
da A. Bianchi - P. Di Giovanni

Biblioteca di Scienze Sociali

vol. 7 - Socializzazione e formazione

Paravia

6. LA VALUTAZIONE

6.1. Obiettivi della valutazione. Che cosa ci si aspetta quando si valuta l'apprendimento degli allievi? Il livello di apprendimento raggiunto da un allievo in un dato momento dipende sia dal lavoro svolto da lui, sia da quello svolto dal formatore. Perciò il formatore può orientare la valutazione sugli allievi o sulla didattica, può essere interessato cioè ad accertare se gli allievi hanno raggiunto i traguardi formativi stabiliti o a verificare l'efficacia delle strategie di formazione adottate per aggiustare di conseguenza il lavoro successivo. Nel primo caso si parla (la distinzione risale a M. Scriven, 1967) di *valutazione sommativa* (è una sorta di bilancio dei risultati ottenuti fino a quel momento), nel secondo caso di *valutazione formativa* (serve a monitorare e indirizzare il processo formativo).

La valutazione sommativa è quella più tradizionale, l'unica in uso in passato. La pratica della valutazione formativa si è diffusa nella seconda metà del Novecento, via via che dal modello tradizionale di processo formativo centrato sulle competenze e più rigido si è passati a un modello centrato sul contesto, più flessibile e a circuito chiuso, con continuo controllo a feedback (§ 1.4). A promuovere la valutazione formativa ha contribuito anche l'esperienza delle formazioni a distanza, rese possibili dalle telecomunicazioni e oggi in rapida crescita con Internet. Nella formazione faccia a faccia il formatore ha l'impressione di tenere sotto controllo la didattica anche senza predisporre apposite prove di valutazione formativa, perché raccoglie feedback durante le lezioni, parla con gli allievi tra una lezione e l'altra, ecc. Nella formazione a distanza invece prove apposite sono indispensabili, dato che mancano le impressioni del rapporto faccia a faccia.

Solitamente il formatore stabilisce in partenza quando adoperare la valutazione in ottica sommativa e quando adoperarla in ottica formativa. Può capitare però che una valutazione programmata per uno scopo si riveli utile per l'altro. Ad esempio, dopo una prova (un compito di matematica o una prova oggettiva di lingua straniera) fatta per scopi sommativi, il docente può scoprire che la distribuzione dei risultati nella classe è bimodale, cioè i punteggi si addensano intorno a due valori distanti tra loro (fig. 27). Un risultato del

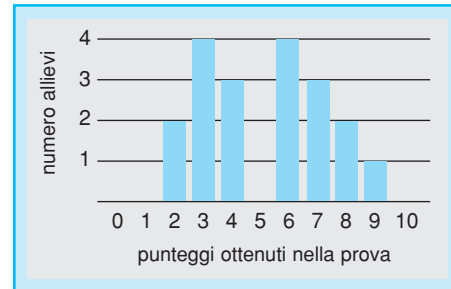


Fig. 27. Distribuzione bimodale dei risultati. Se una curva del genere si ottiene una volta, si può sospettare che si tratti di un caso. Se però si ripresenta in più prove, testimonia che la classe è a due velocità e invita a ripensare l'organizzazione didattica.

genere dice che ci sono allievi in difficoltà e altri in condizione di rendere adeguatamente: c'è – si dice – una classe a due velocità. Specie se si ripete in altre prove, dice anche che fino a quel momento le strategie didattiche adottate non sono riuscite a recuperare gli allievi in difficoltà.

La valutazione può essere utilizzata in vista di obiettivi più specifici. Solitamente si distinguono diversi tipi di valutazione, a seconda dell'uso che se ne fa. Anche se risulta in parte artificiosa, perché nella pratica le differenze sono sfumate, la classificazione aiuta a orientarsi.

1. Valutazione diagnostica. Serve al formatore per raccogliere informazioni utili a individuare problemi e esigenze della formazione e a progettare di conseguenza l'attività. La diagnosi è l'obiettivo più comune delle valutazioni formative. In un corso o nello svolgimento di un modulo l'ideale è prevedere una valutazione diagnostica all'inizio (valutazione d'ingresso) e alcune in itinere (valutazioni intermedie).

2. Valutazione predittiva. Dalla valutazione ci si aspetta di ricavare previsioni sulla riuscita degli allievi in studi e formazioni successive o in attività lavorative o in ambiti di vita. Oggi ci sono molti dubbi sull'utilità delle valutazioni predittive ufficiali fatte a conclusione di un corso o di un ciclo di studi. Si sa infatti che facilmente diventano profezie che si autoadempiono che possono danneggiare gli allievi: chi ha avuto un giudizio predittivo negativo, incontrerà più difficoltà, perché si trova a fare i conti col peso psicologico della profezia. La valutazione predittiva non presenta problemi invece quando è usata in ottica formativa, cioè quando il formatore si chiede come riusciranno gli allievi in un segmento di corso successivo al fine di progettare adeguatamente l'attività futura.

3. Valutazione selettiva. Ci si chiede se le competenze degli individui valutati sono al di sotto o al di sopra di una determinata soglia ritenuta necessaria per accedere a un corso (come nei test di ammissione di Università) o a un'attività lavorativa.

4. Valutazione orientativa. Si usa la valutazione per indirizzare gli allievi negli studi successivi o nelle attività lavorative. Può essere negativa (questa strada non è adatta a te) o proattiva (questo è un cammino per te) o mista (più che a questo dovresti puntare a quest'altro). Anche la valutazione orientativa, come la predittiva, è delicata. C'è egualmente il rischio della profezia che si autoadempie ed è facile sbagliare. È importante evitare gli approcci episodici: conviene non basare il giudizio su valutazioni concentrate in un dato momento, ma servirsi di informazioni accumulate in un lasso di tempo lungo.

6.2. I problemi della valutazione. Valutare è difficile. La docimologia è nata proprio quando si è preso coscienza di quanto fossero seri i problemi della valutazione. Agli inizi del Novecento nei paesi avanzati, alle prese con l'esplosione scolastica e con il fatto che la scolarizzazione diventava sempre più importante per i singoli e per la società (*Unità 3*, § 4.3), si è cominciato a interrogarsi sul valore degli esami. H. Piéron, considerato il fondatore della docimologia, dimostrò che, se una stessa prova veniva fatta valutare da correttori diversi, si avevano discrepanze, anche considerevoli, nei giudizi (*Lecture 1*). Lo stesso correttore a distanza di tempo poteva formulare giudizi diversi.

Come spiegare le discrepanze? Piéron vide che correttori diversi mirano a valutare cose diverse. Ad esempio, in un componimento di filosofia uno può guardare alle abilità espositive, un altro alle abilità di ragionamento, un altro ancora alla conoscenza del pensiero degli autori. Più precisamente ogni correttore pensa a una serie di elementi da valutare e li mette in gerarchia, dal più importante al meno importante, secondo un proprio schema. Ciascun correttore poi utilizza standard diversi di riferimento: quella che è una buona esposizione per uno può essere mediocre per un altro. Inoltre al momento di tradurre in punteggio la valutazione, correttori diversi hanno in mente scale diverse.

L'esperienza di Piéron si basa su un principio semplice: mettere a confronto correzioni diverse di una stessa prova. Tuttavia ha portato alla luce un intero mondo fino ad allora rimasto nascosto: il retroscena della valutazione, le procedure e il lavoro mentale che stanno dietro al giudizio espresso. Finché si considera come un'attività sintetica, presa nel suo insieme, la valutazione appare qualcosa di pacifico. Quando si comincia ad analizzarne il retroscena, si capisce che valutare è un'attività in cui si incontrano problemi ad ogni passo.

6.3. Il nesso prestazione-competenza. Una delle difficoltà che si incontrano nella valutazione è che dobbiamo inferire competenze a partire da prestazioni. Quando valutiamo siamo interessati a stabilire quali competenze hanno gli allievi. Ci chiediamo ad esempio: ha dimestichezza con la grammatica inglese? conosce i fatti e le date principali della seconda guerra mondiale? sa stabilire se un ragionamento condizionale è corretto? Per valutare ricorriamo a prove (traduzioni, temi, problemi, interrogazioni, ecc.). Può sembrare che non ci sia nulla di strano, ma a ben guardare le prove non documentano direttamente le competenze. Una prova ci informa su una *prestazione*, cioè sul *comportamento di un individuo nell'esecuzione di un compito in una determinata circostanza*. Un'interrogazione, ad esempio, ci dirà come quell'allievo ha sostenuto quel colloquio in quella situazione. La competenza è un tratto relativamente stabile dell'individuo, mentre la prestazione è contingente.

Si dirà: le prestazioni sono legate alle competenze. È vero ed è su questo legame che si basa la valutazione. Tuttavia il collegamento prestazioni-competenze non è semplice e scontato. Può accadere che un allievo fornisca una prestazione al di sotto delle proprie competenze, come può accadere il contrario, che la prestazione superi le competenze. Il punto è che a determinare la prestazione intervengono altri fattori oltre le competenze.

I fattori che possono interferire sono molti. Possiamo raggrupparli in quattro categorie.

1. Motivazione. La prestazione di un allievo dipende da quanto e come è motivato a sostenere la prova.

2. Interpretazione della prova. Chi sostiene una prova cerca di capire che cosa gli si chiede. Abitualmente non si limita a interpretare le istruzioni esplicite, ma si sforza anche di andare più in profondità: che cosa si aspetta il valutatore? con quali criteri giudicherà? che cosa intende in special modo valutare? L'interpretazione data alla prova può condizionare la prestazione, facilitando l'allievo o portandolo fuori strada.

3. Altre competenze. Nell'esecuzione di un compito intervengono sempre altre competenze oltre quelle che intendiamo valutare. Ad esempio, un problema di fisica richiede generalmente conoscenze matematiche e conoscenze linguistiche. Un allievo può incontrare difficoltà non perché ignora la fisica, ma perché ha problemi con la lingua o con la matematica. Un altro che conosce meno di lui la fisica potrà essere avvantaggiato e fare una figura migliore grazie alle competenze linguistiche o di matematica.

4. Contesto. Ogni prova si svolge in un contesto concreto che ne influenza l'andamento. Pensiamo a un'interrogazione: non ci sono solo le domande e le risposte, ma anche l'interazione insegnante-allievo, l'interazione con gli altri della classe, il clima che viene a crearsi e via dicendo. Tutti questi fatti possono influire sull'andamento dell'interrogazione, col risultato che, date certe competenze, un allievo può fornire a seconda dei casi una prestazione migliore o peggiore.

Come inferire correttamente le competenze dalle prestazioni? Indubbiamente molto dipende dall'abilità e dall'esperienza del formatore. Alcuni accorgimenti di metodo sono però di aiuto.

1. Motivare alle prove. Ricerche empiriche hanno dimostrato che gli allievi sono più motivati se si sentono coinvolti nel processo di valutazione, ne condividono l'impostazione e il senso e ci credono. Quando l'impegno è basato soltanto sulla mira di evitare l'insuccesso o

ottenere un successo, è generalmente minore. Detto diversamente, per ottenere il massimo dagli allievi nelle prove non è sufficiente un coinvolgimento punitivo-incentivante, ma occorre un coinvolgimento simbolico nell'attività di valutazione. Per questo la strategia migliore è essere trasparenti: parlare apertamente dei risultati delle valutazioni e del loro significato nel percorso didattico e discuterne assieme, curando di orientare i discorsi in ottica costruttiva. È importante anche un clima di fiducia. L'allievo deve avere l'impressione che i giudizi hanno un minimo di coerenza e serietà. Non è importante che li giudichi perfetti. Si ha fiducia anche in istituzioni che presentano limiti (il difetto è umano). Basta che le valutazioni risultino chiare, rispondano a regole riconoscibili e siano ragionevoli.

2. Fornire istruzioni chiare. È un errore pensare che mettere in difficoltà gli allievi (ad esempio con domande ambigue) sia un buon espediente per farne emergere le capacità. Al massimo con sistemi del genere si può valutare la loro destrezza, ma solitamente il formatore non è interessato a questo e comunque per uno scopo del genere sarebbe bene ricorrere a specifici test psicologici. Le prove non devono mettere in difficoltà, al contrario devono risultare il più possibile chiare per gli allievi, tanto che questi non devono sentire il bisogno di mettersi a interpretare le richieste del compito da svolgere.

3. Servirsi di prove il più possibile mirate. Più la competenza che vogliamo saggiare è ampia e vaga, più è difficile risalire ad essa dalla prestazione. L'ideale è servirsi di prove dedicate a valutare competenze ben definite (ad esempio, linguaggio tecnico di Platone). Mirando molto le prove si finisce però per moltiplicarle, cosa che rende onerosa e al limite insostenibile la pratica didattica. Occorre perciò trovare un equilibrio che consenta di saggiare in modo sufficientemente esteso l'apprendimento degli allievi con prove sufficientemente mirate. Un buon sistema è ricorrere a prove articolate, costituite in realtà da insiemi di piccole prove, ciascuna tesa a valutare una specifica competenza.

4. Tener conto delle circostanze in sede di giudizio. Ci sono fattori di interferenza che anche con i migliori sforzi non si possono eliminare. Conviene tenerne conto al momento di formulare il giudizio.

6.4. Valutazione e decisione. A prima vista si può credere che valutare sia un'operazione puramente logica, come risolvere un qualsiasi problema. Si tratterebbe di prendere in esame i dati disponibili, ragionarci su correttamente e arrivare alla conclusione. Invece, analizzata attentamente, la valutazione è in tutto e per tutto una decisione. Come mai valutare è decidere? La cosa diventa chiara se ci fermiamo a considerare alcuni punti su cui il docente che valuta è chiamato a prendere posizione.

1. Programmazione e traguardi. Per valutare occorre essersi dati obiettivi educativi e aver prestabilito un percorso didattico per raggiungerli. La programmazione consiste essenzialmente in questo ed è qualcosa che si decide. Per il docente la decisione è in parte obbligata, perché presa a monte dall'istituzione scolastica. Resta comunque anche per lui uno spazio di manovra che di solito è piuttosto ampio.

2. Prove. Ogni docente si affida a certe prove per cercare di stabilire come procede il lavoro didattico, quali risultati gli allievi stano raggiungendo e via dicendo. Anche se ci sono metodologie sempre più codificate, è inevitabile che si debba decidere il tipo di prova da adottare (se aperta, strutturata, semistrutturata, sotto forma di test oggettivo, se orale, se scritta), quando effettuarla e come inserirla nel complesso dell'attività valutativa.

3. Criteri, termini di confronto e formulazione dei giudizi. Il docente decide in base a quali criteri valutare i risultati. A volte spetta a lui stabilire in quale forma esprimere il giudizio (se con voti, con formule descrittive, con la semplice indicazione di sufficienza/insufficienza).

Sceglie poi quale gruppo di riferimento avere in mente nel suo giudizio. Può limitarsi a confrontare un alunno con un altro o instaurare un confronto tra tutti i membri della classe o tenere presente l'andamento di altre classi, anche degli anni passati o, se dispone dei dati necessari, basarsi sui risultati conseguiti da popolazioni ampie di scolari.

4. Divario tra prestazione e competenza. Il docente deve decidere in quale misura il risultato di una prova è da ascrivere alla prestazione e in quale misura alla competenza da valutare. In altri termini, deve stimare il divario tra prestazione e competenza.

5. Considerazione dei risultati. Il docente è chiamato anche a decidere quanto conta una singola prova nel quadro complessivo del lavoro svolto assieme. Ad esempio, il basso rendimento in uno scritto può essere attribuito a un calo momentaneo, a stress, a fatti contingenti, come può al contrario assumere significato maggiore, di spia del cattivo lavoro svolto dall'allievo, del fatto che non ha metodo di studio o che non ha proprio le abilità necessarie per la materia.

Come in qualsiasi attività decisionale, per valutare occorre formulare *giudizi di valore*, stabilire che cosa è desiderabile e che cosa non lo è, e *giudizi di probabilità*, azzardare ipotesi su come stanno le cose e come possono evolvere. Una valutazione scolastica, anche la più banale, se analizzata sul piano cognitivo, ha dietro un intreccio di molteplici giudizi di valore e di probabilità. Eppure di solito i docenti fanno fatica a rendersene conto: specie i giudizi di probabilità sfuggono alla loro consapevolezza. Il motivo è che queste componenti della complessa attività cognitiva del valutare sono automatizzate.

Giudizi di valore e di probabilità affiorano un po' dappertutto nella valutazione, anche in operazioni di routine per un insegnante, come correggere un compito di italiano. Mettiamo che un alunno abbia adoperato degli indicativi in subordinate che, a essere rigorosi, vorrebbero i congiuntivi. Di fronte a questo dato il docente deve stabilire (giudizio di valore) se l'errore è grave e quanto pesa sul giudizio complessivo, che includerà anche la valutazione dei contenuti, dell'esposizione, la precisione di informazione, l'originalità, la maturità, ecc. I docenti di italiano hanno spesso convinzioni diverse sul peso di un errore del genere. Ancora più importante, però, è il fatto che si rende necessario un giudizio di probabilità. L'alunno ha usato indicativi perché non conosce il buon uso dei congiuntivi? Oppure usa disinvoltamente indicativi e congiuntivi, perché è un profondo conoscitore della lingua, un lettore di romanzi e opere letterarie recenti? Nel primo caso ignora la sintassi, nel secondo caso ha chiaro che la sintassi non è qualcosa di rigido, ma in evoluzione e adattabile al registro che si sta adoperando. In una situazione come questa, di solito gli insegnanti hanno parecchi elementi per arrivare a un giudizio di probabilità ragionevole. Conoscono l'alunno, sanno se è un buon lettore, possono rifarsi ad altri compiti e testi scritti da lui e all'interno dello stesso tema controllare se l'uso del verbo è da conoscitore o da inesperto. Tuttavia non sempre le cose sono così semplici e i giudizi di probabilità possono far nascere dubbi che l'insegnante deve ripromettersi di andare a verificare in seguito, nel tempo.

Il semplice sapere che valutare è decidere, diventare consapevoli delle operazioni mentali che comporta e dei punti critici che presenta è di aiuto all'insegnante. Il valutatore sensibilizzato agli aspetti cognitivi della valutazione è in genere più affidabile e obiettivo.

6.5. I biases nella valutazione. Chi valuta è soggetto a *biases*, a distorsioni e errori sistematici. Non è il caso di meravigliarsi di questo: il docente chiamato a valutare ha una mente come quella di tutte le altre persone, con i normali pregi e difetti della mente umana, ed è un individuo immerso nella vita sociale. La via maestra per rendere più obiettive le valutazioni non è fingere che non ci siano distorsioni di giudizio, ma, al contrario, riconoscerle e cercare di tenerle sotto controllo. Da questo punto di vista la psicologia dell'educazione può fornire un contributo fondamentale, portando gli insegnanti a una nuova sensibilità nei riguardi del problema.

Quasi tutti i *biases* noti intervengono nell'attività di valutazione. Esaminiamone alcuni tra i più importanti.

1. Accentuazione. È il *bias* di categorizzazione che porta a minimizzare le differenze tra gli appartenenti a una stessa categoria esagerando quelle tra appartenenti a categorie diverse. A scuola è importante soprattutto per l'effetto dei sottogruppi che si formano nella classe. Quando un alunno si intende e lega con certi compagni (perché ha interessi in comune al di fuori dello studio o per altre ragioni), i docenti corrono il rischio di attribuirgli caratteristiche di quel gruppo che magari non ha. Se, ad esempio, si tratta di un gruppo di ragazzi con scarsa partecipazione e basso rendimento, è facile che anche quell'allievo sia giudicato così e che i suoi sforzi di dimostrarsi diverso stentino a dare risultati.

2. Errore fondamentale di attribuzione. È la tendenza a sopravvalutare l'importanza dell'individuo nel procurarsi successi e insuccessi e a sottovalutare l'influenza della situazione e dei condizionamenti ambientali. Interviene più che altro al momento di giudicare il divario tra prestazione nelle prove e competenza. I docenti sono portati a considerare poco i fattori motivazionali, cognitivi e situazionali che possono rendere una prova più o meno buona o cattiva e a credere che alla prestazione corrisponda quasi automaticamente la competenza dell'individuo.

3. Self-serving biases. I risultati scolastici dipendono sia dal lavoro degli allievi, sia dalla didattica. Ogni volta che valutano, gli insegnanti devono stabilire quanto è dovuto alla propria responsabilità e quanto alla responsabilità degli allievi. Perciò è naturale che affiori la tentazione di attribuire gli insuccessi agli alunni e i successi alla didattica.

4. Effetto alone. È il meccanismo scoperto da Thorndike, per cui un giudizio in un campo viene esteso indebitamente a campi che non c'entrano. Se il docente, ad esempio, si fa un'idea positiva dell'allievo quanto a profitto, tende a giudicarlo positivamente in tutto, anche nel comportamento, la motivazione, la disciplina, i rapporti sociali e la capacità di lavorare in gruppo.

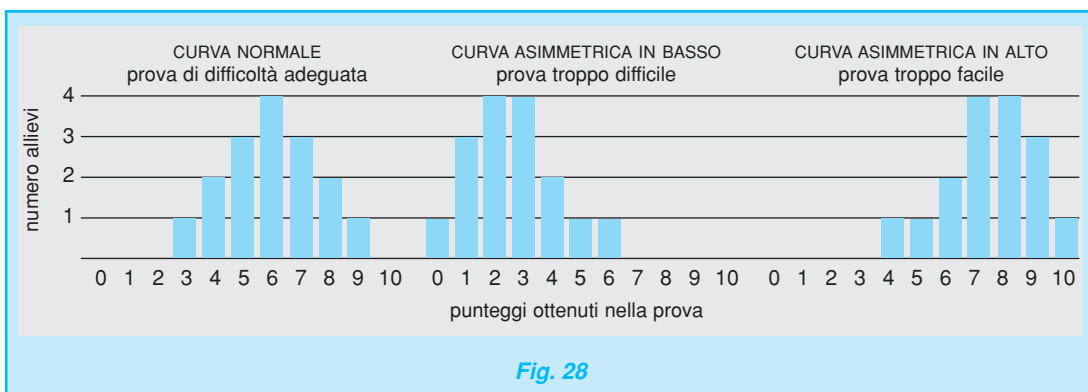
5. Autoconvalida. Consiste nel complesso di meccanismi distorsivi che inducono a conservare i giudizi che si hanno a dispetto delle prove contrarie. È il motivo per cui i bravi restano di solito bravi, anche se fanno poco, mentre i mediocri restano mediocri, a dispetto dei recuperi che tentano di fare.

6.6. La qualità delle prove. Idealmente una prova, per essere un buon strumento di valutazione, dovrebbe possedere una serie di requisiti.

1. Validità. Una prova è valida nella misura in cui fornisce informazioni sulla competenza che vogliamo valutare, cioè se è adeguata ai nostri scopi.

2. Attendibilità. Una prova è attendibile se dà risultati costanti quando viene ripetuta a distanza di tempo o quando viene somministrata da persone diverse o in circostanze diverse. Gli studi pionieristici di H. Piéron evidenziavano un difetto di attendibilità nelle comuni valutazioni. L'attendibilità non è meno importante della validità. Anche se centriamo il bersaglio e valutiamo ciò che ci interessa, servirà a poco un risultato valido solo per quella volta. Non potremo confrontarlo infatti con altri risultati, né dargli significato obiettivo. Servirà al convincimento personale, ma non avrà più valore delle impressioni soggettive ricavate quotidianamente a contatto con gli allievi.

3. Standardizzazione. Il risultato ottenuto dall'allievo in una prova ha senso solo se comparato ai risultati ottenuti in quella stessa prova da altri allievi. Ogni giudizio non può che essere relativo e comparativo. Quando diciamo che un ragazzo è bravo in fisica, solitamente non stiamo esprimendo una stima assoluta. Non appena lo mettiamo a fianco ad Einstein, ci sem-



brerà un ignorante di fisica. Evidentemente lo parliamo con i ragazzi che fanno gli stessi studi di fisica e che si trovano all'incirca allo stesso punto del curriculum. Una prova è standardizzata quando consente di valutare in rapporto a un preciso standard di riferimento. Possiamo scegliere come standard di riferimento la classe o l'insieme di tutti i ragazzi che frequentano lo stesso anno o di tutti i ragazzi che l'hanno frequentato nei dieci anni precedenti. In ogni caso lo standard dev'essere noto e la prova dev'essere concepita in modo tale da permettere il confronto con lo standard.

4. Difficoltà adeguata. Prove troppo facili o troppo difficili non consentono una buona valutazione, in quanto non fanno emergere le differenze tra gli allievi, ma appiattiscono la valutazione in alto o in basso. Un modo semplice per capire se una prova somministrata è troppo facile o difficile consiste nell'esaminare la distribuzione dei risultati nella classe o nel gruppo di persone esaminate (fig. 28).

LA VALUTAZIONE A STEP

Un espediente per accrescere l'attendibilità di una prova di tipo soggettivo, come un tema o un orale, è valutare a step. Quando ha un pacco di compiti da correggere, l'insegnante può trattarli uno per uno. Ne esamina uno, formula il giudizio e mette il voto. Un procedimento comparativo è però decisamente più attendibile. Si procede così: si esaminano tutti i compiti senza formulare giudizi, si ordinano in graduatoria e alla fine di creano gruppi omogenei. Nel fare i gruppi è buona regola usare un numero dispari di fasce (di solito 5) per riprodurre la distribuzione normale di una curva gaussiana. Così c'è una moda centrale, ci sono minusvarianti, plusvarianti e gruppi intermedi a destra e a sinistra. La curva normale rappresenta la distribuzione degli eventi in natura ed è molto probabile che anche le risposte di una popolazione di allievi ai compiti si distribuiscano così. Se non ne teniamo conto, rischiamo di forzare la realtà in uno schema nostro. Un procedimento analogo si può usare per valutare una serie di prove orali: si esamina un certo numero di candidati (ad esempio 5), quindi si valuta stabilendo una gerarchia e poi formulando i singoli giudizi.

6.7. Le prove oggettive di profitto. Molte difficoltà della valutazione possono essere superate con l'impiego di prove oggettive di profitto. Si tratta di test costruiti con tecniche ormai consolidate (*Approfondimenti 1*), che hanno i requisiti della validità, dell'attendibilità, e, in alcu-

ni casi (quando si tratta non di prove di classe, ma di prove istituzionali, fatte da équipes di esperti per la popolazione studentesca), anche della standardizzazione su larga scala.

Le prove oggettive presentano il grosso vantaggio di mettere al riparo da errori e offrono al docente uno strumento prezioso per valutare i risultati effettivi della propria attività didattica. Difficilmente un insegnante può rendersi conto di ciò che i ragazzi stanno imparando solo in base alle impressioni che riceve in classe e con le prove soggettive. C'è sempre una certa tendenza a vedere le cose migliori o peggiori di come sono o, per lo meno, a lasciarsi sfuggire qualche aspetto. La prova oggettiva, se ben costruita, offre uno spaccato nitido e spietato del livello di rendimento raggiunto, dettagliando obiettivo per obiettivo che cosa si è ottenuto e mostrando il profilo di ciascun allievo. Somministrando prove oggettive all'inizio del corso (test di ingresso) durante l'anno (test intermedi) e alla fine (test finali) l'insegnante può orientare strada facendo il suo lavoro e fare un bilancio complessivo dell'attività.

Tuttavia le prove oggettive presentano inconvenienti. Uno, per altro risolvibile, è che mettono in gioco la memoria di riconoscimento, anziché di rievocazione. Per far in modo che la prova risulti davvero oggettiva è necessario evitare che l'allunno formuli le risposte. Deve operare su risposte belle e fatte. Così però non gli si chiede di rievocare, ma solo di riconoscere dati acquisiti. Dagli studi sull'immagazzinamento nella memoria a lungo termine sappiamo che anche con un apprendimento meccanico è possibile incamerare dati in modo tale da riconoscerli. Invece occorre un apprendimento elaborativo (con la codifica, l'organizzazione, la rielaborazione) per incamerare dati che si è poi in grado di rievocare. In teoria perciò in una prova oggettiva, l'allunno che ha studiato pappagallescamente può riuscire bene, al pari di quello che ha capito e approfondito. A questo inconveniente si rimedia facilmente costruendo le prove oggettive in modo tale che non riproducano fedelmente la struttura dei discorsi fatti nel testo e a lezione, sia nelle formulazioni, sia a livello concettuale. Se i quesiti della prova vertono sulle stesse cose studiate, ma sono posti diversamente, l'allunno che ha appreso meccanicamente è tagliato fuori e viene facilmente distinto da quello che ha imparato davvero.

Le prove oggettive presentano un altro limite, cui invece non c'è modo di ovviare. Possono verificare soltanto l'apprendimento di specifici contenuti e non altro. Ad esempio non ci dicono nulla su come l'allievo sa esporre, su come collabora in un lavoro di gruppo e via dicendo. Questo non vuol dire che le prove oggettive siano nozionistiche, come afferma qualche denigratore. Attraverso una prova oggettiva possiamo verificare non solo la semplice acquisizione di dati, ma anche il modo in cui l'allievo organizza e collega i concetti, la padronanza del linguaggio e del lessico, la capacità di applicare le cose imparate a situazioni concrete e di riorganizzare in maniera nuova il materiale appreso.

6.8. Effetti della valutazione. Fino ad ora la maggior parte delle ricerche è stata condotta sugli effetti negativi e si sono trascurati gli effetti positivi che può produrre una valutazione ben fatta. In parte la cosa è dovuta al fatto che negli anni '60 (il periodo della contestazione studentesca) le pratiche di valutazione sono state messe sotto accusa, giudicate disumane, ingiuste e controproducenti. Di conseguenza da una parte e dall'altra, sia tra gli psicologi favorevoli ad abolire o ridimensionare le valutazioni, sia tra quelli contrari si è diffusa la tendenza a cercare di capire se davvero valutare produce danni.

Per gli insegnanti valutare è a volte motivo di ansia. C'è la paura di sbagliare e di essere ingiusti, di venir messi in discussione per il proprio operato, sia nella attività valutativa, sia nella didattica precedente. La reazione più comune all'ansia è l'irrigidimento in procedure stereotipate di valutazione, che tengono in poco conto le trappole psicologiche dell'attività valutativa e si ancorano invece a principi di esattezza formale. Più raro è il caso di reazioni di osti-

lità verso gli allievi oppure, al contrario, di eccessiva bonarietà. Sia l'ostilità, sia la bonarietà tendono a vanificare la sostanza della valutazione, trasformandola in una faccenda emotiva.

Gli effetti negativi riscontrati sugli alunni sono diversi e vanno dallo spostamento della bilancia motivazionale verso l'estrinseco (il tornaconto invece del gusto dell'apprendere in sé), all'angoscia dell'insuccesso, alle tensioni, allo svuotamento del proprio senso del valore come persone, alla creazione di un clima eccessivamente competitivo in classe. Tuttavia se la valutazione è condotta con obiettività, se il rapporto tra docente e alunno è trasparente e cordiale, caldo e soprattutto se, pur facendo di tutto per essere equo, l'insegnante non esagera l'importanza dei propri giudizi, ma li considera fallibili e un elemento tra gli altri nella vita degli allievi, gli effetti negativi si riducono notevolmente e possono anche scomparire del tutto. Insegnante e allievi possono star bene nonostante i giudizi e i voti, se si tengono nel debito conto gli aspetti psicologici della valutazione.



LETTURE

1

H. Piéron

Il problema dell'attendibilità delle valutazioni

H. Piéron, *Esami e docimologia* (1963)

trad. it. di N. Galli, Armando Roma, 1965

Nel brano riportato di seguito Piéron riferisce i risultati ottenuti mettendo a confronto le valutazioni di correttori diversi nelle prove degli esami di baccalauréat in Francia, corrispondenti alla nostra maturità.

...lo *scarto medio* fra correttori era molto elevato [...] Lo scarto più frequente (si tenga conto che veniva usata una scala in ventesimi) variava tra i quattro punti delle prove di inglese, matematica e fisica e i sette punti delle prove di composizione francese e di filosofia. È da notare che nel caso di queste ultime prove lo scarto massimo osservato giungeva rispettivamente a 13 e 12 punti. L'ampiezza dello scarto rivela l'inesistenza di criteri uniformemente condivisi e applicati nella valutazione: in sostanza, ciascun correttore è sensibile a certe caratteristiche delle prove [...] i punteggi attribuiti dai vari correttori davano luogo a *distribuzioni diverse*: ciò significa che i diversi correttori utilizzavano, entro la scala in ventesimi, gamme molto diverse, e ricorrevano con diversa frequen-

za ai singoli voti. È come se dicessimo che ogni correttore interpreta a suo modo una scala di voti, conferendo ai singoli punti un particolare significato. Ciò non deve sorprendere: lo stesso accade comunemente nelle nostre scuole secondarie superiori. Gli insegnanti hanno a disposizione i dieci punti di una scala in decimi, ma ne utilizzano solo una parte, in genere non più di tre o quattro punti (per esempio, dal quattro al sette o dal cinque all'otto). Per di più, questa gamma limitata corrisponde a diverse interpretazioni del valore dei punteggi. Il sette di un insegnante può corrispondere al sei o all'otto di un altro, ma non è detto che il sei dello stesso insegnante comporti una corrispondente diminuzione nei voti espressi dagli altri.



APPROFONDIMENTI

1

Le prove oggettive di profitto

Come costruirle?

Come formulare gli items? Le prove oggettive sono dette anche *prove strutturate*, perché composte da una serie di items, unità in cui figurano già formulate sia domande, sia risposte. Un tipo di

item è il *vero o falso*, che richiede di scegliere tra due alternative semplici e radicali. Per formulare un'affermazione vera conviene cercare una semplice nozione contenuta nella materia e esplicitar-

TIPI DI ITEMS

<p>Cio che si impara con l'imprinting si dimentica presto <input type="checkbox"/>VERO <input type="checkbox"/>FALSO</p> <p>Una prova oggettiva è accettabile anche se non è standardizzata <input type="checkbox"/>VERO <input type="checkbox"/>FALSO <input type="checkbox"/>VERO / FALSO a seconda dei casi</p> <p>Abitualmente nella correzione a step i compiti vengono divisi in <input type="checkbox"/>4 fasce <input type="checkbox"/>5 fasce <input type="checkbox"/>6 fasce</p> <p>Riempi con espressioni e termini appropriati le parti vuote scegliendo tra le voci dell'elenco sotto riportato (la lunghezza degli spazi è indipendente dal numero di parole da inserire). <i>Gli studi sui programmi di rinforzo hanno messo in evidenza che il rinforzo, produce un apprendimento più saldo del rinforzo</i> 1 intermittente 2 arbitrario 3 continuo 4 stabile 5 permanente 6 a step</p> <p>Nella colonna di sinistra trovi alcuni requisiti che le prove devono possedere. In quella di destra sono indicate strategie per migliorare le prove. Collega i requisiti con le strategie pertinenti (per ogni requisito c'è una sola strategia da collegare, due strategie sono in eccedenza e non vanno collegate).</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%; text-align: center; padding: 5px;">attendibilità</td> <td style="width: 50%; padding: 5px;">fissare termini di riferimento</td> </tr> <tr> <td style="width: 50%; text-align: center; padding: 5px;">standardizzazione</td> <td style="width: 50%; padding: 5px;">valutare obiettivi specifici</td> </tr> <tr> <td style="width: 50%;"></td> <td style="width: 50%; padding: 5px;">correggere a step</td> </tr> <tr> <td style="width: 50%;"></td> <td style="width: 50%; padding: 5px;">valutare collegialmente</td> </tr> </table>	attendibilità	fissare termini di riferimento	standardizzazione	valutare obiettivi specifici		correggere a step		valutare collegialmente	<p>VERO O FALSO</p> <p>VERO O FALSO O VERO/FALSO</p> <p>SCELTE MULTIPLE</p> <p>COMPLETAMENTI</p> <p>CORRISPONDENZE</p>
attendibilità	fissare termini di riferimento								
standardizzazione	valutare obiettivi specifici								
	correggere a step								
	valutare collegialmente								

STRUTTURA DEI TEST

<p>Sezione 1 - VERO O FALSO</p> <p>Sezione 2 - VERO O FALSO O VERO/FALSO</p> <p>Sezione 3 - SCELTE MULTIPLE</p> <p>Sezione 4 - COMPLETAMENTO</p> <p>Sezione 5 - CORRISPONDENZE</p>	<p style="text-align: center;">Riscaldamento</p> <p style="text-align: center;">TEST</p> <p style="text-align: center;">DI VALUTAZIONE</p>
---	---


la in forma chiara, mentre per formularne una falsa un buon espediente è esprimere una convinzione di senso comune che il sapere contraddice. Nel tipo di item *vero o falso o vero/falso* c'è una

terza possibilità: che l'affermazione risulti vera o falsa a seconda dei casi. Nelle *scelte multiple* figurano una domanda con una risposta esatta e alcuni *distrattori*, risposte sbagliate da scartare. Di solito

SEZIONE VERO O FALSO O VERO/FALSO


ESEMPIO
La valutazione formativa mira ad accertare se gli allievi hanno raggiunto il livello di competenze previsto dalla programmazione

VERO FALSO VERO / FALSO
a seconda dei casi



RISPOSTA
La valutazione formativa mira ad accertare se gli allievi hanno raggiunto il livello di competenze previsto dalla programmazione

VERO FALSO VERO / FALSO
a seconda dei casi



ci sono una sola risposta esatta e due distrattori, ma sono possibili anche items con due risposte esatte e tre distrattori o con un numero più alto di distrattori. Per costruire questo tipo di item la procedura migliore è partire da un'affermazione esatta che riempie la parte di quesito e di risposta esatta. Poi si trovano i distrattori, badando a fare in modo che siano plausibili, credibili e non ingannevoli. Le prove oggettive che contengono distrattori "a trabocchetto" o astrusi sono mal fatte. Nei **completamenti** c'è un testo con delle parti da riempire. È preferibile elaborare appositamente testi, piuttosto che prenderli dai libri. Non bisogna chiedere di scrivere le parti mancanti, ma di riempirle scegliendole da un elenco di risposte esatte e distrattori elencate in calce. Nelle **corrispondenze** si chiede di connettere elementi logicamente legati per qualche ragione. Per evitare che dopo aver sistemato un certo numero di corrispondenze si proceda meccanicamente nelle ultime, è consigliabile che il numero di termini della seconda serie sia maggiore di quello della prima.

Come organizzare il questionario? Quando si va a predisporre il questionario della prova oggettiva, conviene raggruppare gli items per tipo, non per argomento. L'alunno non può passare da un *vero o falso* a una scelta multipla, a una corrispondenza senza incontrare difficoltà e rallentare, perché cambiano i processi mentali. È bene (anche se non sempre si fa) far precedere gli items su cui basare la valutazione da alcuni (bastano 4 o 5) **items di riscaldamento**, che servono a entrare nel tipo di logica della prova oggettiva e ad attivare le aree di conoscenza da esplorare. Il riscaldamento mette

tutti gli alunni sulla stessa linea di partenza, sia quelli più portati a questo genere di prove, sia quelli meno portati. Di solito per riscaldamento si adoperano i *vero o falso*, sia perché fanno fare una buona ginnastica mentale (soprattutto impongono di stare attenti al dettato esatto delle affermazioni), sia perché, siccome c'è il 50% di probabilità di indovinare rispondendo a caso, sono meno indicati nella valutazione vera e propria.

È importante che il test dia istruzioni chiare e non induca in errore l'alunno. Perciò è buona regola far precedere ciascuna sezione da un esempio del tipo di item su cui operare.

Come calcolare i punteggi? Solitamente si assegna 1 punto a ogni risposta esatta ai *vero o falso* (anche se molti preferiscono non includere questo tipo di item nel calcolo dei punteggi), 1 punto per ciascuna delle integrazioni esatte dei completamenti, 2 punti per ciascuna risposta esatta ai *vero o falso o vero/falso*, 2 punti per ciascun collegamento giusto nelle corrispondenze. In alcuni test si fa uso di punteggi negativi per gli errori, ma oggi la tendenza prevalente è a assegnare 0 punti sia alle risposte errate, sia alle risposte omesse.

Quanti items? Per decidere il numero di items da inserire nel test occorre tener conto da un lato delle cose che si vogliono verificare, dall'altro del tempo a disposizione e trovare un compromesso. In media per un item occorrono 1 o 2 minuti ed è sconsigliabile progettare prove che vadano oltre 60-90 minuti, perché affaticano e comportano un calo di attenzione nella seconda parte. Perciò un questionario dovrebbe contenere al più 30-50 items.

Come ripartire il numero di items stabilito?

Dobbiamo decidere quanti items includere per ciascun tipo, quanti dedicare a verificare un dato argomento e quanti ad accertare che sia stato raggiunto un determinato obiettivo didattico. Per questa parte del progetto del test, la più complessa e delicata, occorre tener presente che ciascun item è contrassegnato da tre parametri: l'**argomento** che verifica, l'**obiettivo educativo** che accerta e il **tipo**. Quanto agli argomenti, la cosa più semplice è suddividere la materia in *chunk*, blocchi che per noi (stando al corso svolto e a ciò che si è studiato) sono di pari importanza. Ad esempio, possiamo decidere di suddividere il tema della costruzione delle prove oggettive in tre parti egualmente estese: tipi di items, struttura dei test, progettazione. Si fissano poi gli obiettivi didattici e si stabilisce quanto vale ciascuno rispetto agli altri, quanto "pesa" in percentuale. Di solito si ricorre a una classificazione ispirata alla tassonomia di B.S. Bloom. Si possono considerare cinque obiettivi: conoscenza delle informazioni, competenza linguistica, capacità di organizzare i concetti, capacità di rielaborare e

obiettivi didattici	peso relativo
dati	?%
competenza linguistica	?%
organizzazione concettuale	?%
pensiero produttivo	?%
applicazioni	?%

produrre idee nuove, capacità di applicare.

È facile a questo punto calcolare quanti items assegnare a ciascun *chunk* e a ciascun obiettivo. Basta partire dal numero complessivo deciso e tener presente che ogni *chunk* dovrà avere lo stesso numero di items e che gli items dedicati a ciascun obiettivo sono proporzionali al peso dell'obiettivo. Per quanto riguarda i tipi di items è consigliabile adoperare in prevalenza le scelte multiple, per tutti gli argomenti e per tutti gli obiettivi. I completamenti vanno riservati all'accertamento della competenza linguistica, mentre per accertare l'organizzazione concettuale è bene introdurre alcune corrispondenze.

N. ITEMS	DATI	COMPETENZA LINGUISTICA	ORGANIZZAZIONE CONCETTUALE	PENSIERO PRODUTTIVO	APPLICAZIONI
TIPI DI ITEMS	?	?	?	?	?
STRUTTURA DEI TEST	?	?	?	?	?
PROGETTO	?	?	?	?	?

7. LA QUALITÀ DELLA FORMAZIONE

7.1. Valutare la formazione. Finora ci siamo occupati della valutazione dei risultati ottenuti lavorando con gli allievi. Si può pensare però di valutare la formazione nel suo complesso. Ci si può chiedere, ad esempio, qual è la qualità di un corso di formazione professionale o dei corsi tenuti in una scuola o di un ordine scolastico o dell'intero sistema scolastico di un paese.

L'interesse per la qualità della formazione è maturato negli ultimi decenni. Le prime esperienze di analisi di qualità in questo settore sono state condotte negli anni '60 negli Stati Uniti, quando si diffuse la preoccupazione che a causa di un'istruzione inefficiente il paese potesse restare indietro nella competizione tecnologica con l'URSS (*Unità 3*, § 1.3). A partire dagli anni '80, passata l'ondata della contestazione studentesca, anche i governi di altri paesi avanzati hanno cominciato a porsi il problema della qualità dei sistemi scolastici. Nel 1987, su iniziativa del Ministero dell'Educazione degli Stati Uniti, si è tenuta la Conferenza Internazionale di Washington sugli indicatori di qualità dell'istruzione. In quella sede è stato affidato all'OCSE (Organizzazione per la Cooperazione e lo Sviluppo Economico), in particolare alla sua emana-

zione in campo educativo, il CERI (Centro per la Ricerca e l'Innovazione dell'Insegnamento), il compito di portare avanti un'attività di studio e ricerca sulla qualità dei sistemi formativi.

Il fermento intorno al problema della qualità della formazione ha spinto gli studiosi a interrogarsi sul piano teorico, a mettere a punto metodi e tecniche, a sperimentare e a confrontarsi. È nato un filone di studi, che va solitamente sotto il nome di *Evaluation Research* (ricerca valutativa). Come si spiega l'esplosione dell'interesse per la qualità della formazione?

Lo sviluppo delle scienze della formazione ha accresciuto la consapevolezza dei processi formativi e ha reso gli interessati (utenti, amministratori, formatori) più esigenti, meno disposti a prendere le formazioni a scatola chiusa e più desiderosi di avere informazioni precise in merito. Varie ragioni hanno spinto i governi a interrogarsi sulla qualità dei sistemi scolastici. Hanno avuto un certo peso i motivi economici, di contenimento della spesa pubblica: nelle analisi di qualità si è visto un mezzo per razionalizzare gli investimenti per l'istruzione, prestando più attenzione al rapporto costi-benefici. I governi si sono trovati poi nella necessità di prendere rapide decisioni per adeguare i sistemi scolastici a esigenze nuove (aggiornamento dei contenuti, adeguamento dell'offerta a nuovi segmenti di popolazione, ecc.). Le analisi di qualità potevano dare il polso della situazione e consentire politiche più consapevoli e meditate. In parte l'interesse manifestato dai governi per le analisi di qualità è legato alla concorrenza tra i vari paesi e all'incremento degli scambi e delle comunicazioni internazionali. Ciascun paese si sforza di avere un sistema formativo competitivo o per lo meno all'altezza degli altri, che assicuri ai propri cittadini di circolare nel mondo a sviluppo avanzato in condizioni che non siano di inferiorità.

7.2. Obiettivi dell'analisi di qualità della formazione. L'analisi di qualità di una formazione o di un sistema formativo può essere fatta per scopi diversi, riconducibili per lo più a tre fondamentali.

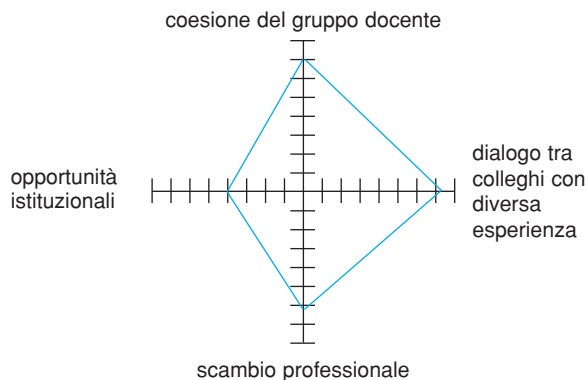
1. Certificazione. L'obiettivo può essere sapere qual è il livello della formazione e dell'istituzione che la eroga in rapporto a determinati standard. La certificazione del risultato servirà per accreditare formazione e sistema formativo presso enti, amministratori, utenza e per rendere trasparente l'offerta formativa.

2. Miglioramento dell'attività. Da un'analisi di qualità si possono trarre indicazioni per migliorare l'attività formativa utili a vari livelli: dai suggerimenti per i formatori a quelli per chi amministra, organizza, pianifica, decide le politiche (fig. 29).

3. Promozione di una cultura della qualità tra gli operatori. La cultura della qualità si contrappone solitamente alla cultura burocratica. In un sistema burocratico le attività sono rigidamente programmate, si lavora per procedure prestabilite, mentre in un sistema basato sulla cultura della qualità le attività sono elastiche e si adattano continuamente alle esigenze che emergono. Nel sistema burocratico ci si assicura che le persone seguano le procedure controllandone l'operato dall'alto. In un sistema basato sulla qualità si cerca di creare invece un clima di cooperazione e di miglioramento continuo: tutti collaborano con l'obiettivo di fare sempre meglio. Cultura burocratica e della qualità rispondono a due filosofie diverse della gestione delle organizzazioni. In generale oggi si tende a preferire la cultura della qualità, perché più aperta e più adatta a gestire personale istruito, come sono i lavoratori di oggi. Nella formazione poi è essenziale che operatori e formatori non si comportino da burocrati, ma si sentano partecipi di un lavoro collettivo teso al miglioramento continuo. Il modo di pensare e le motivazioni dei formatori sono infatti decisive per la riuscita della formazione. Le analisi di qualità possono promuovere la cultura della qualità in un'organizzazione formativa, se vengono ripetute nel tempo, fino a diventare un vero e proprio monitoraggio dell'attività e del

Fig. 29. Dati utili per migliorare la formazione.

Il grafico riproduce i risultati di una rilevazione fatta nel quadro di un'analisi di qualità degli istituti di formazione di una forza armata. Si riferisce ai dati su come i docenti percepiscono le proprie opportunità di crescita professionale. Il grafico fa vedere che i docenti pensano di poter contare a sufficienza sul rapporto con i colleghi, improntato a coesione e dialogo, ma considerano carenti le opportunità che l'istituzione offre di crescere professionalmente. Sentono che il sistema non investe abbastanza sulla loro crescita professionale di insegnanti. Risultati del genere possono indurre l'istituzione a dare maggior impulso alla formazione dei formatori o a organizzarla diversamente. Un docente che si sente considerato e valorizzato dal sistema renderà con ogni probabilità di più.



sistema e se i risultati vengono resi noti agli operatori della formazione e discussi con loro in incontri periodici. L'analisi di qualità diviene allora un'occasione per riflettere assieme agli altri sul lavoro che si porta avanti, per pensare interventi di miglioramento e per avere feedback sugli interventi che si attuano.

7.3. Analisi di qualità e ricerca scientifica. Quando si fa un'analisi di qualità di una formazione o di un sistema formativo si porta avanti uno studio, con una riflessione teorica e con indagini empiriche. A volte (è la via generalmente seguita quando si tratta di analizzare interi sistemi scolastici di paesi) ci si limita a lavorare su dati esistenti e a elaborarli statisticamente nell'ottica dell'analisi di qualità. In altri casi si conducono ricerche sul campo mediante questionari, test, interviste o altri mezzi. Anche se capita che amministratori o formatori cerchino di portare avanti da soli analisi di qualità, a svolgere il lavoro sono in genere professionisti ed esperti, che padroneggiano lo strumentario concettuale e tecnico delle scienze sociali. L'analisi di qualità è ricerca scientifica?

Sulla questione si è discusso a lungo. Per alcuni (P.H. Rossi 1982) l'analisi di qualità è un campo delle scienze sociali, per altri (H. Talmage 1982), anche se si avvale di presidi scientifici, è essenzialmente un'attività politico-decisionale, manageriale, per altri ancora (D.L.

RICERCA SCIENTIFICA	ANALISI DI QUALITÀ
<p>è orientata alla scoperta mira a verificare ipotesi ottiene risultati di interesse generale è più neutrale</p>	<p>è orientata alle decisioni mira a confrontare il caso in esame con uno standard ottiene risultati di interesse specifico è più legata alle politiche gestionali</p>

Fig. 30

Stuffledeam 1967, M. Scriven 1967) si colloca a metà strada tra ricerca e management. Tutto sommato sembra ragionevole dire che l'analisi di qualità non è ricerca scientifica, sebbene richieda competenze da ricercatore (almeno in chi guida il lavoro) e abbia molto in comune con la ricerca nelle scienze sociali. Per rendersi conto che non si tratta di vera e propria ricerca basta passare in rassegna alcune differenze tra analisi di qualità e ricerca scientifica (fig. 30).

7.4. Dilemmi dell'analisi di qualità. È un errore pensare che l'analisi di qualità della formazione sia un lavoro essenzialmente tecnico, che basti cioè conoscere le procedure e metterle in pratica. I problemi principali di un'analisi della qualità della formazione sono quelli di impostazione, cioè riguardano i presupposti teorici di fondo.

C'è innanzitutto la *questione della neutralità dell'analisi*: l'indagine può limitarsi a raccogliere dati in maniera distaccata senza fornire giudizi e senza entrare nel merito delle decisioni da prendere? Possiamo pensare che l'analista della qualità che, ad esempio, studia un'istituto scolastico fornisca una descrizione e al più una spiegazione di ciò che avviene e che spetti poi al preside o ad altri decisori istituzionali trarre le conclusioni. L'esperienza ha dimostrato che un approccio del genere è poco sostenibile. È difficile, se non impossibile, come sempre nelle scienze sociali, che chi conduce lo studio si limiti a una ricostruzione neutrale dei fatti. Tenderà comunque a introdurre il proprio punto di vista. C'è però un altro problema. Se a chi deve istituzionalmente decidere si mettono in mano dati senza giudizi e suggerimenti, è facile che i risultati vengano fraintesi. Anche senza volerlo il decisore istituzionale tenderà a leggerli in modo da trovarvi conferma a scelte che ha già in animo di fare. Così l'analisi di qualità rischia di essere semplicemente uno strumento di legittimazione delle politiche, delle pianificazioni e delle programmazioni decise per altre ragioni.

Si pone a questo punto la *questione dei valori di riferimento*. Se le analisi di qualità non sono neutrali, ma esprimono giudizi, in base a quali criteri valutano? A quali valori si ispirano? Che cosa considerano desiderabile e che cosa no? Quale punto di vista tengono in considerazione? Quello dei politici, degli amministratori, dei manager, dei docenti o quello delle persone da formare o quello delle realtà (mondo del lavoro, altre istituzioni formative, ecc.) che fruiranno del bagaglio acquisito dagli individui in quella formazione?

A riguardo due approcci hanno fatto scuola, entrambi con pregi e difetti.

1. La valutazione per obiettivi. Si prende per riferimento ciò che il corso si prefigge o ciò che l'istituzione formativa mira a ottenere con l'attività formativa. Il modello classico di valutazione per obiettivi è il *rationale di Tyler* (1942, 1949). Come dice il nome, lo sforzo è di dare un fondamento razionale alla valutazione (in inglese *rationale* = base logica). Il principio in effetti è semplice e di buon senso: per valutare un corso occorre vedere se alla fine gli allievi hanno raggiunto gli obiettivi previsti. Occorre perciò prima precisare gli obiettivi, classificandoli e definendoli uno per uno in termini di comportamenti e contenuti cognitivi, poi individuare tecniche per misurare i risultati, misurarli e infine confrontarli con gli obiettivi. Il *rationale* di Tyler è rigido, perché considera il prima e il dopo e non il cammino formativo. Non ammette perciò che nel corso della formazione gli obiettivi iniziali vengano rivisti. Altri autori hanno proposto modelli più dinamici e flessibili, in cui ci sono più momenti di definizione degli obiettivi e di determinazione dei risultati.

La valutazione per obiettivi ha il pregio di far uscire la formazione dal vago e di costringere a riflettere su ciò che si vuole ottenere, specificandolo in termini precisi e concreti. Presenta però seri limiti. Come ha messo in evidenza Scriven (1967), quando precisiamo gli obiettivi di una formazione, facciamo riferimento a un progetto ideale che di solito non trova rispondenza, o non trova piena rispondenza, nella realtà. Di fatto la formazione è diversa e di solito più ricca dei progetti. Nel corso della formazione si raggiungono obiettivi non previsti e di

cui a volte non ci si rende neanche bene conto, ma che sono di grande importanza. Una valutazione basata esclusivamente sugli obiettivi prestabiliti rischia di essere parzialmente cieca, di non tener conto di ciò che effettivamente fa la formazione.

Un altro problema è che nella valutazione per obiettivi hanno voce in capitolo soltanto gli erogatori della formazione. Si parte dal presupposto che ciò che l'istituzione formativa giudica una meta desiderabile sia davvero tale per le persone da formare. Valutazioni così condotte privilegiano il punto di vista dell'istituzione, mettendo in secondo piano le persone e i loro bisogni.

2. La goal-free evaluation. È l'approccio elaborato da M. Scriven (1967). Invece di far riferimento agli obiettivi istituzionali previsti dal progetto formativo, la valutazione si basa sui bisogni delle persone coinvolte nella formazione (allievi, formatori, responsabili, committenti, datori di lavoro che utilizzeranno il personale formato, ecc.). Scriven li definisce tutti **consumatori della formazione** e chiama **diretti** quelli che prendono parte all'attività formativa e **indiretti** quelli che se ne giovano a distanza. L'approccio teorico è di tipo economico: la formazione è un prodotto offerto a un ambiente sociale che lo consuma. La valutazione va fatta sulla base delle esigenze dell'ambiente sociale che ne fruisce. Perciò si parla anche di **approccio centrato sui consumatori**.

La *goal-free evaluation* risolve i problemi della valutazione per obiettivi, ma ne fa nascere di altri. Il più grave è che il consumatore diviene sovrano e c'è il rischio di derive etiche, di finire per valutare cioè sulla base di bisogni che i consumatori certamente hanno, ma che non sono nobili e più che essere assecondati vanno scoraggiati. Ad esempio, per ragioni diverse i consumatori di un sistema scolastico possono trovarsi d'accordo nel trascurare insegnamenti poco utili in pratica, ma che favoriscono la crescita personale. Anche se c'è soddisfazione di tutti, dal punto di vista etico la cosa è inaccettabile. Spetta all'istituzione salvaguardare l'eticità dell'insegnamento. Se la valutazione per obiettivi è sorda ai bisogni delle persone, la *goal-free evaluation* rischia di perdere di vista l'ethos istituzionale.

L'ideale è integrare i due approcci e condurre analisi in cui si valuta facendo riferimento sia agli obiettivi, sia ai consumatori.

ETICA DELLE CONSEGUENZE E ETICA DEI PRINCIPI

L'alternativa fra valutazione per obiettivi e *goal-free evaluation* rinvia a una questione fondamentale della riflessione filosofica sulla morale: il dilemma etica delle conseguenze/etica dei principi. Nell'etica delle conseguenze un'azione è giusta se gli effetti che produce sono positivi, mentre nell'etica dei principi è giusta se in sé, positivi o negativi che siano gli effetti che produce, risponde a determinati doveri che abbiamo. È un'etica delle conseguenze l'**eudemonismo** (dal greco *eudaimonía* = felicità), che ha dominato il pensiero antico e che si può sintetizzare nella massima «il bene è una vita felice». L'eudemonismo antico ha il difetto di essere individualistico, cioè cade dinanzi all'obiezione che ciò che rende felice il singolo può essere dannoso per la collettività. La versione moderno-contemporanea dell'etica delle conseguenze è l'**utilitarismo**, che non è più individualistico, ma sociale. Nasce nel Settecento, con Hutcheson, Helvetius, Beccaria, si consolida nell'Ottocento, con i padri fondatori Jeremy Bentham e John Stuart Mill, ispira il pensiero di vari filosofi contemporanei detti neoutilitaristi (R.B. Brandt, N. Smart, J.J. Smart, J. Flechter, M. G. Singer, R. Hare, J. Harsanyi). Si basa sul principio dell'utile o della massima felicità: il bene è, nelle parole di Hutcheson, «la maggior felicità per il maggior numero di persone».

L'etica dei principi si ritrova in San Tommaso e nella tradizione tomistica, nel giusnaturalismo (che fa appello alla legge naturale), ma ha la sua massima espressione nell'etica di Kant. Per Kant il bene è agire secondo i comandi incondizionati della ragione che ci guida nelle scelte pratiche. Al momento di stabilire che cosa è giusto fare bisogna chiedersi: che cosa accadrebbe se tutti faces-

sero così? Se non è ammissibile che l'azione diventi generale, è nostro dovere astenercene.

La *goal-free evaluation* si ispira all'etica utilitaristica delle conseguenze: se gli interessati (docenti, alunni, ecc.) sono contenti, la scuola va bene. In particolare riprende il neoutilitarismo, che considera la felicità alla stregua di un bene di consumo sul quale ognuno ha diritto ad avere le proprie preferenze. Così l'ottica è quella di un'economia sociale della felicità. La valutazione per obiettivi può ispirarsi sia all'etica delle conseguenze, sia all'etica dei principi. Gli obiettivi che un'istituzione si pone infatti possono essere dettati da principi ritenuti inderogabili come dal calcolo sugli effetti delle azioni. Va tenuto presente che spesso oggi le politiche sociali istituzionali si ispirano all'etica utilitaristica delle conseguenze. È importante però il fatto che nella valutazione per obiettivi possano entrare in gioco anche i principi.

Un importante limite dell'etica delle conseguenze sta nel fatto che l'utile non coincide necessariamente col giusto. In particolare un'azione può produrre l'utile generale, la felicità di tutti secondo le preferenze di ciascuno, e ciononostante risultare ingiusta. Come ha notato il filosofo contemporaneo A.C. Ewing «i principi utilitaristici, logicamente sviluppati, darebbero luogo a molti più imbrogli, menzogne e azioni inique di quanti un qualsiasi uomo buono possa tollerare». Ad esempio, mettiamo che nello sforzo di aprire un sistema scolastico a tutti si faccia una politica di abbassamento del livello di istruzione, semplificando i contenuti di insegnamento. Teoricamente potrebbe accadere che tutti siano soddisfatti per preferenze (cioè calcoli delle conseguenze) diverse: chi perché riesce dove non riusciva, chi perché ha poco da sforzarsi, ecc. Tuttavia si può sospettare che una politica scolastica del genere inganni gli interessati: se per vivere nel mondo attuale c'è bisogno di maggiore istruzione, è come se la scuola stesse vendendo merce scadente col consenso del consumatore. Se chi conduce l'analisi di qualità valuta nell'ottica utilitaristica della *goal-free evaluation*, una scuola così fatta è sicuramente di qualità. Nella valutazione per obiettivi invece l'analista ha modo di appellarsi al principio e può mettere in evidenza quei difetti che diversamente non emergerebbero.

D'altra parte l'etica dei principi richiede una coerenza a volte eccessiva e rischia di scivolare nella rigidità e nella mancanza di senso pratico. Inoltre, mentre alcuni principi (ad esempio il rispetto della vita umana) sono chiaramente universali, altri sono espressione di una tradizione tra le tante possibili. Restare ancorati a certi principi può voler dire allora incapacità di cambiare e di adattarsi alle novità. Ad esempio, l'analista che valuta una scuola che ha semplificato i contenuti dovrebbe chiedersi se un insegnamento di questo tipo non possa essere un valido adattamento dell'istruzione tradizionale al mondo di oggi: quali vantaggi presenta? quali limiti? le carenze possono essere compensate incentivando la formazione extrascolastica? e via dicendo.

7.5. Un modello ecosistemico per lo studio dei sistemi formativi. Per ragioni pratiche le analisi di qualità devono tradursi in resoconti facilmente consultabili e confrontabili l'uno con l'altro. Perciò anche se sono moltissime le cose che potremmo indagare in un sistema formativo, dobbiamo concentrarci su alcuni elementi ritenuti fondamentali. Si cerca anche di esprimere con numeri i risultati. L'esigenza di limitare il campo e l'esigenza di resoconti quantitativi portano a far uso di *indicatori*, cioè a lavorare su fatti empirici che fanno da spia ai fenomeni da indagare e che possono essere misurati con opportuni metodi.

Un modo efficace di classificare gli indicatori di qualità di un sistema formativo consiste nel collocarli in un modello di tipo ecosistemico. Si parte dal presupposto che la scuola (come per altro si fa oggi per tutte le organizzazioni) sia un'*unità produttiva in rapporto dinamico con l'ambiente*. La scuola si colloca in un ambiente materiale, economico, sociale, storico e culturale con cui interagisce. Da un lato incamera risorse e stimoli dalla realtà circostante, dall'altro emette prodotti che hanno effetti sulla realtà circostante. Perciò trasforma il mondo e ne viene trasformata, inserendosi in un circuito continuo (fig. 31). In quest'ottica gli indicatori di qualità si distinguono a seconda del momento, del punto in cui andiamo a esaminare l'interazione scuola-ambiente.

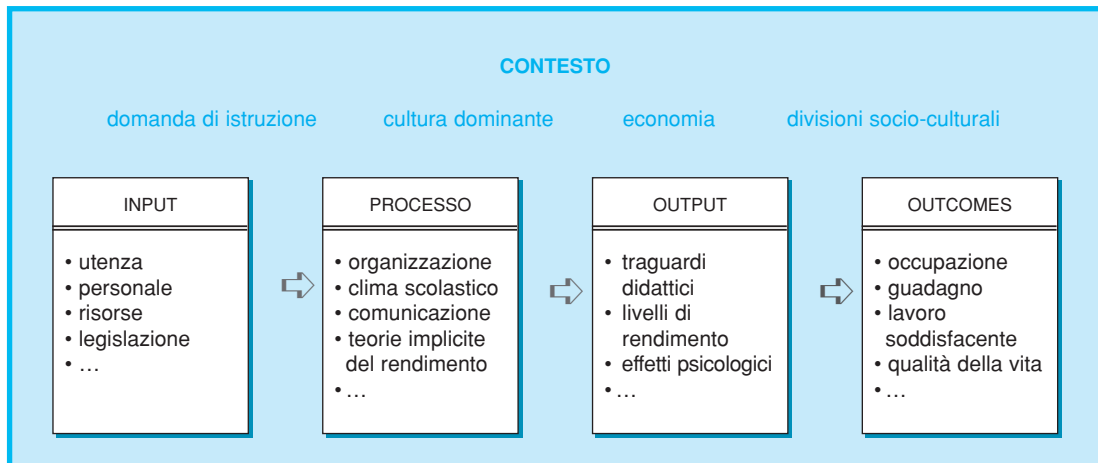


Fig. 31. Modello sistemico di scuola come unità produttiva (da Jaap Scheerens 1992).

1. Indicatori di input. Ci dicono su che cosa può contare la scuola per lavorare, in termini di alunni, insegnanti, risorse finanziarie, attrezzature, provvedimenti legislativi, ecc.

2. Indicatori di output. Fanno riferimento al prodotto immediato della scuola, cioè al raggiungimento degli obiettivi didattici e ai livelli di rendimento degli allievi.

3. Indicatori di outcomes. Guardano agli esiti, agli effetti a distanza della scolarizzazione, come l'occupazione, il guadagno, la soddisfazione nel lavoro, la qualità della vita.

4. Indicatori di processo. Sono i più interessanti, perché non si limitano a guardare ciò che entra e ciò che esce dalla scuola, ma entrano nel merito del suo funzionamento interno.



1. Bandura A. (1977) *Social learning theory*; Prentice-Hall, Englewood Cliffs, N.J.
2. Barbier J.M. (1987) *L'évaluation en formation*, PUF, Paris; trad. it. *La valutazione nel processo formativo*, Loescher, Torino, 1989
3. Becker G.S. (1964) *Human capital*, New York
4. Cambi F., Frauenfelder E. (1994) *La formazione*, Unicopli, Milano
5. Dewey J. (1929) *The sources of a Science of Education*, Livering Publ. Co, New York; trad. it. *Le fonti di una scienza dell'educazione*, La Nuova Italia, Firenze, 1951
- Dweck C.S. & Benpechat J. (1983) *Childrens theories of intelligence: consequences for learning*, in S.G. Paris e. all. (a cura di) *Learning and motivation in the classroom*, Erlbaum, Hillsdale
6. Ferster C.B., Skinner B.F. (1957) *Schedules of reinforcement*, Appleton-Century-Crofts, New York
7. Gage N.L. (1964) *Theories of learning and instruction*, University of Chicago Press, Chicago
8. Green T.F. (1964) *A topology of the teaching concept*, in «Studies in Philosophy of Education», n. 3
9. Harlow F.H. (1949) *The formation of learning sets*, in «Psychological Review», n. 56
10. Köler W. (1917) *Intelligenzprüfungen an Antropoiden*, in «Atti dell'Accademia Prussiana delle Scienze»; trad. it. in *L'intelligenza delle scimmie antropoidi*, Editrice Universitaria, Firenze, 1960
- Lepper M.R., Greene D., Nisbett R.E. (1973) *Undermining children's intrinsic interest with extrinsic rewards: a test of the overjustification hypothesis*, In «Journal of Personality and Social Psychology», n. 28
- Lewin K. (1935) *A dynamic theory of personality*, McGraw-Hill, New York; trad. it. *Teoria dinamica della personalità*, Giunti-Barbera, Firenze, 1965
11. Maehr M.L. (1976) *Continuing motivation: a analysis of a seldom considered educational outcome*, in «Review of Educational Research», n. 14
- Miller N.E., Dollard J. (1941) *Social learning and imitation*, Yale University Press, New Haven
12. Pavlov I. P. (1923) *Uslovnyi refleks*, trad. it. *Il riflesso condizionato*, Editori Riuniti, Roma, 1968
13. Polanyi M. (1966) *The tacit dimension*; trad. it. *La conoscenza inespresa*, Armando, Roma, 1979
14. Rossi P.H. (1982) *Standard for evaluation practice*, Jossey Bass, San Francisco
15. Schherens J. (1992) in OECD (ed.), *L'OECD et les indicateurs internationaux de l'enseignement. Un cadre d'analyse*, OECD, Paris; trad. it. *Il sistema degli indicatori del funzionamento della scuola*, in OECD, *Gli indicatori internazionali dell'istruzione*, Armando, Roma, 1994
16. Scriven M. (1967) *The methodology of evaluation*, in Tyler R.W., Gagné R.M., Scriven M.; *Prospective of curriculum evaluation*, Rand McNally, Chicago,
17. Skinner B.F. (1953) *Scienze and human behavior*, The Macmillan Company, New York; trad. it. *Scienza e comportamento*, Franco Angeli, Milano, 1971
18. Stufflebeam D.L. (1967) *Evaluation ad enlightenment for decision making*, ristampato in W.H Beatty, A.B. Walcott (eds.), *Improving educational assessment and an inventory of measures of affective behavior*, Supervision and Curriculum Development, Washington
19. Tyler R.W. (1942) *General statement on evaluation*, in «Journal of Educational Research», n. 35
20. Tyler R.W. (1949) *Basic principles of curriculum and instruction*, University of Chicago Press, Chicago
21. Talmage H. (1982) *Evaluation of programs*, in H.E. Mitzel (ed.) *Encyclopedia of educational research*, The Free Press, New York
22. Thorndike E.L. (1898) *Animal intelligence*, in «Psychological Review Monograph Supplement», n.8; trad. it. parziale *L'intelligenza degli animali* in N. Dazzi e L. Mecacci (a cura di), *Storia antologica della psicologia*, Guinti-Barbèra, Firenze, 1982
23. Vanderberg D. (1969) *Theaching and learning*, University of Illinois Press, Urbana-London