285 (sull'infanzia), tutto quello che riguarda i servizi sociali, in particolare se vengono dati in outsourcing all'esterno, non si fa altro che chiedere partecipazione, dalla decisione alla valutazione, in un modo completamente irrealistico. La cosa importante invece sarebbe quella di consentire alle persone che sono all'interno di determinate situazioni di dotarsi di strumenti di conoscenza che le mettano in grado di applicare creativamente quello che hanno visto fare altrove o di migliorare quello che sanno fare: in sostanza, tra i tanti possibili modi di fare qualcosa adattare alla propria situazione ciò che può essere più utile e più congeniale, al fine di ottenere il miglioramento delle politiche, o un maggior rendimento delle amministrazioni. Quindi occorre modestia – come auspica Martini – nel senso che so che devo migliorare, però devo cercare di migliorare a seconda del contesto nel quale mi trovo e delle differenze di situazioni nelle quali mi trovo.

Interventi (sintesi redazionale)

CHAINIAN - *Riferisce della esperienza sul terreno in corso da parte della Camera di commercio di Treviso, di cui è Segretario generale, nella valutazione degli effetti prodotti dalle iniziative e dagli interventi dell'Ente, in termini di miglioramento -secondo le loro finalità istituzionali- dello*

sviluppo economico del territorio, osservando come, pur con la prudenza dettata dalla non sempre facile individuazione di tali effetti in concreto, una valutazione per quanto imperfetta e perfettibile resta pur sempre preferibile all'alternativa di rinunciarvi del tutto.

TARADEL- *Come Direttore della Rivista Trimestrale di Scienza dell'Amministrazione, fondata da Giuseppe Cataldi -che Caianiello ha ricordato come Presidente della Corte dei conti noto anche all'estero quale studioso e promotore fin dagli anni cinquanta del 1900 delle nuove metodologie di controllo- comunica l'intenzione di tenere un convegno celebrativo del cinquantenario della rivista, centrando gli interventi proprio su Cataldi e sugli studiosi che lo hanno circondato formando la galassia della mancata riforma dell'amministrazione (Lucifredi, Lessona, fino a Benvenuti e Giannini, oltre a M. Amendola). Concorda con i richiami di Martini alla prudenza, perché non si veda nella valutazione lo strumento per risolvere tutti i problemi, gravissimi in un'amministrazione pubblica come quella italiana che vanta un ritardo più che secolare rispetto ad altre amministrazioni, proprio nelle cose modeste. Ritardo riconducibile al declino delle concezioni e delle prassi pur presenti -anche nella Corte dei conti- nei primi anni dell'Unità nazionale, declino sul terreno dell'organizzazione, cui possono attribuirsi anche disastri militari, come quello di Adua, fino al marasma attuale, essendo fra l'altro scomparsa l'essenziale autovalutazione che spetta agli stessi dirigenti, ed apparendo perciò ancora necessario un autentico rivolgimento di cultura e mentalità*

CAIANIELLO *rileva che neppure la Corte dei conti sembra finora essersi dotata di un consistente sistema di autovalutazione.*

MELONI, *Consigliere della Corte dei conti- Rifacendosi alla recente riforma costituzionale di potenziamento delle autonomie, ritiene che la valutazione spetti essenzialmente ai livelli locali, con riferimento in particolare alla customer satisfaction. Tipico, a suo parere, il caso della pubblica istruzione, specialmente per le esperienze rese possibili dall'autonomia universitaria, mentre a renderlo preoccupato e un po' scettico, a livello centrale, sta fra l'altro una perdurante interferenza della politica nell'amministrazione, a danno dell'autonoma responsabilità da assicurare alla dirigenza.*

Professoressa STAME

Io avrei una obiezione rispetto al problema sollevato dal dott. Meloni, ossia al dove si valuta, se in alto o in basso. Secondo me si valuta dappertutto, ma si valutano cose diverse. Quando parliamo di valutare delle politiche e dei programmi dobbiamo fare una valutazione a livello nazionale, addirittura a livello europeo. La questione riguarda cosa valutiamo e con quale intenzione. Possiamo valutare una politica di incentivi per l'occupazione nel sud, nella zona obiettivo 1, oppure in quella obiettivo 2: queste sono valutazioni a livello nazionale. A livello locale, invece, ci sono altri programmi. Il problema poi è darsi gli strumenti conoscitivi adatti. Uno di questi è la *customer satisfaction*, e il problema è di usarla al posto giusto e nel modo giusto.

Il dottor Meloni identifica valutazione e *customer satisfaction*: io non sono d'accordo. La *customer satisfaction* è uno dei tanti modi per cercare di capire se un intervento ha avuto un risultato desiderato, e la si fa quando si valuta la erogazione di una prestazione ad un utente. Certamente a livello centrale non si farà la *customer satisfaction* perché non si valuta l'erogazione di un servizio; ma si useranno altri strumenti per valutare gli effetti di un programma. Io penso che la *customer satisfaction* sia uno strumento molto delicato, poco adatto perfino nel suo uso canonico, perché è fatta in modo tale per cui normalmente la gente dice che è contenta dei servizi anche quando tutti sanno che sono molto scadenti (il noto "paradosso della soddisfazione"). E' uno dei tanti modi di valutare gli effetti, e come ha detto Martini nella sua conclusione - anche se non vi si è troppo addentrato - "se vi parlano di valutare effetti chiedete se si rendono conto di quanto sia difficile". Ora, quando una cosa è difficile, le strade sono due: o l'abbandoniamo o cerchiamo di farla con i migliori mezzi che abbiamo. Al congresso dell'AIV di quest'anno abbiamo premiato una tesi di laurea che metteva in luce i limiti della *customer satisfaction* e suggeriva metodi per migliorare l'attendibilità dei dati.

Professor MARTINI

Soprattutto, per stare in questo ambito della *customer satisfaction* ci sono due atteggiamenti, c'è quello di dire: faccio la *customer satisfaction* perché così mi salvo l'anima, un'altra cosa è fare la *customer satisfaction* per uno scopo di miglioramento. Però bisogna individuare un meccanismo, non basta chiedere agli utenti se sono contenti oppure no. Nel caso della nostra *customer satisfaction* all'Università abbiamo individuato un meccanismo, che poteva funzionare o non poteva funzionare, ma noi avevamo un'ipotesi che facendo vedere ai singoli docenti i risultati della loro prestazione rispetto a quella di tutti i loro colleghi, quelli che erano sotto la media si vergognassero.

Invece *customer satisfaction* spesso si fa semplicemente così, facendo un questionario, e per giunta fatto in modo tipicamente per ricevere sempre le stesse risposte, il 30% dice che è molto soddisfatto, il 50% che è abbastanza soddisfatto e il 20% che non è per niente soddisfatto. E poi tutto finisce lì. Ma a che serve? Quello che c'è da chiedersi è: perché facciamo la *customer satisfaction*? la facciamo perché abbiamo individuato un meccanismo di miglioramento o la facciamo perché è un adempimento, o un abbellimento? Io non sto dicendo che le cose bisogna farle perfettamente o non farle, sto dicendo una cosa molto diversa: che bisogna rendersi conto perché si fanno, e mettere in evidenza come, anche fatte imperfettamente, possano servirci a migliorare, se no diventa un alibi "ho fatto *customer satisfaction*", e ci prendiamo in giro a vicenda.

Professoressa STAME

Posso approfittare per fare una pubblicità? Vi vorrei presentare questo libro, che si chiama "*International Atlas of Evaluation*" ("*Atlante internazionale della valutazione*") ed è stato scritto

da un gruppo di persone che costituiscono un network internazionale, International Evaluation Research Group. Questo volume è stato curato da Jan Eric Furubo, Ray Rist e Rolf Sandhal. Furubo è un membro della Corte dei Conti svedese, così come Sandhal; Rist è stato al General Accounting Office, e in questo momento lavora nell' Operation Evaluation Department della Banca Mondiale. Io personalmente ho scritto il capitolo sull'Italia.

Vengono presentati 25 casi di paesi diversi che sono divisi in questo modo. La prima parte riguarda paesi che hanno introdotto la valutazione grazie a una forte pressione esterna, ma anche a una forte pressione interna: sono paesi che hanno cominciato da soli ma poi hanno ricevuto una ulteriore pressione esterna, come Danimarca, Finlandia, Francia, Germania, Olanda, Svezia e Regno Unito. La seconda parte riguarda i paesi con una debole pressione esterna e una forte pressione interna: sono Australia, Canada, Corea (sì, anche la Corea, per qualche motivo che chi leggerà l'articolo avrà piacere di capire), Norvegia e Stati Uniti. Sono sostanzialmente i paesi in cui si è cominciato e da cui poi si è sviluppato tutto il resto. La terza parte riguarda i paesi – tra cui rientra anche l'Italia - che hanno una forte pressione esterna e una più debole interna, e questi sono i paesi che sono stati spinti o dall'Unione Europea (come Irlanda, Italia e Spagna), oppure dalle agenzie di aiuti internazionali, come la Banca Mondiale (come Cina e Zimbabwe). Infine ci sono paesi con una debole spinta esterna e una debole interna, come Israele, Giappone, Nuova Zelanda e Svizzera. Questi 25 casi danno poi luogo a dei confronti interessanti. Mi piacerebbe che questo libro fosse conosciuto all'interno del vostro dibattito, sia perché esso rappresenta una delle presentazioni più complete dello stato dell'arte internazionale sulla istituzionalizzazione della valutazione, sia perché chi lo ha scritto proviene prevalentemente da una esperienza analoga a quella di molti appartenenti al Cogest; all'inizio erano prevalentemente membri delle Corti dei Conti, e auditors; poi si sono aggiunti altri studiosi interessati a questi temi.