

## Valutazione e pedagogia. Un paradigma semiologico

### Assessment and pedagogy. A semiological paradigm

**Massimo Baldacci**

Full professor of General and Social Pedagogy | Department of humanities studies | University of Urbino | massimo.baldacci@uniurb.it

OPEN ACCESS

**Siped**  
Società Italiana di Pedagogia

Double blind peer review

**Citation:** Baldacci, M. (2023). Assessment and pedagogy. A semiological paradigm. *Pedagogia oggi*, 21(1), 17-25. <https://doi.org/10.7346/PO-012023-02>

**Copyright:** © 2023 Author(s). This is an open access, peer-reviewed article published by Pensa MultiMedia and distributed under the terms of the Creative Commons Attribution 4.0 International, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. *Pedagogia oggi* is the official journal of Società Italiana di Pedagogia ([www.siped.it](http://www.siped.it)).

#### Journal Homepage

<https://ojs.pensamultimedia.it/index.php/siped>

Pensa MultiMedia / ISSN 2611-6561

<https://doi10.7346/PO-012023-02>

#### ABSTRACT

This paper describes a new paradigm for assessment in the field of education. We first analyze some limitations of the metrological model of docimology, and then propose a semiological model, suitable for the assessment of intellectual and moral habits. This model is qualitative, but can also include quantitative elements.

Questo articolo descrive un nuovo paradigma per la valutazione nel campo dell'educazione. Prima analizza alcuni limiti del modello metrologico della docimologia e poi propone un modello di tipo semiologico, adeguato alla valutazione degli abiti intellettuali e morali. Questo modello è di tipo qualitativo, ma può includere anche elementi di tipo quantitativo.

**Keywords:** education; assessment; measurement; semiology

**Parole chiave:** educazione; valutazione; misurazione; semiologia

Received: March 22, 2023

Accepted: May 19, 2023

Published: June 30, 2023

#### Corresponding Author:

Massimo Baldacci, [massimo.baldacci@uniurb.it](mailto:massimo.baldacci@uniurb.it)

## Premessa

La valutazione rappresenta un momento necessario di qualsiasi intervento formativo. Essa interviene nel corso di tutte le fasi di questo: in quella iniziale, come diagnosi della situazione di partenza (necessaria per definirne o calibrarne il progetto); in quella dell'intervento, per regolarne lo svolgimento; e in quella conclusiva-sommativa, per fare un bilancio dei risultati.

L'impostazione tradizionale della valutazione si concentra però sul momento sommativo, e rappresenta una sorta di estensione della dottrina gentiliana: così come la perizia didattica è risolta nella padronanza della disciplina d'insegnamento, lo stesso accade per la valutazione. Chi sa la propria materia sa anche valutare le prestazioni dell'alunno. I limiti del paradigma tradizionale appaiono evidenti: in assenza di cautele metodologiche, la soggettività del docente gioca un ruolo troppo elevato. La ricerca docimologica ha messo in evidenza tutta una serie di inadeguatezze che sono troppo note perché le si debba riepilogare. La *docimologia* ha avuto il merito di operare una svolta fondamentale nella cultura della valutazione, dotandola di un profilo scientifico. Esagerare tale merito è difficile.

Il *paradigma metrologico* della docimologia assegna un ruolo fondamentale alla *misurazione*. Senza misure valide e attendibili dell'apprendimento, la valutazione poggia nel vuoto e diventa perciò arbitraria. Come si diceva, questa posizione ha avuto il merito di delegittimare pratiche di valutazione impressionistiche, che tendono ad essere – se non arbitrarie – poco affidabili. Tuttavia, rispetto a questo paradigma rappresentano un problema quegli aspetti dell'educazione che per loro natura sono difficilmente assoggettabili a operazioni di misurazione. Nei riguardi di questi aspetti, si corrono allora alcuni rischi. Il primo è quello di un approccio riduzionistico nei loro confronti, che per renderli misurabili finisce per stravolgerne il senso. Il secondo è quello di decretare che in quanto non misurabili non sono importanti, delegittimando così aspetti in realtà rilevanti dell'educazione. Il terzo è quello di abbandonarli a una valutazione impressionistica, magari rivalutata da un atteggiamento di marca ermeneutica.

La questione che intendo porre in questo saggio è se è possibile una valutazione basata su un paradigma diverso da quello metrologico, un paradigma che pur restando nell'ambito epistemologico – e quindi senza concedere spazio a soluzioni impressionistiche – eviti di cadere nel riduzionismo. A questo proposito, prima compirò qualche osservazione sul paradigma metrologico, poi cercherò di tratteggiare a grandi linee un diverso paradigma.

## 1. Osservazioni sul paradigma metrologico

Per *paradigma metrologico* intendo il modello docimologico che pone la misurazione come fondamento necessario della valutazione.

Come ha osservato Davidson (1992, pp. 300-302) la misurazione ha due presupposti (cfr. anche Boniolo, Vidali, 1999, pp. 317-331). In primo luogo, occorre una relazione a due posti che sia transitiva e asimmetrica; ossia, una *relazione d'ordine*:  $(x > y) \ \& \ (y > z) \rightarrow (x > z)$ .

Questa relazione indica però soltanto che rispetto a una certa *proprietà* qualcosa è maggiore (o minore) di qualcos'altro, ma non specifica la natura di questa proprietà e di questo qualcosa. In secondo luogo, occorre perciò "l'esistenza di qualche contenuto empirico" (ivi, p. 301), tale da dare un significato specifico a tale relazione. Ciò a sua volta richiede una teoria dell'oggetto e della proprietà in questione, ossia: così come "non possiamo assegnare intellegibilmente una lunghezza a un oggetto qualunque se non c'è una teoria comprensiva che valga per gli oggetti di quel genere" (ivi, p. 302), allo stesso modo non possiamo attribuire una misura all'apprendimento al di fuori di una teoria plausibile dell'apprendimento stesso (e di una concezione circa l'educabilità degli esseri umani).

Intendo compiere alcune osservazioni sul paradigma metrologico muovendo da questi due presupposti.

Iniziamo dal secondo presupposto, la misurazione degli apprendimenti implica una teoria dell'apprendimento medesimo. Raramente i lavori docimologici si soffermano su questo aspetto. Ritengo, però, che questo non dipenda da un difetto teorico, bensì dall'adesione tacita alla teoria canonica dell'apprendimento, definito come una modificazione del comportamento, o – in termini cognitivisti – come una modificazione della struttura cognitiva. Dare per scontata questa concezione dell'apprendimento può, però, portare ad esiti discutibili per la pedagogia. A questo proposito, occorre porsi almeno due domande.

La prima è se l'apprendimento può rappresentare la base empirica della teoria dell'educazione, o se vi sono aspetti dell'educazione che non sono riducibili ad esso. A questo proposito, due posizioni emblematiche sono quelle di Laporta e di Biesta. Laporta (1996) ha sostenuto l'esigenza di operare una *riduzione empirica* del concetto di educazione – troppo vago – a quello di apprendimento, più preciso e che si riferisce a risultati misurabili, per fondare una scienza empirica dell'educazione. A questo proposito, si deve sottolineare che una riduzione empirica è diversa da una *riduzione ontologica*. Quest'ultima asserirebbe che l'educazione non è altro che apprendimento, mentre la prima comporta soltanto che l'apprendimento sia un *indicatore* dell'educazione. Tuttavia, il rischio di slittare dal piano empirico a quello ontologico è evidente, e di fatto percorrendo la strada riduzionista si finisce spesso per parlare soltanto di apprendimento, obliando il concetto stesso di educazione. Per motivi di questa natura, Biesta (2022) ha criticato questa riduzione dell'educazione a mero apprendimento, denominandola come una *learnification* dell'educazione che porta a smarrirne gli aspetti più rilevanti, soltanto perché non sono assoggettabili a operazioni di misurazione. Tuttavia, se le critiche di Biesta sono convincenti, le sue indicazioni alternative sono invece vaghe e poco concrete. Dal canto mio, ho cercato di prendere seriamente sia la posizione di Laporta, sia i rischi di riduzionismo ontologico. L'ipotesi che ho formulato circa il rapporto tra educazione e apprendimento fa perno sull'idea della *sopravvenienza* (cfr. Davidson, 1992): l'educazione sopravviene sull'apprendimento, ossia presuppone l'apprendimento ma non si riduce ontologicamente ad esso, è caratterizzata da un sovrappiù di senso (per i particolari rinvio a Baldacci, 2012). Ciò implica che osservando l'andamento dell'apprendimento si possa comunque capire qualcosa circa l'andamento del processo educativo. Quindi la risposta alla domanda posta è che l'apprendimento può essere assunto come la base empirica della teoria dell'educazione non in forma immediata, ma soltanto nel quadro di una serie di condizioni, di cautele e di limiti.

La seconda domanda che occorre porsi è relativa alla teoria dell'apprendimento stesso. La teoria canonica a cui abbiamo fatto riferimento, secondo la quale l'apprendimento è una modificazione del comportamento, è idonea a fornire il contenuto empirico alla teoria dell'educazione nel suo complesso? Mi sembra legittimo dubitarne. Bateson (1997) ha asserito che tale teoria è unilaterale, poiché coglie soltanto un tipo di apprendimento. La teoria dei livelli logici dell'apprendimento formulata da questo studioso ne indica tre. Li accenniamo soltanto, rinviando ad altre sedi per un loro esame organico. L'apprendimento di primo livello o *protoapprendimento* corrisponde alla definizione canonica: rappresenta una modificazione del comportamento; rispetto al campo della formazione i suoi prodotti tipici sono conoscenze e abilità inerenti a un certo campo. L'apprendimento di secondo livello o *deuteroapprendimento* rappresenta una modificazione del processo di protoapprendimento, i suoi prodotti tipici sono abiti intellettuali e/o etico-sociali. Per abito s'intende una disposizione duratura, un atteggiamento consolidato. Questo tipo d'apprendimento si struttura in modo collaterale a quello di primo livello, ma secondo una scala temporale più estesa. Infine, l'apprendimento di terzo livello rappresenta una modificazione del deuteroapprendimento, ossia degli abiti precedentemente strutturati. Poiché questi ultimi costituiscono una seconda natura del soggetto, il loro cambiamento è però infrequente e difficoltoso.

Sulla base di questa teoria più sofisticata dell'apprendimento, possiamo tentare di riarticolare la questione del contenuto empirico della teoria dell'educazione. L'ipotesi che ho avanzato è quella di assumere i diversi tipi di apprendimento descritti come domini estensionali di concetti pedagogici distinti, che vengono così a differenziarsi sulla base del loro riferimento empirico (si veda Baldacci, 2012). In questa direzione, si può azzardare che l'*istruzione* si riferisca (prevalentemente) al *protoapprendimento* (all'acquisizione di conoscenze e abilità relative ai vari campi del sapere); mentre il dominio estensionale dell'*educazione* sia costituito (essenzialmente) dal *deuteroapprendimento* (dalla strutturazione di abiti intellettuali ed etico-sociali). L'apprendimento di terzo livello sarebbe invece il riferimento della *metabetica*, delle pratiche di auto-cambiamento tipiche dell'età adulta. Ovviamente, queste corrispondenze non sono assolute, ma valgono soltanto *per lo più*. Tuttavia, chiariscono che l'apprendimento può rappresentare la base empirica della teoria pedagogica soltanto in senso generale. Quando s'intende passare a processi di controllo empirico e di valutazione, occorre selezionare specifici domini estensionali, e operare secondo i vincoli intrinseci alla logica interna di ciascuno di essi. Vediamo le conseguenze.

Il paradigma metrologico della docimologia è pensato in rapporto alla logica intrinseca al protoapprendimento, all'istruzione concernente conoscenze e abilità disciplinari. E rispetto a questo dominio estensionale risulta pienamente plausibile, e anzi – come si è accennato – ha indubbi meriti. L'errore sarebbe quello di generalizzare arbitrariamente questo paradigma agli altri domini estensionali, agli altri livelli del-

l'apprendimento. Oppure, poiché si percepisce che questa generalizzazione sarebbe arbitraria, abbandonare questi altri domini a una valutazione meramente impressionistica.

Da ciò consegue che la valutazione inerente al dominio del deuteroapprendimento, a cui si riferisce essenzialmente l'educazione, debba assumere un paradigma calibrato sulla logica interna a questo livello, e che tale paradigma non può quindi essere quello metrologico. Quest'ultimo è mirato ad apprendimenti specifici che si realizzano in tempi circoscritti (nel giro di una unità didattica o d'apprendimento). Al contrario, gli abiti intellettuali ed etico-sociali che si strutturano nel processo educativo (che a scuola è solitamente collaterale ai percorsi di istruzione) sono di tenore generale, e la loro strutturazione avviene in tempi estesi. Date queste caratteristiche degli abiti, la loro valutazione esige l'osservazione della condotta in situazioni ordinarie diverse e per lunghi periodi. Ma di questo diremo più avanti. Prima preveniamo un paio di possibili obiezioni.

Si è detto che gli abiti sono definibili come atteggiamenti duraturi. Si potrebbe allora osservare che la psicologia sociale prevede procedure di misurazione degli atteggiamenti (Manganelli Rattazzi, 1991). Non voglio escludere che tali procedure possano svolgere un ruolo ausiliario rispetto alla valutazione degli abiti in sede educativa. Tuttavia, si deve osservare che nella psicologia sociale l'operationalizzazione del costrutto "atteggiamento" lo rende almeno in parte diverso dal costrutto di "abito" a cui sto facendo riferimento. In linea di principio, si accetta che l'atteggiamento sia uno stato latente misurabile solo indirettamente attraverso l'osservazione del comportamento. Tuttavia, poiché risulta impossibile seguire i soggetti nelle varie situazioni ordinarie, si opta per un artificio operativo. La misurazione avviene tramite *scale di atteggiamento* che consistono nel chiedere di esprimere il grado di accordo o di disaccordo rispetto a questioni specifiche. In questo modo si misurano però atteggiamenti specifici, piuttosto che generali. Si potrebbe ribattere che è possibile costruire scale composte di molte domande che possono sondare atteggiamenti di raggio più ampio. Tuttavia, è stata ampiamente constatata la bassa validità predittiva delle misure così ottenute rispetto ai comportamenti effettivi dei soggetti, la bassa correlazione tra le une e gli altri (Tajfel, Fraser, 1979). Ciò viene spiegato ipotizzando che sul comportamento non incida soltanto l'atteggiamento ma anche le diverse situazioni specifiche e i significati loro attribuiti dai vari soggetti. Ma ciò equivale ad ammettere la scarsa *validità ecologica* di questa procedura di misurazione degli atteggiamenti, e a indicare la maggiore appropriatezza dell'osservazione delle condotte in differenti situazioni ordinarie (di cui la vita scolastica offre un campionario significativo, anche se non pienamente rappresentativo).

Questo però innesca una seconda obiezione. Se nella vita scolastica è possibile osservare i comportamenti in situazioni diverse, allora (come si indicava sopra) è possibile misurare indirettamente gli atteggiamenti attraverso tale osservazione. Si individueranno quali comportamenti possono rappresentare gli indicatori degli atteggiamenti che s'intende valutare; quindi, – attraverso apposite *check list* o tavole strutturate di osservazione – si registrerà la loro occorrenza, conteggiandone poi la frequenza. Si tratta cioè di una misurazione compiuta a livello di *scala nominale* (per presenza-assenza), che – come tale – autorizza il conteggio delle frequenze. Sulla base di queste ultime si può poi valutare in quale grado sia presente un dato atteggiamento. Anche in questo caso, non intendo escludere la almeno parziale utilità di questo modo di procedere. Tuttavia, occorrono almeno un paio di precisazioni. Innanzitutto, si deve considerare che (in modo analogo alla competenza linguistica), una delle caratteristiche di un abito è la *produttività*, cioè la sua capacità di produrre molti comportamenti diversi a seconda delle situazioni e delle circostanze. Qualsiasi lista chiusa di rilevazione dei comportamenti ha perciò bisogno di essere integrata da un'osservazione di carattere più ampio e informale, e quindi la misurazione ottenibile rappresenta solo un elemento di una considerazione più estesa. Per altro, questo è tipico anche delle valutazioni mediche, che si giovano di osservazioni cliniche oltre che di dati quantitativi. Si potrebbe però asserire che anche questa osservazione più fluida ricade entro la logica della misurazione su scala nominale: la constatazione della presenza o dell'assenza di certi comportamenti, e la stima della loro frequenza – sebbene compiuta in modo approssimativo. E, almeno sotto un certo profilo, questa asserzione è corretta. Tuttavia, questo ci riconduce alla questione del primo presupposto del concetto di misurazione: la necessità di una *relazione di ordine*.

Tra quelle che vengono indicate come le possibili *scale di misurazione* (nominale, ordinale, a intervalli, di rapporti), una tale relazione sussiste soltanto a partire da quella *ordinale*, che è basata appunto su relazioni del tipo "maggiore di" e "minore di", senza presupporre un intervallo costante tra un grado e l'altro. L'esempio tipico è rappresentato dalla scala di durezza dei minerali di Mohs (dal talco al diamante). Sartori (2011) ha perciò asserito che di misurazione si può parlare sensatamente soltanto a partire dalla scala ordinale, e

che la scala nominale non è quindi un'effettiva scala di misurazione, bensì un'operazione di categorizzazione in cui l'eventuale uso di numeri ha una mera funzione di codificazione, non di quantificazione. Si obietterà che è possibile conteggiare le occorrenze di una certa categoria, ma – come ha osservato Agazzi (1979) – questo non sopprime la differenza tra variabili autenticamente quantitative (p.e. la statura) e variabili intrinsecamente qualitative (p.e. il colore degli occhi). Pertanto, includere la categorizzazione nominale tra le scale di valutazione rappresenta una mera convenzione, per quanto ciò possa essere utile in molti casi. Per considerarla una misurazione in senso proprio sarebbe necessario uno stiramento concettuale (*conceptual stretching*; Sartori, 2011) della nozione di misurazione, o un suo indebolimento conseguente alla rinuncia al requisito della *relazione d'ordine*. Si hanno allora versioni deboli o stirate del concetto di misurazione, come quella proposta da Stevens (1971, p. 89): “Misurare è assegnare, conformemente a una regola, numeri a oggetti o eventi” (che però egli considera come una definizione convenzionale (Ivi, p. 87)). Se si prende questa definizione in modo sostanziale, allora quando un allenatore di calcio decide la formazione della squadra, assegnando a ogni giocatore una maglia con un certo numero, sulla base di una regola di corrispondenza tra numero e ruolo ricoperto (cioè, 1 = portiere; 2 = terzino destro; 3 = terzino sinistro ecc.) compie una misurazione? In base a tale definizione si dovrebbe rispondere positivamente. Perché allora a nessuno verrebbe in mente dire che l'allenatore ha fornito la misurazione della squadra, anziché dire che ne ha annunciato la formazione? Semplicemente perché non sarebbe di alcuna utilità porre la questione in questi termini, mentre è utile ed economico codificare la disposizione dei giocatori nei vari ruoli attraverso i numeri delle maglie.

In base a questo controesempio, la definizione debole di Stevens è allora fallace? Perché stirare la definizione di misurazione se questo la porta a coprire casi per i quali parlare di misurazione appare palesemente inadeguato? Rispetto al motivo di questo stiramento, credo che innanzitutto si possa chiamare in causa la tendenza di ogni paradigma a incasellare nel proprio reticolo il più vasto ambito di fenomeni che sia possibile. Kuhn (1978) ha definito questa pratica come “soluzione di rompicapo”, connettendola alla difesa di un paradigma (nei periodi di “scienza normale”). Infatti, ciò che non è incasellabile nella sua griglia rappresenta per esso una “anomalia” che ne limita la validità o la mette addirittura in crisi. Così, la *scala nominale* è un artificio per risolvere il rompicapo di sussumere le categorizzazioni qualitative sotto il paradigma metrologico.

Lo stesso Stevens presenta però tale artificio come una *convenzione*. Come ha chiarito Preti (1975) il convenzionalismo si serve di stipulazioni, ma queste non sono arbitrarie, sono motivate dalla loro efficacia rispetto a certi problemi. Trattare certi fenomeni nei termini di una loro misurazione su *scala nominale* (con le conseguenti operazioni matematiche e statistiche che essa autorizza) si è rivelato utile in molti casi (anche nell'ambito pedagogico). E ciò legittima questa pratica, sebbene non ci si dovrebbe mai scordare (come invece talvolta avviene) del suo carattere convenzionale. Appare però chiaro che usare in questo senso nominale il concetto di misurazione non riposa su un sostrato ontologico, su proprietà intrinsecamente quantitative, bensì su una operazione convenzionale la cui giustificazione è puramente pragmatica. Cioè, è praticata perché è utile, ed è accettabile solo finché risulta tale. Quando non lo è (come nel caso della formazione calcistica) sarebbe vano e inappropriato porre le cose in questi termini. E se ciò porta addirittura a un modo svantaggioso di vedere le cose (che ne fa perdere aspetti rilevanti), allora sarebbe miope volerle incasellare per forza secondo il punto di vista metrologico. Sarebbe piuttosto auspicabile un cambiamento dell'angolo visuale, ossia un mutamento di paradigma. Credo che questo sia proprio il caso della valutazione portata nel campo dell'educazione, riferita ai processi di deuteroapprendimento connessi alla formazione degli abiti intellettuali ed etico-sociali. Invece di porre la questione nei termini della misurazione su scala nominale dei comportamenti assunti come indicatori degli abiti, la si può vedere secondo un diverso paradigma, più appropriato a questi fenomeni formativi. Senza per questo negare una parziale utilità delle operazioni di misurazione, ma considerandole come una via ausiliaria.

## 2. Un paradigma semiologico della valutazione

Come si è detto, considerare la rilevazione della presenza o dell'assenza di certi comportamenti come una misurazione su scala nominale è giustificato dall'utilità di questo modo di procedere. Ma quando questa utilità sia dubbia, non vi è alcun vincolo epistemologico che obbliga ad agire così.

Come è noto, le tavole di rilevazione per presenza e assenza sono state originariamente ideate da Bacone (nel suo *Novum Organum*, del 1620), nel contesto di un modello qualitativo dell'osservazione volto a determinare le forme delle nature semplici delle cose, le loro qualità irriducibili (Rossi, 2004, pp. 382-390). L'idea che il mondo fisico sia un contesto di fenomeni quantitativi e misurabili si afferma con Galilei (dal *Saggiatore* del 1623 al *Dialogo sui massimi sistemi* del 1632) (Geymonat, 1969), ed è all'origine della scienza moderna.

Tuttavia, per quanto concerne l'ambito delle scienze umane, compiere opzioni aprioristiche tra gli approcci quantitativi e quelli qualitativi, non costituisce una buona soluzione epistemologica. A questo proposito, mi limito a tre rapide annotazioni. In primo luogo, la scelta tra approcci quantitativi e approcci qualitativi dovrebbe essere compiuta rispetto al tipo di problema da affrontare (nel nostro caso in rapporto alla natura dell'oggetto da valutare) secondo il criterio di appropriatezza rispetto ad esso. In secondo luogo, il confine tra il quantitativo e il qualitativo è meno netto e rigido di quanto comunemente si ritenga, e vi sono perciò zone di sovrapposizione (Thom, 1981, pp. 461-476): per esempio, l'apprezzamento di una qualità è graduabile secondo una scala ordinale d'intensità (poco/molto) (cfr. anche Gattullo, 1990). Infine, la vera questione – che riguarda sia il quantitativo che il qualitativo – è quella del *rigore* della rilevazione, della severità e della scrupolosità delle sue modalità (nei limiti in cui ciò è possibile nel campo delle scienze umane).

Il paradigma che intendo proporre è di tipo prevalentemente qualitativo, ma senza alcuna preclusione verso modalità quantitative, che vi sono ricomprese in forza della suddetta sovrapposizione. Le vere questioni sono quella dell'appropriatezza rispetto all'oggetto da valutare, e quella del rigore delle rilevazioni. A questo proposito, nel discutere il rapporto tra modelli quantitativi e modelli qualitativi, Granger (1981, p. 95) asserisce che – se per un certo periodo si è ritenuto che si dia scienza solo dell'universale, e in un periodo successivo solo del misurabile – oggi dovremo piuttosto dire che “si ha scienza solo dello *strutturabile*”. La questione è dare una struttura appropriata e sufficientemente rigorosa alla valutazione in educazione.

In ogni caso, intendo restare nel quadro dei paradigmi di natura epistemologica, senza saltare nel terreno di quelli di tipo ermeneutico. Mi sto riferendo alla distinzione posta da Rorty (1986): l'epistemologia consiste nel ritenere che le asserzioni possano essere giustificate in base a criteri di razionalità; l'ermeneutica non ritiene questo possibile, e crede che l'unica fonte di legittimazione sia l'accordo intersoggettivo stipulato attraverso la conversazione. Rispetto a ciò, sono d'accordo con Putnam (1993) nel ritenere basilare la distinzione tra asserzioni migliori e asserzioni peggiori, e quindi nel considerare irrinunciabile un atteggiamento epistemologico. Nel campo delle scienze umane, questa distinzione non va però irrigidita, e si deve piuttosto guardare a una predominanza relativa, che vede la necessità di una componente ermeneutica anche in un modello epistemologico (e viceversa).

Il paradigma che intendo proporre rimane quindi sul terreno epistemologico (pur contemplando una componente ermeneutica) ed è di tipo essenzialmente qualitativo (pur potendosi giovare di apprezzamenti quantitativi in via ausiliaria). Tale paradigma vuole dare una struttura appropriata e sufficientemente rigorosa alla valutazione dei processi di deuterioapprendimento, ossia alla formazione degli abiti intellettuali ed etico-sociali. Come si è detto, tale formazione richiede tempi estesi, e il suo andamento è perciò meno evidente rispetto agli apprendimenti di primo livello. Sono però necessarie informazioni per regolarne lo svolgimento (nonché per tracciare bilanci conclusivi). Pertanto, l'oggetto di osservazione e valutazione è per lo più il gruppo classe nel suo insieme, gli abiti che appaiono in via di strutturazione diffusa in tale gruppo.

Il paradigma in questione non è una mia creazione, e non è nato in ambito pedagogico. Nella sua impostazione generale è stato tracciato da Ginzburg (2004), che lo ha presentato come un *paradigma indiziario*. Io ho soltanto cercato di trasporlo sul terreno pedagogico della valutazione degli abiti, rafforzandone la connessione con altri due riferimenti importanti per dare una struttura epistemologica a tale valutazione: la *semiotica cognitiva* di Peirce (1980), e la *logica dell'incerto* di De Finetti (1989). Preferisco, perciò, l'espressione *paradigma semiologico*. Ho messo a punto questa ipotesi paradigmatica nel corso del progetto Prin 2017 sul *curricolo di educazione etico-sociale*, e ne ho dato conto nell'omonimo volume (Baldacci, 2020), al quale rinvio per una sua illustrazione più esauriente. Qui mi limiterò a descriverne alcune linee essenziali.

Il paradigma epistemologico che Ginzburg (2004) presenta come “indiziario” concerne “discipline emi-

mentemente qualitative, che hanno per oggetto casi, situazioni e documenti individuali, *in quanto individuali*” (p. 111), come le scienze storiche. Ritengo che anche la pedagogia possa essere inclusa tra le discipline “eminamente qualitative”, senza che ciò implichi la rinuncia ad avvalersi pure di approcci quantitativi, quando questi sono utili e appropriati.

In tali discipline l’osservazione è diretta a cogliere *segn*, *indizi*, *sintomi* che rinviano a stati, fenomeni o eventi non direttamente osservabili (perché latenti o già passati). Ne è un esempio tipico la *semeiotica medica*, sulla cui base il medico cerca i sintomi che permetteranno di diagnosticare la malattia (avvalendosi sia dei dati qualitativi dell’esame clinico obbiettivo, sia degli esami quantitativi di laboratorio). Un altro esempio è quello dell’*investigazione poliziesca*, nel corso della quale il detective cerca gli *indizi* che lo conducano all’autore di un delitto (anche in questo caso, sulla base di reperti sia qualitativi che quantitativi). Sintomi e indizi sono *segn naturali* che permettono di orientare l’indagine. Nella semiotica cognitiva di Peirce ricadono sotto la categoria dell’*indice*, ossia di una tipologia di segni che hanno una connessione reale e oggettiva con l’oggetto che indicano (Peirce 1980, p. 140), anziché convenzionale (come nel caso del segno linguistico). In base a questa impostazione, si tratta di considerare le condotte degli scolari come possibili *segn* (indici) dei loro atteggiamenti (cioè, dei loro abiti in via di strutturazione). Questo non è altro che l’ordinario processo con cui stimiamo gli atteggiamenti o le qualità personali (etico-sociali e intellettuali) degli altri. A questo proposito, ai fini della valutazione educativa, faccio mia l’affermazione di Dewey (1974) secondo cui l’indagine scientifica non è altro che l’indagine di senso comune resa più attenta e sorvegliata (cioè, più rigorosa). Di conseguenza, si tratta di osservare le condotte degli scolari, interpretandole come indizi dei tipi di abiti che sono in via di strutturazione, nel quadro di una struttura epistemologica che renda cauto e sorvegliato questo processo. Qui mi limito a due sole indicazioni. In primo luogo, per serbare *validità ecologica*, questa osservazione deve essere compiuta nel corso ordinario della vita scolastica, senza prevedere prove artificiose. In secondo luogo, una volta rilevati determinati indizi comportamentali si pone la questione del passaggio inferenziale da questi all’atteggiamento latente di cui sono espressione. La semiotica cognitiva di Peirce fornisce il quadro inferenziale per categorizzare il passaggio dall’Indice all’Oggetto (latente o assente) che indica. Ovviamente, nel nostro caso I = condotta; O = atteggiamento (latente). Si tratta dell’*inferenza abduttiva* (distinta da quella deduttiva e da quella induttiva). La sua forma è la seguente:

O → I      (il fuoco implica il fumo);  
 I            (si rileva la presenza di fumo);  
 dunque O   (dunque, vi è un fuoco).

Da un punto di vista strettamente logico, questa inferenza all’indietro (detta anche *retroduzione*) da I ad O non è valida. Infatti, il fumo potrebbe anche essere dovuto a circostanze diverse dal fuoco. Ciò che si può legittimamente inferire dalla premessa O → I, è che dato O deve seguire necessariamente I, e che se non si dà I non vi sono motivi di ritenere che si dia O. L’inferenza abduttiva da I a O ha perciò un carattere congetturale, e quindi la conclusione è puramente *probabile* (cioè, se vedo fumo posso dire che probabilmente vi è un fuoco) (Peirce, 1978). Nonostante questo limite, questo è il modo di procedere dell’indagine clinica medica e della stessa investigazione poliziesca (il cui modello è Sherlock Holmes, sebbene Conan Doyle la descrivesse erroneamente come “deduzione” (cfr. Eco, Sebeok, 2004)). In queste ricerche si mira a raccogliere indizi fino a che non si consideri raggiunta la soglia dell’evidenza probatoria (in ciò entra quindi in gioco anche una componente induttiva (Peirce, 1978)). Nel nostro caso, se le condotte di tipo I sono indice dell’atteggiamento O, in prima approssimazione l’inferenza dall’osservazione di I alla presenza di O è suffragata dalla frequenza con cui si rileva I e dalla varietà delle situazioni nelle quali si dà I (Strawson, 1975).

Si può notare come questo processo sia eminentemente qualitativo, ma comprenda anche aspetti di natura quantitativa, sebbene stimati in modo approssimativo. Le categorizzazioni (nominali) del tipo di atteggiamento e del tipo di condotta sono qualitative, e la constatazione della presenza di un certo tipo di condotta ha anche una componente ermeneutica (per considerare un certo gesto come indice di generosità occorre averlo interpretato come un gesto generoso). Tuttavia, gli indici comportamentali rimangono evidenze empiriche (l’osservazione è sempre carica di teoria (Boniolo, Vidali, 1999, pp. 177-196)). La frequenza di certi indici viene inoltre stimata quantitativamente, sebbene in modo approssimativo, del tipo:

poco, abbastanza, molto. Si tratta di una misurazione su scala ordinale? Mi sembra più felice considerarla come una *stima* approssimativa (di tipo ordinale). Prima di giudicarla ascientifica è bene considerare che il passaggio dal pressappoco alla precisione (Koyrè, 1967) non è un dogma assoluto, e viene piuttosto regolato in base a criteri pratici. Quando un medico chiede al paziente se ha avuto tosse, non desidera una risposta del tipo “123 colpi di tosse”, considera sufficiente sapere se questi ha tossito poco o molto frequentemente. Se l’arte medica ha un fondamento scientifico non è perché il medico conta con precisione il numero di colpi di tosse, ma perché colloca la tosse in un quadro teorico (clinico-semeiotico) entro il quale assume il valore di sintomo di determinate patologie. Per diagnosticare la presenza di una di queste cercherà però anche altri sintomi. La valutazione degli atteggiamenti sulla base dell’osservazione delle condotte ha un profilo scientifico se è resa cauta e sorvegliata dalla consapevolezza di questi limiti e di queste clausole. La specificazione di queste ultime non è però ancora conclusa.

Un altro aspetto notevole è rappresentato dal citato carattere probabilistico dell’inferenza dalle condotte all’atteggiamento soggiacente. Questo rende necessaria una teoria della logica probabilistica dell’inferenza inversa (retroductiva). Una tale logica è stata messa a punto dai cosiddetti modelli neo-bayesiani, tra i quali spicca la *logica dell’incerto* di De Finetti (1989). Qui ne darò solo un cenno, rinviando per una più ampia illustrazione a Baldacci (2020, pp. 139-149).

Come si è detto, il nesso tra l’atteggiamento (O) e la condotta (I) non è di tipo deterministico, in quanto gioca anche il tipo di situazione. Pertanto, tale nesso è di natura probabilistica, ossia dato un atteggiamento (O) vi è una certa probabilità che si verifichi la condotta (I). Possiamo esprimere questa probabilità con  $P(I/O)$ , con / come segno per “dato”: la probabilità di I dato O.

Il nostro problema è però inverso a questo: poiché è possibile osservare soltanto I, si tratta di stimare le probabilità della presenza di O in base alla rilevazione di I, ossia  $P(O/I)$ . E le probabilità  $P(I/O)$  e  $P(O/I)$  non si equivalgono, così come la probabilità che un italiano abbia gli occhi bruni è diversa da quella che un individuo con gli occhi bruni sia italiano. Tenendo conto di questo, come si può stimare la probabilità della presenza dell’atteggiamento O muovendo dalla rilevazione delle condotte di tipo I? Appare intuitivo che se I si manifesta frequentemente e in situazioni diverse è maggiormente probabile che sia presente O. Tuttavia, questo criterio intuitivo, per quanto sensato, non è sufficiente. Il tipo di inferenza di cui stiamo trattando ha una logica controintuitiva, e si rischia facilmente di rimanere vittime di *biases* che portano a stime non corrette (Giroto, 1994, pp. 117-124). Il punto è che nella determinazione della probabilità di  $(O/I)$  incide anche  $P(O)$ , la cosiddetta *probabilità di base* di O (che in prima approssimazione può essere fatta corrispondere alla percentuale di diffusione di O entro l’Universo di riferimento), la quale viene invece spesso trascurata. Il modello di De Finetti, adottato entro un certo quadro di clausole, indica come stimare e comporre questi diversi aspetti e quindi come arrivare a valutare in modo rigoroso la  $P(O/I)$ . E la sua logica induce ad essere cauti e a non trarre conclusioni affrettate. Infatti, essa porta a ribadire che la stima di  $P(O/I)$  richiede osservazioni ripetute nel tempo (e in situazioni diversificate) della presenza di I, che portano a correggere progressivamente la stima di tale probabilità. Tenere conto dell’entità e dell’andamento (crescente o decrescente) di quest’ultima costituisce un criterio per regolare l’intervento educativo. In Baldacci (2022, pp. 177-194) ho cercato di mostrare che la logica del modello di De Finetti regge anche se invece di operare con valori numerici di probabilità (p.e.  $p = 0,25$ ), si ragiona secondo stime ordinali (del tipo: poco, abbastanza, molto probabile). Qui non posso entrare negli aspetti tecnici della questione.

La logica generale del paradigma semiologico è però quella illustrata da Ginzburg: assumere gli elementi osservabili come *segni* di stati latenti, e seguire ragionevoli cautele metodologiche nel compiere l’inferenza dai segni agli stati. Se si ragiona in questo modo, i segni possono essere sia qualitativi sia quantitativi, e la valutazione tenderà a tenere conto di una pluralità di segni. Il dominio del paradigma semiologico appare quindi più esteso di quello metrologico.

## Riferimenti bibliografici

- Agazzi E. (1979). Analogicità del concetto di scienza. In V. Possenti. *Epistemologia e scienze umane* (pp. 57-66). Milano: Mario.
- Baldacci M. (2012). *Trattato di pedagogia generale*. Roma: Carocci.
- Baldacci M. (2020). *Un curriculum di educazione etico-sociale*. Roma: Carocci.



- Baldacci M. (2022). *Praxis e concetto. Il linguaggio della pedagogia*. Milano: Franco Angeli.
- Bateson G. (1997). *Verso un'ecologia della mente*. Milano: Adelphi.
- Biesta G.J.J. (2022). *Riscoprire l'insegnamento*. Milano: Raffaello Cortina.
- Boniolo G., Vidari P. (1999). *Filosofia della scienza*. Milano: Bruno Mondadori.
- Davidson D. (1992). *Azioni ed eventi*. Bologna: Il Mulino.
- De Bartolomeis F. (1974). *Valutazione e orientamento*. Torino: Loescher.
- De Finetti B. (1989). *La logica dell'incerto*. Milano: Il Saggiatore.
- Dewey J. (1974). *Logica, teoria dell'indagine*. Torino: Einaudi (ed. or. 1938).
- Eco U., Sebeok T.A. (eds.) (2004). *Il segno dei tre. Holmes, Dupin, Peirce*. Milano: Bompiani.
- Gattullo M. (1990). Quantitativo e qualitativo in educazione e nella ricerca educativa. In V. Telmon, G. Balduzzi. *Oggetti e metodi della ricerca in campo educativo* (pp. 239-48). Bologna: Clueb.
- Geymonat L. (1969). *Galileo Galilei*. Torino: Einaudi.
- Ginzburg C. (1979). Spie. Radici di un paradigma indiziario. In U. Eco, T.A. Sebeok (eds.) (2004), *Il segno dei tre. Holmes, Dupin, Peirce* (pp. 95-136). Milano: Bompiani.
- Giroto V. (1994). *Il ragionamento*. Bologna: Il Mulino.
- Granger G. (1981). Modelli qualitativi, modelli quantitativi nella conoscenza scientifica. In AA.VV., *Il sapere come rete di modelli* (pp. 89-102). Modena: Panini.
- Hintikka J., Hintikka M.B. (2004). Sherlock Holmes e la logica moderna. In U. Eco, T.A. Sebeok (eds.), *Il segno dei tre. Holmes, Dupin, Peirce* (pp. 183-202). Milano: Bompiani.
- Koyré A. (1967). *Dal mondo del pressappoco all'universo della precisione*. Torino: Einaudi.
- Kuhn T. (1978). *La struttura delle rivoluzioni scientifiche*. Torino: Einaudi.
- Laporta R. (1996). *L'assoluto pedagogico*. Firenze: La Nuova Italia.
- Manganelli Rattazzi A.M. (1991). La ricerca empirica sugli atteggiamenti: problemi e tecniche. In R. Trentin, *Gli atteggiamenti sociali* (pp. 203-42). Torino: Bollati Boringhieri.
- Peirce C.S. (1980). *Semiotica. I fondamenti della semiotica cognitiva*. (1895-1910). Torino: Einaudi.
- Peirce C.S. (1978). *La logica dell'abduzione*. In Id., *Scritti di filosofia*. Bologna: Cappelli (1901).
- Preti G. (1975). *Praxis ed empirismo*. Torino: Einaudi.
- Putnam H. (1992). *Un realismo dal volto umano*. Bologna: Il Mulino.
- Rorty R. (1986). *La filosofia e lo specchio della natura*. Milano: Bompiani.
- Rossi P. (2004). *Francesco Bacone*. Bologna: Il Mulino.
- Sartori G. (2011). *Logica, metodo e linguaggio nelle scienze sociali*. Bologna: Il Mulino.
- Stevens S.S. (1971). La quantificazione della sensazione. In D. Lerner, *Qualità e quantità e altre categorie della scienza* (pp. 83-103). Torino: Boringhieri.
- Strawson P.F. (1975). *Introduzione alla teoria logica*. Torino: Einaudi (ed. or. 1952).
- Tajfel H., Fraser C. (1979). *Introduzione alla psicologia sociale*. Bologna: Il Mulino.
- Thom R. (1981). Qualità/quantità. In *Enciclopedia Einaudi*, Vol. 11. Torino: Einaudi.