

[dal Blog di Giorgio Israel, 20.11.2010]

## A PROPOSITO DELLA VALUTAZIONE DI SCUOLE E DOCENTI

*Sono stati proposti a livello ministeriale due modelli di valutazione da sperimentare in gruppi di scuole per valutare scuole e docenti.*

*Mi astengo al momento da una analisi specifica di questi modelli, su cui ho le più grandi perplessità: lo farò in altra prossima occasione.*

*Mi limiterò, per ora, a proporre un documento che presenta una serie di idee e che non è stato degnato di alcuna considerazione.*

*dal Blog di Giorgio Israel, 20.11.2010*

Uno dei problemi principali che si pone nel progettare un sistema di valutazione è quello delle fonti da cui deve esso attingere nelle sue varie articolazioni (input) e quello della destinazione e dell'uso del suo lavoro (output). A questo va aggiunto il problema dell'uso degli input, inclusi quelli che il sistema di valutazione costruisce con il suo stesso lavoro. Al riguardo non condivido l'idea diffusa secondo cui «la scelta e la definizione delle prove spetta all'autonomia scientifica dell'Istituto» (in buona sostanza, all'Invalsi). Essa è corretta in termini stretti, ma ritengo che non ci si possa esimere da una definizione precisa dei metodi che debbono essere applicati dall'Istituto (altrettanto dicasi per la componente esterna di valutazione, ovvero per il sistema di ispezione). Altrimenti si rischia di costruire una struttura di cui non è chiara l'utilità e che può benissimo risolversi nella creazione di un gigantesco e dispendioso carrozzone che non assolve alcuna funzione utile nella valutazione del sistema dell'istruzione, acquisendo dati di scarsa utilità, elaborando valutazioni inattendibili e, in fin dei conti, fornendo elementi che rischiano di essere persino fuorvianti.

Un esempio di carattere generale è rappresentato dal dato della dispersione scolastica, degli abbandoni o delle “bocciature”. A mio avviso, tale dato viene usato in alcune procedure di valutazione in modo assolutamente superficiale e poco responsabile, poiché esso non ha soltanto una fortissima dipendenza sociale e territoriale che rende i dati spesso non omogenei e inconfondibili. Ma, soprattutto, è banale che la consapevolezza di essere valutati mediante il fattore “abbandoni” o “insuccessi scolastici” produca un effetto immediato, che è quello di abbassare l'asticella della valutazione, promuovendo tutti. Ho potuto sperimentare in modo diretto questo effetto all'università, dove l'introduzione del parametro “laurea in tempo” ha immediatamente prodotto un effetto di promozione di massa. Pertanto, trovo necessario che sulla questione delle informazioni di tipo amministrativo si diano indicazioni generali che escludano parametri che interferiscano con la qualità dell'insegnamento.

Vorrei tuttavia soffermarmi su una questione più di merito che riguarda i test, ovvero il tipo di test da usare e i limiti che ci possiamo generalmente attendere da test a risposta chiusa, gli unici che possono essere adatti a una valutazione quantitativa, poiché nel caso di risposte aperte è evidente che la valutazione non può andare oltre la classica scala del voto.

Qui sono costretto a tornare su una questione che ho sollevato più volte, ovvero che non bisogna farsi troppe illusioni circa la possibilità di raggiungere valutazioni “standard”, “oggettive”, “esatte”, basate su “misurazioni di prestazioni qualitative”, ecc. Prego di riflettere sul fatto che non si tratta di una diatriba

accademica o ideologica, ma di una questione che ha implicazioni pratiche estremamente concrete, come cercherò di mostrare con qualche esempio. Ritengo che non si possa passare sopra le modalità della valutazione, finendo con l'ignorare che - come dicevo prima - rischiamo di avere in mano risultati privi di qualsiasi utilità.

Uno dei pochi casi in cui si può parlare di una valutazione "oggettiva" e grossolanamente quantificabile delle capacità dell'alunno, è quello dell'ortografia e della grammatica e, entro certi limiti, anche della sintassi. Non appena si va oltre, ovvero sul terreno della capacità espressiva e dello stile, il ventaglio delle capacità si allarga a un punto tale che è difficile costruire una scala di merito esente da giudizi qualitativi con forti componenti soggettive.

Ho raccolto scetticismo riguardo la mia tesi secondo cui anche nel caso della matematica le indicazioni fornite dai test sono estremamente limitate. Voglio motivare con alcuni esempi questa affermazione: le implicazioni sono rilevanti, perché mostrano a cosa possono servire i test.

Chiunque abbia una sufficiente esperienza di insegnamento della matematica sa che la risoluzione corretta di un problema proposto (ottenimento del risultato esatto con procedure corrette) è una componente di base minima nella valutazione. Diciamo che essa rappresenta niente più che la "sufficienza". Al di là di questo livello minimo, vi sono molti fattori: e, in particolare, la scelta della metodologia più adeguata ed efficace. Questa efficacia è quella che in matematica viene spesso chiamata "eleganza" nella procedura che non è affatto un fattore meramente estetico, bensì espressione delle capacità di uso autonomo e persino originale delle conoscenze acquisite e riflette la creatività e autonomia dello studente (proprio quel livello generalmente incluso nelle "competenze"). Il ventaglio di queste capacità può essere molto ampio. Sarebbe un grave errore ritenere che ciò si manifesti soltanto a livelli superiori. Al contrario, esso si manifesta fin dai primi contatti con l'apprendimento della matematica. Due esempi e alcune osservazioni.

Esempio 1. - Un problema posto a un gruppo di bambini della scuola elementare (quarta): un sommozzatore è sceso 8 metri sott'acqua; scende ancora di 4 metri e poi risale di 9; a quanti metri sott'acqua si trova dopo queste manovre? Un problema del genere può ovviamente essere dato come test a risposta chiusa. La risposta esatta è stata conseguita dalla maggioranza dei bambini con procedure assai diverse: c'è chi ha semplicemente tradotto il problema in somme e sottrazioni; c'è chi addirittura ha ragionato in termini di numeri negativi, avendo evidentemente acquisito tale conoscenza in altra sede; c'è chi ha invece fatto ricorso a una rappresentazione grafica, del tipo ordinata cartesiana, pur non avendone esplicita consapevolezza, il che è ancor più significativo. Tralascio altre soluzioni. In generale, osservo che in molti casi, semplicissimi problemi danno luogo a un numero impressionante di risposte diverse. La valutazione della risposta al problema precedente in termini di correttezza numerica dice certamente qualcosa circa l'esistenza di conoscenze di base tali da portare alla soluzione esatta, ma poco assai circa le capacità (o "competenze") nella risoluzione del problema. Personalmente considererei il procedimento migliore quello geometrico-intuitivo perché dimostra capacità autonome e un'inventiva notevole da parte del bambino, e porrei al secondo livello il secondo e all'ultimo il primo. È probabile che altri insegnanti la pensino, legittimamente, in modo diverso. Ma quel che è certo è che una "somministrazione" (come si dice con linguaggio lassativo)

del problema sotto forma di test, occulta in modo completo gran parte degli aspetti che vengono fatti rientrare comunemente sotto la voce “competenze”, perché ignora il modo con cui si è giunti al risultato corretto, e produce una valutazione sommaria, elementare e appiattita sulla voce “conoscenze di base” e capacità elementari di applicarle. Potrei produrre molti altri esempi di questo tipo ma mi limito a questo tanto per dare un’idea.

Esempio 2. - Nel corso di una procedura di valutazione sui “fondamenti e didattica della matematica” per un’abilitazione di insegnanti della scuola primaria che, dato il numero molto elevato di concorrenti, si è dovuto svolgere con test a risposta chiusa, si è svolta una approfondita discussione circa le modalità da seguire; discussione che ha visto la partecipazione di docenti universitari specialisti di tecniche docimologiche. Pur ammettendo la necessità inevitabile di ricorrere ai test chiusi, il parere unanime è che essi non potevano fornire altro che indicazioni estremamente rozze e di base tali da non accertare altro che il possesso di conoscenze elementari e di capacità elementari di farne uso, tali da non poter garantire altro che il futuro maestro conoscesse i concetti elementari della matematica e non commettesse errori di ragionamento clamorosi. Un esempio di test del genere: si consideri l’insieme dei numeri pari e l’insieme  $A$  dei numeri  $\{1, 2, 3, 4, 5, 6\}$ . Qual è l’intersezione? Risposte possibili:  $A$ ,  $2$ ,  $\{2, 4, 6\}$ , l’insieme vuoto. Lascio a voi giudicare quale livello di conoscenza del sistema dei numeri interi si ricavi da un simile test, salvo la verifica indiretta che il maestro sappia cosa sia un numero pari.

Osservazione. - Nel corso di una discussione approfondita con esperti di valutazione a livello istituzionale si è concordemente ammesso che i test a risposta chiusa in matematica forniscono indicazioni estremamente limitate. Tali indicazioni si riducono, ai vari livelli scolastici, nella rilevazione di un possesso elementare delle nozioni richieste dai “programmi” e da una loro capacità basilare di applicarle. Es: nelle elementari, le tabelline, la conoscenza delle operazioni numeriche e le loro regole fondamentali, la conoscenza di alcune definizioni geometriche, di alcune formule delle aree o dei perimetri, ecc. Ma non riflettono in alcun modo le capacità autonome e quindi quei livelli che sono significativi per individuare una significativa qualità degli apprendimenti. Peraltro, non appena si sale a livelli superiori (licei) un problema matematico può essere risolto in moltissimi modi diversi difficili da collocare in una scala di qualità, se non con un giudizio qualitativo la cui traduzione quantitativa (rozza) non va oltre il voto. È superfluo inoltre dire che, se la valutazione del test può essere standardizzata in modo completo, la scelta dei test contiene un elemento soggettivo fortissimo, come mostra l’esempio 2 (siamo tutti d’accordo che un livello del genere e una simile domanda sia utile a stabilire la adeguatezza a insegnare la matematica nelle primarie?).

Una discussione analoga sul sistema dei test per altre materie molto più “qualitative” mostrerebbe criticità ancor più evidenti.

Tutto ciò per concludere che il sistema dei test non è da proscrivere ma bisogna essere perfettamente consapevoli che esso non può andar oltre il rilevamento di un livello di “literacy” e “numeracy” minimale. Esso può consentirci di verificare se il sistema scolastico forma dei giovani al di sopra di una soglia minima sul piano soprattutto delle “conoscenze”, molto meno sul piano delle “competenze” - adotto questa

distinzione che non mi appassiona affatto, e quanto precede può far intuire perché. Sarebbe un grave errore attendersi qualcosa di più.

Nello specifico ritengo che l'Invalsi si sia mosso recentemente in modo accettabilmente "moderato", utilizzando il sistema dei test cum grano salis. Ho anche apprezzato l'approccio ragionevole dei tecnici che vi lavorano e, in particolare, il non attribuire un valore salvifico e assoluto alle statistiche su un tema così intriso di aspetti soggettivi e qualitativi qual è il sistema dell'istruzione. L'auspicabile potenziamento dell'Invalsi, che può consentirgli di estendere la rilevazione mediante test all'intero sistema, non deve - a mio avviso - modificare la struttura istituzionale e l'approccio metodologico seguito. Anzi, è bene che contenga qualsiasi propensione ad ambizioni totalizzanti.

Se quanto precede è vero nel caso della valutazione degli apprendimenti degli allievi, esso è ancor più vero nel caso della valutazione dei docenti, come già osservato nel contesto dell'Esempio 2. Per quel che riguarda i compiti dell'Invalsi, ritengo che esso debba restare rigorosamente fuori da una valutazione dei docenti, se non, al limite, anche qui con test di base volti a individuare la presenza di elementi di minima decenza. Può fornire invece elementi di valutazione delle scuole nel loro complesso, integrando i test di base (nei limiti anzidetti) con tutte le informazioni provenienti dal livello amministrativo, ma con rigorosa esclusione di parametri fasulli e fuorvianti come abbandoni scolastici, e "insuccessi" formativi.

Da quanto precede deriva una conseguenza concreta: la valutazione dei docenti - elemento di grave criticità e di notevole urgenza, come osservato nel documento - deve essere affidata alle ispezioni. Queste ultime non debbono avere soltanto il compito di valutare il singolo istituto scolastico e il suo dirigente, ma anche i singoli insegnanti.

Anche qui occorre liberarsi di alcune scorciatoie illusorie che possono rendere il sistema di valutazione semplice quanto inefficace e fonte di veri e propri errori. Alludo, in particolare, all'idea circolante di fare degli "utenti" - studenti e famiglie - i principali attori della valutazione della scuola e dei docenti. È, ripeto, una soluzione facile: basta distribuire schede di valutazione. Lo si fa già da tempo in diverse università, per fortuna senza effetti pratici ovvero a scopo puramente sperimentale, con risultati tra l'inquietante e il grottesco; soprattutto se tale sistema viene fatto in regime di anonimato. Non intendo soffermarmi sui numerosissimi esempi negativi che è facile immaginare. Vale qui la considerazione generale che ci troviamo in un contesto che favorisce quella mentalità secondo cui troppi genitori si fanno "sindacalisti" dei figli e hanno come principale intento ottenere il massimo della valutazione in cambio del minimo impegno, costringendo talora a cedimenti i dirigenti scolastici soggetti a maggiori pressioni. Potrei produrre un'ampia documentazione di lettere di insegnanti e anche di dirigenti oppressi da questa situazione. Essa deriva da un'idea banalmente sbagliata e cioè che la scuola sia un'azienda fornitrice di beni e servizi e studenti e famiglia l'utenza. Ma il processo di formazione che la scuola realizza non è né un bene né un servizio che possa essere valutato in termini di "customer satisfaction" come si giudica il barattolo di conserva fornito dal supermercato o la qualità del servizio fornito da uno sportello bancario. Se la conserva è deteriorata è mio diritto chiedere il risarcimento, ma la bocciatura non equivale a una conserva deteriorata: può anzi riflettere un'attività formativa rigorosa e valida.

Pertanto, il sistema migliore di valutazione dell'istituto scolastico e dell'insegnante è quello condotto dai competenti in materia, non da chi può aver interesse a ottimizzare una funzione di utilità che mira a ottenere il massimo risultato col minimo sforzo. Come è stato giustamente osservato da un imprenditore a un recente convegno confindustriale cui sono intervenuto, una buona scuola è quella che stimola il bravo studente a studiare di più e quindi a trovare nel successo motivi per andare avanti e non quella che stabilisce "benchmark" standardizzati e persegue in tal modo l'appiattimento, magari verso il basso, per ottenere il successo formativo garantito. Questa non è una scuola meritocratica e risponde a criteri caratteristici di una società non liberale.

Siamo pertanto ricondotti al sistema delle ispezioni. Da quanto precede, discende anche una risposta alla questione se occorra seguire il modello Ofsted oppure no. Ritengo che il corpo degli ispettori non debba essere un'autorità indipendente a quella maniera, se pure deve godere di caratteristiche di indipendenza indiscusse. Penso che il modello migliore sia quello di un corpo composto da ispettori e dagli stessi insegnanti (beninteso quelli qualificati allo scopo), e da altre eventuali figure come insegnanti pensionati, universitari, ecc. - anche se molte questioni sono aperte e da discutere in dettaglio: come conferire il "patentino", chi lo deve conferire, come costituire le commissioni d'ispezione, ecc.

Poiché la funzione delle ispezioni deve assolvere proprio a quei compiti non facilmente standardizzabili mediante test, questionari ed altre procedure del genere è bene che essi siano definiti in modo preciso, ma non mediante griglie meccaniche che riproducano surrettiziamente la procedura dei test. È necessario che venga definito in modo preciso e analitico l'insieme di tutti gli aspetti che l'ispezione deve esaminare, che venga fatto un elenco preciso ed analitico delle risposte che essa deve dare, ma non è affatto opportuno che il tipo di risposte sia standardizzato, altrimenti ci priviamo ancora una volta della possibilità di ottenere un giudizio autenticamente di merito. In altri termini, penso ad una relazione dettagliata e "libera" nello stile e nei contenuti, anche se, evidentemente, rispondente a un insieme di quesiti che si ritengono necessari per avere tutti gli elementi che sono stati stabiliti come significativi per una valutazione completa. È esattamente il tipo di procedura che viene seguito, per esempio, da una commissione di valutazione di un'università di tecnologia francese, di cui faccio parte e che ha iniziato i suoi lavori in questi giorni, e non vedo alcun motivo per cui una procedura del genere non possa essere applicata a un istituto scolastico.

Quanto ai destinatari del processo di valutazione - espresso quindi in relazioni d'ispezione dettagliata, prevalentemente verbali e di contenuto - penso che i referenti debbano essere: a) scuole stesse; b) Ministero; c) assessorati regionali. Trovo anche ragionevole l'idea della costituzione di una Direzione generale per la valutazione che però non abbia né composizione meramente burocratico-amministrativa né di "esperti" non meglio qualificati. In un sistema dell'istruzione come il nostro soltanto in questo contesto può prendere forma un sistema di premialità del merito. Tuttavia, anche questo è un aspetto estremamente delicato da approfondire.

Quel che mi sembra fondamentale assumere come punto di vista è che il processo di valutazione deve essere inteso come un processo culturale e non come un processo manageriale. Il secondo prevede la definizione di standard o obiettivi predeterminati su cui gli "esperti" misurano i comportamenti virtuosi o meno del

sistema. È persino da discutere se un processo del genere sia davvero efficiente persino nel caso di un'azienda: anzi, è ben noto che molti esperti del settore lo considerino del tutto inefficiente anche dal punto di vista aziendale, con buona pace degli ex-ingegneri McKinsey. Ma quel che è certo è che esso è totalmente inadeguato in un sistema i cui contenuti sono culturali, non misurabili, non passibili di una definizione oggettiva affidabile alla gestione di "esperti" esterni. D'altra parte, perché sia realmente efficiente, e non una foglia di fico, la valutazione deve investire con la massima energia e il massimo rigore l'intero sistema in un processo di feed-back di giudizi che, a loro volta, rimettano continuamente in discussione i criteri stessi di valutazione e permettano il loro continuo affinamento. Un processo aperto, dunque, che rivitalizzi l'intero sistema sottraendolo ad ogni forma di gestione burocratica e standardizzata che, in questo specifico contesto, rischia di andare incontro a un rapida sclerotizzazione.

Viene ora la questione del terzo pilastro del sistema di valutazione prospettato in varie proposte, ovvero la formazione in servizio. Su questo per ora non mi soffermo perché mi pare che vi sia già abbastanza carne al fuoco. Come unico elemento generale mi sentirei di sottolineare il principio che, anche in questo caso, è da evitare la concentrazione dell'intero processo nelle mani di un organismo autoreferenziale. Esso dovrebbe avere piuttosto una funzione di coordinamento, questa sì necessaria, dei soggetti inevitabilmente coinvolti nel processo di formazione in servizio e che sono, in primo luogo, il mondo stesso della scuola e l'università. Al riguardo, osservo che il fatto che l'università non si sia comportata bene manifestando scarso interesse per questa tematica, non è un buon motivo per escluderla da una funzione che deve essere inclusa nei suoi compiti istituzionali. L'università è una sede elettiva per contribuire al processo di aggiornamento agli sviluppi più recenti e significativi. Non si vede, del resto, perché il ruolo dell'università debba essere considerato ovvio e scontato sul terreno metodologico - con la partecipazione finora massiccia di pedagogisti e cultori di didattica disciplinare - e debba invece essere esclusa sugli altri piani (con una divisione di piani ancora una volta sbagliata).

A me pare che la funzione di un ente come quello prospettato debba essere soprattutto quella di selezionare e garantire la qualità dei contributi alla formazione in servizio, il loro livello culturale. Non si vede perché debba esistere - com'è giusto - una forma di accreditamento delle commissioni di ispezione, e invece i contributi alla formazione in servizio siano fuori controllo. Posso ben comprendere tutte le legittime diffidenze nei confronti di una gestione statalista. Ma sarebbe bene prendere atto del fatto che la sbilenca autonomia di cui gode la scuola - troppa su certi terreni, nulla su altri - ha già aperto la strada a esperienze di aggiornamento selvagge che esprimono i peggiori difetti di una situazione del tutto fuori controllo. Anche qui non voglio affliggere con esempi, che potrei produrre a iosa. Ma è sufficiente un esame anche superficiale per constatare che molte scuole consigliano assai insistentemente - per la pressione congiunta di alcune associazioni sindacali e professionali - la "consulenza" di certi "esperti" o gruppi di "esperti" che vanno per tutta Italia a tenere corsi di aggiornamento a pagamento. Si da il caso di corsi di aggiornamento di matematica - mi scuso per l'esempio, ma è il campo che conosco meglio - tenuti al costo di 15-20 euro per partecipante e un ciclo di 6-8 lezioni i cui programmi e contenuti costituiscono un autentico scandalo e un contributo massiccio all'opera di degrado culturale della nostra scuola. È evidente che le persone che

gestiscono queste iniziative sono anche le più attive nel promuoverle. Riuscire a contenerne gli effetti negativi è estremamente faticoso e difficile.

È quindi necessario che l'autorità prospettata abbia una funzione di coordinamento e accreditamento garantita da commissioni di alta e indiscussa qualità scientifica e culturale, soggette a rinnovamento periodico e che non siano mera espressione di gruppi d'interesse. A mio avviso, il principio che deve essere seguito - e che è costantemente capovolto nella nostra sbilenca autonomia - è: massima libertà metodologica e minima licenza sul piano dei contenuti. Non vedo perché un istituto scolastico non debba poter scegliere liberamente entro la più vasta offerta di contributi all'aggiornamento e formazione in servizio, ispirata alle metodologie e alle impostazioni culturali preferite - e poi sui risultati si valuterà l'opportunità della sua scelta. La condizione minima è che l'"offerta" sia sempre di qualità accertata e indiscussa sul terreno di quei requisiti di serietà, conoscenza che sono richiesti a qualsiasi studente per procedere nei suoi studi o a qualsiasi insegnante per abilitarsi ed esercitare la professione.