

VALUTARE O MISURARE I RISULTATI? IL CASO DELL'UNIVERSITÀ*

Eliana Minelli, Gianfranco Reborà, Matteo Turri

“Perocché di qual giudizio voi giudicherete, sarete giudicati; e della misura che voi misurerete, sarà altresì misurato a voi” (San Matteo, VII)

Ambiguità della misurazione

L'etimologia stessa del termine “misura” appare carica di ambiguità. L'italiano misura e l'inglese *measure* derivano palesemente dal latino *mensura* e quindi dal verbo *metiri* e dal suo participio *mensus*. Ciò sembra significare giudizio, misura in senso psicologico più che fisico, e anche prudenza. Quindi non tanto un richiamo a un metro obiettivo di calcolo ma l'esercizio ponderato e riflessivo di una valutazione. La radice dello stesso verbo *metiri* e dell'equivalente greco *metrein* rimandano però anche a *me-s*, *mese* e *month*, dove viene evocata una unità basilare di misura del tempo, che richiama anche *moon*, termine indicativo del ciclo lunare [Mari, 2003].

Del resto il titolo del dramma di Shakespeare *Measure for measure*, rende compiutamente questa ambiguità, rimandando anche implicitamente alla frase del Vangelo di Matteo riportata in epigrafe, dove i termini “giudizio, giudicherete, sarete giudicati” (piano di riferimento soggettivo) e “misurerete, sarà misurato” (piano di riferimento oggettivo) sembrano al tempo stesso contrapporsi, combinarsi e completarsi.

Questa consapevolezza di ambiguità è presente nell'ambito della stessa scienza delle misurazioni, al punto che un recente articolo su *Epistemology of measurement*, pubblicato da un nostro collega su una rivista internazionale, pone esplicitamente il problema: <<la misurazione è una valutazione soggettiva o una descrizione oggettiva?>> [Mari, 2003].

L'ambiguità è del resto condizione normale del funzionamento di tutte le organizzazioni, proprio perché diversi significati e interpretazioni possono essere attribuiti ai loro input e output, come alle correnti di attività che le attraversano [Weick, 1979]. E ancora di più lo è per almeno una parte delle sfere di attività che sono di pertinenza del management pubblico [Noordegraaf e Abma, 2003].

* Il presente paper è stato presentato dagli autori al VI Workshop dei Docenti e Ricercatori di Organizzazione Aziendale “Organizzare a Misura d’Uomo”. Università Cattolica del Sacro Cuore Milano 3-4 febbraio 2005

Scopo del nostro paper è quello di affrontare la tematica della misurazione dei risultati delle organizzazioni, con particolare attenzione per quei contesti caratterizzati da una limitata importanza del mercato e da condizioni di forte complessità e di valenza relazionale del prodotto o servizio. L'università, e in particolare le sue attività didattiche e di ricerca, sono anzi il riferimento empirico primo che si intende affrontare. Questo in ragione anche dell'attenzione che sta riscuotendo il tema della valutazione dei risultati come strumento di *governance* e *management* degli atenei.

Valutazione e misurazione

Per entrare correttamente nel tema è utile partire dal punto di vista della scienza della misurazione, nel cui ambito si possono incontrare riflessioni più approfondite su alcuni presupposti concettuali di quanto avvenga normalmente nelle trattazioni proprie delle discipline aziendali e manageriali.

Scontata l'ambiguità del termine misura e anche l'evoluzione storica di alcune concezioni di fondo, oggi sembra ragionevole parlare di misurazione nell'ambito di una scienza empirica quando siano presenti alcune condizioni e precisamente, in estrema sintesi, <<che le misurazioni siano valutazioni dotate di carattere intersoggettivo e oggettivo>> [Mari, 2003].

Occorre in sostanza che:

- i risultati della misurazione forniscano la stessa informazione a diversi osservatori;
- i risultati della misurazione forniscano informazioni legate solo alla cosa misurata e non al suo ambiente, che può includere anche l'osservatore, e che quindi misurazioni della stessa cosa in condizioni differenti e in presenza di differenti osservatori producano gli stessi risultati [Mari, 2003].

Questo significa considerare il ruolo dei sistemi di misurazione non solo come strumenti operativi ma come componenti essenziali che danno fondamento epistemologico alla misurazione.

In questo modo la misurazione si distinguerebbe dalla generica valutazione, che resta fondamentalmente una forma di giudizio soggettivo e assumerebbe, rispetto a essa, un superiore valore, in quanto valutazione i cui risultati possono in modo adeguato essere riconosciuti oggettivi. Secondo questa impostazione, la scienza e la teoria della misurazione hanno un futuro proprio nell'affrontare i problemi di formalizzazione della struttura e della funzione operativa dei sistemi di misura riferiti a fenomeni sempre più complessi.

Management, misurazione e valutazione

Il management però non è scienza, è un'attività pratica che contiene anche aspetti di arte. Si differenzia da altri tipi di attività intellettuale, perchè configura "un'arte pratica", più che "una scienza applicata" [Carlsson, 1951; Mintzberg, 1973]. E dare un'adeguata e oggettiva descrizione di un lavoro artistico è sempre arduo. Non si tratta solo di applicare principi ben

definiti a determinati casi, come può fare per esempio un ingegnere progettista. L'arte pratica non si può basare su principi ben definiti e sull'applicazione di una conoscenza sistematica, perché non ha un completo e sicuro fondamento di questo tipo. La conoscenza su cui si basa è frammentaria, non completamente definita, continuamente in lotta con la sfera del non conosciuto. Le regole di riferimento richiedono un'attiva interpretazione da parte del soggetto e quindi giudizio, intuizione e abilità personale di tipo creativo. I manager agiscono disordinatamente per assorbire l'ambiguità del contesto e per consentire al resto dell'azienda di operare con sistematicità; agiscono in un contesto di caos calcolato, dove occorre dare un'interpretazione convincente di un compito che non ha i confini fissati, i cui termini sono aperti (*open ended*) [Mintzberg, 1991]. La superficialità è il problema centrale che il management deve fronteggiare, la mancanza di profondità di pensiero e di azione che deriva dalla molteplicità di pressioni, dal carico di lavoro elevato, dalla necessità di gestire contemporaneamente diversi ruoli, progetti, contatti, informazioni. E in questo i manager non trovano aiuto nella concezione prevalente della professionalità manageriale come insieme di pratiche, di concetti, di strumenti spendibili indipendentemente dalla diversità dei contesti aziendali [Mintzberg, 1991; Mintzberg, 2004]. Questo riguarda sempre più anche l'analisi e l'interpretazione dei risultati aziendali, come mostra la diffusione di strumenti di analisi e valutazione dei risultati di carattere multidimensionale, che riflettono una poliedricità di riferimenti [Coda, 1989; Kaplan e Norton, 1996]. È vero infatti che per svolgere i propri compiti i manager hanno bisogno sempre più di disporre di strumenti di analisi, che riflettano i criteri di oggettività e intersoggettività cari ai cultori della scienza delle misurazioni. Ma hanno anche bisogno di prendere decisioni in condizioni di informazione imperfetta e incompleta; e non sempre la condizione migliore è di contare esclusivamente sui requisiti dell'intuizione, o del fiuto. In questo ordine di idee acquisiscono rilievo strumenti improntati alla logica della valutazione, che non è misura in senso scientifico ma frutto di un giudizio ragionato ed espresso, a volte, anche in modo formale. Da questo punto di vista, nell'ottica del management, non sembra accettabile considerare la misurazione uno strumento di qualità superiore alla valutazione, una sorta di riferimento cui la seconda dovrebbe tendere. In realtà, per molti fenomeni dove la componente umana e soprattutto l'interazione interpersonale assumono un ruolo centrale, dove si opera in regime di "caos calcolato", la valutazione, con tutto il suo carico di soggettività e di non perfetta comunicabilità, o di non trasferibilità in altri contesti, risulta priva di alternativa nella misurazione e anzi è ad essa preferibile. È lo stesso concetto espresso in altro ambito, quello della determinazione del valore economico del capitale aziendale, da un acuto studioso: <<i>metodi concettualmente più fondati e perciò più razionali risultano di fatto maggiormente influenzati da elementi soggettivi<>> [Paganelli, 1990].

Questo risulta ancor più fondato nel caso del management pubblico e, in modo particolare in alcune delle sfere di attività del settore pubblico, caratterizzate da processi al tempo stesso non routinari, sfocati (*fuzzy*), innovativi e conflittuali [Noordegraaf e Abma, 2003]. È proprio in questi contesti, dove ci si basa su pratiche definibili come "non canoniche", che la pretesa propria del movimento del *new public management* di risolvere i problemi di incertezza

attraverso metodi gestionali imperniati sulla misura della *performance* incontra limiti insuperabili. In questi casi, infatti, la definizione di obiettivi perde il proprio significato di riferimento e le relazioni tra fini e mezzi si dissolvono, al di là anche delle limitazioni chiarite dalla classica analisi di Simon su giudizi di fatto e giudizi di valore [Simon, 1947]: in sfere di azione come l'intervento per sviluppare l'occupazione, o salvaguardare l'ambiente naturale, o ancora limitare la congestione del traffico, l'interconnessione tra fatti e valori è talmente stretta da impedire una definizione di obiettivi e una conseguente misurazione di risultati che non sia sensibile a una divergenza di interpretazioni ed esposta a un pluralismo di visioni. Le relazioni tra input e outcome non possono essere rappresentate in modo stabile, i risultati ricercati sono spesso indiretti rispetto ai processi operativi attivati, alcune componenti fondamentali di outcome sono "invisibili", perché di lungo termine o interpretabili in maniere contrastanti e quindi la misurazione della performance diviene fortemente arbitraria. Le diverse misurazioni non sono comunque confrontabili nel tempo e nello spazio perché le pratiche sono uniche e in movimento, sono instabili.

Si dovrà discutere sulla presenza di queste condizioni nelle attività oggetto del nostro interesse. Ma in generale si può affermare, sulla scorta delle analisi svolte nell'ambito delle scienze manageriali, che il management non può fondarsi solo sulla misurazione, e nemmeno solo su questa più l'intuizione creativa e spontanea; ha bisogno del contributo anche di momenti di interpretazione e di valutazione, formalmente espressi e comunicati a livello intersoggettivo, anche se privi del requisito dell'oggettività: osservatori diversi non trarranno le stesse conclusioni elaborando gli stessi dati attraverso le stesse procedure.

La valutazione delle università e la misurazione dei risultati

L'università costituisce un ambito interessante per una riflessione sui rapporti tra management, valutazione e misurazione. Negli ultimi anni un movimento forte di riforma si è diffuso in tutta Europa e ha portato all'istituzione di organi, sistemi e procedure finalizzati alla valutazione dei principali risultati dell'attività universitaria, quali l'insegnamento, la ricerca e i diversi servizi aggiuntivi offerti agli studenti e agli altri interlocutori. Il concetto di valutazione dei risultati e quello di assicurazione della qualità sono divenuti riferimenti obbligati per i processi di riforma dei sistemi universitari, anche come necessario contrappeso dell'autonomia riconosciuta agli atenei. Molti paesi hanno istituito agenzie nazionali al fine di valutare l'output delle attività accademiche, sia pure in diverse forme [Felt, 2004]. Le politiche di regolazione delle università promosse dai vari governi si sono naturalmente combinate e integrate con sistemi di *performance management* e altre pratiche connesse al movimento del *new public management*, a partire da esperienze guida come quelle della Gran Bretagna.

Nell'ambito delle attività di valutazione, pur nella molteplicità dei metodi proposti e sperimentati, largo spazio è stato assunto dai tentativi di misurazione di aspetti specifici dei prodotti e dei servizi erogati.

Da più parti si è sostenuto il rilievo degli indicatori, variamente composti e combinati, come misure oggettive di efficienza ed efficacia delle diverse attività. Gli indicatori sono considerati anche possibile e naturale riferimento per un *ranking* degli atenei per livello di qualità e quindi per l'allocazione delle risorse pubbliche destinate alla formazione superiore.

In Italia a partire dal 1989 diversi interventi legislativi hanno valorizzato e promosso l'autonomia delle università, modificando gradualmente il preesistente assetto fortemente centralizzato. Queste stesse norme hanno progressivamente costruito un sistema nazionale di valutazione delle università. Il Comitato nazionale per la valutazione del sistema universitario (CNSVU) è l'organismo centrale della configurazione di tale sistema, cui spetta il compito di indirizzare le attività di valutazione. Autonomia e valutazione sono state quindi concepite come aspetti complementari, destinati a svolgere il ruolo di idee guida sulle quali impennare un complessivo processo di trasformazione.

In questa logica, la valutazione è concepita come strettamente collegata a un processo di programmazione e controllo dell'intero sistema e l'organismo nazionale ha cercato in primo luogo di attivare un ampio patrimonio informativo, sistemicamente concepito e comprensivo di banche dati quantitative e di indagini qualitative. Lo sforzo è stato di mettere a disposizione degli organi di governo del sistema, ma anche degli atenei, degli studenti e delle loro famiglie, dei terzi interessati, uno strumento di informazione trasparente e affidabile, al tempo stesso tempestivo e alimentato in continuazione.

Le stesse informazioni sono state utilizzate anche al fine del riorientamento del sistema di finanziamento degli atenei, segnatamente in una logica di riequilibrio delle dotazioni rispetto ai parametri considerati più significativi del volume e della qualità dell'attività svolta. Questa finalità, imperfettamente e solo parzialmente realizzata in una prima fase, è divenuta uno degli obiettivi fondamentali dello sviluppo dell'attività valutativa. Il CNSVU ha elaborato documenti che propongono una radicale revisione del sistema di finanziamento degli atenei, e sta sempre più orientando i suoi sforzi verso l'esempio offerto dai sistemi di finanziamento e di governo delle università inglesi che, riprendendo le parole di uno dei membri del CNSVU, <<sono minuziosamente governate attraverso un'oculata distribuzione di fondi basata su parametri resi noti in anticipo>> [Figà Talamanca, 2001]. In questo senso sono del resto orientate anche le indicazioni del Ministro dell'università e del Governo nazionale.

È così che in un numero speciale della rivista del MIUR dedicato alla valutazione e direttamente curato del CNSVU si afferma: <<...raccogliere i dati non risponde più solo alla necessità inderogabile di avere una conoscenza globale del sistema che prima non si aveva (...) il compito prioritario è quello di capire come i dati possano servire a una efficace gestione dell'autonomia e al governo centrale delle risorse pubbliche>> [CNSVU, 2002a].

Lo stesso CNSVU tende però a presentare il proprio ruolo anche in un diverso modo, non come quello di <<controllori, o di valutatori esterni, ma come accompagnatori, e questo dovrebbe sgombrare il campo da sospetti, da accuse di essere coloro che danno voti, che fissano i paletti entro i quali gli altri debbono lavorare>> [CNSVU, 2002b].

D'altra parte, l'esperienza dei paesi che hanno affrontato da più tempo e con maggiore determinazione questa strada, l'Inghilterra in primo luogo, mette in luce la difficoltà e la patente contraddizione di rendere compatibili le due diverse logiche della spinta per l'efficienza attraverso forme di regolazione e intervento dall'alto e lo schema dello sviluppo professionale e organizzativo gestito nelle realtà di base e tendente al miglioramento della qualità e dell'efficacia [Pollitt, 1987; Barzelay, 1996].

In ottica di confronto tra le diverse esperienze internazionali, si riconosce ormai che i processi di assicurazione della qualità non funzionano allo stesso modo dentro le università e all'interno delle imprese [Felt, 2004]. In linea generale i risultati dell'università non restano al suo interno ma vengono considerati a livello politico e hanno un'ampia ripercussione [Felt, 2004]. A maggior ragione questo avviene quando questi stessi risultati vengono collegati a processi formali che influiscono sull'allocazione delle risorse, oppure a forme di *ranking* che posizionano in modo differenziato le diverse università in classifiche che vengono portate a conoscenza del pubblico. È più che naturale in tale contesto che le università sviluppino strategie per nascondere le proprie debolezze e per enfatizzare i propri punti di forza: ma si tratta di comportamenti che ostacolano e vanificano i presupposti di un aperto processo di miglioramento della qualità [Felt, 2004].

In Inghilterra, il sistema che sembra divenire riferimento primo per le più recenti esperienze italiane ha suscitato violente critiche; e viene considerato espressione di uno stato regolatore e di una società dei controlli che, per ragioni di efficienza e di economia di risorse, cercano di erodere la tradizionale autonomia delle istituzioni accademiche [Brown, 2004]. Comunque si voglia valutarlo, il sistema inglese rappresenta un tentativo di regolare l'istruzione superiore inducendo una sorta di mercato simulato: in questo l'uso estensivo di parametri e misurazioni svolge il proprio ruolo, che è anche quello di massimizzare l'informazione diffusa sui risultati delle diverse università con trasparenza e oggettività.

Tuttavia, questo modo di procedere sembra incapace di superare alla fine una contraddizione di fondo: l'esame attento delle metodologie e degli strumenti di valutazione impiegati mostra che nessuno di questi in pratica corrisponde al duplice criterio di avere una valenza di vera misurazione, nel senso indicato da Mari, e di avere un significato pienamente difendibile dal punto di vista finalistico.

È chiaro infatti che una buona parte delle tecniche utilizzate nelle diverse esperienze europee per la valutazione di didattica e ricerca si traduce nell'espressione formale di parametri quantitativi che però non costituiscono misure dotate di carattere intersoggettivo e oggettivo. Non succede in sostanza che i risultati della misurazione forniscano la stessa informazione a diversi osservatori e che misurazioni della stessa cosa in condizioni differenti e in presenza di differenti osservatori producano gli stessi risultati [Mari, 2003].

Questo vale ad esempio per tutta la famiglia delle metodologie imperniate sulla *peer review*, applicata sia alla didattica che alla ricerca, che hanno ormai larghissimo impiego in tutta Europa. Gli esperti esterni, che compiono visite e ispezioni, esaminano dati e pubblicazioni, assistono a momenti rappresentativi dell'attività da valutare, possono seguire istruzioni

dettagliate, valersi di linee guida omogenee, utilizzare manuali di riferimento, partecipare a momenti di formazione comuni, rispondere a organi e ruoli di coordinamento. Tutto questo esprime una tensione a oggettivare le valutazioni ma il cuore stesso di questo metodo consiste nell'affidarsi in ultima istanza all'esperienza e capacità di giudizio delle persone selezionate come *peer* e quindi esiste un limite al grado di coerenza dei diversi giudizi [Clark, 1996]. L'analisi dell'esperienza empirica inglese ha mostrato che le contromisure adottate non bastano a garantire l'uniformità di giudizio, a maggior ragione quando vengono valutate sfere disciplinari diverse, nella ricerca o nella didattica. Non vi è garanzia che tali riferimenti portino a costituire sistemi di misurazione in senso proprio, al massimo si può arrivare ad ottenere una soddisfacente coerenza delle valutazioni espresse dallo stesso gruppo di *peers*, o che riguardano la stessa istituzione [Underwood, 2000; Brown, 2004].

Tutto questo è ancora più palese per l'altra rilevante famiglia di metodologie, quella imperniata sulle valutazioni, in questo caso dichiaratamente soggettive, da parte degli studenti, o altri destinatari delle diverse attività. Nel prossimo paragrafo si tratterà in particolare delle problematiche sottese alla valutazione della didattica da parte degli studenti, probabilmente la più diffusa applicazione di questo metodo.

I rischi insiti nell'attribuire, con una forzatura, una valenza oggettiva a valutazioni che hanno diversa natura sono ben rappresentati dalla polemica sviluppatasi in Inghilterra sulla scelta di preferire una logica di gradazione ad una di soglia. Se il giudizio, riguardante un'istituzione accademica o un suo programma o progetto, si limita a riconoscere il conseguimento o meno di un livello-soglia di qualità considerato come requisito necessario, i problemi di oggettività della misurazione vengono contenuti se non minimizzati. L'insistere invece, come si è fatto in Inghilterra, su una graduazione dei giudizi, che porta di conseguenza a istituzionalizzare forme di *ranking*, enfatizza i problemi di attendibilità e difendibilità dei giudizi così espressi. Quando si perviene a giudizi sommativi, univocamente espressi su una scala quantitativa, si finisce inevitabilmente per attirare l'attenzione su quella parte della valutazione che ha meno valore e validità, per semplici ragioni tecniche, con effetti ben prevedibili sulla credibilità del sistema adottato presso gli stessi valutatori, presso i valutati e presso gli altri possibili destinatari dei report di valutazione [Sharp, 1995].

Per quanto riguarda, invece, quella parte delle metodologie adottate che rispondono ai requisiti di una vera e propria misurazione, emergono i limiti di un'eccessiva semplificazione. I parametri e gli indicatori quantitativi, comunque scelti, non si rivelano sufficientemente significativi come espressioni generalizzabili di quelle caratteristiche che gli organi regolatori vorrebbero valorizzare.

Vale in questo senso la critica di Trow (1994), secondo cui:

may well be that vigorous effort by agencies of central government to assist the quality of university work lead to his decline, as more and more energy is spent on bureaucratic reports, and as university activities themselves begin to adapt to the simplifying tendencies of the quantification of outputs. Our research suggests that departments and individuals shape their activities to what 'counts' in the assessments, to the impoverishment of the life of university, which is always more complex and varied than assessment of 'outputs' can capture.

Alla fine risulta convincente la conclusione di Power (1997), che vede la messa in atto di pratiche formali di assessment dal punto di vista della legittimazione degli organi decisionali. L'estensione di queste pratiche nei differenti contesti e soprattutto nelle attività professionali più complesse, come istruzione e sanità, più che una risposta tecnica reale a problemi di responsabilità e di management, contribuisce a razionalizzare e rinforzare un'immagine pubblica di controllo. La presentazione del fondamento di queste pratiche in misurazioni oggettive risponde in pieno a questa esigenza.

E se queste critiche valgono per esperienze consolidate e perseguite con determinazione, come quella inglese, a maggior ragione calzano per approcci ancora iniziali e ingenuamente condotti come quelli italiani degli ultimi anni. Qui è patente il compromesso e il tentativo di mascherare le contraddizioni attraverso comportamenti degli organi centrali di valutazione che oscillano tra lo stile autoritario e quello paternalistico.

Le valutazioni degli studenti sulla didattica

Quanto affermato al paragrafo precedente vale in misura accentuata per le valutazioni espresse dagli studenti sulla qualità della didattica (SRT, *Student's Rating of Teaching*), ormai sancite in Italia anche dalla legge. Si tratta del principale strumento affermatosi in tutto il mondo per la valutazione della didattica e di quello maggiormente enfatizzato nella comunicazione verso gli studenti e verso la pubblica opinione.

D'altra parte, l'adozione di questi metodi, almeno nella loro versione *hard* come fondamento di un sistema premiante e di una gestione dei docenti con criteri manageriali, è stata oggetto di critiche approfondite nella letteratura internazionale. Le pratiche di SRT comporterebbero la "*commodification*" degli studenti, la reificazione della relazione insegnante-studente e una concezione molto ristretta della professionalità docente e della responsabilità accademica [Lawrence e Sharma, 2002; Singh, 2002]. Al fondo, si constata la volontà di assimilare gli studenti a "consumatori" con una visione alquanto riduttiva dell'esperienza universitaria.

Una parte di queste reazioni deriva probabilmente da ragioni ideologiche, o da forme di conservatorismo accademico ma è comunque difficile accettare la proiezione dello SRT al di fuori dell'ambito della singola università verso l'intero sistema. Ciò che può avere senso come pratica manageriale, con il suo carico ragionevolmente accettabile di soggettività, perde sicuramente significato in un ambito più vasto.

Destano quindi stupore gli intenti del Ministro italiano dell'Università, dichiarati con vasta eco di stampa, di assumere i risultati dello SRT come riferimento importante per l'allocazione dei finanziamenti alle diverse università e per altri interventi ministeriali. Il bollettino informativo quindicinale del MIUR del 30 giugno 2004 così si esprime in proposito:

<<il complesso delle valutazioni, che sarà disponibile quanto alla didattica entro settembre, consentirà al MIUR di acquisire importanti e significative indicazioni circa la rispondenza dei servizi erogati dalle università ai bisogni degli studenti nel nuovo quadro di rilevazione del

grado di soddisfazione degli studenti. (...). I risultati periodici di tali rilevazioni – ha spiegato il Ministro Letizia Moratti – offriranno un quadro oggettivo della qualità dei servizi universitari in modo da consentire, anche attraverso misure di sostegno, d'incentivazione e di disincentivazione, di adottare concrete iniziative per il governo del sistema universitario in una ottica di perseguimento della qualità ed efficacia dei servizi>> [MIUR, 2004].

La raccolta delle schede attraverso le quali gli studenti esprimono le loro valutazioni sulla didattica, prevista dalla legge e orientata da un modello ufficiale elaborato dal CNSVU, manca infatti di qualsiasi presupposto che giustifichi tale scelta.

Da un lato tutto il processo non offre alcuna seria garanzia di affidabilità, lasciato come è all'autonoma organizzazione da parte degli atenei e dei loro nuclei di valutazione.

Dall'altro lato, e questa è l'obiezione fondamentale, non rimediabile con modalità organizzative più accurate, non si vede come l'espressione formale di opinioni da parte degli studenti possa soddisfare i requisiti per una attendibile misurazione, data la diversità dei contesti che si verrebbero arbitrariamente a confrontare nell'ambito di una scala di punteggi.

Le conseguenze di un orientamento manageriale per la valutazione delle università

La ricerca di un fondamento in misurazioni oggettive per la valutazione delle università deve fare i conti non soltanto con la debolezza e la scarsa persuasività, proprio dal punto di vista della misurazione, delle metodologie più frequentemente adottate, ma anche e soprattutto con le differenti esigenze che premono oggi sul management universitario.

La stessa autonomia riconosciuta alle università implica che ciascuna di queste si impegni in relazioni con la realtà esterna, stabilendo legami e acquisendo una posizione nell'ambiente, fino a definire un proprio specifico profilo strategico e a orientare di conseguenza le proprie azioni nei diversi settori di interesse. Altre tendenze, come la competizione, l'internazionalizzazione, l'acquisizione di commesse di ricerca finanziate, l'inserimento in reti di collaborazione sia a livello nazionale che a livello locale, spingono verso la costruzione di un profilo e di una identità specifici di ciascuna università.

Nel rispondere a queste pressioni e opportunità, gli atenei non possono evitare di gestire i rapporti di confine e di affrontare processi di negoziazione con altri attori. L'autonomia si rivela sempre più il risultato di una rete di negoziazioni che coinvolge partner di diverso peso, organismi e soggetti differenziati. Il management universitario deve così definire strategie che orientino la selezione delle relazioni considerate meritevoli di investimento, dei legami nei quali è utile e ha valore impegnarsi [Felt, 2004].

Non deve meravigliare quindi che tutto ciò risulti in una differenziazione dei profili delle università, in una forte eterogeneità delle rispettive caratteristiche e linee di comportamento.

Sembra ancor più difficile in questo contesto far convergere l'attenzione su scale di misura dell'attività universitaria e dei suoi risultati che siano valide indipendentemente dai contesti.

Si pensi al caso, frequente in Italia, di università di recente istituzione, sorte in contesti locali periferici e finalizzate soprattutto ad una funzione di stimolo culturale, di offerta di

servizi formativi innovativi, di animazione di realtà sociali ed economiche abbastanza isolate. Una valutazione comparativa dei risultati di ricerca, rispetto per esempio a università metropolitane di forte tradizione, sembra avere poco significato, data la palese differenza di obiettivi. Il problema non consiste tanto nella probabile penalizzazione di queste università periferiche, ma nei prevedibili effetti distorsivi, per l'orientamento all'impegno nella ricerca. In molti settori tematici, università di questo tipo, per conseguire i propri obiettivi, dovrebbero orientarsi su tematiche di rilievo locale, privilegiare l'inserimento in reti di rapporti propri del tessuto socio-economico contiguo. L'orientamento dei docenti, spinto anche dalle priorità della valutazione nazionale, verso il riferimento a circuiti di altro tipo, anche di rilievo nazionale e internazionale, produrrebbe uno sradicamento dal territorio, sostituito da una probabilmente velleitaria e inconcludente ricerca di affermazione in ambiti più vasti, oppure da una "colonizzazione" da parte di gruppi di ricercatori espressi da atenei più forti, che questi docenti finirebbero per assumere come riferimento prioritario. In tutti i casi con scarso beneficio per la strategia di inserimento nel contesto territoriale.

In realtà, il concetto stesso di qualità tende a cambiare significato in funzione dello specifico assetto e delle scelte di posizionamento compiute da ciascuna università [Felt, 2004]. Questo del resto corrisponde perfettamente alle più evolute interpretazioni della valutazione dei risultati di gestione delle imprese, secondo una prospettiva strategica e multidimensionale.

Si riconosce ormai che l'identico risultato espresso da uno stesso indicatore per due aziende differenti, o anche per due situazioni di mercato e competitività differenti attraversate dalla stessa azienda, non comporta la medesima valutazione.

Per esempio, indicatori riferiti ai costi degli accordi contrattuali e ai costi della conflittualità aziendale con riferimento alle relazioni sindacali non sembrano costituire elementi rappresentativi di un valore generato prescindendo dal contesto. In diversi contesti la corretta lettura può assumere un segno ben diverso.

Per leggere correttamente i risultati di un processo e per comprendere quindi il valore generato, occorre sempre focalizzare l'attenzione sugli aspetti più critici ed essenziali. Ma questi sono *firm specific*, sono connessi alla tipicità del processo operativo di un'azienda e anche alle griglie di clienti, concorrenti, fornitori, intermediari effettivamente presenti e coinvolti in un tempo e in un luogo.

Anziché parlare di misurazione sembra corretto riferirsi a valutazione dei risultati ottenuti in relazione a un processo; valutazione implica giudizio, attribuzione di valore; è frutto di analisi e ragionamento e non pretende di essere oggettiva.

C'è in questo, soprattutto, il riferimento a una prospettiva di *management*, che è diversa da quella di "misurazione": in un'ottica manageriale il livello (il valore espresso quantitativamente) di una variabile interessa relativamente. Pensiamo ad una rilevazione della soddisfazione del personale, per esempio su una scala 1-5: sapere che si è raggiunto un livello medio, poniamo di 3,6, a fronte di un livello medio per altre aziende del settore di 3,8, dice poco in una prospettiva manageriale. A prescindere dalla dubbia attendibilità di rilevazioni di questo tipo condotte in differenti contesti, non sono neppure chiari i nessi di questa variabile

con i livelli di produttività del lavoro e di competitività aziendale: difficilmente si può dare un significato positivo o negativo al divario così accertato. L'utilizzo di misurazioni di questo tipo in chiave manageriale richiede attenzione non solo per il livello o il valore rilevato ma soprattutto per la relazione, il nesso causale tra il risultato analizzato e gli obiettivi che l'azienda persegue [Becker, Huselid e Ulrich, 2001].

Si tratta di dare significato ai numeri e alle rilevazioni dal punto di vista del finalismo aziendale. E per dare significato, occorre collocare il risultato che si cerca di definire in un contesto, quello che è rilevante per l'azienda, mettendo in luce la relazione con un prospettiva di miglioramento cui si assegna un valore positivo.

Una considerazione in chiave strategica dei processi critici per un'impresa, o per un'istituzione come l'università, rende quindi molto difficile individuare riferimenti generali: la predominanza di elementi qualitativi e immateriali tende a produrre situazioni *firm specific*. Al punto che le stesse esperienze di *benchmarking* hanno scarso valore, se orientate verso il semplice confronto di parametri [Becker, Huselid e Ulrich, 2001].

Come combinare valutazione e misurazione con riferimento alle università

La proposta metodologica che si intende delineare cerca di recepire l'istanza verso il miglioramento della qualità, che non può essere contestata al di là dei limiti e dei difetti delle applicazioni correnti della valutazione e delle pratiche e tecniche ispirate dal *new public management*. Non vi è dubbio infatti che le università devono dare risposte credibili alle richieste di responsabilità e di affidabilità che vengono dalla società e dagli studenti, uscendo da un orgoglioso stato di autoreferenzialità [Hoecht, 2004].

Questo non significa che si debba seguire unicamente la via della misurazione e degli indicatori, dimostratasi sterile, almeno come principale riferimento, in tutti i casi nei quali ci si è dovuti confrontare con oggetti complessi, influenzati da relazioni intersoggettive plurime, incertezza e variabilità del contesto. E tanto meno che si debba insistere nella ricerca di false misurazioni solo per rispondere a livello di facciata a un generico bisogno di sicurezza e di oggettività.

Riteniamo utile invece proporre una traccia di riferimento per l'analisi e valutazione della qualità universitaria che tenga conto e lasci spazio all'ambiguità e all'esigenza di interpretazione, che sono inerenti al cuore delle attività accademiche, la ricerca e l'insegnamento.

È proprio la ricerca di rigore concettuale che ci spinge a mettere in evidenza questo aspetto. Bisogna sforzarsi di comprendere ciò che veramente qualifica il *core business* dell'università, quindi la didattica e la ricerca, considerate anche nella loro connessione, per capire quale spazio può essere occupato da misurazioni, o rilevazioni statistiche, anche sofisticate, e quale ruolo possono giocare altri modi di esprimere attenzione verso i risultati di queste attività.

La buona dotazione di strutture, di servizi, di risorse professionali, l'analisi dei volumi di attività sviluppati e gli altri aspetti che si prestano a una vera misurazione costituiscono in

questa logica un fondamento necessario, un basamento, una sorta di piedestallo o di piattaforma. È compito del management universitario garantire queste condizioni, anche attraverso idonei sistemi di controllo di gestione e di assicurazione della qualità.

Ma occorre anche avere chiaro in mente che una volta garantite queste condizioni, attraverso opportune misurazioni quantitative curate da organismi tecnicamente affidabili, la problematica di promuovere la vera qualità degli studi universitari si incontra con altri aspetti, che chiamano in causa più direttamente le persone, l'interazione attiva e creativa tra i diversi soggetti, i docenti, gli studenti, i tecnici e i professionisti dei vari servizi. È al di sopra di questa "soglia di accettabilità delle prestazioni" che si incontrano le problematiche critiche per il management delle università e per la stessa valutazione.

La vera qualità è dinamica, interattiva, evolutiva, nasce da una combinazione di rapporti, richiede di essere continuamente rinnovata [Pirsig, 1981 e 1992]. È influenzata da fattori di contesto e di ambiente, spesso incomprensibili e privi di significato al di fuori di questo. Come si è documentato in altra sede, didattica e ricerca si prestano ad una misurazione, che è molto utile per i fattori che ne costituiscono il piedestallo, la piattaforma di base [Rebora, 2003; Minelli, Rebora e Turri, 2002]. La qualità statica si presta a misurazione. La qualità dinamica appartiene a un altro ordine concettuale, si presta ad essere valutata, secondo metodi e logiche che restano inevitabilmente intrisi di soggettività. È una sfera di soggettività che è possibile temperare, come avviene adottando la famiglia di metodologie della *peer review* [Minelli, Rebora e Turri, 2004]. Ma occorre avere coscienza che a questo punto si opera nell'orizzonte della valutazione, espressione della facoltà umana e soggettiva del giudizio, dell'intelligenza profonda di cose e situazioni. Una facoltà che conviene sostenere e raccordare con tutte le misurazioni possibili ma che non ha rispetto a queste minore rilevanza e valore, come modalità autonoma di sviluppo della conoscenza e della capacità di governo e management di sistemi complessi.

L'applicazione di questi concetti è utile a entrambi i livelli rilevanti per l'intervento sulla qualità, che è opportuno distinguere con chiarezza: il livello del sistema di università nazionale o internazionale, nel cui ambito hanno senso organi centrali di promozione e assicurazione della qualità; e il livello della singola università e dell'azione dei corrispondenti organi di governo e management.

Il sistema complessivo di valutazione dei risultati e di assicurazione della qualità dovrebbe prevedere un limitato ricorso a regole, standard e requisiti, imperniati su oggettive e incontrovertibili misurazioni, imponendone il rispetto generalizzato. Vanno in questo senso anche le recenti proposte di Brown (2004) riferite al riorientamento futuro del sistema inglese di assicurazione della qualità:

the way to raise quality is to make accountability – in the sense of compliance with standards and rules – only a small part of quality assurance, and to appeal instead to the professional motivation of staff to do better by engaging in a constructive, professional dialogue...

Per tutto il resto occorrerebbe quindi affidarsi alla capacità di management delle diverse università, che gli organi centrali possono solo promuovere e assecondare con interventi soft, di sostegno, stimolo e accompagnamento.

Ma anche a questo secondo livello emerge l'esigenza di combinare in modo opportuno gli interventi che garantiscono il piedistallo o la piattaforma della qualità statica, attraverso anche opportune misurazioni e il presidio di standard e regole anche rigidamente stabilite. A questo livello sarà importante per esempio garantire il rispetto degli orari per le diverse attività, monitorare la frequenza effettiva degli studenti, garantire le corrette procedure di esame. Per il resto, l'affidamento alla professionalità dei docenti potrà trovare stimolo nell'adozione di strumenti di autovalutazione ed anche di raccolta di opinioni dei vari *stakeholders*, e quindi anche di strumenti come lo SRT. Potranno avere un ruolo importante le *peer review*. Ma questi restano strumenti che hanno rilievo all'interno di un'università e nell'ambito dei suoi rapporti con l'ambiente di riferimento. Se divengono passaggi e strumenti di una regolazione esterna, questi metodi vengono snaturati e perdono rilevanza per i processi di miglioramento e apprendimento organizzativo.

In conclusione, è giusto affermare che confrontarsi con i numeri e con le statistiche è importante per le università, perché costringe a fare i conti con i fatti e con dati di realtà. E anche lo sforzo per realizzare misurazioni dotate di un buon grado di attendibilità e oggettività nel senso in precedenza descritto è di grande importanza per istituzioni non abituate ad applicare al proprio stesso funzionamento quel metodo scientifico che utilizzano e diffondono invece attraverso la ricerca e l'insegnamento. Non è questo che si intende mettere in discussione. Ma occorre che le misurazioni siano collocate e comprese nel loro giusto spazio e non confuse con altri strumenti, che misurazioni non sono, e che rappresentano diverse modalità per incontrarsi con la realtà di riferimento. Dal punto di vista del management, proprio questo secondo tipo di strumenti è destinato ad assumere rilievo nei contesti come l'università, che svolgono una funzione difficilmente riconducibile a standard univoci.

Bibliografia

- Barzelay, M., *Performance Auditing and the New Public Management: Changing Roles and Strategies of Central Audit Institutions*, in PUMA, *Integrating People Management into Public Service Reform*, Parigi, OECD, 1996a.
- Becker, B.E.-Huselid, M.A.- Ulrich 2001, D., *The HR Scorecard*, Boston, Harvard Business School Press, 2001.
- Brennan, J. - Shah, T., *Quality assessment and institutional change: experiences from 14 country*, «Higher Education», n. 40 (2000).
- Brook R., *The Evaluation of Research: the Benefit, the Risks and Some Simple Rules*, «Atti del Convegno: La valutazione della ricerca», Roma, CIVR, 2000.
- Brown, R., *Quality Assurance in Higher Education. The UK Experience Since 1992*, London, RoutledgeFalmer, 2004.
- Carlsson, S., *Executive Behavior: A Study of Work Load and the Working Methods of Managing Directors*, Stoccolma, Strombergs, 1951.
- Clark, P., *Quality Assessment in English Higher Education: Achievement and Prospects*, paper presented at the Eight International Conference on Assessing Quality in Higher Education, Australia, Broadbeach, 14-16 Jul 1996.
- CNSVU, *L'evoluzione del sistema di valutazione negli ultimi 10 anni*, «Atenei», n. 1 (2002a).
- CNSVU, *L'architettura della valutazione*, «Atenei», n. 2 (2002b).
- Coda, V., *L'orientamento strategico dell'impresa*, Milano, UTET, 1989.
- Davies, A. – Thomas, R., *Managerialism and Accountability in Higher Education: the Gendered Nature of Restructuring and the Cost to Academic Service*, «Critical Perspectives on Accounting», 13, (5-6) (2002).
- Dillard, J.F., *Dialectical Possibilities of Thwarted Responsibilities*, «Critical Perspectives on Accounting» 13, (5-6) (2002).
- Felt, U., *University Autonomy in Europe: a background Study*, in *Observatory for Fundamental University Values and Rights, Managing University Autonomy. Collective Decision making and Human Resources Policy*, Bologna, Bononia University Press, 2002.
- Felt, U., *University Autonomy in Europe: Shifting Paradigms in University Research*, in *University Autonomy: Shifting Paradigms in University Research, Proceedings of the Seminar of the Magna Charta Observatory, 15 September 2003*, Bologna, Bononia University Press, 2004.
- Figà Talamanca, A., *Autonomia e valutazione nel sistema universitario*, Università Cattolica del Sacro Cuore, Conferenza del 7 giugno 2001.
- Goldhaber, D., *The Mystery of Good Teaching, Education Next*, Hoover Institution, 2002.
- Kaplan, R.S. – Norton, D.P., *The Balanced Scorecard. Translating Strategy into Action*, Harvard, Harvard Business School Press, 1996.
- Hoecht, A., *Accountability, Trust and Professional Autonomy in UK Higher Education Quality Management*, paper presented in the workshop of EIASM in Siena “process of reform of the university across Europe”, May 24-26 2004.
- Mari, L., *Epistemology of Measurement*, «Measurement», vol. 34 (2003).
- Lawrence, S. – Sharma, U., *Commodification of Education and Academic Labour*, «Critical Perspectives on Accounting», 13 (5-6) (2002).
- Minelli, E. – Reborà, G. – Turri, M., *Il valore dell'università. La valutazione della didattica, della ricerca, dei servizi negli atenei*, Milano, Guerini, 2002.
- Minelli, E. – Reborà, G. – Turri, M., *Evaluation in Universities: the Organizational Impact*, paper presented at the EES Sixth Conference, Governance, Democracy and Evaluation, Berlin, 30 september-2 October 2004.

- Mintzberg, H., *The Nature of Managerial Work*, New York, Harper & Row, 1973.
- Mintzberg, H., *Managerial Work: Forty Years Later, Uppsala Lectures, May 24-25, 1989*, Uppsala, Textgruppen, 1991.
- Mintzberg, H., *Managers not MBAs. A Hard Look at the Soft Practice of Management Development*, London, Prentice Hall, 2004.
- MIUR, *Puntualità alle lezioni, chiarezza didattica e disponibilità. Gli studenti universitari giudicano i loro docenti*, «Università & Ricerca Informazioni», 30 giugno 2004.
- Neave, G., *The Evaluative State Reconsidered*, «European Journal of Education», 33 (3) (1998).
- Noordegraaf, M.- Abma, T., *Management by Measurement? Public Management Practices amidst Ambiguity*, «Public Administration», 81 (4) (2003).
- Paganelli, O., *Valutazione delle aziende. Principi e procedimenti*, Torino UTET, 1990.
- Pirsig, R.M., *Lo Zen e l'arte della manutenzione della motocicletta*, Milano, Adelphi, 1981.
- Pirsig, R.M., *Lila*, Milano, Adelphi, 1992.
- Power, M., *The Audit Society. Rituals of Verification*, Oxford University Press, 1997.
- Pollitt, C., *The Politics of Performance Assesss: Lessons for Higher Edutation?*, «Studies on Higher Education», 12 (1) (1987).
- PUMA, *Integrating People Management into Public Service Reform*, Parigi, OECD, 1996a.
- PUMA, *Performance Auditing and the Modernisation of Government*, Parigi OECD, 1996b.
- PUMA, *In Search of Results, Performance Management Practices*, Parigi, OECD, 1997.
- Rebora, G., *La valutazione dei risultati nelle amministrazioni pubbliche*, Milano, Guerini, 1999.
- Rebora, G., *La valutazione degli atenei: alla ricerca della qualità*, «Azienda pubblica», n. 5-6 (2003).
- Rowan, B. – Correnti, R. – Miller, R.J., *What Large-Scale, Survey Research Tells Us About Teacher Effects on Student Achievement, Consortium for Policy Research in Education, University of Pennsylvania, Graduate School of Education, Research Report Series, November 2002.*
- Sharp, S., *The Quality of Teaching and Learning in Higher Education: Evaluating the Evidence*, «Higher Education Quarterly», 49 (4) (1995).
- Simon, H., *Administrative Behaviour*, New York, 1947, trad. it. Bologna, Il Mulino, 1958.
- Singh, G., *Educational Consumers or Educational Partners: a Critical Theory Analysis*, «Critical Perspectives on Accounting», 13 (5-6) (2002).
- Turri, M., *Lo stato dell'arte, limiti ed evoluzione nella valutazione a livello di sistema e di singolo ateneo*, «Azienda Pubblica», N.5-6 (2003).
- Trow, M., *Managerialism and the Academic Profession: Quality and Control*, «Higher Education Report», 2, Quality Support Centre, May 1994.
- Underwood, S., *Assessing the Quality of Quality Assessment: the Inspection of Teaching and Learning in British Universities*, «Journal of Education for teaching», 26 (1) (2000).
- Van Damme, D., *Higher Education in the Age of Globalization*, in AA.VV., *Globalization and the market in higher education*. Paris, UNESCO, 2002.
- Vedung, E., *Public Policy and Program Evaluation*, New Brunswick, Transaction Publishers, 1997.
- Weick, K., *The Social Psychology of Organizing*, Newbery Award Records, 1979, trad. it. Organizzare, Torino, ISEDI, 1993.