

L'attualità degli studi docimologici

The actuality of docimological studies*

Marika Calenda

Department of Human Sciences, University of Basilicata, marika.calenda@unibas.it

Francesco Milito

Department of Human, Philosophical and Educational Sciences, University of Salerno, fmilito@unisa.it

Scholastic evaluation is an issue that has long engaged research in the fields of humanities, social sciences and education. The evaluation involves psychological variables and is inextricably linked to learning and teaching: monitoring the achievement of the objectives on which the teaching action is based necessarily involves measuring the results achieved. In recent years the theme of evaluation has taken on a central role in the scholastic debate, enriching itself with reflections and connotations that today give us a very different idea of evaluation that is certainly more complex than in the past. To reach these conclusions, the step was not short nor without problems in the scientific community. Compared to traditional models of evaluation, a new positioning had to be sought, reasoning based on objective measurement criteria, promoting an idea of mature evaluation and above all an evaluation of one's own training potential. The present work aims to carry out an examination of the peculiar problems of docimology, tracing its developments over time and to propose a reflection on the relevance of docimological studies.

Keywords: Evaluation; Docimology; Measurement; Evaluation criteria; Objectivity.

La valutazione scolastica è una questione che impegna da tempo gli studiosi nel campo delle scienze umane, sociali e dell'educazione. La valutazione coinvolge variabili psicologiche ed è indissolubilmente legata all'insegnamento e alla didattica: monitorare il raggiungimento degli obiettivi su cui si basa l'azione didattica implica necessariamente la misurazione dei risultati raggiunti. Negli ultimi anni il tema della valutazione ha assunto un ruolo centrale nel dibattito scolastico, arricchendosi di riflessioni e di connotazioni che ci restituiscono oggi un'idea di valutazione molto diversa e sicuramente più complessa rispetto al passato. Per giungere a tali conclusioni il passo non è stato breve, né privo di problematizzazioni in seno alla comunità scientifica, si è dovuto ricercare un nuovo posizionamento rispetto a modelli tradizionali di valutazione, ragionare su criteri oggettivi di misurazione e promuovere un'idea di valutazione matura e soprattutto consapevole del proprio potenziale formativo. Il presente lavoro si pone come obiettivo quello di effettuare una disamina delle peculiari problematiche della docimologia ripercorrendo, a grandi linee, i suoi sviluppi nel corso del tempo e di proporre una riflessione sull'attualità degli studi docimologici.

Parole chiave: Valutazione; Docimologia; Misura; Criteri di valutazione; Oggettività.

Il contributo rappresenta il risultato di un lavoro congiunto degli autori, tuttavia Francesco Milito ha scritto i §§ 1-2 e Marika Calenda i §§ 3-4-5.

L'attualità degli studi docimologici

1. Introduzione

Lo studio scientifico dei problemi psicopedagogici della valutazione delle conoscenze acquisite a scuola è sistematicamente avviato negli anni Venti del Novecento da Piéron, ricordato soprattutto per aver pubblicato nel 1963 i risultati di uno studio, ormai pietra miliare della docimologia, sull'analisi degli esami finali della maturità francese, il noto Baccalauréat (informalmente detto anche Bac), che rappresenta il titolo di studio della scuola secondaria.

La docimologia nasce come critica verso un sistema di valutazione prettamente soggettivo per arricchirsi nel tempo di riflessioni costruttive attraverso la proposta di tecniche e metodi più rigorosi in modo da assicurare una valutazione più equa (De Landsheere, 1974). Fin dall'inizio questa scienza utilizza le tecniche statistiche come strumento principale per meglio isolare quelle variabili che, inevitabilmente, si presentano in qualunque processo valutativo, ma che non sono legate all'oggetto di misurazione. Una caratteristica dell'atteggiamento docimologico è l'impiego di tecniche statistiche come strumento per giungere alla massima obiettività possibile delle misurazioni, tradotte in punteggi affidabili e attendibili (Notti, 2002). Le indagini condotte in Francia dalla fine degli anni Sessanta, nella direzione di identificare i fattori di instabilità e la loro articolazione in un modello esplicativo globale, hanno contribuito ad estendere nel tempo il campo della docimologia e aperto nuove prospettive per migliorare la valutazione scolastica. Lo stesso Piéron (1965) osserva che negli esami il caso ha un ruolo predominante e che è necessario studiare misure capaci di attenuarlo.

I primi lavori docimologici evidenziano l'instabilità delle valutazioni dal punto di vista delle differenze interindividuali e intraindividuali, della validità e della precisione. In un primo momento, lo sviluppo delle ricerche porta soprattutto a raccomandare misure volte ad attenuare le divergenze di natura più o meno sistematica nella valutazione scolastica, misure che includono l'accordo tra gli esaminatori sui criteri da prendere in considerazione nella valutazione delle prove, l'uso di prove standardizzate, la correzione oggettiva. Agli inizi del Novecento,

Edgeworth, nel saggio *The statistics of examinations* (1923), sottolinea lo scarso valore insieme alla limitata validità e attendibilità dei giudizi attribuiti agli esami.

L'evoluzione degli studi docimologici va letta alla luce dei profondi cambiamenti che interessano i sistemi scolastici: "Le tematiche docimologiche sono strettamente collegate all'aumento della richiesta di formazione, al crescere della scuola di massa che [...] dagli anni '60 in poi incomincia a richiedere forme di valutazione più oggettive e diverse modalità di accertamento del rendimento scolastico" (Benvenuto, 2008, p. 2). L'interesse per i problemi e per le disfunzioni della valutazione cresce in relazione a una richiesta più ampia di formazione, fino ad arrivare agli anni Settanta, importantissimi per le riflessioni sulle tematiche docimologiche: insieme alla scolarizzazione di massa, le contestazioni giovanili, le tensioni sociali riportano al centro dell'attenzione il tema degli esami e della selezione sociale (Benvenuto, 2003).

2. La valutazione: un concetto in trasformazione

Le prospettive della docimologia come scienza che studia gli esami, i sistemi di votazione e il comportamento degli esaminatori, vengono delineate ulteriormente da De Landsheere (1971), con lo sforzo principalmente di ridurre l'approssimazione, l'imprecisione e il soggettivismo nella valutazione, ponendo in risalto la necessità di definire obiettivi misurabili, capaci di quantificare i comportamenti dei soggetti e di determinare in modo quanto più oggettivo possibile in che misura sono raggiunti nell'insegnamento gli obiettivi ad esso assegnati dalla comunità educativa. Nel senso più ampio, valutare significa giudicare, un evento, un individuo, un oggetto, in base a un criterio: nel caso specifico della valutazione degli alunni, e principalmente a partire dagli anni Cinquanta, il criterio è definito in termini di obiettivi educativi.

È proprio in questi anni che si ha, di fatto, come reazione all'educazione progressiva, la nascita della pedagogia per obiettivi. Figura di spicco in questo campo è Benjamin Samuel Bloom (1956), che nella sua opera *Taxonomy of educational objectives*, getta le basi della messa a punto di una tassonomia di comportamenti affettivi e cognitivi. Le basi della pedagogia per obiettivi sono rintracciabili tuttavia in Tyler (1932), che negli anni precedenti, riteneva che per una valutazione appropriata fosse indispensabile definire preliminarmente gli obiettivi in termini di comportamenti attesi.

La classificazione di obiettivi di Bloom prende spunto dall'attività pratica degli insegnanti, dal loro intento di raccogliere in modo empi-

rico i criteri di valutazione presi a riferimento nell'esaminare i loro allievi e i relativi percorsi di apprendimento. Gli obiettivi pedagogici permettono di definire un'attività precisa del discente e di precisare i criteri che serviranno alla valutazione. Bloom sostiene, infatti, che la scuola deve agire sulle caratteristiche dello studente, sulla qualità dell'istruzione e sui risultati dell'apprendimento, ma anche sulla tassonomia di obiettivi educativi e capacità generali e specifiche da acquisire e utilizzare durante l'apprendimento (Notti & Tammaro, 2015, p. 46-47).

La definizione operativa e la specificazione in termini di comportamenti osservabili, l'inserimento in un modello classificatorio gerarchico di categorie di comportamento, danno luogo alla cosiddetta tassonomia. La tassonomia di obiettivi educativi più frequentemente citata, e che Landsheere ritiene sia servita da ispirazione o come modello per la maggior parte di quelle che seguirono (de Landsheere & de Landsheere, 1976), è appunto quella elaborata da Bloom e dai suoi collaboratori (Bloom *et al.*, 1956; Krathwohl, Bloom & Masia, 1964), e riguarda i domini cognitivi e affettivi. La tassonomia rispetta il principio strutturale della complessità crescente, e le categorie, organizzate in modo gerarchico, includono comportamenti dal più semplice al più complesso (dominio cognitivo) e dal meno interiorizzato al più interiorizzato (dominio affettivo). La pedagogia per obiettivi, ispirata a Tyler e Bloom, è stata poi perfezionata dal contributo di Mager (1975), che nell'opera *Gli obiettivi didattici* individua tre caratteristiche che definiscono un obiettivo efficace: performance (osservabile e misurabile), condizione (circostanze in cui eseguire la performance), criterio (abilità per definire una performance idonea).

La concezione tradizionale della valutazione come misurazione il più oggettiva possibile viene messa in discussione sul finire degli anni Sessanta, quando alla funzione elitaria della scuola si preferisce una funzione egalitaria, finalizzata ad assicurare uguali opportunità educative a tutti, indipendentemente dalla provenienza sociale. La valutazione non si identifica più con la misurazione del sapere e con la conseguente selezione dei possessori del sapere, ma va assumendo un carattere non selettivo, formativo, orientativo poiché diviene strumento della formazione per orientare l'alunno, non per giudicarlo o condannarlo (Falcinelli, 1999).

Nell'ambito degli studi sull'insegnamento individualizzato del *mastery learning* (apprendimento per padronanza) di Bloom (1971) iniziarono ad emergere man mano le nuove funzioni della valutazione. Bloom, attraverso questo metodo, sollecita gli insegnanti a focalizzare le difficoltà degli studenti, per poter intervenire con strumenti specifici e rispondenti alle caratteristiche di ogni alunno (Tammaro, 2018).

Il concetto di valutazione formativa ha origine proprio in relazione all'esigenza di ottenere informazioni necessarie ad individualizzare l'istruzione nell'ottica di una strategia della padronanza (Airasian, 1972). Sulla base di questi studi, Scriven (1991) introduce il termine di *meta-valutazione*, per descrivere la valutazione di un piano per poter valutare i prodotti educativi e il termine *assessment*, per esprimere il senso di predisporre un disegno per farlo. Precisa, inoltre, sulla base dei riferimenti di Bloom sulla funzione formativa, una distinzione della stessa rispetto alla funzione sommativa: la prima finalizzata al miglioramento, la seconda alle esigenze di rendicontabilità.

Non era più sufficiente disporre di dati sul rendimento solo alla fine di un ciclo scolastico o di una unità di istruzione, era più utile e necessario disporre di informazioni nel corso dell'attività di apprendimento. Iniziò ad affermarsi con forza il principio che gran parte della valutazione scolastica doveva avere carattere formativo.

Cardinet (1977), ragionando sulla valutazione, andò ad associare ad essa una funzione regolativa, in quanto costituisce uno strumento informativo che permette di raccogliere informazioni e di conseguenza di dare feedback allo studente, ma anche al docente, diventando così un elemento che orienta e corregge l'intervento didattico. La valutazione diviene così occasione di riflessione sul proprio fare e stimolo per la ricerca, momento di lettura del proprio modo di porsi nell'azione e dunque stimolo per il cambiamento personale (Guasti, 1996). Con il passare del tempo, il dibattito sul concetto di valutazione ha posto sempre di più al centro della riflessione il contesto, i sistemi di riferimento, i progetti, i programmi, i materiali e, in generale, le situazioni specifiche in cui si situano i processi di insegnamento-apprendimento e le soluzioni date ai problemi della valutazione *del e nel* sistema di istruzione si sono configurate come vero e proprio costrutto politico-educativo (Domenici, 1993).

L'impatto della valutazione nei contesti educativi, mutevoli e diversificati, ha determinato una fitta trama di studi e ricerche con focalizzazioni diversificate. Negli anni Settanta, un importante settore di ricerca sulla valutazione è inaugurato dagli studi di Campbell, Cronbach e Stake che problematizzano la relazione tra fattori qualitativi oggetto della valutazione e strumenti quantitativi centrali nelle indagini condotte nel periodo precedente, in particolare le loro ricerche si contraddistinguono per: una maggiore apertura all'aspetto qualitativo, l'attenzione allo studio dei processi educativi, la ricerca di una possibile unificazione tra conoscenza quantitativa e conoscenza qualitativa, la consapevolezza dell'importanza delle interazioni tra processi sperimentali e caratteristiche dei soggetti (De Landsheere, 1988).

L'approccio *goal-free* ha, invece, come scopo principale quello di descrivere gli effetti dell'intervento, a prescindere dagli intenti che hanno guidato l'azione, il focus è sulla situazione osservata, sugli esiti effettivamente prodotti e non sul raggiungimento di obiettivi prefissati (Scriven, 1972). L'approccio *responsive* ha lo scopo di far emergere le diverse istanze dei soggetti coinvolti, come interessi, problemi, esigenze, necessità, domande e punti di vista attraverso uno scambio di informazioni che diventano il focus di questo tipo di *responsive evaluation* (Stake, 1975). Dagli anni Novanta fino ad oggi, parallelamente all'emergere di nuove teorie sui processi di apprendimento, il tema della valutazione è stato oggetto di profonde riflessioni che riguardano le forme, le funzioni, il significato della valutazione. Partendo dalla definizione di Beeby (1997), il quale afferma che la valutazione consiste nella raccolta sistematica e nell'interpretazione dei dati e conduce, come parte integrante del processo, ad un giudizio di valore teso verso l'azione, è possibile inquadrare un campo epistemico della valutazione educativa in cui poter distinguere alcune caratteristiche del valutare in *senso* educativo. Gli aspetti essenziali della valutazione riguardano l'uso di strumenti rigorosi per raccogliere dati informativi su un determinato evento educativo e l'attivazione di un percorso dinamico, che implica un processo di interpretazione dei dati raccolti con strumenti scientificamente adeguati, che ha delle ricadute sulla situazione di partenza, nel senso di essere orientato all'azione e al cambiamento per produrre un miglioramento (Bondioli & Ferrari, 2000).

Anche in Italia i primi studi docimologici si rivolgono principalmente all'individuazione di metodologie e interventi per la riduzione della soggettività nella valutazione scolastica grazie alla riflessione aperta da Visalberghi (1955) sulla distinzione tra i termini *misurazione* e *valutazione*, fasi da intendersi non come antitetiche o contrapposte, ma come momenti di uno stesso processo che si conclude con la formulazione di giudizi. In particolare, Visalberghi condannava i pregiudizi in base ai quali si presumeva di rifiutare ogni discorso in termini di misurazione. Il dibattito sulla docimologia continua a svilupparsi e ad arricchirsi grazie alla proficua produzione di Calonghi sulla validazione di prove nelle varie materie scolastiche (Calonghi & Boncori, 1967; Calonghi & Coggi, 1982) e di Gattullo che in particolare ha sostenuto con insistenza la preminenza dell'approccio quantitativo e della misurazione (1989). Per Calonghi (1990), l'adeguatezza della valutazione è strettamente legata agli obiettivi dell'azione didattica, da intendersi come criteri in base ai quali raccogliere i risultati alla fine di un programma di insegnamento.

Il contributo della docimologia alla qualità dell'istruzione è eviden-

ziato da Vertecchi (1991) attraverso il collegamento valutazione-decisione e la razionalità della decisione. Misurazione e valutazione sono i due momenti della verifica, uno finalizzato all'altro; la valutazione si basa su criteri autonomi rispetto ai metodi e agli strumenti che vengono impiegati nella misurazione, essa si basa infatti sugli obiettivi prefissati e sul loro raggiungimento (Coggi & Notti, 2002). “L'abito stesso del misurare, implicando l'attitudine a vedere un più ed un meno dove il giudizio affrettato scorge qualità assolute, è esso stesso un abito di riflessività, di moderazione e di prudenza” (Visalberghi, 1955, p.11). Nella prospettiva visalberghiana alla misurazione è affidato un ruolo importante ma intermedio, considerato che i risultati della misura non sono esclusivi e l'atto decisivo resta comunque la valutazione (Corsini, 2018).

La cultura docimologica ci ha abituati a una definizione di valutazione come “procedimento sistematico che mira a determinare in quale misura gli obiettivi sono stati raggiunti dagli allievi” (Giovannini, 1994, p.48), benché il suo campo di azione sia molto più esteso e vada ben oltre la valutazione degli apprendimenti.

3. La valutazione educativa

La valutazione educativa è un campo disciplinare complesso e ancora in divenire che ha alle spalle una robusta e complessa tradizione di ricerca. Nell'ambito della ricerca PRIN 2006¹, *Learning objects e struttura ontologica: messa a punto ed analisi di metodologie per la fruizione di Learning objects inseriti all'interno di ontologie di dominio*, la valutazione educativa è stata definita come una “disciplina finalizzata ad emettere giudizi sulle azioni formative e di insegnamento (o complesso di azioni organizzate come programmi o corsi), intenzionalmente progettate o svolte per guidare e sviluppare apprendimenti (individuali, collaborativi, organizzativi) nei destinatari, con effetti sul sistema formativo, economico e sociale, e fondata sull'uso di metodi e strumenti propri della ricerca empirica e sperimentale in educazione” (Galliani, 2009, p. 35).

La ricerca valutativa in educazione riguarda attività che mirano a produrre sia nuove conoscenze sugli oggetti di studio, sia decisioni per

- 1 Progetto di ricerca intrapreso per rispondere a interrogativi riguardo la gestione e lo scambio di informazioni e i processi di costruzione collaborativa della conoscenza, finalizzato alla costruzione di una ontologia sul concetto di valutazione educativa all'interno del dominio conoscitivo delle scienze umane.

il miglioramento e l'innovazione. Nella prospettiva di natura pedagogica adottata, la valutazione, insieme alla progettazione e alla comunicazione, è una componente ontologica e metodologica della scienza didattica la cui qualificazione pedagogica è dovuta al fatto che l'esperienza educativa appunto "si qualifica, da un lato, come *transazione relazionale* tra persone e, dall'altro, come *mediazione didattica* sui saperi disciplinari/professionali, per cui le azioni del progettare, del comunicare, del valutare, competono, seppure in misura diversa, agli stessi soggetti coinvolti, in primo luogo i docenti/formatori e gli allievi/formandi" (Galliani, 2015, p. 11).

Nell'ottica di produrre una visione sistemica e organica dell'evoluzione i paradigmi sono assunti in senso semiotico-fenomenologico e, conseguentemente si riferiscono a specifici *modelli empirici, immagini metaforiche e categorie descrittive*, ben descritte nella dimensione epistemologica dell'Ontologia per/della valutazione educativa (Galliani, 2014, p.49). Lo scopo è quello esplicitato nell'Ontologia per/della Valutazione Educativa di realizzare "un modello formativo orientato a produrre una cultura didattica praticata quotidianamente nei processi di insegnamento-apprendimento e pertanto ambisce a diventare strumento di ricerca per l'apprendimento a partire dalla condivisione di linguaggi comuni relativi al dominio scientifico della valutazione" (Notti, 2014, p. 7). Questa visione unitaria vuole integrare sistematicamente per l'appunto, "i tre paradigmi, (*positivista, pragmatista, costruttivista*) con altrettante coerenti e allineate categorie che riguardano l'insegnamento (*espositivo, dialogico, cooperativo*), la comunicazione didattica (*unidirezionale, interattiva, relazionale*), le tipologie psico-sociali della conoscenza (*dichiarativa, procedurale, performativa*), i processi cognitivi / emotivi di apprendimento (*adattivo, reattivo, regolativo*), le forme categoriali della valutazione (*sommativa, diagnostica, formativa*)" (Notti, 2014, pp.49-50).

La valutazione educativa esplora diversi ambiti: dal profitto, ai contesti, ai sistemi che implicano operatività che vanno oltre l'attribuzione di un voto e che hanno a che fare con differenti attori sociali coinvolti a diverso titolo nel processo valutativo e nell'atto stesso del valutare. L'atto del valutare è fondamentalmente connesso all'atto del conoscere (Dewey, 1976) e siccome la conoscenza per essere attendibile e completa necessita dell'uso congiunto di una pluralità di fonti, la valutazione deve spingersi oltre le distinzioni suggerite dai diversi approcci e sfruttare la possibile sinergia che si può ottenere dal loro utilizzo integrato per far luce su determinati aspetti dell'azione formativa. Il processo valutativo si configura, dunque, come un fatto pedagogico complesso che tocca le dinamiche di insegnamento/apprendimento e che richiede ai docenti

di compiere scelte consapevoli, inserite all'interno di specifici progetti, nell'ambito di programmazioni che ne definiscono obiettivi, regole e confini. L'atto valutativo, in quanto espressione di un giudizio di valore orientato all'azione e al cambiamento, produce degli effetti dal punto vista relazionale, sociale, personale (Notti, 2014). Valutare significa esprimere un giudizio di valore, un apprezzamento su un'azione educativa in relazione a specifici criteri e parametri di riferimento, esplicitati e condivisi da tutti i soggetti coinvolti.

Il giudizio di valore non è mai fine a sé stesso: si valuta per conoscere, si valuta per comprendere, si valuta per correggere, si valuta per cambiare, si valuta, dunque, per migliorare (Notti, 2017). La valutazione, per riuscire ad assumere il ruolo di strumento di ricerca volto ad indagare i processi didattici, per comprenderne a fondo le dinamiche e per progettare, gestire, orientare il loro cambiamento in termini di efficienza ed efficacia, dovrà essere intesa sempre più marcatamente come una guida preziosa nei processi decisionali. E se la capacità di decidere rispetto a tutti questi elementi è fortemente condizionata dal poter disporre di informazioni valide e affidabili, allora sarà sempre più importante scegliere con accuratezza le tecniche e gli strumenti da impiegare per la rilevazione dei dati valutativi. Per recuperare pienamente il ruolo che la valutazione occupa nel processo educativo, il punto cruciale è privilegiare le strategie dell'educazione che si ispirano alla metodologia scientifica sostenendo l'importanza di costruire modelli didattici caratterizzati da una flessibilità "possibile solo a condizione di poter disporre, nel corso dell'attività didattica, di copiose informazioni di ritorno, ottenibili solo attraverso un uso consapevole degli strumenti valutativi" (Vertecchi, 1976, p. 10). La qualità della didattica risulta così strettamente connessa all'adeguatezza delle verifiche, alla loro coerenza con gli obiettivi, all'attendibilità dei dati raccolti, in altre parole alla stabilità e precisione delle valutazioni ed è per questo che non si può prescindere dal dato quantitativo, ovvero dal valore della misura per decidere, informare, comprendere.

4. Il valore della misura per decidere, informare, comprendere

Con il passare del tempo, l'importanza e l'attualità degli studi docimologici sono temi che stanno interessando non solo il mondo della scuola, in relazione alla delicata e complessa questione della valutazione degli apprendimenti degli alunni, stanno diventando argomenti portanti all'interno del dibattito sul ruolo della formazione e dell'educazione in generale. La valutazione è un argomento "caldo" proprio per

l'importanza che riveste dentro e fuori dalla scuola, tuttavia, mentre una cultura "critica" della valutazione fatica a farsi strada (Giovannini, 2012) nel nostro Paese il confronto sulla valutazione si sta estendendo sempre di più a soggetti non direttamente coinvolti nelle pratiche educative. Si pensi a esponenti della politica e del giornalismo, al personale delle istituzioni scolastiche e universitarie, agli impiegati nelle ditte incaricate alla preparazione dei test.

In ambito educativo, la cultura della valutazione si è andata imponendo nell'ultimo ventennio come funzione complementare all'autonomia scolastica e universitaria con il compito di controllare la gestione delle risorse, rilevando l'efficacia degli interventi educativi. La misurazione dei risultati, basti pensare ai punteggi nelle prove INVALSI o nella valutazione della qualità della ricerca, è divenuta oramai centrale all'interno del nostro sistema di istruzione, formazione e ricerca, rapportandosi costantemente a quell'insieme di indicatori da considerare nella compilazione di importanti documenti come il Rapporto di Autovalutazione o il Riesame dei corsi di laurea (Corsini, 2018). La misura, nella sua veste oggettiva e autorevole, è altresì invocata dall'opinione pubblica per la facile lettura delle informazioni che fornisce (Corsini & Zanazzi, 2015).

Da tempo, gli argomenti valutativi rappresentano, per la stampa, per la società, una notizia rilevante riguardo all'efficacia scolastica italiana, al confronto con ricerche nazionali ed internazionali; purtroppo, però, la crescente esigenza valutativa o "valutomania" (Coniglione, 2015) che si sta diffondendo sembra essere, troppo spesso, funzionale a denunciare il degrado della scuola italiana o a produrre effetti controproducenti per la qualità della produzione scientifica universitaria.

Nella valutazione, come sostiene Vertecchi (2008), sarebbe opportuno comprendere la luce di quei dati in relazione al contesto di riferimento e non solo alla pura analisi statistica degli stessi. Le finalità perseguite con la valutazione nei e dei processi di istruzione si configurano come un indicatore della interpretazione sia delle funzioni e del ruolo che una data società assegna alla formazione. La valutazione può divenire un mezzo utile per acquisire il massimo di informazioni significative affidabili sui problemi formativi, al fine di ridurre l'ambiguità interpretativa e incrementarne perciò la probabilità di risolverli. È necessario sviluppare la consapevolezza che l'attività valutativa sia parte costitutiva di ogni azione umana intenzionale, e a maggior ragione di quella di apprendimento (Domenici, 2011).

Test e misurazione sono diventate ormai parole di moda, spesso presenti nei vari documenti ministeriali (De Luca & Lucisano, 2011), che alimentano quella diffusa tendenza a misurare e ad utilizzare i risultati

delle misure per decidere, valutare, fare le classifiche di Atenei, scuole e studenti.

È vero ciò che afferma Visalberghi (1955): “la misurazione nasce dalla valutazione e nella valutazione confluisce” (p. 18). Nel caso di misure di abilità, a monte c’è un giudizio, sulle abilità che i misurandi debbono possedere, a questo segue la misura alla quale segue un altro giudizio che conferma o smentisce il giudizio di partenza. La misura diretta è quella operazione che si effettua confrontando la grandezza da misurare con un’altra grandezza a essa omogenea, presa come campione: cioè, misurare una grandezza significa trovare un numero che dica quante volte tale grandezza è più grande o più piccola del campione di riferimento. Per quanto riguarda le abilità a cui si riferiscono le prove utilizzate in ambito educativo è però evidente che non è possibile effettuare misure dirette. Per quanto la misurazione sia indispensabile alla valutazione, dal punto di vista formativo si riscontrano limiti evidenti della stessa in quanto risulta incompleta rispetto alla capacità di controllo di tutti gli innumerevoli comportamenti cognitivi, artificiale, in quanto l’osservazione degli allievi non avviene nella spontaneità della loro attività, e eccessivamente esemplificativa, in quanto il processo si riduce al solo punteggio (Ancona, 2007). La qualità della misurazione può essere incrementata se viene curata la validità degli stimoli, la costanza e la validità dei risultati. La progettazione di una prova richiede una serie di decisioni teoriche ed operative, di cui il primo e più decisivo passo da compiere è quello di esplicitare la finalità complessiva della prova stessa (“che cosa vogliamo misurare con questa prova?”, “per quale motivo stiamo utilizzando questo tipo di prova?”) (Vertecchi, 2003).

Corsini (2018) dimostra efficacemente come in Visalberghi sia presente un’impostazione dialettica del rapporto tra misurazione e valutazione e di come emerga con forza l’idea di utilizzare i test come strumento funzionale all’interno di un processo didattico incentrato su apprendimenti di tipo collaborativo. Misurazione e valutazione costituiscono i passaggi fondamentali di una didattica che Visalberghi modella sulle fasi deweyane della ricerca in cui il processo di attribuzione di un valore a dati raccolti in relazione al livello di apprendimento è funzionale all’assunzione di decisioni didattiche volte a migliorare l’apprendimento attraverso l’insegnamento.

“Il test viene così posto da Visalberghi al servizio della didattica come mezzo valutativo e non come suo fine, e non solo in funzione di verifica dell’efficacia dell’insegnamento svolto, ma anche per contrastare la tendenza a rafforzare [...] una motivazione estrinseca all’apprendimento” (Corsini, 2018, p. 15).

Il fenomeno dei valutatori improvvisati non è, purtroppo, circoscritto ai problemi valutativi che riguardano il contesto scolastico, basti pensare al fatto che, nei vari contesti dell'apprendimento, si sta assistendo ormai da anni a una diffusione massiccia e a un proliferare di prove strutturate di conoscenza utilizzate a scopi selettivi. Insieme al proliferare di test si avverte, però, anche la necessità di sottolineare l'importanza di un approccio rigoroso nella costruzione di tali prove che sia basato su un sapere esperto (Bottani, 2011), su un insieme articolato di conoscenze e pratiche valutative quale presupposto per lavorare scientificamente alla rilevazione e misura di conoscenze e capacità.

Come ricordano De Luca e Lucisano (2011), gli strumenti utilizzati in ambito educativo hanno lo scopo di misurare tratti che in realtà sono latenti, e lo fanno associando concetti astratti a indicatori empirici, così le prove, ad esempio, misurano abilità non direttamente osservabili. Per questo è indispensabile procedere con metodo, seguendo alcuni passaggi fondamentali: approntare un modello teorico che contenga la definizione operativa del tratto latente; definire la relazione tra tratto latente e indicatori; prevedere un metodo per la misura degli indicatori. Inoltre, siccome la relazione tra concetti e indicatori è difficile da esplorare, gli autori affermano che l'analisi dei soli indicatori empirici può portare a inferenze non corrette e a conclusioni fuorvianti. Questo significa che partendo da domande formulate male, prive di senso, che prevedono più risposte corrette o nessuna corretta non è possibile avviare alcuna riflessione sulle risposte e quindi sui risultati ottenuti. "I controlli statistici non rendono intelligente una prova oggettiva che già non lo sia, mentre una prova impostata in modo larvatamente tendenzioso costituisce, anche se dotata di ottime qualità statistiche, un invito al conformismo non meno grave di quello che può venire da certe interrogazioni orali" (Visalberghi, 1955, p. 92).

Il proliferarsi di test, indicatori, comparazioni tra i differenti sistemi formativi, ha sempre vissuto una forte opposizione tra controllo e miglioramento. Gli approcci valutativi legati all'autoanalisi scolastica enfatizzavano lo scopo migliorativo della valutazione interna e, per contro, stigmatizzavano la funzione burocratica del controllo esterno (Tessaro, 2015, p. 276). Con l'espandersi dei programmi internazionali di valutazione, si afferma una nuova idea concettuale che comprende i due costrutti non più in opposizione, ma in interazione reciproca, retroattiva e proattiva: il controllo si trasforma in *accountability*, il miglioramento in *improvement*. Si valuta, dunque, primariamente per rendere conto in modo trasparente e comprensibile del sistema (*accountability*) e, conseguentemente per regolare il suo sviluppo in un'ottica di evoluzione e miglioramento (*improvement*). L'*accountability*

indica pienamente il significato che ha in sé la valutazione, il suo legame con il termine *improvement*; non si valuta solo per controllare e giudicare, bensì per rendere quel controllo frutto di miglioramento. Domenici (2007), argomentando sul concetto di *accountability* ha scritto che “il termine rinvia alla necessità di far conoscere ai principali attori dei processi formativi (allievi, docenti, dirigenti, famiglie), nonché all’opinione pubblica locale, in forme comunicative e rappresentative diverse a seconda del pubblico cui l’informazione è diretta (genitori, amministrativi, ...), ma sempre leggibili e utilizzabili per il giudizio e l’azione, e perciò in modi differenziati e non contraddittori, le scelte compiute e i vincoli entro i quali sono state fatte, i processi attivati e gli esiti conseguiti, ma anche i problemi aperti e le possibili soluzioni adottabili” (p. 48). Insieme, integrati, i processi di *accountability* e di *improvement*, perfezionano la valutazione di sistema assicurando l’accompagnamento delle politiche educative nella spirale virtuosa dello sviluppo (Tessaro, 2015, p. 278). La valutazione, in tale ottica, esce dall’ambito della pura arbitrarietà e del soggettivismo per divenire modalità rigorosa, processo di ricerca, in cui l’elemento soggettivo e l’elemento probabilistico, comunque presenti nella stessa azione didattica, sono controllati attraverso una scelta attenta degli strumenti e dei criteri adottati (Coggi & Notti, 2002).

Lucisano (2011) chiama Sindrome del Figlio dell’Uomo una malattia che ha pervaso la cultura della valutazione, nata dall’idea che qualcuno, depositario del bene per l’umanità, possa distinguere i buoni dai cattivi, esprimendo giudizi di merito per giustificare e avallare le decisioni prese, comprese quelle di perseguire coloro i quali si discostano da questo modello di “presunto” bene. Giovannini (2012) descrive come ancora problematica la condizione della valutazione nel nostro Paese, dove il prevalere di concezioni altalenanti e restrittive e di crescenti derive valutative impedisce, di fatto, la creazione e la diffusione di una cultura critica della valutazione. In tal senso, un aspetto che non può essere certamente tralasciato riguarda il fatto che “da un punto di vista metodologico e tecnico, va salvaguardata una serie di procedure, pena l’inaffidabilità dei dati (p. 208).

5. Conclusioni

Gli studi in campo docimologico nascono da problemi come il controllo della validità delle prove e dell’affidabilità nell’esaminare, temi considerati ormai come “temi forti”, ossia presenti sia nelle prime ricerche che nei recenti campi di interesse della docimologia (Benvenuto,

2003). Con il passare del tempo, il termine assume ulteriori significati includendo lo studio dei comportamenti dell'esaminatore e dell'esaminato e acquisendo sempre di più un significato pedagogico. La docimologia comincia a configurarsi come parte della didattica che studia, su basi scientifiche, i criteri della votazione scolastica al fine di elaborare tecniche di esame e di valutazione del profitto degli allievi (Notti & Tammaro, 2015).

Nonostante la sua importanza, il tema della misurazione non è affrontato con il dovuto rigore, trovando paradossalmente difficoltà, resistenze e spesso incompetenza nelle argomentazioni quando si tratta di esprimere pareri sul livello di accuratezza, precisione, coerenza delle prove e, invece, eccessiva tolleranza e semplicità quando si tratta di leggere, commentare, interpretare i risultati. Il discorso non ruota soltanto intorno alla questione dell'imparzialità nella valutazione di studenti, concorrenti, aspiranti specializzandi, dirigenti, ecc., sostanzialmente è prioritario riuscire a contrastare il fenomeno dilagante della non corretta misurazione. Partendo dal presupposto che le prove oggettive sono comunque elaborate soggettivamente da qualcuno, è condivisibile affermare che chi costruisce tali prove dovrebbe conoscere le regole di costruzione e gli elementi di cui esse si compongono, sapendo, inoltre, scegliere, come ricorda Domenici (1993), lo strumento di verifica omologo e congruente con la rispettiva funzione valutativa e con gli obiettivi che essa intende raggiungere. Una prova costruita male, sbagliata nella formulazione dei quesiti, nella scelta dei distrattori non sarà mai adeguata alla verifica degli obiettivi prefissati, l'oggettività da sola non garantisce automaticamente la precisione e la funzionalità del test. Visalberghi (1955) nella prefazione di quella lui stesso definisce una trattazione essenziale "intesa a considerare i problemi del testing in rapporto ai problemi educativi, culturali e sociali in generale", sosteneva che tale trattazione doveva "dar posto sufficiente all'informazione [...], sugli aspetti tecnici della «misurazione» del profitto scolastico" (pp. 7-9) con l'ambizione di promuovere un dibattito e ancor di più la sperimentazione per contribuire a una più aggiornata impostazione dei problemi educativi.

L'attualità degli studi docimologici è racchiusa in questa finalità ambiziosa in vista della quale chi, a vario titolo, si occupa di costruire test dovrebbe essere in grado di perseguire: costruire strumenti in grado di misurare scientificamente. Il problema della tecnica diventa così uno degli aspetti più urgenti e spinosi da affrontare per riuscire a garantire fondatezza e affidabilità a importanti decisioni politico-istituzionali che riguardano la promozione, il superamento, l'ammissione a fasi successive nei concorsi di centinaia di migliaia di persone che ogni anno

affrontano prove articolate in item a risposta multipla. È auspicabile ritornare a lavorare con quell'atteggiamento scientifico che affonda le sue radici nei problemi e nelle domande generate dalle condizioni di contesto (Dewey, 1939), senza dimenticare che anche l'esaminare è un'arte ed è "come tutte le arti, una sintesi di tecnica e di atti originali d'intuizione" (Visalberghi, 1955, p. 15).

Solo se si è in possesso di informazioni affidabili si possono assumere decisioni pertinenti ed efficaci e solo se si è in grado di costruire prove valide e attendibili, essendo inoltre in grado di riflettere sulle loro caratteristiche e analizzarle partendo dai dati prodotti, si riuscirà a promuovere il vero cambiamento nella teoria e nella pratica della valutazione, quello che riguarda la correttezza delle misure. Necessità di misure accurate, consapevolezza dell'uso che si intende fare delle misure, necessità di dotarsi di strumenti con una forte coerenza interna sono problemi attuali e tutt'altro che risolti in ambito educativo e il contributo che gli studi docimologici possono dare alla ricerca su questi temi tutt'altro che esaurito.

Riferimenti bibliografici

- Airasian, P. (1972). Il ruolo della valutazione nel mastery learning. In J. Block (Ed.), *Mastery learning. Procedimenti scientifici di educazione individualizzata*. Torino: Loescher.
- Ancona, D. (2007). Verifica e valutazione dell'apprendimento scolastico. Estratto da: <http://anconadonato.interfree.it/>.
- Beeby, C.E. (1997). The meaning of evaluation. *Current Issues in Education*, 4: 68-78.
- Benvenuto, G. (2008). La docimologia: dalla critica agli esami all'analisi dei principi della valutazione. In G. Benvenuto & A. Giacomantonio (Eds.), *Un po' di storia della valutazione scolastica: letture e riflessioni* (pp. 1-6). Estratto da: https://www.academia.edu/36650138/Un_po_di_storia_della_valutazione_scolastica.
- Benvenuto, G. (2003). *Mettere i voti a scuola. Introduzione alla docimologia*. Roma: Carocci.
- Besse, J.M. (1977). Vers une pédagogie par objectifs? *Bulletin de la Société Alfred Binet et Théodore Simon*, 556, 77 (III): 114-147.
- Bloom, B.S. (1971). Mastery learning and its Implications for Curriculum Development. In J.H. Block (Ed.), *Mastery Learning: Theory and Practice* (pp. 47-63). New York: Holt, Rinehart & Winston.
- Bloom, B.S. (Ed.), Engelhart, M.D., Furst, E.J., Hill, W.H., & Krathwohl, D.R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain*. New York: David McKay.
- Bondioli, A., & Ferrari, M. (Eds.) (2000). *Manuale di valutazione del contesto*

- educativo: teorie, modelli, studi per la rilevazione della qualità della scuola.* Milano: FrancoAngeli.
- Bottani, N. (2011). Concorso per dirigente, non si scherza sui questionari. Tuttoscuola, 13 settembre 2011. Estratto da: <http://www.tuttoscuola.com/cgi-local/disp.cgi?ID=26489>.
- Calonghi, L. (1990). *Finalità e obiettivi nella scuola.* Teramo: Giunti & Lisciani.
- Calonghi, L., & Boncori L. (1967). *Prove oggettive di lingua italiana.* Roma: Ateneo Salesiano.
- Calonghi, L., & Coggi, C. (1982). *Diagnosi delle difficoltà e proposte didattiche in geometria piana.* Salerno: Rispes.
- Cardinet, J. (1977). *Objectifs pédagogiques et fonctions de l'évaluation.* Neuchâtel: I.N.R.P.
- Coggi, C., & Notti, A.M. (2002). *Docimologia.* Lecce: Pensa MultiMedia.
- Coniglione, F. (2015). Dalla valutazione distribuita a quella amministrata. Il caso italiano. *I problemi della pedagogia*, 2: 345-376.
- Corsini, C., & Zanazzi S. (2015). Valutare scuola e università: approccio emergente, interventi e criticità. *I problemi della pedagogia*, 2: 305-334.
- Corsini, C. (2018). Sull'utilità e il danno di "Misurazione e valutazione" in educazione. In C. Corsini (Ed.), *Rileggere Visalberghi* (pp. 13-28). Roma: Edizioni Nuova Cultura. Estratto da: https://www.academia.edu/363-84424/Rileggere_Visalberghi.
- De Landsheere, G, & De Landsheere, V. (1976). *Definir les objectifs de l'Education.* Paris: PUF.
- De Landsheere, G. (1971). *Elementi di docimologia. Valutazione continua ed esami.* Firenze: La Nuova Italia.
- De Landsheere, G. (1974). *Evaluation continue et examens. Précis de Docimologie.* Paris: F. Nathan.
- De Landsheere, G. (1988). *Storia della pedagogia sperimentale.* Roma: Armando.
- De Luca, A.M., & Lucisano, P. (2011). Item analysis tra modello e realtà. *Giornale Italiano della Ricerca Educativa*, IV, 7: 85-96.
- Dewey, J. (1976). *Teoria della valutazione.* Firenze: La Nuova Italia.
- Dewey, J. (1939). Unity of science as social problem. In O. Neurath, R. Carnap, & C. Morris (Eds.), *Foundations of the unity of science. Toward an international encyclopedia of Unified science.* Chicago-London: University of Chicago Press, trad. it. In P. Lucisano (2000), L'unità della scienza come problema sociale, *Cadmo*, n. 22, 33-37).
- Domenici, G. (1993). *Manuale della valutazione scolastica.* Bari: Laterza.
- Domenici, G. (2007). Centralità, autonomia, valutazione. In G. Domenici, & F. Frabboni (eds.), *Indicazioni per il curriculum. Scuola dell'infanzia, primaria e secondaria di primo grado.* Erickson: Trento.
- Domenici, G. (2011). Valutazione e autovalutazione come risorse aggiuntive nei processi di istruzione. *Education Sciences & Society*, 2, 2: 69-82. Disponibile da: http://riviste.unimc.it/index.php/es_s/index.

- Edgeworth, F. Y. (1923). The statistics of examinations. *Journal of the Royal Statistical Society*, 51, 3: 599-635.
- Falcinelli, F. (1999). *La valutazione dell'azione formativa. Intelligenza pedagogica e intervento didattico*. Roma: Seam.
- Galliani, L., & Notti A.M. (2014). *Valutazione educativa*. Lecce-Brescia: Pensa MultiMedia.
- Galliani, L. (2009). *Web ontology della valutazione educativa*. Lecce: Pensa MultiMedia.
- Galliani, L. (2015) (Ed.). *L'agire valutativo. Manuale per docenti e formatori*. Brescia: La Scuola.
- Gattullo, M. (1989). Qualitativo e quantitativo in educazione e nella ricerca educativa. *Scuola e Città*, 40, 4: 158-162.
- Giovannini, M.L. (2012). Una cultura «critica» della valutazione: un lusso che non possiamo permetterci? *Journal of Educational, Cultural and Psychological Studies*, 6: 207-217.
- Giovannini, M.L. (1994). *Valutazione sotto esame*. Milano: Ethel Mondadori.
- Guasti, L. (1996). *Valutazione e innovazione*. Novara: De Agostini.
- Krathwohl, D.R., Bloom, B.S., & Masia, B.B. (1964). *Taxonomy of educational objectives: The classification of educational goals. Handbook II: The affective domain*. New York: David McKay.
- Lucisano, P. (2011). La sindrome del figlio dell'uomo. *Journal of Educational, Cultural and Psychological Studies*, 1: 155-167.
- Mager, R.F. (1975). *Gli obiettivi didattici*. Teramo: Giunti & Lisciani.
- Notti, A. M. (2017). La funzione regolativa della valutazione. In A. M. Notti, *La funzione educativa della valutazione. Teoria e pratiche della valutazione educativa* (pp. 9-24). Lecce-Brescia: Pensa MultiMedia.
- Notti, A.M., & Tammaro, R. (2015). Autori e storia della docimologia. In L. Galliani, *L'agire valutativo* (pp. 41-56). Brescia: La Scuola.
- Notti, A.M. (2002). *Strumenti per la ricerca educativa*. Salerno: Edisud.
- Notti, A.M. (Ed.) (2014). *A scuola di valutazione*. Lecce-Brescia: Pensa MultiMedia.
- Piéron, H. (1965). *Esami e docimologia*. Roma: Armando.
- Scriven, M. (1972), Pros and Cons about Goal Free Evaluation. *Evaluation comments*, 3: 1-4.
- Scriven, M. (1967). The methodology of Evaluation. In R. Tyler, R. Gagné & M. Scriven (Eds.), *Perspectives of Curriculum Evaluation*, AERA Monograph Series on Curriculum Evaluation, 1, Chicago: Rand McNally.
- Scriven, M. (1991). *Evaluation Thesaurus*. Thousand Oaks: Sage.
- Stake, R. (1975). *Evaluating the Arts in Education: A Responsive Approach*. Columbus, Ohio: Merrill.
- Tammaro, R. (2018). *Valut/Azione*. Lecce: Pensa Multimedia.
- Tessaro, F. (2015). La valutazione di sistema e le politiche educative tra accountability e improvement. In L. Galliani (Ed.), *L'agire valutativo* (pp. 265-279). Brescia: La Scuola.
- Vertecchi, B, & Agrusti, G. (2008). *Laboratorio di valutazione*. Roma-Bari: Laterza.

- Vertecchi, B. (1976). *Valutazione formativa*. Torino: Loescher.
- Vertecchi, B. (1991). *Origini e sviluppo della docimologia*. Teramo: Giunti & Lisciani.
- Vertecchi, B. (1993). *Decisione didattica e valutazione*. Firenze: La Nuova Italia.
- Vertecchi, B. (2003). *Manuale della valutazione*. Milano: Franco Angeli.
- Visalberghi, A. (1955). *Misurazione e valutazione nel processo educativo*. Milano: Edizioni di Comunità.