

La formazione delle competenze valutative degli insegnanti

Cristiano Corsini, Bruno Losito

Abstract – Assessment and evaluation are two of the weakest areas of the Italian school system, both with reference to the national regulations governing them and to the assessment practices of teachers. In both initial teacher preparation and in their in-service training, a marginal importance is given to assessment and evaluation. This is a consequence of the educational policies as well as of the weakness of the educational research in the field of assessment and evaluation. The construction and the development of teachers' competences in the field of assessment and evaluation needs to be linked to the development of the empirical research in the field and to the reflection on the assessment and evaluation practices actually implemented at a school level. The critical analysis of the national assessments carried out by INVALSI could be an effective starting point for enhancing teachers' awareness of the complexity of the assessment and evaluation tasks and for the development of their competences in this field.

Riassunto – L'ambito della valutazione rappresenta uno degli aspetti di maggiore debolezza dell'istruzione scolastica nel nostro paese, sia per le norme che la regolano, sia dal punto di vista delle pratiche adottate dagli insegnanti. Sia nella formazione iniziale, sia in quella in servizio degli insegnanti, i temi legati alla valutazione occupano uno spazio marginale. Questa marginalità è da mettere in relazione sia alle scelte di politica educativa, sia alla debolezza della ricerca in ambito valutativo, particolarmente evidente nel campo della valutazione delle competenze. La costruzione e lo sviluppo delle competenze valutative degli insegnanti richiede un rapporto stretto con la ricerca valutativa e con la riflessione sulle pratiche valutative effettivamente utilizzate. L'analisi critica delle rilevazioni nazionali condotte dall'INVALSI e delle caratteristiche delle prove utilizzate in queste rilevazioni può rappresentare un punto di partenza efficace per incentivare la consapevolezza degli insegnanti della complessità della valutazione e per lo sviluppo delle loro competenze valutative.

Keywords – assessment, evaluation, competences, teacher training and professional development, national assessments

Parole chiave – valutazione, valutazione delle competenze, formazione degli insegnanti, rilevazioni nazionali

Cristiano Corsini è Professore Ordinario di Pedagogia sperimentale presso il Dipartimento di Scienze della Formazione dell'Università Roma Tre. I suoi principali temi di ricerca riguardano l'impiego della valutazione come strategia trasformativa e lo sviluppo professionale degli insegnanti. Tra le sue pubblicazioni: *Rileggere Visalberghi* (Roma, Nuova Cultura, 2018); *La valutazione che educa. Liberare insegnamento e apprendimento dalla tirannia del voto* (Milano, FrancoAngeli, 2023);

Bruno Losito è stato Professore Ordinario di Pedagogia sperimentale presso il Dipartimento di Scienze della Formazione dell'Università Roma Tre. I suoi principali temi di ricerca riguardano la valutazione, la ricerca valutativa, l'educazione civica e alla cittadinanza, lo sviluppo professionale degli insegnanti. Tra le sue pubblicazioni: *Origini e sviluppi della ricerca valutativa* (in coll. con B. Vertecchi e G. Agrusti, Milano, FrancoAngeli, 2010); *Becoming Citizens in a Changing World. IEA International Citizenship Education Study 2016 International Report* (in coll. con W. Schulz, J. Ainley, J. Fraillon, G. Agrusti, T. Friedman, Springer Open, 2018).

1. Farsi carico dell'incertezza

Qualsiasi discorso sulla formazione delle competenze valutative degli insegnanti deve necessariamente partire da due elementi di consapevolezza. Il primo riguarda il nostro sistema di istruzione, il secondo il nostro sistema di formazione degli insegnanti.

Nei corsi di Scienze della formazione primaria, la valutazione occupa un posto secondario, come emerge con chiarezza da una analisi dei corsi offerti agli studenti. Va, inoltre, rilevata l'assenza di un investimento in grado di garantire in maniera strutturata un adeguato accompagnamento del personale docente della scuola primaria alla nuova valutazione periodica e finale introdotta nel dicembre del 2020. Nella formazione iniziale per la scuola secondaria è a oggi del tutto assente un percorso in grado di far dialogare dialetticamente l'esperienza in aula con i saperi teorici. Per quanto riguarda la formazione in servizio, una parte rilevante delle azioni a livello d'istituto dovrebbe far riferimento a priorità indicate dalle scuole nei Rapporti di Autovalutazione e nei Piani di Miglioramento. Tuttavia, prendendo in considerazione le notevoli difficoltà che hanno incontrato le scuole nella determinazione di tali priorità¹, appare lecito nutrire qualche perplessità sul fatto che tali interventi risultino generalmente significativi.

L'ambito della valutazione rappresenta uno degli aspetti di maggiore debolezza dell'istruzione scolastica, sia dal punto di vista delle norme che la regolano, sia dal punto di vista delle pratiche valutative adottate dagli insegnanti. È una debolezza che ha origini lontane e che è stata più volte analizzata e denunciata dagli studiosi che hanno dedicato larga parte della propria attività di ricerca ai temi della valutazione. Sorprende l'attualità che ancora oggi caratterizza i contributi di alcuni decenni fa di studiosi come Mario Gattullo e Aldo Visalberghi². Una attualità che deriva dalla sostanziale incapacità di affrontare e risolvere i problemi da essi indicati. Una incapacità di carattere in primo luogo politico, di politica educativa e scolastica, ma anche di carattere teorico e metodologico. Nel corso degli anni, il nostro sistema di istruzione è stato interessato da molteplici cambiamenti, molti dei quali hanno riguardato la valutazione degli studenti e il sistema di formazione iniziale e in servizio degli insegnanti. Allo stesso tempo, notevoli passi in avanti sono stati compiuti nell'ambito della ricerca docimologica e valutativa. Nonostante questo, molti dei problemi individuati alcuni decenni fa sono tuttora non risolti e – anzi – ad essi se ne sono aggiunti altri in relazione a nuovi compiti posti al nostro sistema di istruzione.

In *Didattica e docimologia*, Gattullo sottolineava la debolezza della preparazione degli insegnanti in ambito valutativo. Ne conseguiva che “il patrimonio professionale comune alla gran maggioranza degli insegnanti italiani, compresi quelli universitari” finiva per coincidere da un lato con un atteggiamento di sostanziale tolleranza (se non acquiescenza) nei confronti degli studenti da esaminare, dall'altro con una sorta di pregiudizio nei confronti di tutte quelle procedure valutative che cercavano di limitare la soggettività e l'aleatorietà dei giudizi valutativi, in primo luogo le cosiddette “prove oggettive”. Di fatto, nelle loro pratiche valutative gli insegnanti tendevano a riprodurre procedure e pratiche che avevano caratterizzato le loro espe-

¹ D. Robasto, *Autovalutazione e piani di miglioramento a scuola*, Roma, Carocci, 2017.

² Si vedano, a titolo di esempio, A. Visalberghi, *Misurazione e valutazione nel processo educativo*, Milano, Edizioni di Comunità, 1955 e M. Gattullo, *Didattica e docimologia*, Roma, Armando, 1967.

rienza pregresse di studenti (esami e interrogazioni). Un'ottica di tipo "riproduttivo" che ricorda molto le analisi e le interpretazioni di Pierre Bourdieu³.

In *Misurazione e valutazione nel processo educativo* Aldo Visalberghi sottolineava, tra gli altri aspetti, come il prevalere di approcci e procedure valutativi di carattere intuitivo e routinario (le 'interrogazioni' orali) sottraessero alla attività didattica tempo prezioso che avrebbe potuto essere molto più proficuamente dedicato ad attività più ricche e più stimolanti per gli studenti.

Entrambi gli studiosi, sottolineavano con forza non soltanto la necessità di affrontare questi problemi sul piano delle politiche educative, ma anche quella di sostenere queste politiche con una seria ricerca docimologica e valutativa.

Molto è stato fatto in questo ambito da quando Gattullo e Visalberghi sostenevano questa necessità, anche grazie alla partecipazione del nostro paese alle indagini comparative internazionali promosse prima dall'IEA (*International Association for the Evaluation of Educational Achievement*) e poi dall'OCSE (Organizzazione per la cooperazione e lo sviluppo economico)⁴.

Questa partecipazione ha dato impulso al diffondersi di una consapevolezza della possibilità e della utilità di una ricerca scientificamente e metodologicamente fondata in ambito valutativo (ancorché in una dimensione di tipo comparativo su larga scala). Inoltre, la partecipazione a queste indagini ha anche contribuito a formare negli anni studiosi e ricercatori nel campo della valutazione, attraverso il confronto con i modelli teorici, gli approcci metodologici e gli apparati strumentali che le caratterizzano e attraverso il dibattito scientifico e metodologico che le ha accompagnate. Le stesse indagini valutative nazionali dell'INVALSI affondano le loro radici in questa partecipazione⁵. Nonostante questo, è innegabile che l'impatto della partecipazione a queste indagini sulle politiche educative e scolastiche sia risultato abbastanza limitato, al di là del riconoscimento formale di alcune istanze di carattere generale.⁶ In parte, questo mancato impatto è riconducibile a diffidenze di vario tipo nei confronti di queste indagini e nei confronti degli strumenti di rilevazioni utilizzati. Ma è molto probabile che esso sia anche una espressione del difficile rapporto tra ricerca educativa (e in modo specifico valutativa) e politiche scolastiche, che ha motivazioni storiche e culturali consolidate nel nostro paese⁷, imputabile non soltanto ai cosiddetti "decisori politici", ma anche alla debolezza persistente della pedagogia italiana nel campo della ricerca empirica e sperimentale. Basta scorrere le riviste di settore per rilevare come non siano frequenti i contributi che presentano risultati di ricerche empiriche (per non dire 'sperimentali') su problemi specifici legati alla valutazione del rendimento scolastico.

Se dal piano della ricerca si passa a quello della formazione degli insegnanti, è abbastanza evidente che sia nella formazione iniziale, sia in quella in servizio la valutazione occupa un

³ Cfr. P. Bourdieu, J. C. Passeron, *La reproduction. Eléments pour une théorie du système d'enseignement*, Paris, Les Editions de Minuit, 1970.

⁴ Partecipazione a cui sia Visalberghi sia Gattullo hanno dato un contributo fondamentale.

⁵ Sulle caratteristiche e i limiti delle rilevazioni condotte dall'INVALSI si veda il paragrafo 2.

⁶ Si veda, su questo, N. Bottani, *Il difficile rapporto tra politica e ricerca scientifica sui sistemi scolastici*, Programma Education FGA, Working Paper n. 17, Fondazione Giovanni Agnelli, 2009.

⁷ Si veda E. Becchi, A. Visalberghi, *Introduzione*, in E. Becchi, A. Visalberghi (a cura di), *Manuale critico della sperimentazione e della ricerca educativa*, Milano, FrancoAngeli, 1984, pp.11-33.

posto abbastanza marginale. Di fatto, si assiste a una sostanziale rinuncia, salvo sporadiche eccezioni, a un esame critico metodologicamente e scientificamente fondato delle pratiche valutative prevalenti nel nostro sistema di istruzione.

Questi due aspetti (debolezza della ricerca in ambito valutativo e caratteristiche della formazione degli insegnanti), si saldano a un'altra caratteristica del nostro sistema di istruzione, la scarsa coerenza tra indicazioni curriculari, didattica e valutazione. La mancanza di questa coerenza appare abbastanza evidente in relazione ai compiti che alle scuole e agli insegnanti vengono attribuiti, in particolare in riferimento alla valutazione delle competenze.

Pur riconoscendo l'opportunità e la necessità di una riflessione critica sul concetto stesso di competenza⁸ e sulle possibili implicazioni dell'adozione di un approccio didattico 'per competenze' (in termini, concettuali, organizzativi e didattici), la valutazione delle competenze richiederebbe non soltanto procedure differenziate e contestualizzate di raccolta delle evidenze necessarie per poter arrivare alla espressione di giudizi valutativi sufficientemente fondati (attendibili e validi), ma anche un'idea stessa della valutazione che si faccia carico della complessità di questo compito.

In primo luogo, è necessario essere consapevoli dello scarto in una certa misura inevitabile tra "prestazioni" effettivamente rilevabili in ambito scolastico e competenze di cui esse sarebbero manifestazione o "indizio". Ma per far questo sarebbe necessario organizzare la didattica e la valutazione intorno a problemi reali e porre gli studenti "in situazione", proponendo loro di affrontare compiti differenziati, di diversi livelli di complessità, che consentano loro di confrontarsi con problemi il più possibile simili a quelli che potrebbero doversi trovare ad affrontare nella vita reale – cioè compiti autentici⁹ – mettendo in atto comportamenti "competenti".

In secondo luogo, per poterle in qualche modo valutare, è necessario tenere conto delle diverse componenti di una competenza (cognitive, metacognitive, affettivo-comportamentali, valoriali) e – conseguentemente – raccogliere una *documentazione* il più possibile articolata ("evidenze" valutative di diversa natura, raccolte in contesti diversi, in momenti diversi, con strumenti diversi).

Come ha sottolineato Michele Pellerey:

È possibile, dunque, inferire la presenza di una competenza non solo genericamente, ma anche in maniera articolata sulla base di una famiglia di prestazioni, che svolgono il ruolo di indicatori di esistenza o di livello raggiunto. Solo nel caso di competenze elementari che mettano in gioco schemi d'azione di tipo ripetitivo, oppure assai semplici applicazioni di regole e principi, è possibile valutarne l'acquisizione osservando una o poche prestazioni¹⁰.

In terzo luogo, l'espressione di un giudizio valutativo comporta una ineliminabile attività (il più possibile collettiva) di analisi e di interpretazione delle evidenze valutative raccolte¹¹.

⁸ Si veda su questo L. Benadusi, S. Molina (a cura di), *Le competenze. Una mappa per orientarsi*, Bologna, il Mulino, 2018.

⁹ G. Wiggins, *Educative Assessment: Designing Assessments to Inform and Improve Student Performance*, San Francisco, Jossey-Bass, 1998

¹⁰ Cfr. M. Pellerey, *Le competenze individuali e il portfolio*, Milano, RCS, 2004, p. 114.

¹¹ Non affrontiamo qui il problema delle modalità di espressione del giudizio valutativo, ma è evidente l'incongruenza tra la complessità che caratterizza la valutazione dei livelli di competenza degli studenti e la espressione

Qualsiasi “evidenza”, qualsiasi “dato” dovrebbe essere interpretato in riferimento agli obiettivi che ci si propone di conseguire, ai livelli di partenza degli studenti, alle caratteristiche delle tecniche utilizzate per raccogliervi, all’insieme delle variabili che potrebbero in qualche modo averli influenzati.

In che modo costruire processi inferenziali in “maniera articolata”? In quali contesti didattici? Con quali procedure e con quali strumenti è possibile/utile raccogliere le evidenze valutative necessarie? Come interpretare le “evidenze” raccolte per arrivare ad esprimere un giudizio sul livello di competenza raggiunto da uno studente? In che modo articolare e formulare questa interpretazione in un giudizio valutativo che comunicabile in modo sia sufficientemente chiaro agli studenti, in modo che possa essere punto di partenza per consolidare e migliorare il proprio livello di competenza? Come tenere conto, in un’ottica valutativa, della ‘dimensione evolutiva’ di una competenza? Come tradurre, cioè, in pratiche valutative adeguate la consapevolezza che una competenza nel tempo (e in riferimento a compiti e contesti diversi) può certamente consolidarsi e svilupparsi, ma anche regredire?

Non sono molti gli studi che hanno cercato di trovare delle risposte a queste domande attraverso percorsi di ricerca metodologicamente rigorosi. Per due ragioni di fondo. In primo luogo perché ancora prevale un’idea di valutazione ancorata (spesso implicitamente) a categorie concettuali e metodologiche che richiederebbero invece di essere ridiscusse in relazione ai nuovi contesti educativi e didattici e alle nuove richieste che la valutazione delle competenze pone agli insegnanti. Un esempio può essere quello relativo al concetto di “padronanza”, riferito agli obiettivi di apprendimento, che ha avuto un ruolo importante nel dibattito e nelle pratiche didattiche e valutative negli scorsi decenni. Dal punto di vista valutativo, la ‘padronanza’ di un obiettivo determina la necessità di rilevare la presenza-assenza del possesso di una determinata conoscenza o abilità. Nell’ottica della costruzione e dello sviluppo delle competenze, diventa importante rilevare quale sia piuttosto il “livello” a cui si manifesta una competenza, in maniera contestualizzata e in riferimento alle caratteristiche del compito da affrontare. non solo la sua presenza o assenza. Inoltre, i livelli raggiunti da uno studente in una stessa competenza possono variare in riferimento – appunto – ai contesti e ai compiti e possono essere influenzati dal diverso coinvolgimento affettivo e motivazionale nei confronti di tali compiti. In secondo luogo, la didattica per competenze e la valutazione delle competenze è di fatto non solo resa più difficile, ma spesso ostacolata, dalle caratteristiche dei contesti organizzativi della nostra scuola. La costruzione e lo sviluppo delle competenze e la loro valutazione richiedono contesti organizzativi e didattici che consentano la “mobilitazione” delle conoscenze e delle abilità (le “risorse”) degli studenti, in contesti di lavoro che sollecitino la soluzione di problemi. Ma questo risulta molto difficile da realizzare all’interno dei contesti didattici e organizzativi so-

del giudizio valutativo in voti numerici, che rafforza un’idea errata della misurazione in campo educativo e incentiva procedure semplificate e prive di fondatezza metodologica (quale, ad esempio, il ricorso al calcolo della media dei voti attribuiti a prestazioni diverse). La stessa normativa relativa all’espressione della valutazione tramite giudizi descrittivi nella scuola primaria non è esente da elementi di ambiguità (primo fra tutti la non considerazione dei diversi livelli di complessità dei compiti, che dovrebbe essere elemento importante nella valutazione dei livelli di competenza raggiunti dagli studenti). D’altro canto, una simile innovazione avrebbe richiesto un’adeguata e una sistematica e tempestiva azione di monitoraggio.

stanzialmente rigidi che caratterizzano la nostra scuola (tanto più rigidi quanto più si passa dalla scuola primaria alla scuola secondaria di primo e poi di secondo grado).

Nell'ottica della costruzione e dello sviluppo delle competenze, la valutazione si caratterizza come un'attività di tipo essenzialmente interpretativo. Questo non significa mettere in discussione la necessità di disporre di "evidenze" sufficientemente attendibili e valide o di contrapporre l'attività di interpretazione a procedure rigorose (anche di misurazione). Al contrario, qualsiasi interpretazione sufficientemente fondata non può non fare affidamento su dati valutativi affidabili e attendibili. È questo il solo modo per poter affrontare e farsi carico della inevitabile "incertezza" che accompagna qualsiasi intervento educativo in tutti i suoi momenti (quello che Lawrence Stenhouse chiamava il "peso dell'incertezza"¹²) e che – allo stesso tempo – rende possibile il confronto tra colleghi che consente di arrivare ad esprimere giudizi valutativi il più possibile condivisi.

Farsi carico dell'inevitabile "incertezza" che accompagna ogni intervento educativo ripropone l'esigenza di una riflessione continua e sistematica sulle pratiche didattiche e valutative. La valutazione, in un'ottica di questo tipo, non riguarda solo gli 'apprendimenti' degli studenti, ma anche i processi didattici messi in atto per suscitargli e le stesse procedure utilizzate per valutarli.

All'interno di un'ottica di questo tipo, la valutazione si caratterizza essa stessa come "apprendimento". Per gli studenti, perché li aiuta a riflettere sui loro processi di apprendimento e sui risultati che conseguono. Per gli insegnanti, perché li aiuta a considerare l'adeguatezza e l'efficacia delle pratiche educative e didattiche messe in atto per aiutare gli studenti a costruire e sviluppare il proprio apprendimento.

2. Agire tra conformità e tradimenti

Chi insegnerà dovrà operare in contesti contrassegnati dalla propensione del nostro sistema educativo a presentare contemporaneamente dispositivi normativi fondati dal punto di vista pedagogico e docimologico e altri che, come vedremo, tradiscono (anche in maniera piuttosto grossolana) quanto acquisito dalla prassi e dalla ricerca valutativa. Orientarsi in un simile scenario senza aver sviluppato uno sguardo consapevole e critico nei confronti dei dispositivi valutativi realmente operanti può risolversi nella tendenza a considerare la valutazione una mera incombenza da espletare allo scopo di fotografare la realtà per riprodurla e non per conferire a essa il valore necessario a trasformarla.

La formazione di competenze valutative non può dunque che legare dialetticamente quanto elaborato dallo sviluppo della ricerca valutativa negli ultimi decenni con elementi che caratterizzeranno le concrete esperienze scolastiche. Da questo punto di vista, tanto la conformità quanto l'incoerenza tra approcci, strumenti e prassi operanti nei contesti e presenti nella normativa e concetti come quelli di validità e competenze rappresentano preziose opportunità formative. Un'ottima occasione per analizzare elementi di conformità e di incoerenza è fornita dalle prove INVALSI. Da un lato, esse possono essere a buon diritto considerate un valido

¹² Si veda L. Stenhouse, *An introduction to Curriculum Research and Development*, London, Heinemann, 1975.

strumento per la valutazione di sistema, rappresentando un elemento fondamentale per ottenere stime della diffusione nella popolazione delle conoscenze e delle abilità testate. Da oltre mezzo secolo (dalle prime indagini IEA alle attuali rilevazioni condotte dalla stessa IEA e da OCSE e INVALSI¹³) è attraverso l'impiego di prove standardizzate che riusciamo a ottenere informazioni valide e affidabili non solo rispetto ai risultati rilevati, ma anche ai loro legami con variabili di sfondo, come il retroterra sociale, economico e culturale di studentesse e studenti o il contesto nel quale sono inseriti i diversi istituti.

Se dal punto di vista della valutazione di sistema appare davvero inconcepibile rinunciare al patrimonio conoscitivo rappresentato da indagini come quelle condotte dall'INVALSI, va segnalato anche come, da una quindicina di anni, si sia progressivamente affermata la tendenza a considerare le prove nazionali misura non solo delle conoscenze e delle abilità effettivamente testate, ma anche delle competenze. In aggiunta a ciò, le prove vengono impiegate per misurare gli apprendimenti di ogni singolo individuo (e, ancor più recentemente, la sua fragilità) e, a fini rendicontativi (DPR 80/2013), l'efficacia di ciascun istituto. Si tratta di scelte del tutto incoerenti con quanto indicato dalla letteratura di riferimento.

Partiamo dalla prima scelta, ovvero da quella relativa all'impiego delle prove INVALSI per valutare competenze. Se, come abbiamo visto nel paragrafo precedente, la competenza è caratterizzata da dimensioni cognitive, metacognitive, affettivo-comportamentali, valoriali e dal riferimento a situazioni inserite in contesti reali, è del tutto inappropriato ritenere che delle prove standardizzate possano offrire misure sufficienti a valutarle. Le prove INVALSI, non diversamente da quelle utilizzate per le indagini OCSE o IEA, offrono informazioni rispetto a conoscenze e abilità: ritenere che rilevino competenze è un portato dell'erronea tendenza a considerare la competenza mera somma di alcune conoscenze e alcune abilità.

Anche la seconda scelta, relativa all'uso delle prove come misura dell'apprendimento di singoli individui, appare del tutto inappropriata. Infatti, questi strumenti nascono per stimare la distribuzione di conoscenze e abilità su gruppi più o meno ampi di soggetti (nazioni, regioni e, entro certi limiti, singoli istituti). Quanto più tali gruppi sono estesi, tanto più le stime che offrono sono al riparo del margine d'errore che inevitabilmente caratterizza ogni misurazione. Al contrario, se lo scopo è quello di valutare gli apprendimenti in Italiano o Matematica di un singolo studente, piuttosto che basare l'accertamento su simili strumenti è necessario fare ricorso anche a strumenti di differente natura¹⁴, in modo da triangolare fonti diverse, stimolare diverse intelligenze e ridurre il margine d'errore.

Infine, la terza scelta impone a ciascun istituto di impiegare i risultati ottenuti alle prove INVALSI a fini di valutazione interna e *accountability*, in modo da rendere conto a sistema, famiglie e territorio dell'efficacia delle azioni intraprese. Questo però, come vedremo, incide negativamente sulla validità di costruito e sulla validità delle conseguenze delle misure. L'INVALSI fornisce a ogni istituto l'Effetto Scuola (ES), ovvero una stima dell'incidenza netta della scuola sul rendimento di studentesse e studenti alle prove. L'ES viene calcolato tenendo conto del punteggio ottenuto nelle prove precedentemente svolte e di una serie di fattori individuali e

¹³ C. Corsini, *Pregi e limiti delle indagini internazionali su grandi campioni*, in G. Domenici, P. Lucisano, V. Biasi (a cura di), *Ricerca sperimentale e processi valutativi in educazione*, Milano, McGraw Hill Education, 2020, pp. 321-345.

¹⁴ F. Batini, *Insegnare e valutare per competenze*, Torino, Loescher, 2016.

sociali relativi al contesto personale e scolastico (come il retroterra socioeconomico). Il problema è che tale indicatore palesa, esattamente come ogni misura del valore aggiunto¹⁵, lacune nella validità di costrutto¹⁶, ovvero nella coerenza tra risultati e modello teorico di riferimento. Infatti, sebbene l'ES possa essere considerato un'utile misura del miglioramento nei risultati ottenuti alle prove di studentesse e studenti, va osservato che la scelta di incentrare la definizione dell'efficacia di una scuola sulla base di test di apprendimento è di dubbia validità, sia perché non è affatto scontato che i risultati ai test rappresentino adeguatamente l'obiettivo della didattica sia perché su di essi agiscono tutte le variabili esogene che non sono state inserite nella funzione di produzione educativa. E così, anche nel nostro caso sembra ripresentarsi il problema, tipico delle indagini del valore aggiunto, dell'insufficiente coerenza tra costrutto teorico (le caratteristiche di una scuola efficace) e i dati raccolti. Come evidenziato in un lavoro condotto sui dati pugliesi relativi all'ES calcolato da INVALSI¹⁷, le scuole con ES alto mostrano di essere più "selettive" delle scuole con ES basso, che risultano caratterizzate da maggiore eterogeneità di contesto. Sostanzialmente, il rischio (ampiamente prevedibile) è che l'Effetto Scuola non individui le scuole più efficaci ma, al contrario, quelle meno inclusive. D'altro canto, come indicato da Campbell¹⁸, se alle prove vengono associate determinate conseguenze (come effetti di natura reputazionale per dirigenti e docenti), la loro validità viene compromessa, mentre peggiora la qualità dei processi di apprendimento e insegnamento oggetto di rilevazione. Questo può risolversi in comportamenti opportunistici come il *cheating* (imbrogliare), il *teaching to the test* (addestrare alle prove) e lo *skimming* (selezionare in modo da ben figurare alle prove).

Se l'obiettivo è quello di valutare la qualità e l'efficacia delle scuole, è necessario osservare e descrivere i processi organizzativi e didattici interni e usare anche prove di diversa natura (più complesse, ricche e informative di quelle standardizzate) e, più in generale, adottare modelli meno semplicistici¹⁹.

Infine, va anche considerato che il fatto di somministrare le prove all'intera popolazione, reso inevitabile dalla scelta di usarle per la valutazione degli apprendimenti di singoli individui e dell'efficacia di ciascun istituto, minaccia anche la validità di contenuto degli strumenti. Sappiamo che uno strumento di misurazione ha validità di contenuto se copre adeguatamente gli obiettivi oggetto di misurazione. Una prova non mostra validità di contenuto quando una specifica dimensione del costrutto indagato non è adeguatamente rappresentata. Questo avviene quando tale dimensione viene rilevata attraverso un numero insufficiente di quesiti. Se analizziamo i quesiti di comprensione della lettura somministrati dall'INVALSI fino al 2017²⁰ (ultimo

¹⁵ C. Corsini, *Il valore aggiunto in educazione*, Roma, Nuova Cultura, 2009.

¹⁶ L. Cohen, L. Manion, K. Morrison, *Research methods in Education*, London-New York, Routledge-Taylor and Francis, 2007.

¹⁷ L. Perla, V. Vinci, P. Soletti, *L'impatto delle variabili didattiche e organizzative sull'Effetto-Scuola INVALSI: le risultanze pugliesi*, in "Giornale Italiano della Ricerca Educativa", XIV, 26, 2021, pp. 67-87.

¹⁸ D. T. Campbell, *Assessing the impact of planned social change*, in "Evaluation and Program Planning", 2, 1979, pp. 67-90.

¹⁹ B. Creemers, L. Kyriakides, L., *The dynamics of educational effectiveness*, London, Routledge, 2008.

²⁰ C. Corsini, *Accountability e ricerca educativa: una riflessione sulla validità delle prove Invalsi somministrate fino al 2017*, in G. Elia, S. Polenghi, V. Rossini (a cura di), *La scuola tra saperi e valori etico-sociali. Politiche culturali e pratiche educative*, Lecce, Pensa, 2019, pp. 993-1005.

anno in cui tutte le prove vengono pubblicate), notiamo, come evidenziato nella Tabella 1, che gli item relativi a una dimensione considerata fondamentale nel quadro teorico alla base delle prove, ovvero quelli che hanno l'obiettivo di rilevare abilità di riflessione e valutazione partendo dal testo, non hanno un'adeguata copertura (e risultano assenti in gran parte delle prove).

Aspetto	% item
2 Comprendere il significato, letterale e figurato, di parole ed espressioni e riconoscere le relazioni tra parole	13,6%
1 Individuare informazioni date esplicitamente nel testo	19,8%
3 Fare un'inferenza diretta	8,9%
4 Cogliere le relazioni di coesione e di coerenza testuale	11,2%
5a Ricostruire il significato di una parte più o meno estesa del testo	25,3%
5b Ricostruire il significato globale del testo	10,7%
6 Sviluppare un'interpretazione del testo	9,5%
7 Riflettere sul testo e valutarne il contenuto e/o la forma alla luce delle esperienze personali	0,8%
Totale	100,0%

Tabella 1 – Aspetti della comprensione del testo (Quadro di Riferimento INVALSI per Italiano) e loro rappresentazione nelle prove somministrate tra il 2010 e il 2017. Il legame tra aspetto e item è quello indicato dall'INVALSI nella Guida alle prove.

Somministrare all'intera popolazione rende difficile impiegare quesiti a risposta aperta e complessa, la cui valutazione, a differenza dei quesiti a risposta chiusa o aperta e univoca, farebbe lievitare moltissimo i costi dell'operazione. Senza tali quesiti vengono però a mancare dimensioni essenziali indicate nel Quadro di Riferimento, dato che pensare di rilevare informazioni valide sulla capacità di riflettere e valutare un testo senza chiedere a studentesse e studenti di argomentare le proprie risposte è davvero poco realistico. Se osserviamo l'evoluzione metodologica delle indagini su vasta scala sulla comprensione della lettura condotte da IEA (PIRLS) e OCSE (PISA), notiamo come da quasi un quarto di secolo è un dato acquisito il fatto che a tali fini i quesiti a scelta multipla forniscono indicazioni molto meno valide rispetto a quelle garantite dall'impiego di domande a risposta aperta e complessa. Infatti, prendendo in considerazione indagini che, a differenza delle prove INVALSI, presentano un'adeguata validità di contenuto, rileviamo che gran parte delle domande relative ai processi di riflessione e valutazione sono a risposta aperta e complessa (il 73% nell'edizione 2009 dell'OCSE-PISA). L'insufficiente presenza (o la totale assenza) di domande simili incide negativamente sulla possibilità di rappresentare adeguatamente il costrutto di riferimento e tale lacuna rende ancor più inappropriata la scelta di impiegare le prove INVALSI per valutare competenze.

3. Considerazioni conclusive

Riteniamo che, nella formazione docenti, sia estremamente utile lavorare sulle incoerenze esistenti tra visioni valutative pedagogicamente e docimologicamente fondate e scelte di politica scolastica che rappresentano una conferma su larga scala della legge di Campbell sugli effetti corruttivi di misure che da mezzi divengono fini. Un simile lavoro, incentrato sull'impiego appropriato di ciò che viene effettivamente misurato attraverso le rilevazioni nazionali, consente di sviluppare competenze progettuali e valutative. Impiegare le prove INVALSI e le loro interpretazioni come strumento formativo (facendo lavorare futuri docenti sui rapporti tra singoli item, dimensioni dei modelli teorici di riferimento, finalità delle somministrazioni, analisi e interpretazione dei dati) conferisce significatività all'apprendimento di concetti come quelli di validità del contenuto, del costruito, delle conseguenze di uno strumento di misurazione e può orientare efficacemente nelle scelte da compiere nella valutazione delle competenze.

In questa prospettiva, le ricerche che integrano rigore metodologico e studio sul campo con una chiara impostazione critico-riflessiva possono davvero rappresentare un prezioso strumento formativo, restituendo lacune e punti di forza dei dispositivi indagati. Limitandoci all'ambito rendicontativo, citiamo a titolo esemplificativo lavori come quelli di Robasto²¹ sulle potenzialità e i limiti dei Piani di Miglioramento e di Landri²² sul punto di vista gruppi di docenti impegnati nei processi di *accountability*. Concludiamo rilevando che proprio l'emersione della prospettiva docente rappresenta un elemento irrinunciabile per la formazione di competenze valutative. L'alternativa è quella di razzolare diversamente da come si predica, senza coinvolgere attivamente corsiste e corsisti nella valutazione dell'efficacia della didattica e rinunciando così a elementi utili a orientare le scelte successive, col rischio di riprodurre gli stessi errori che hanno caratterizzato le precedenti esperienze formative.

4. Bibliografia di riferimento

Batini F., *Insegnare e valutare per competenze*, Torino, Loescher, 2016.

Becchi E., Visalberghi A., *Introduzione*, in E. Becchi, A. Visalberghi (a cura di), *Manuale critico della sperimentazione e della ricerca educativa*, Milano, FrancoAngeli, 1984, pp.11-33.

Benadusi L., Molina S. (a cura di), *Le competenze. Una mappa per orientarsi*, Bologna, il Mulino, 2018.

Bottani, N., *Il difficile rapporto tra politica e ricerca scientifica sui sistemi scolastici*, Programma Education FGA, Working Paper n. 17, Fondazione Giovanni Agnelli, 2009.

Bourdieu P., Passeron J.C., *La reproduction. Eléments pour une théorie du système d'enseignement*, Paris, Les Editions de Minuit, 1970; tr. it. *La riproduzione. Per una teoria dei sistemi di insegnamento*, Rimini-Firenze, Guaraldi, 1972.

Campbell D.T., *Assessing the impact of planned social change*, in "Evaluation and Program Planning", 2, 1979, pp. 67-90.

²¹ D. Robasto, *Autovalutazione e piani di miglioramento a scuola*, Roma, Carocci, 2017.

²² P. Landri, *To resist, or to align? The enactment of data-based school governance in Italy*, in "Educational Assessment, Evaluation and Accountability", 33, 3, 2021, pp. 563-580.

Cohen L., Manion L., Morrison K., *Research methods in Education*, London-New York, Routledge-Taylor and Francis, 2007.

Corsini C., *Il valore aggiunto in educazione*, Roma, Nuova Cultura, 2009.

Corsini C., *Accountability e ricerca educativa: una riflessione sulla validità delle prove Invalsi somministrate fino al 2017*, in G. Elia, S. Polenghi, V. Rossini (a cura di), *La scuola tra saperi e valori etico-sociali. Politiche culturali e pratiche educative*, Lecce, Pensa, 2019, pp. 993-1005.

Corsini C., *Pregi e limiti delle indagini internazionali su grandi campioni*, in G. Domenici, P. Lucisano, V. Biasi (a cura di), *Ricerca sperimentale e processi valutativi in educazione*, Milano, McGraw Hill Education (Italy), 2020, pp. 321-345.

Corsini C., Losito, B. *Le rilevazioni Invalsi: a che cosa servono?*, in "Cadmò", XXI, 2, 2013, pp. 55-76.

Creemers B., Kyriakides L., *The dynamics of educational effectiveness*, London, Routledge, 2008.

Gattullo M., *Didattica e docimologia*, Roma, Armando, 1967.

Landri P., *To resist, or to align? The enactment of data-based school governance in Italy*, in "Educational Assessment, Evaluation and Accountability", 33, 3, 2021, pp. 563-580.

Pellerey M., *Le competenze individuali e il portfolio*, Milano, RCS, 2004.

Perla L., Vinci V., Soletti P., *L'impatto delle variabili didattiche e organizzative sull'Effetto-Scuola INVALSI: le risultanze pugliesi*, in "Giornale Italiano della Ricerca Educativa", XIV, 26, 2021, pp. 67-87.

Robasto D., *Autovalutazione e piani di miglioramento a scuola*, Roma, Carocci, 2017.

Stenhouse L., *An introduction to Curriculum Research and Development*, London, Heinemann, 1975.

Visalberghi A., *Misurazione e valutazione nel processo educativo*, Milano, Edizioni di Comunità, 1955.

Wiggins G., *Educative Assessment: Designing Assessments to Inform and Improve Student Performance*, San Francisco, Jossey-Bass, 1998.

Data di ricezione dell'articolo: 17 marzo 2023

Date di ricezione degli esiti del referaggio in doppio cieco: 11 e 23 aprile 2023

Data di accettazione definitiva dell'articolo: 10 maggio 2023