

LA “VALUTAZIONE AUTENTICA”

di Mario Comoglio

0. INTRODUZIONE¹

Non è raro imbattersi in articoli e in interventi ministeriali che affrontano la questione di come valutare l'apprendimento degli studenti. La valutazione, è noto, è un tema “spinoso” che non trova in genere d'accordo gli insegnanti, le famiglie degli studenti e gli studenti stessi. I motivi di insoddisfazione sono numerosi, ma comune e diffusa in tutti è la difficoltà a trovare giustificazioni e elementi di trasparenza nell'espressione dei propri giudizi. Molti insegnanti, poi si chiedono anche quale valore reale abbia la valutazione scolastica. C'è una ricerca intensa per individuare non solo i criteri ma anche gli strumenti che permettano di esprimere giudizi maggiormente fondati. Uno di questi strumenti è il portfolio. Mentre alcuni tendono a sottolineare del portfolio gli elementi di piena novità rispetto agli strumenti tradizionali, altri ne evidenziano le caratteristiche di relativa innovazione in quanto già presente e utilizzato in vari contesti professionali e comunque già oggetto di ricerca nell'ambito dello studio sulla valutazione.

I problemi evidenziati e incontrati nella valutazione dell'apprendimento sono stati a lungo e ampiamente discussi negli ultimi anni da una riflessione che per il numero dei contributi e per gli sviluppi pratici che ha avuto ha assunto le dimensioni di un vero e proprio “movimento di pensiero”: la *valutazione autentica* o *alternativa* che può essere considerata come contrapposta alla valutazione tradizionale² o, in modo più preciso, alla valutazione misurata attraverso forme standardizzate. Vi accenniamo qui non per sintetizzarne tutti i suoi sviluppi e radici teoriche, ma per cogliere alcuni aspetti che possono contribuire a chiarire alcuni aspetti problematici e precisare alcuni termini per migliorare a molti insegnanti questo compito così importante dell'attività scolastica.

1. LA VALUTAZIONE AUTENTICA

In genere la valutazione del profitto scolastico è stabilita dal confronto dei risultati ottenuti dagli studenti con i risultati attesi (obiettivi). E in base alla loro vicinanza o distanza che si traggono inferenze sul grado di apprendimento. Quando è emersa l'esigenza di una misura che fosse il più possibilmente precisa, si è fatto soprattutto ricorso a prove standardizzate.

¹ Tratto da: M. Comoglio (2002). La valutazione autentica. *Orientamenti Pedagogici*, 49(1), 93-111

² Il movimento della *valutazione autentica* o *alternativa* è sorto negli Stati Uniti agli inizi degli anni '90 come contrapposizione critica alla forma di valutazione diffusa di orientamento comportamentista fondata soprattutto su tesi standardizzate per lo più a scelta multipla. Questa contrapposizione non ha gli stessi riferimenti per quanto riguarda il contesto italiano. La valutazione tradizionale italiana, pur essendo molto variegata e diversificata, va qui intesa non nel senso degli strumenti, ma nel senso delle intenzioni e degli scopi: verificare l'apprendimento da parte dello studente di una conoscenza trasmessa dall'insegnante.

Questo sistema, il cui scopo iniziale è stato quello di constatare soltanto il successo oppure l'insuccesso dell'apprendimento per suggerire interventi di rinforzo o di aiuto, per molti è diventato anche un sistema di giudizio selettivo. Il limite maggiore della valutazione tradizionale sembra collocarsi "in ciò che" essa intende e riesce a valutare. Valutando ciò che un ragazzo "sa", si controlla e si verifica la "riproduzione" ma non la "costruzione" e lo "sviluppo" della conoscenza e neppure la "capacità di applicazione reale" della conoscenza posseduta.

Una valutazione che voglia essere maggiormente autentica dovrebbe consentire di esprimere un giudizio più esteso dell'apprendimento e cioè della capacità di pensiero critico, di soluzione dei problemi, di metacognizione, di efficienza nelle prove, di lavoro in gruppo, di ragionamento e di apprendimento permanente» (Arter & Bond, 1996, p. 1).³

La prospettiva di una "valutazione alternativa" in sostituzione di quella tradizionale è stata proposta da *Grant Wiggins* (1993)⁴ e sta a indicare una valutazione che intende verificare non solo ciò che uno studente sa, ma ciò che "sa fare con ciò che sa" fondata su una *prestazione reale e adeguata* dell'apprendimento. Perché una verifica attraverso una prestazione e non attraverso un test?

A volte capita che alcuni studenti riescano bene nei test a scelta multipla ma, quando viene richiesto loro di dimostrare ciò che sanno in una prestazione concreta, sembrano confusi e dimostrano una competenza da principianti e non da esperti come hanno dimostrato nei test. Si pensi a come sia possibile conoscere molto bene come funziona una macchina, ma poi non si sappia come guidarla nel traffico della città. Questo indica come la valutazione di concetti e di fatti isolati non dimostri le reali capacità di ragionamento, di creatività e di soluzione di problemi in situazioni concrete di vita. Perciò, sapere che uno studente è in grado di operare in contesti reali con prestazioni in grado di conseguire certi obiettivi dice molto di più sul suo apprendimento che non ciò che egli dimostra in prove di riconoscimento della verità o di affermazioni o il punto in cui si colloca rispetto ai compagni.

Vi sono però anche altri motivi che si rifanno a riflessioni recenti sui processi di insegnamento e di apprendimento che giustificano altre forme più attendibili di valutazione. Teorie dell'apprendimento autentico, dell'insegnamento inteso come inculturazione, della cognizione situata, del costruttivismo o del costruttivismo sociale dimostrano che gli studenti comprendono e assimilano in misura maggiore quando hanno a che fare con situazioni reali rispetto a quanto devono apprendere in situazioni decontestualizzate.⁵ Perciò, visto che si apprende di più in questo modo, è giusto anche valutare l'apprendimento non in modi astratti e artificiali, ma con

³ J. Arter, & L. Bond (1996). *Why is assessment changing*. In R. E. Blum, & J. A. Arter (Eds.), *A handbook for student performance assessment in an era of restructuring*, (1-3 - 1-4). Alexandria, VA: Association for Supervision and Curriculum Development.

⁴ G. Wiggins (1993). *Assessing student performance: Exploring the purpose and limits of testing*, San Francisco, CA: Jossey-Bass.

⁵ L. Resnick (1987). *Learning in school and culture*. *Educational Researcher*, 16, 13-20; J. S. Brown, A. Collins, & P. Duguid (1996). *Situated cognition and the culture of learning*. In H. McLellan (Ed.), *Situated learning perspectives*. (pp. 19-44). Englewood Cliffs, NJ: Educational Technology Publications; A. Collins, & J. S. Brown, & S. E. Newnan (1989). *Cognitive apprenticeship; Teaching the crafts of reading, writing, and mathematics*. In L. B. Resnick (Ed.), *Knowing, learning, and instruction; Essays in honor of Robert Glaser* (pp. 453-494). Hillsdale, NJ: Erlbaum.

prestazioni creative, contestualizzate. Parliamo quindi di valutazione autentica.

“quando ancoriamo il controllo al tipo di lavoro che persone concrete fanno piuttosto che solo sollecitare risposte facili da calcolare con risposte semplici. La valutazione autentica è un vero accettazione della *prestazione* perché da essa apprendiamo se gli studenti possono in modo intelligente usare ciò che hanno appreso in situazioni che in modo considerevole li avvicinano a situazioni di adulti e se possono rinnovare nuove situazioni” (Wiggins, 1998, p. 21).⁶

Oppure possiamo ancora definire la valutazione autentica come:

“la valutazione che ricorre continuamente nel contesto di un ambiente di apprendimento significativo e riflette le esperienze di apprendimento reale... L’enfasi è sulla riflessione, sulla comprensione e sulla crescita piuttosto che sulle risposte fondate solo sul ricordo di fatti isolati.

L’intento della “valutazione autentica” è quello di coinvolgere gli studenti in compiti che richiedono di applicare le conoscenze nelle esperienze del mondo reale. La “valutazione autentica” scoraggia le prove “cata-e-penna” sconnesse dalle attività di insegnamento e di apprendimento che al momento avvengono. Nella “valutazione autentica”, c’è un intento personale, una ragione a impegnarsi. e un ascolto vero al di là delle capacità/doti dell’insegnante” (Winograd & Perkins, 1996, 1-8: 2).⁷

La valutazione autentica o alternativa si fonda quindi anche sulla convinzione che l’apprendimento scolastico non si dimostra con l’accumulo di nozioni, ma con la capacità di generalizzare, di trasferire e di utilizzare la conoscenza acquisita a contesti reali. Per questo nella valutazione autentica le prove sono preparate in modo da richiedere agli studenti di utilizzare processi di pensiero più complesso, più impegnativo e più elevato.

Verificando con maggiore autenticità l’apprendimento si possono far raggiungere livelli più elevati di prestazione e preparare meglio gli studenti a un inserimento di successo nella vita reale. Non avendo propriamente lo scopo della classificazione o della selezione, la valutazione autentica cerca di promuovere e di rafforzare tutti, dando opportunità a tutti di compiere prestazioni di qualità. Essa offre la possibilità sia agli insegnanti che agli studenti di vedere a che punto stanno, di autovalutarsi e, in conformità a ciò, migliorare il processo di insegnamento o di apprendimento: gli uni (gli insegnanti) per sviluppare la propria professionalità e gli altri (gli studenti) per diventare autoriflessivi e assumersi il controllo del proprio apprendimento. In questo modo, gli uni (gli insegnanti) scoprono il ruolo come “mediatori” dell’apprendimento, gli altri (gli studenti) si scoprono esaminatori di se stessi.

Wiggins (1998, pp. 22.24) indica come le *caratteristiche* della valutazione autentica:

1. *E’ realistica*

Il compito o i compiti replicano i modi nei quali la conoscenza della persona e le abilità sono “controllate” in situazione di mondo reale

2. *Richiede giudizio e innovazione*

Lo studente deve usare la conoscenza e le abilità saggiamente e in modo efficace per risolvere problemi

⁶ G.Wiggins (1998). *Educative assessment. Designing assessment to inform and improve student performance*, San Francisco, CA: Jossey-Bass.

⁷ P.Winograd, & F D Perkins 1996). *Authentic assessment in the classroom: Principles and prtices*. In R. E. Blum, & J. A. Arter (Eds.), *.A handbook for student performance assessment in an era of restructuring* (I-8: I-11). Alexandria, V.A.: Association for Supervision and Curriculum Development

non strutturati, ad esempio progettare un piano, la cui soluzione richiede di più che seguire una routine, una procedura stabilita o l'inserimento di una conoscenza.

3. *Richiede agli studenti di “costruire” la disciplina*

invece di ridire, di riaffermare o di replicare attraverso una dimostrazione ciò che gli è stato insegnato o ciò che già conosce, lo studente deve portare a termine una esplorazione e lavora “dentro” la disciplina di scienze, di storia e dentro ogni altra disciplina.

4. *Replica o simula i contesti nei quali gli adulti sono “controllati” sul luogo di lavoro, nella vita civile e nella vita personale*

I contesti richiedono situazioni specifiche che hanno costrizioni, finalità e spettatori particolari. I tipici test scolastici sono senza contesto. Gli studenti hanno bisogno di sperimentare che cosa vuol dire fare un compito in un posto di lavoro e in altri contesti di vita reale che tendono ad essere disordinati e poco chiari: in altre parole i compiti veri richiedono un buon giudizio. I compiti autentici capovolgono quella segretezza, quel silenzio che alla fine sono dannosi e quell'assenza di riserie e di feedback che segnalano il testing tradizionale.

5. *Accerta l'abilità dello studente a usare efficacemente e realmente un repertorio di conoscenze e di abilità per negoziare un compito complesso*

La maggior parte degli item del test convenzionale sono elementi isolati di una prestazione – simile agli esercizi pre atletici svolti dagli atleti prima di entrare in gara piuttosto che l'uso integrato di abilità che una gara richiede. Anche qui è richiesto un buon giudizio. Sebbene ci sia uno spazio per gli esercizi pre-gara, la prestazione è sempre più della somma di questi esercizi.

6. *Permette appropriate opportunità di ripetere, di praticare, di consultare le risorse e di avere feedback su e di perfezionare la prestazione e i prodotti.*

Per essere educativa una valutazione deve tendere a migliorare la prestazione degli studenti. Il classico test convenzionale manca di questa prerogativa in quanto mantiene le domande segrete e i materiali di risorsa lontani dagli studenti fino a che dura la prova. Se dobbiamo focalizzarci sull'apprendimento degli studenti attraverso cicli di *prestazione - feedback - revisione - prestazione*, sulla produzione di prodotti e di standard conosciuti di qualità elevata, e se dobbiamo ancora aiutare gli studenti ad apprendere ad usare le informazioni, le risorse e le annotazioni per eseguire una prestazione reale in un contesto, i testi convenzionali non sono utili allo scopo”.

2. **GLI STRUMENTI DI VALUTAZIONE**

La valutazione alternativa intende correggere i limiti di quella tradizionale non solo offrendo una prospettiva diversa da cui vedere l'apprendimento ma anche suggerendo strumenti diversi per accertarlo. Per realizzare questi obiettivi, la valutazione autentica utilizza l'integrazione di varie strategie come quelle scritte (portfolio, rubriche, saggi), visive (osservazione diretta durante lo svolgersi delle prestazioni oppure durante lo sviluppo

della conoscenza), e orali (colloquio insegnante studente) per raccogliere le necessarie informazioni.

Essa ha luogo durante il processo di istruzione piuttosto che dopo, e fornisce risultati più diretti che aiutano gli insegnanti a perfezionare il loro percorso educativo. Essa serve a controllare costantemente il progresso dello studente al fine di adattare meglio l'istruzione. A differenza della valutazione tradizionale, è "una valutazione fondata sull'osservazione e sul giudizio"; vale a dire: si osservano i risultati di un'attività autentica (i compiti e le prestazioni sono scelti per la loro analogia con i compiti reali) e si dà valore alla sua qualità (attraverso indicatori che descrivono la bontà o qualità della prestazione). Il riferimento ad una prestazione reale, la scelta di compiti autentici contraddistingue la valutazione alternativa da quella tradizionale secondo diverse caratteristiche in modo da renderla più "predittiva" della reali capacità possedute da chi si sottometta alla prova (si veda la Tabella 1).

Tabella 1

Tipici test⁸	Compiti autentici	Indicatori di autenticità
Richiedono solo una risposta corretta	<ul style="list-style-type: none"> ▪ Richiedono un prodotto di qualità e/o una prestazione e una giustificazione 	<ul style="list-style-type: none"> ▪ Accertano se lo studente può spiegare, applicare, auto – adattarsi o giustificare le risposte, non solo la correttezza delle risposte utilizzano fatti e algoritmi
Non deve essere conosciuto in anticipo per assicurare la validità	<ul style="list-style-type: none"> ▪ Deve essere conosciuto il più possibile in anticipo; richiede esigenze di eccellenza e compiti essenziali; non sono esperienze "di fortuna" 	<ul style="list-style-type: none"> ▪ I compiti, i criteri e gli standard attraverso i quali il lavoro sarà giudicato sono prevedibili o conosciuti – come un pezzo di recitazione di una rappresentazione, il motore che è aggiustato, la proposta ad un cliente, ecc.
Sono disconnesse da un contesto e da costrizioni realistici	<ul style="list-style-type: none"> ▪ richiedono l'utilizzo della conoscenza del mondo reale: lo studente deve "fare" storia, scienze, ecc. in situazioni realistiche o di uso reale 	<ul style="list-style-type: none"> ▪ il compito è una sfida e un insieme di costrizioni che sono autentiche- probabilità che sono incontrate da un professionista, da un cittadino o da un consumatore (è richiesto un "saper come", non solo una improvvisazione)
Contiene item isolati che richiedono un uso o un	<ul style="list-style-type: none"> ▪ sono sfide integrate nelle quali la conoscenza e il giudizio 	<ul style="list-style-type: none"> ▪ il compito ha molti aspetti e non routine, anche se c'è una risposta

⁸ qui si fa riferimento alle varie forme di prove oggettive a scelta multipla e no, ma con le dovute differenze si può pensare a tutti i tipi di prove e interrogazioni tradizionali che in modo più o meno formale avvengono in classe o a prove di esami

ricoscimento di risposte o di abilità conosciute	devono essere usati in modo innovativo per confezionare un prodotto di qualità o una prestazione	“corretta2. Ciò richiede la chiarificazione di un problema, tentativi e errori, adattamenti e adattarsi al caso o ai fatti che si hanno in mano
Sono semplificati in modo da poter essere esaminati in modo facile e sicuro	<ul style="list-style-type: none"> ▪ implicano compiti complessi e non arbitrari, criteri e standard 	<ul style="list-style-type: none"> ▪ il compito richiede aspetti importanti di prestazione e/o sfide sostanziali del campo di studio, non facilmente analizzato; non sacrifica la validità per l’affidabilità
Sono eseguite in un momento	<ul style="list-style-type: none"> ▪ sono iterativi: contengono compiti essenziali ricorrenti, generi e standard 	<ul style="list-style-type: none"> ▪ il lavoro è programmato per rivelare se lo studente ha conseguito una padronanza o comprensione versus solo familiarità nel tempo
Dipendono da correlazioni tecniche elevate	<ul style="list-style-type: none"> ▪ offre una evidenza diretta, coinvolgendo compiti che sono stati valicati rispetto a ruoli essenziali adulti e sfide fondate sulla disciplina 	<ul style="list-style-type: none"> ▪ il compito è valido e giusto nel suo presentarsi. Per questo richiama l’interesse e la persistenza dello studente e sembra adatto a sfidare gli studenti e l’insegnante
Offre un’opportunità di punteggio	<ul style="list-style-type: none"> ▪ offre un feedback utilizzabile, diagnostico (alle volte alternativo): lo studente è capace di confermare i risultati e autoadattarsi nella misura in cui è necessario 	<ul style="list-style-type: none"> ▪ la prova è programmata non solo per verificare la prestazione, ma anche per migliorare la prestazione futura. Lo studente è considerato come il “cliente” primario dell’informazione

Da G. Wiggins (1998). *Educative assessment. Designig assessments to inform and improve student performance* San Francisco, CA: Jossey Bass, p. 23

La differenza tra la valutazione autentica e tradizionale è così profonda da assumere una posizione fortemente critica verso la valutazione tradizionale considerata come “inautentica”. Per essere più oggettiva la valutazione autentica si presenta non solo critica nella scelta dei compiti e delle forme nelle quali si predispose la prova, ma ha sviluppato anche forme di misurazione (le rubriche) e di controllo

(pubblicità, autovalutazione, progresso dimostrato) della qualità della prestazione viene valutata o misurata per mezzo di criteri di prestazione e di rubriche. Esaminiamo per ora quello che è un elemento cruciale della valutazione autentica: la predisposizione delle rubriche di valutazione della prestazione.

3. LA DETERMINAZIONE DELLE RUBRICHE: DIMENSIONI, CRITERI, SCALE

La valutazione esprime un giudizio fondato su criteri stabiliti o su attese da parte di colui che valuta riguardo la qualità, il valore o il pregio di una risposta, di un prodotto o di una prestazione e dovrebbe essere guidato da tre principi:

1. Lo scopo della valutazione è quello di trarre informazioni per migliorare l'insegnamento e l'apprendimento
2. Il giudizio deve esprimere una sintesi fatta su un ampio arco di informazioni disponibili
3. Il giudizio deve esprimere queste tre caratteristiche: essere valido, attendibile e imparziale

Poiché le valutazioni fondate su prestazioni non possono avere un giudizio univoco (sì/no, giusto/sbagliato), per esprimere un giudizio che abbia le tre caratteristiche sopra enunciate si fa riferimento ad una *rubrica*. La rubrica è lo "strumento"-per identificare e chiarificare le aspettative specifiche relative a una prestazione e indica come si sono raggiunti gli obiettivi prestabiliti.

«Si tratta di uno strumento generale di valutazione impiegato per valutare la qualità dei prodotti e delle prestazioni in un determinato ambito. La rubrica con punteggi prefissati (ad esempio, una scala composta da quattro punteggi) e di una lista di criteri che descrivono le caratteristiche di ogni punteggio della scala. Le rubriche sono frequentemente accompagnate da esempi di prodotti o di prestazioni che hanno lo scopo di illustrare ciascuno dei punteggi. Tali esempi sono detti *ancore*» (Mc Tighe & Ferrara, 1996, p.8)⁹

⁹ J. Mc Tighe & S. Ferrara (1996). *Performance-based assessment in the classroom: A Planning framework*. In R. E. Blum, & J. A. Arter (Eds.) *a handbook for student performance assessment in an of restructuring* (1-5: 1-9). Alexandria, VA: Association for Supervision and Curriculum Development.

Una *rubrica*¹⁰ e un elenco di linee guida che specificano contraddistinguono la qualità (buona, debole o scarsa) di una prestazione.

Alla determinazione e alla precisazione di una rubrica concorrono diversi elementi: dimensioni o tratti, criteri (attese o aspettative), indicatori, descrittori. Punti di ancoraggio, esemplari: standard di prestazione, punti di riferimento (*benchmark*); scale (qualitative, numeriche-qualitative). Le *dimensioni-tratti* indicano le “caratteristiche peculiari” che contraddistinguono una particolare prestazione. Ad esempio in una composizione scritta potremmo indicare queste dimensioni o tratti: idee e contenuto, organizzazione, espressione, scelta delle parole, scorrevolezza della convenzioni. Per laboratorio di scienze potrebbero essere: la progettazione dell’esperimento (definizione dell’ipotesi, procedura per la ricerca, registrazione e organizzazione e la presentazione del resoconto dell’esperimento (qualità delle osservazioni, grafico, calcoli, conclusioni). Herman, Aschbacher e Winters (1992, p. 58)¹¹ suggeriscono alcune domande che un insegnante potrebbe porsi per identificare le dimensioni:

- 1) Ho già fatto questo compito altre volte?
- 2) Che cosa mi aspetto di vedere in questo compito che mi permetterà di contraddistinguere questa prestazione da un'altra?
- 3) Quali sono le caratteristiche che qualificano questa prestazione?
- 4) Quali caratteristiche peculiari avrebbero i compiti migliori?

I *criteri* di una prestazione possono essere considerati gli “strumenti di misurazione” con i quali si monitora l’abilità degli studenti. Mentre le dimensioni indicano la prospettiva a partire dalla quale può essere determinata la qualità della prestazione, i criteri indicano ciò che viene assunto come indicatore della qualità del prodotto. «I criteri di prestazione definiscono ciò che si apprezza nel lavoro dello studente, definiscono gli scopi e gli obiettivi che ci si propone di raggiungere, e sono utili per determinare quando sono stati raggiunti».¹² Inoltre, essi aiutano gli insegnanti e gli studenti a definire chiaramente gli obiettivi da raggiungere, in modo da coinvolgere gli studenti stessi nella valutazione. Herman, Aschbacher e Winters (1992, p. 58) definiscono i criteri come:

¹⁰ La parola “rubrica” deriva alla parola latina “ruber” (= rosso). Anticamente una rubrica indicava un insieme di istruzioni scritte in rosso per una legge o per un servizio liturgico. Una rubrica istruisce su come delle persone “legalmente” possono giudicare una prestazione.

¹¹ J. L. Herman, P.R. Aschbacher, &L. Winters (1992). *A practical guide to alternative assessment*, Alexandria, V.A: Association for Supervision and Curriculum Development

¹² J. Arter (1996). *Using performance assessment in an era of restructuring*, (IV-10:49. Alexandria, V.A: Association for Supervision and Curriculum Development

“gli obiettivi di priorità in termini di contenuti che devono essere affrontati, di conoscenze o di abilità che devono essere dimostrate e di contesto nel quale queste devono esprimersi”.

I criteri scelti per descrivere una prestazione dovrebbero avere alcune caratteristiche: raccordarsi con gli obiettivi, essere sensibili, essere significativi, chiari e convincenti, essere imparziali (non dipendenti dal background socioeconomico, dal genere o dalla cultura dello studente), essere generalizzabili (non servire in modo molto circoscritto a un compito o a una disciplina, ma a una abilità), essere fattibili (avere un numero limitato di dimensioni o di tratti).

I criteri scelti per descrivere una prestazione dovrebbero avere alcune caratteristiche: raccordarsi con gli obiettivi, essere sensibili, essere significativi, chiari e convincenti, essere imparziali (non dipendenti dal background socioeconomico, dal genere o dalla cultura dello studente), essere generalizzabili (non servire in modo molto circoscritto a un compito o a una disciplina, ma a un'abilità), essere fattibili (avere un numero limitato di dimensioni o di tratti).

I *descrittori* indicano che cosa si deve osservare di una prestazione mentre gli altri *indicatori* offrono misure specifiche (esempi, segnali, manifestazioni concrete e *definite* di una prestazione). Ambedue indicano ciò da cui si inferisce la presenza o meno di un criterio e di una dimensione.

Ad esempio:

- (1) Dà una risposta completa con una spiegazione chiara, coerente, non ambigua ed elegante (*criterio*).
Si sofferma ad analizzare le richieste della domanda e risponde ad ognuna di esse in modo pertinente ed esauriente (*descrittore*). Ad esempio: nella risposta usa il vocabolario specifico attinente alla domanda (...), indica chiaramente quali sono le premesse (...) e quali le conseguenze (...), si esprime con frasi complete e corrette (...) (*indicatore*)

- (2) Comunica con chi lo ascolta efficacemente (*criterio*). Non interrompe e ascolta con attenzione chi ha il turno di parola. Prende il turno di parola e si esprime dopo aver controllato di aver ben compreso chi ha parlato prima di lui (*descrittore*). Ad esempio: guarda negli occhi chi parla, ha una pausa di silenzio prima di cominciare a parlare dopo il suo interlocutore, riassume ciò che ha detto chi ha parlato prima di lui (*indicatore*)

- (3) Identifica gli elementi importanti del problema (*criterio*). Evidenzia quali sono le informazioni importanti del problema su cui fermare la propria riflessione, esamina la connessione tra le informazioni per rilevare quali sono più significative (*descrittore*). Ad esempio: Dopo aver riflettuto circa la scelta di come variabile più significativa. Tra le varie connessioni cheha con altri elementi del problema ... giustifica la rrealzione con ...come maggiormente determinante rispetto alle altre (*indicatore*)
- (4) Include un diagramma chiaro e semplificato (*criterio*). La rappresentazione delle relazioni tra i fattori sono non solo complete, ma anche evidenti nei loro rapporti di causa ed effetto e nella loro successione temporale; le parole scelte sono essenziali ed efficaci (*descrittore*). Ad esempio: Il diagramma è presentato al momento giusto dello sviluppo del ragionamento e si rivela molto efficace per precisare meglio i risultati ottenuti. E corredato anche da chiare spiegazioni che facilitano la sua comprensione (*indicatore*)

L'ancora è un esempio concreto, scelto tra i molti disponibili, che serve precisare ulteriormente un indicatore o un descrittore. Può anche essere descritta come un esempio di prestazione con la quale possono essere messe a confronto altre prestazioni. Ecco alcuni esempi, in corsivo di àncora dei criteri sopra riportati:

- 1) Dà una risposta completa con una spiegazione chiara, coerente, non ambigua ed elegante (Ad esempio, usa espressioni come: *“Si potrebbe considerare il problema da almeno tre punti di vista. In base al primo.. in base al secondo... in base al terzo... Tuttavia queste prospettive sembrano non essere completamente soddisfacenti perché... Secondo una mia opinione [àncora]”*
- 2) Comunica efficacemente con chi lo ascolta (ad esempio, usa espressioni come: *“Dopo aver presentato questa tesi in generale e astratto, penso che potrei essere più comprensibile con una esemplificazione e con una metafora...” [àncora]*)
- 3) Identifica gli elementi importanti del problema (Ad esempio, si esprime con espressioni come: *“Dopo aver letto più volte il problema penso che questi siano le informazioni sulle quali si debba riflettere per trovare una soluzione. Tuttavia è importante non solo identificare le informazioni significative, ma anche porle secondo un ordine di importanza rispetto all’obiettivo clic deve essere raggiunto... “[àncora]”*)

4) Include un diagramma chiaro e semplificato (Ad esempio. dopo aver descritto con parole un fenomeno o un sistema di relazioni fra variabili, afferma: “*Potremmo sinteticamente presentare quanto è stato fin qui detto con questa rappresentazione che illustra come le varie forze siano tra loro correlate e come ognuna di esse vari col variare delle altre... “ [àncora])*

Nel sistema di valutazione in generale alcuni esempi potrebbero assumere una posizione particolare per indicare con concretezza gli standard che la scuola o il processo formativo intende raggiungere a una determinata età, alla fine di un ciclo o a conclusione di un ciclo evolutivo. Questi esempi vengono indicati come *standard di prestazioni di riferimento* (in inglese indicati con il termine: *benchmark*).

Le prestazioni (in particolari quelle complesse) non sono quasi mai eseguite dicotomicamente bene o male. Poiché richiedono l'integrazione di molte dimensioni o tratti, sono soggette a un apprendimento più o meno lungo, richiedono più o meno difficoltà per essere messe in atto. In ogni caso, nel momento in cui sono messe in atto presentano variazioni individuali o richiedono uno sviluppo da un livello di non presenza (assenza di competenza) fino a un livello di competenza (livello di eccellenza). Un insieme di criteri distribuiti in modo da rappresentare un “continuum” di sviluppo da un livello minimale ad un livello di perfezione è una *scala*.

Le scale passano ad essere numeriche e o qualitative oppure numeriche-qualitative. Le scale *numeriche* sono scale che ne utilizzano numeri su un continuum. In questo caso è importante decidere il numero dei punti di una scala. Non vi è una indicazione precisa al riguardo invece tenere conto dei vantaggi e degli svantaggi che hanno a seconda che la scala sia troppo ampia o troppo ristretta. Quanto più ampia è la scala tanto più è difficile stabilire una differenza tra *punto e punto* della scala stessa. In una *valutazione* di 10 o di 30 punti. qual è la differenza tra 1 e 2 o tra 9 e 10 oppure tra 29 e 30? Qualora però si riuscissero a stabilire le differenze con chiarezza, è evidente una scala più ampia permette un giudizio più preciso. Una forma per mediare i vantaggi e svantaggi è quella di utilizzare contemporaneamente l'una e l'altra suddivisione come nell'esempio riportato alla Tavola 2.

Tavola 2

RIFLETTERE CRITICO (1-30)

Livello-1: Successo minimo (1-6)

- Dimostra una comprensione scarsa e solo una comprensione limitata dello scopo del problema o dei problemi.
- Impiega solo le parti più fondamentali delle informazioni fornite.
- Mescola fatto e opinione nello sviluppo di un punto di vista.
- Arriva a conclusioni dopo uno sguardo veloce e superficiale soltanto a uno o due elementi di informazione.
- Non prende in considerazione le conseguenze.

Livello-2: Successo a livello rudimentale (7-12)

- Dimostra solo una comprensione molto generale dello scopo del problema.
- Si focalizza su un singolo problema.
- Utilizza solo l'informazione fornita.
- Potrebbe includere opinioni ed anche fatti nello sviluppo di una posizione.
- Arriva a conclusioni dopo un esame limitato dell'evidenza con poco interesse per le conseguenze.

Livello-3: Successo lodevole (13-18)

- Dimostra una comprensione generale dello scopo del problema e più di uno dei problemi implicati.
- Utilizza i punti principali di una informazione dai documenti e almeno una idea generale dalle conoscenze personali per sviluppare una posizione.
- Costruisce una conclusione sull'esame delle informazioni e su qualche considerazione delle conseguenze.

Livello-4: Successo superiore (19-24)

- Dimostra una comprensione chiara dello scopo del problema e almeno di due problemi centrali.
- Utilizza i punti principali dell'informazione dai documenti e la conoscenza personale che è rilevante e coerente nello sviluppo di una posizione.
- Costruisce una conclusione su un esame dell'evidenza maggiore.
- Considera almeno un'azione alternativa e le conseguenze possibili.

Livello-5: Successo eccezionale (25-30)

- Dimostra una comprensione chiara e accurata dello scopo del problema e delle ramificazioni dei problemi implicati.
- Impiega tutte le informazioni dai documenti e le conoscenze personali estese che sono di fatto rilevanti, accurate e coerenti nello sviluppo di una posizione.
- Basa la conclusione su un esame completo dell'evidenza, sull'esplorazione di alternative ragionevoli e sulla valutazione delle conseguenze.

(Da: J. Herman, P. R. Aschbacher. & L. Winters (1992) *A practical guide to alternative assessment*, p. 46).

Le scale *qualitative* utilizzano invece aggettivi per descrivere diversi livelli di prestazione. Tali scale indicano diversi livelli di prestazione senza stabilire dei punti. Si usano espressioni come: Ha dimostrato un buon livello di competenza... dimostra una buona maturità... Pur non avendo ancora raggiunto il livello richiesto si può dire che sia ormai ad un livello avanzato... E in via di sviluppo... Può essere definito agli inizi, non ancora sufficienti... ecc. Ad esempio, nella Tabella 4 si ha una scala qualitativa: minimo, impegnato, attivo, efficace, riflessivo

Talvolta si possono usare scale *numeriche- qualitative* come nell'esempio riportato alla Tavola 3:

Tavola 3

Risposta esemplare

Valutazione = 6

Dà una risposta completa con una spiegazione chiara, coerente, non ambigua ed elegante; include un diagramma chiaro e semplificato; comunica efficacemente con chi lo ascolta; mostra di comprendere i processi e le idee di matematica di problemi aperti; identifica tutti gli elementi importanti del problema; è possibile che includa esempi e controesempi; presenta argomentazioni di sostegno.

Piccoli difetti ma soddisfacente

Valutazione = 4

Completa il problema in modo soddisfacente, ma le spiegazioni possono essere pasticciate; l'argomentazione può essere incompleta; il diagramma può essere poco appropriato o non chiaro; comprende le idee matematiche sottostanti; utilizza con efficacia le idee matematiche.

Non è capace di cominciare in modo efficace

Valutazione=1

Le parole non riflettono il problema; i disegni travisano la situazione del problema; copia parte del problema, ma senza tentare una soluzione; manca di indicare quale informazione è appropriata al problema.

(Da: J. Herman, P.R. Aschbacher, & L. Winters (1992). *A practical guide to alternative assessment*, p. 8

Le rubriche possono essere *olistiche*, *analitiche* o *del tratto primario*. Le prime sono pensate per fornire un'idea generale degli elementi di qualità e dei livelli di competenza nelle prestazioni di uno

studente. Così le definiscono Taggart e Wood (1998, p. 62):¹³

«mostrano come uno studente conosce, comprende o in grado di fare in relazione agli obiettivi specifici di prestazione del programma di istruzione. Forniscono una impressione

globale e sono intese come una valutazione sommativa del compito. I vantaggi delle rubriche olistiche risiedono nella velocità di analizzarle e nel fatto che offrono una visione generale del successo dello studente...».

Tavola 4

Molto esperto

- Ho considerato gli interessi di chi mi ascoltava nella mia introduzione
- Sono stato espressivo con la mia voce
- Il contenuto delle mie conoscenze è stato evidente
- Ho utilizzato in modo accurato diverse fonti di informazione
- Mi sono servito di frasi chiare con buoni passaggi
- Ho organizzato e presentato le informazioni in un modo sequenziale
- Mi sono servito di vari supporti visivi per illustrare il mio contenuto
- Ho appreso nuovi fatti che posso utilizzare dopo la mia relazione orale
- ...

Competente

- Ho fatto una introduzione che era secondo me interessante
- Talvolta sono stato espressivo nella mia voce
- Il contenuto della mia conoscenza mancava di dettagli
- Ho utilizzato poche fonti di informazione in modo accurato
- Sono stato chiaro nelle mie frasi
- Talvolta mi sono espresso con frasi che non riguardavano l'argomento
- Mi sono servito di strumenti visivi senza connetterli al contenuto
- Ho appreso fatti nuovi riguardanti la mia relazione orale
- ...

Principiante

- La mia introduzione è stata senza interesse
- Ho parlato in modo monotono per la maggior parte del tempo
- Ho utilizzato un contenuto di conoscenza scarso
- Ho utilizzato una sola fonte di informazione
- Le mie frasi erano stentate o incomplete
- Sono uscito parecchie volte fuori dell'argomento
- ...

¹³ G.L. Taggart, & M. Wood (1998). *Rubrics: A cross – curricular approach to assessment*. In G. L. Taggart et alii (Eds.), *Rubrics. A handbook for construction and use* (pp. 57-74). Lancaster, Technomic Publishing.

Segna le indicazioni che meglio identificano la tua preparazione

— ESPERTO

— COMPETENTE

— P~NCIPIANTE

(Da: Taggart, & M. Wood (1998). *Rubrics: A cross – curricular approach to assessment*. In G. L. Taggart, S. J. Phifer, J.A, Nixon, & M. Wood (Eds.), *Rubrics. A handbook for construction and use* (pp. 57-74). Lancaster, Basel: Technomic Publishing, p. 62.

Le rubriche *analitiche* servono per individuare il livello della prestazione rispetto a due o più elementi di qualità.

La Tavola 5 può essere un esempio di rubrica analitica. In essa vengono scelte tre dimensioni precisate **Con** diversi indicatori: *Espansione* (numero e complessità delle relazioni, informazioni del contenuto/legami concettuali, interdisciplinarietà e connessioni tematiche), *Chiarezza* (concetti generali e varietà di dettagli, concetti irrilevanti e informazioni non fondate) e *assimilazione* (concetti vecchi e nuovi, idee personali e creative e prospettive diverse). Degli indicatori non vengono indicati i criteri e le àncore:¹⁴

Tavola 5

**RUBRICA DI VALUTAZIONE PER UNA PRESTAZIONE
NELLA COSTRUZIONE DI UNA MAPPA SEMANTICA**

Minimo: *Dimostra un livello di comprensione semplicistica del contenuto e/o un impegno limitato*

ASSIMILAZIONE

- Descrizione del contenuto di una conoscenza attraverso associazioni causali senza riferimenti semantici.
- Ricordo meccanico e livellato di informazioni con mescolanza disordinata di informazioni vecchie e nuove.

CHIAREZZA

- Elementi di informazione isolati e disorganizzati.
- Introduzione di informazioni irrilevanti.
- Una certa quantità di errori.

ESPANSIONE

- Pochissimi concetti e quasi nessuna connessione tra gli elementi di informazione.

¹⁴ Nella Tabella 5 relativa ad una Rubrica di valutazione della padronanza di un contenuto rilevato attraverso una mappa semantica si trovano la *Scala* (da Minimo a Riflessivo), le *Dimensioni* (assimilazione, chiarezza, espansione) gli *Indicatori* e i *Criteri* per ciascuna dimensione. Mancano le *àncore* o gli esempi di *benchmark*. Se la si leggesse integrando le tre diverse dimensioni in una sola e unica valutazione le si considererebbe come una *rubrica solistica*.

Impegnato: *Si dimostra impegnato nel compito e manifesta di afferrare gli elementi contenuto e delle informazioni*

ASSIMILAZIONE

- Ricostruzione della struttura della conoscenza con recupero degli elementi di cui è costituita attraverso una memorizzazione meccanica del contenuto.

CHIAREZZA

- Vari tipi di informazione con alcuni dettagli in relazione a concetti generali.
- Manca la chiarezza del tipo di relazione.
- Alcuni errori.

EPANSIONE

- Connessioni molteplici con pochi concetti, nessuna qualificazione delle connessioni e alcune di esse errate.

Attivo: *È attivamente impegnato e riflessivo sul contenuto dell'argomento e dimostra di cominciare a integrare e sviluppare nuove idee*

ASSIMILAZIONE

- Integrazione corretta delle nuove informazioni con quelle già possedute.
- Fraintendimenti precedenti risolti.
- Facilità e sicurezza di movimento sulla rappresentazione.

CHIAREZZA

- Concetti generali pienamente chiarificati con dettagli rilevanti.
- Concetti principali e dettagli sono scelti con attenzione e collocati al posto corretto..

ESPANSIONE

- Molti concetti con molti dettagli, buona strutturazione della conoscenza e qualche errore di connessione nei dettagli.

Efficace: *Sa sintetizzare strategicamente le informazioni e dimostra una capacità di sapersi focalizzare in modo significativo sulle idee centrali e sui dettagli*

ASSIMILAZIONE

- Modificazioni delle imprecisioni delle conoscenze previe.
- Recupero veloce, organizzato e significativo della struttura della conoscenza.
- Capacità di spiegazione dei nodi e delle connessioni.
- Rappresentazione come risultato della elaborazione profonda e significativa della conoscenza.

CHIAREZZA

- Nodi della struttura della conoscenza accuratamente scelti e posti secondo un chiaro ordine gerarchico.
- Connessioni presentate qualificando il tipo di relazione.
- Idee centrali evidenziate e presenza di qualche elemento attribuibile ad uno sviluppo personale.

ESPANSIONE

- Connessioni tematiche, chiarezza del punto di vista, ricchezza e corretta qualificazione delle connessioni, accenni di connessioni interdisciplinari.

Riflessivo: *Cerca sempre una maggiore comprensione della conoscenza scoprendo interpretazioni*

molteplici e implicazioni personali e limiti del lavoro

ASSIMILAZIONE

- Possesso di diverse prospettive della rappresentazione della conoscenza.
- Connessioni chiare con altre aree di conoscenza a cui alcuni nodi possono essere collegati.
- Competenza nell'applicazione della rappresentazione.
- Possesso di una buona autovalutazione della propria conoscenza.

CHIAREZZA

- La rappresentazione è utilizzata avendo chiaro il punto di vista che si assume.
- I limiti di una rappresentazione sono superati utilizzando altre tecniche di rappresentazione o rilevando i limiti della rappresentazione utilizzata.
- Buoni la relazione con le conoscenze precedenti e gli arricchimenti personali.

ESPANSIONE

- Ricchezza di concetti e ottima qualificazione tra i> nodi sia in termini gerarchici che qualificanti.
- Sono registrate anche implicazioni personali e interpersonali.

(Adattamento da: D. Hyerle (1996). Thinking Maps. *Educational Leadership*, 53(4), 85-89).

Vi sono infine rubriche relative alle *caratteristiche primarie*. Sono rubriche create per compiti particolari o prestazioni particolari dei quali vengono scelti sono alcune dimensioni o alcuni tratti più essenziali

Gli *standard* stabiliscono riferimenti di un livello di successo tale da far ritenere la eccellente prestazione. Nel caso di un atleta, gli standard stabiliscono la lunghezza del percorso da raggiungere entro un determinato tempo e entro certe condizioni. Gli standard differiscono dai criteri perché questi «sono *inferiti* dallo scopo e dalla natura della prestazione, mentre gli standard sono *selezionati* per rappresentare l'eccellenza nella prestazione» (p. 2).¹⁵ Gli standard sono fissati in base alle migliori prestazioni, cioè usano esempi di prestazioni appropriate come punto di riferimento e possono variare secondo il contesto. I criteri descrivono le caratteristiche di una prestazione, la rubrica indica i livelli di esecuzione di una prestazione, gli standard dicono “quanto bene” il compito è eseguito, cioè la sua qualità.

4. COSTRUIRE I PROPRI CRITERI DI PRESTAZIONE

¹⁵ G. Wiggins (1996). *What is a rubric? A dialogue on design and use*. In R. E. Blum, & J.A. Arter (Eds.), *A handbook for student performance assessment in an era of restructuring* (VI-5: 1-13). Alexandria, VA, Association for Supervision and Curriculum Development.

Come si procede per la creazione di una rubrica? Arter (1994) consiglia queste diverse fasi:

PRIMA FASE: *Raccogliere esempi di prestazione dello studente*

L'insegnante può raccogliere dal suo archivio o dall'archivio di colleghi lavori o compiti di anni precedenti che possono rappresentare o esprimere la prestazione che si vuole valutare. Poi si scelgano quei compiti che sembrano illustrare la migliore esecuzione. Poi ci si chieda:

- “Questa una prestazione di _____ (soluzione di problema, di composizione scritta, di saggio, ecc.). Perché ho scelto questa esecuzione?
- In che cosa è migliorata questa prestazione rispetto ad una analoga dell'anno precedente?
- In che cosa potrà ancora migliorare?

Quando si raccolgono esempi, è importante rappresentare un arco di livelli di abilità.

SECONDA FASE: *Fare un brainstorming di un elenco di caratteristiche*

Raccolto un certo numero di prestazioni, si può dividere in tre gruppi: eccellenti, medi e scarsi. Poi ci si chieda:

- Quali sono le ragioni di questa divisione?
- Che cosa manifestano queste prestazioni? (criteri! dimensioni)

Si tenti di creare un elenco tanto ampio e ricco quanto possibile. Una strategia utile è l'uso del brainstorming. Ci si preoccupi di sviluppare un vocabolario molto ricco di sfumature come, ad esempio: dettagliato, comprensivo di molte cose, perspicace, fornisce esempi, motivazione, osserva più di una cosa, considera il contenuto come anche il processo, è accurato, la discussione si relaziona ai criteri, crescita sostenuta da esempi, ben organizzata, sentimenti condivisi, bada ai punti forti e a quelli deboli, profondo nell'analisi, buone ragioni e spiegazioni, aperto, espressivo, facile da leggere, bada al miglioramento dell'abilità, padronanza, reazione personale, specifico, guarda in avanti, completo, sintetizza idee, leggibile, chiaro, ecc.

TERZA FASE: *Categorizzare*

Si raggruppino gli indicatori rispetto ai tratti o alle dimensioni.

QUARTA FASE: *Scrivere una definizione con valore neutrale di ciascun tratto*

Si dia una definizione di ciascun tratto. Questa definizione dovrebbe avere un valore neutrale, cioè non dovrebbe esprimere un giudizio di valutazione.

QUINTA FASE: *Descrivere una prestazione Buona, Media e Debole per ciascun tratto/dimensione*

Si preparino descrizioni di prestazione eccellente, media e appena sufficiente per ciascun tratto. Si usino aggettivi e frasi descrittive. Ci si riferisca all'elenco dei commenti che si sono fatti quando si sono scelti i lavori. I criteri di prestazione esigono che si descriva la prestazione così bene da poter confrontare quello che si osserva nel lavoro dello studente con le descrizioni fornite.

SESTA FASE: *Reperire lavori che servano da riferimento (benchmark)*

Si trovino lavori di studenti che siano esempi di "eccellenza", di "accettabilità", di "appena sufficienza" di .abilità. Ci si assicuri di averne diverse affinché: forniscano una gamma tipica di prestazione a vari livelli di valutazione. Questa può essere utilizzata per illustrare agli studenti che cosa fare.

SETTIMA FASE: *Migliorare quello che si è fatto*

I criteri evolvono con l'uso. Probabilmente applicandoli alla pratica si scoprirà che alcuni pezzi dei criteri funzionano bene e altri no. Si aggiungano e si modifichino le descrizioni in modo che esse comunichino meglio. Si scelgano (benchmark) lavori di riferimento e saggi che facciano da esempi più adeguati. Se fosse necessario, si rivedano i tratti. Ci si faccia aiutare dagli studenti questo è un espediente per apprendere.

5. CONCLUSIONE

Quanto qui si è riferito non esaurisce tutta la riflessione che il movimento della valutazione autentica ha prodotto. La riflessione si protrae da più di dieci anni ed è ricca di sviluppi teorici e di applicazioni pratiche. Quelle che seguono sono brevi puntualizzazioni che mirano a evidenziare alcune caratteristiche importanti del tema discusso.

Non è sufficiente perfezionare il processo di valutazione in modo che sia sempre più attendibile e oggettivo. E' necessario dare un fondamento più profondo al processo stesso attraverso un'analisi dei compiti proposti agli studenti. Se i compiti sollecitano semplicemente il richiamo di contenuti o "anche" di competenze connessi alla/e disciplina/e senza un esame preliminare della loro natura, la valutazione potrà non assumere alcun valore predittivo della capacità di saper affrontare problemi o situazioni reali, ma risolversi semplicemente nella rivelazione di ciò che la scuola ha insegnato "senza che la scuola stessa sia valutata su ciò che ha insegnato". Questo renderebbe veramente precario e discutibile lo scopo a cui essa mira.

La valutazione autentica per essere tale deve anche essere compiuta in situazione analoghe o simili a quelle in cui viene eseguita la prestazione cioè in condizioni reali. In caso diverso, non indicherebbe come il soggetto esaminato dimostrerebbe le sue capacità di muoversi in situazioni di vita ordinaria.

La valutazione autentica sottolinea il valore principalmente educativo del processo della valutazione. Spesso la valutazione dell'insegnante è la dimostrazione di un potere acquisito, uno strumento per legittimare il passaggio o meno dello studente ai livelli scolastici successivi, oppure può anche prestarsi a essere utilizzato per accrescere il grado di motivazione dello studente. Ma lo scopo e il fine della valutazione non è questo, anche se può diventarlo. Lo scopo è semplicemente quello di individuare le debolezze e i limiti dello studente come anche del processo di istruzione, un sistema di controllo per focalizzare meglio l'attività dell'insegnante e dello studente. Il fatto di essere utilizzata come "punizione" o come "sistema di selezione di chi può procedere e di chi deve rimanere fermo" non fanno di certo della valutazione uno strumento significativo. Questo appare tanto più ambiguo nella misura in cui il processo esprime un giudizio solo rispetto a "ciò" che è stato insegnato, non su "come" è stato insegnato, né sulla predittività dei compiti di fornire un giudizio sulle competenze richieste dal mondo della vita reale.

Da ultimo, la valutazione autentica sottolinea l'inadeguatezza strumentale tradizionale per valutare il progresso e lo sviluppo dell'apprendimento. Una valutazione non può limitarsi a considerare la vicinanza o la lontananza dagli standard attesi e una valutazione a campionatura con interrogazioni o compiti saltuari

non è né affidabile, né in grado di fornire elementi di valutazione del progresso dimostrato dallo studente. Non tutti gli studenti, si trovano sulla stessa linea di partenza rispetto agli obiettivi finali stabiliti dall'insegnante. Queste "differenze individuali" non possono essere trascurate e essere attribuite semplicemente a mancanza di impegno personale. Per questo, una valutazione che trascuri le differenze o il progresso individuali non può non definirsi "inautentica" e anche — se non soprattutto — "ingiusta". Una valutazione, per essere davvero autentica, deve proporsi l'obiettivo di valutare il "progresso" dell'apprendimento. Come tale, deve essere continuativa, su tempi lunghi (sia per essere evidente, sia perché lo sviluppo di competenze o il recupero di lacune non richiedono tempi brevi), individualizzata, deve mantenere la memoria del passato e del presente, deve far riferimento a un progetto personale di apprendimento, deve essere anche autovalutazione dello studente. Una valutazione autentica che risponda a tutte queste esigenze non sembra possibile utilizzando strumenti tradizionali (per esempio, questionari, interrogazioni orali che difficilmente permettono un controllo preciso delle proprie capacità). Per tali motivi l'attenzione e la riflessione si sono rivolti a un nuovo strumento: il portfolio, di cui si parlerà successivamente.