

Criticità e potenzialità nella costruzione di un sistema nazionale di valutazione

Antonio Calvani^a

^a *Università degli Studi di Firenze*, antonio.calvani@unifi.it

Abstract

Realizzare e gestire un buon sistema nazionale di valutazione (SNV) delle scuole rappresenta una sfida di grande complessità con la quale si confronta la maggior parte dei Paesi dell'area OECD, proponendo soluzioni di diversa efficacia.

In questo lavoro si prova a rispondere sinteticamente a due domande: quali sono le criticità principali che un'operazione del genere dovrebbe evitare e quali sono le azioni più importanti da compiere, in particolare nel contesto italiano. Si aggiunge un esempio indicativo della strada che occorrerebbe seguire. A giudizio dell'autore occorre mettere al centro la natura formativa della valutazione, cioè la possibilità di fornire alle scuole un feedback immediato, concretamente spendibile per un miglioramento constatabile dagli stessi soggetti interessati, senza il quale la valutazione continuerà a rappresentare un'operazione percepita come estranea e invasiva della sfera personale dell'insegnante.

Parole chiave: sistemi di valutazione delle scuole; valutazione formativa; EBE; Invalsi.

Abstract

Creating and managing a National System for Evaluating school is a complex problem the majority of OECD countries is nowadays facing with variously effective solutions. The aim of this work is to answer two questions: which are the main critical points to be avoided and the main important actions to be performed, particularly inside the Italian context. An example to show the direction to be followed is added.

According to the author it is necessary to put in evidence the formative nature of the evaluation, that is the possibility of providing schools with an immediate feedback, useful to demonstrate to the subjects involved how the improvement can be achieved. On the contrary the evaluation will remain perceived by the teachers as an invasive intrusion inside the personal sphere of the teacher.

Keywords: evaluating school system; formative evaluation; EBE; Invalsi - Italian national evaluating school system.

1. Premessa

La recente direttiva MIUR di settembre 2014¹ attiva un complesso processo che intende integrare autovalutazione ed eterovalutazione, interventi sulla valutazione della dirigenza scolastica, sulle rilevazioni nazionali, per arrivare, infine ad una valutazione di sistema.

In questo lavoro intendiamo avanzare alcune proposte relative all'attuazione di un Sistema Nazionale di Valutazione (SNV) nell'ottica di dare alla valutazione il ruolo che le spetta come elemento essenziale per l'apprendimento².

Riprendendo una linea di riflessione che tende ad estendere l'approccio dell'evidence based education (EBE) all'analisi dei sistemi di valutazione (Calvani & Vivanet, 2014), riteniamo che una corretta comprensione ed applicazione del concetto di valutazione formativa (Scriven, 1967) rappresenti il dato più rilevante, del resto già in forte risalto nelle esperienze internazionali, di particolare importanza per il contesto italiano, dove ci si scontra anche con particolari "diffidenze" verso ogni forma di valutazione³.

Il lavoro si suddivide in tre parti. Nella prima ci si chiede quali siano le criticità fondamentali che emergono dalla letteratura internazionale, nella seconda quali avvertenze si dovrebbero mettere al centro in particolare nel contesto italiano, nella terza si presenta un esempio operativo, attinto da una esperienza neozelandese, di gestione di un processo di valutazione formativa in accordo con le Indicazioni Nazionali (MIUR, 2012), attraverso un sistema informatico. Una sintesi delle criticità e delle azioni da tenere soprattutto presenti per il sistema italiano è riportata in Figura 1.

2. Quali sono le criticità più comuni in cui si imbattono i SNV?

Come noto, il tema delle politiche della valutazione degli apprendimenti scolastici negli ultimi anni ha ricevuto un'enfasi crescente accompagnandosi ad una rilevante mole di quadri di riferimento e guideline⁴. Per una efficace sintesi sulla responsabilità di chi

1

http://www.istruzione.it/allegati/2014/DIRETTIVA_SISTEMA_NAZIONALE_DI_VALUTAZIONE.pdf.

² Per la documentazione internazionale ci si avvale del contributo di Vivanet (2014b).

³ Non è intento di questo lavoro affrontare qui un'analisi storica e critica dettagliata delle iniziative che MIUR e Invalsi hanno attuato e stanno attuando, alcune delle quali anche pregevoli. Mentre il lavoro è in attesa di stampa è stata pubblicata la lista degli indicatori, un quadro che a prima vista si presenta organicamente strutturato. Tutto ciò nulla toglie alla rilevanza delle criticità, relative alla sottovalutazione del carico gestionale dell'operazione, alla non chiara indicazione su come l'informazione sarà spesa per consentire effettivi avanzamenti degli apprendimenti scolastici, superando così il distacco psicologico tra le scuole e il sistema di valutazione stesso, aspetti su cui questo lavoro vuol richiamare l'attenzione, fornendo anche qualche piccolo esempio propositivo, alla luce di quanto già esiste nel contesto internazionale.

⁴ I lineamenti quadro sottolineano in particolare la necessità di dare attenzione ad un approccio olistico, allineare il sistema di valutazione con gli obiettivi di apprendimento, mettere al centro lo studente, mirare la valutazione al miglioramento delle pratiche didattiche (focus su valutazione formativa e didattica efficace), rafforzare la professionalità del docente, sviluppare competenze di

valuta, dell'oggetto della valutazione, degli obiettivi della valutazione e delle modalità della valutazione si rimanda a Vivanet (2014b). Limitandoci agli obiettivi della valutazione in un SNV, possiamo individuarne due principali: la rendicontazione (accountability) e lo sviluppo (development)⁵. Nel perseguimento di tali funzioni numerose sono le criticità e i vincoli in cui i sistemi si imbattono e che devono dunque essere aggirati. Per comodità li raccoglieremo in tre macro categorie: (i) implementazione e gestione, (ii) distorsioni (ideologiche e culturali) e (iii) carenze metodologiche e formative. Le prime due sono prevalentemente connesse al primo aspetto (accountability), la terza al secondo (development).

2.1. Implementazione e gestione

La letteratura ha messo in evidenza l'onere complessivo delle rilevazioni per un sistema scolastico su scala nazionale, il complesso apparato burocratico-gestionale che esso richiede, le ingenti risorse economiche necessarie per la raccolta e la successiva analisi dei dati, la sfasatura dei tempi tra raccolta e consegna dei risultati, le difficoltà tecniche legate alla gestione informatizzata e alla loro accessibilità da parte di utenti diversi (pubblico, scuole, famiglie), le difficoltà operative e carenze professionali da parte dei soggetti (docenti, dirigenti scolastici) (McNeil, 2000; Toch, 2006).

Una particolare attenzione va prestata alle modalità tecniche della raccolta dei dati e del loro utilizzo: un problema tipico riguarda il fenomeno dell'accumulo di dati "inerti" (Hattie, 2013; Popham, 2008; Reeves, 2008). Si reputa ingenuamente che più dati si hanno, migliore sarà la valutazione; la paura di non essere esaustivi spinge spesso i compilatori degli indicatori ad aggiungere continuamente ulteriori richieste, anziché limitare l'attenzione solo a pochi essenziali indicatori, facilmente verificabili e ben corrispondenti a precisi processi di cambiamento già ipotizzabili. In questo modo, una volta raccolti i dati, non si sa poi che farne e il sistema tende a ingolfarsi.

2.2. Distorsioni (ideologiche e culturali)

Possiamo segnalarne alcune, tra cui soprattutto:

1. la pressione esercitata sui sistemi scolastici dal dibattito pubblico-politico che accompagna le rilevazioni comparative internazionali (Eurydice Network, 2009; Hopkins, Pennock, Ritzen, Ahtaridou & Zimmer, 2008) tende a generare forme di competitività e di tutela di immagine a livello politico nazionale ed internazionale ("sindrome OECD PISA", cfr. Fondazione Giovanni Agnelli, 2014). Ciò comporta il rischio di stravolgere il vero fine dell'educazione, concentrandosi sul

valutazione, prevenire/evitare distorsioni (teach to the test, cheating), favorire la credibilità della valutazione nella scuola, valorizzare la leadership, attuare politiche che informino adeguatamente sul miglioramento del sistema, gestire e dare risposte ai bisogni locali e speciali, progettare con attenzione, creando consenso e partecipazione (OECD, 2013). Per una discussione più dettagliata si veda Vivanet (2014b).

⁵ Per accountability possiamo intendere l'acquisire conoscenze aggiornate dello stato del sistema scuola per certificarle (agli occhi del pubblico, del ministero, in un contesto internazionale); per development si intende l'insieme delle azioni volte a favorire i processi di miglioramenti a partire dalla scuola stessa rispettando anche le diversità (per questo aspetto nel proseguo del lavoro ci soffermeremo solo su alcune componenti, quelle legate agli atteggiamenti e conoscenze nei riguardi della valutazione).

risultato a scapito del processo e sull'applicazione di logiche premianti o penalizzanti, anche in termini economici, per i soggetti coinvolti (scuole, insegnanti) (Wang, Beckett & Brown, 2006);

2. i sistemi standardizzati possono favorire forme di diseguità: ci si focalizza verso competenze di lettura, scrittura, matematica e scienze con la marginalizzazione di discipline/competenze difficilmente testabili quali ad esempio dimensioni estetiche, ludiche, socio-relazionali (EPPI, 2002; Guilfoyle, 2006; Hout & Elliott, 2011; Phelps, 2009). A fianco di ciò, c'è il rischio rappresentato dalla possibile trascuratezza verso ogni tipo di differenza (bisogni educativi speciali, diversità linguistiche, culturali, sociali)⁶;
3. il cheating (il barare) è un altro fenomeno diffuso, ancora più forte laddove i risultati sono presi in considerazione per la presa di decisioni importanti, ad esempio per le carriere degli insegnanti e/o degli studenti (Wang, Beckett & Brown, 2006) (aspetto sui cui torneremo);
4. il teaching to the test, cioè una subordinazione del curriculum al conseguimento stretto degli obiettivi definiti dagli standard da conseguire, è un'altra delle comuni "distorsioni"; in tal modo rilevanti opportunità educative che possono insorgere anche spontaneamente nel percorso formativo, possono essere abbandonate unicamente perché non troveranno riscontro nella prova finale (Popham, 2001).

2.3. Carenze metodologico-formative

Ci riferiamo in particolare alle criticità legate a cattivi atteggiamenti verso la valutazione e al non aver chiaro come il momento valutativo possa diventare un punto di forza per il miglioramento. Un ruolo particolare lo svolge la scarsa conoscenza tecnica della meccanica della valutazione e di come anche le critiche, pur legittime, al sistema del testing possano essere aggirate⁷. La valutazione è percepita in sostanza come un momento estraneo all'azione didattica. Si tratta di mettere al centro l'importanza della valutazione formativa (Popham, 2011), aspetto che sta conseguendo sempre maggiore attenzione all'interno delle politiche internazionali⁸, anche se su questo tema la stessa OECD (2013) sottolinea che a fronte di politiche e dichiarazioni esplicite nelle normative, si riscontra ancora una diffusa genericità nella specificazione di pratiche efficaci e sistematiche all'interno delle scuole e/o classi. Investire su strategie di valutazione formativa impone un forte impegno in primo luogo nella professionalizzazione di dirigenti scolastici e insegnanti.

⁶ Con la possibilità che la standardizzazione della valutazione, applicata su scala nazionale, tenda ad ignorare questi aspetti anche per questioni di immagine politica dei risultati (Greaney & Kellaghan, 1995; OECD, 2012).

⁷ Su questo aspetto vedi in particolare il contributo di Trinchero (2014).

⁸ In Australia, Corea, Spagna e Messico lo sviluppo di competenze di valutazione formativa è parte obbligatoria del curriculum di formazione iniziale degli insegnanti. In Corea e Spagna è anche obbligatorio l'aggiornamento professionale in questo ambito. Esempi di importanti programmi di formazione/aggiornamento professionale rivolti agli insegnanti per lo sviluppo di competenze di valutazione formativa sono stati avviati anche in Canada, Irlanda, Norvegia (OECD, 2013).

3. Il contesto italiano: quali avvertenze?

Prendendo in considerazione il contesto italiano, quali considerazioni possiamo avanzare? Ci sono almeno due aspetti da cui non si dovrebbe prescindere:

- il primo riguarda la condizione economica e politica sicuramente non facile che l'Italia sta attraversando, il che dovrebbe portare, di per sé, a diffidare di piani di lunga durata. Del resto, qualunque operazione di valutazione, per essere efficace, sia a livello micro che macro, richiede una contrazione dei tempi: il ciclo raccolta del dato-comunicazione agli interessati, miglioramento, deve essere veloce, pena la inefficacia, un punto sul quale convergono da un lato il mondo della didattica efficace e della valutazione formativa, dall'altro quello del Total Quality Management⁹. È importante dunque ridurre al minimo il rischio dell'insuccesso derivante da sovraccarico di dati e della dilatazione dei tempi portando a livello di pratica gestibilità ed efficienza la tempistica dell'intervento (ottimizzare l'implementazione);
- il secondo aspetto riguarda il fatto che il contesto italiano, verosimilmente più di altri¹⁰, si caratterizza per una radicata diffidenza degli insegnanti e delle scuole verso la valutazione (Fondazione Giovanni Agnelli, 2014), vissuta a livello emozionale e ideologico, come un'intrusione in uno spazio personale e una possibile minaccia alla propria autostima¹¹. L'esperienza mostra come ogni richiesta di autovalutazione è destinata a riprodurre i ben noti rituali di occultamento, in senso autogiustificativo ("non abbiamo mezzi...") o in senso autocelebrativo ("tutto è andato bene, i risultati sono stati raggiunti")¹². Le richieste avanzate alle scuole secondo quanto indicato dalla Direttiva, quali la stesura RAV, gli audit a campione, la somministrazione di questionari standardizzati, sono operazioni comunque imposte dall'esterno, che verosimilmente, rischiano di consolidare questo rapporto di estraneità e diffidenza. In questo contesto diventa prioritario agire per migliorare l'atteggiamento delle scuole verso la valutazione (superare la diffidenza verso la valutazione).

Passando alle azioni per il primo dei due aspetti, sembra ragionevole allora:

⁹ Da questo punto di vista, dalla lettura della Direttiva di settembre 2014, si ricava la sensazione di una sottostima del carico informativo che un sistema del genere va a richiedere, dei tempi (anche tecnici) per la messa in piedi del sistema informativo per la raccolta ed elaborazione dei dati.

¹⁰ Non si tratta ovviamente di un fenomeno solo italiano. Sono fenomeni di distorsione del comportamento che, in accordo anche alle recenti indicazioni dell'OECD (2013), possono essere prevenuti sviluppando adeguate strategie di coinvolgimento delle scuole e degli insegnanti (Milanowski & Heneman, 2001), anche attraverso una definizione e comunicazione chiara degli obiettivi (Olsen, 1989).

¹¹ La valutazione è stata anche oggetto di un gioco di fraintendimenti di vario tipo, che hanno caratterizzato le vicende degli ultimi 15 anni sul piano delle politiche educative (Barzanò & Grimaldi, 2013).

¹² Il cheating, in tal senso, è una delle espressioni più tipiche di questo distacco (Greaney & Kellaghan, 2008; Koretz, 2010), presentandosi non solo in situazioni di testing in aula, ma in ogni situazione in cui si chiede al soggetto una autovalutazione su quello che è stato o sarà il proprio operato (trovando peraltro sponda in un certo genericismo proprio del linguaggio didattico).

- Adottare un approccio campionario anziché censitario.

Questa strada è da preferire per ridurre ad una scala sostenibile la dimensione dei costi mantenendo affidabilità dei dati, come già indicato da Vertecchi (2014). Un approccio campionario, regolabile secondo le necessità, rispetto a quello censuario, serve per individuare il giusto punto d'incontro tra risorse disponibili e qualità del risultato: una adeguata strategia di campionamento può consentire un quadro affidabile complessivo dello stato del sistema scolastico italiano e delle criticità e problematiche che lo caratterizzano con considerevole risparmio di energie.

- Allestire e testare un sistema telematico efficiente per la raccolta dei dati e limitare gli indicatori solo funzionali ad obiettivi strettamente operazionalizzabili.

Senza un buon sistema informatico di supporto qualunque operazione di dimensione nazionale rimane ingestibile (Vivanet, 2014b). La sua messa a punto richiede una buona progettazione ed un collaudo preliminare con un lavoro fianco a fianco tra tecnici e progettisti del sistema di valutazione. Nella definizione di indicatori o descrittori di risultato si dovrebbe vincolare la richiesta dei dati a specifici obiettivi operazionalizzabili, a cui cioè devono corrispondere chiari e riscontrabili criteri di verifica o di performance. Ciò ha come effetto, oltre ad alleggerire il sistema con rilevante abbattimento del fenomeno dei “dati inerti”, quello di favorire una cultura più seria della valutazione mettendo al bando le retoriche elusive cui abbiamo accennato, che trovano ampio spazio quando si offrono campi descrittivi aperti¹³.

- Investire in un impianto docimologico preliminare.

La messa a punto di un set di strumenti di valutazione coerenti con i traguardi riportati nelle Indicazioni Nazionali (MIUR, 2012) dovrebbe essere un'operazione preliminare, aspetto su cui torneremo¹⁴.

A fronte del secondo aspetto sembra ragionevole dare prioritaria importanza a

- Formare alla definizione di obiettivi chiari e operazionalizzati.

Sono sotto gli occhi di tutti le difficoltà ed ambiguità linguistiche che avvolgono la didattica ordinaria. Basta fare un esame del linguaggio impiegato dalle scuole per i POF, o in qualunque progetto didattico (quasi mai rendicontato) per rendersene conto. Senza una formazione preliminare sul linguaggio, natura degli obiettivi e criteri di operazionalizzazione da compiere in primo luogo sui dirigenti scolastici, rimane assai facile prevedere che si riprodurranno i noti frasari, ai quali in verità la genericità del lessico didattico diffuso, offre facili sponde.

- Guidare le scuole alla definizione di obiettivi e criteri di valutazione con riscontro diretto dei risultati.

¹³ L'esperienza in questi anni all'interno dell'Anvur per la scheda SUA richiesta alle università è illuminante circa la complessità di definizione di un format di raccolta e gestione dei dati realmente gestibile e fruibile a vantaggio delle università stesse: da queste è pervenuta nel corso degli anni una costante richiesta di semplificazione dei dati richiesti.

¹⁴ Anche su questo aspetto rimandiamo al lavoro di Vivanet (2014b). Non ci riferiamo solo a prove a scelta multipla ma anche a produzione aperta e prove complesse (con modelli di valutazione a rubrica) ed in particolare ad esempi e modelli volti a semplificare il grado di difficoltà o ad accrescerlo.

L'unica strada possibile per superare la diffidenza verso la valutazione è quella di riuscire a dimostrare in modo convincente che essa serve davvero. Qualunque valutazione, a livello micro o macro, deve saper dare un buon feedback, cioè un'informazione chiara che fa capire a che punto si trova l'interessato, gli ricorda quale è l'obiettivo da raggiungere e gli indica quale è l'azione successiva da fare per raggiungerlo (livello micro, Hattie, 2011); alla stessa stregua ciò ha valore a livello di sistema (livello macro, Total Quality Management); in ogni caso la valutazione serve per far capire subito cosa fare per migliorare in vista di un traguardo; se il sistema di valutazione non è incisivo su questo, fallisce.

Occorre responsabilizzare, entro limiti circoscritti, le scuole al conseguimento di precisi obiettivi, facilmente constatabili, abbinati a prove di performance che non rimangano relegate a generiche dichiarazioni di intenti o alla valutazione soggettiva del docente interessato. La libertà della progettazione e valutazione è un puro inganno se la scuola non è adeguatamente supportata da uno scaffold di repertori adeguati a cui possa appoggiarsi. Il sistema deve offrire un set di modelli e strumenti di valutazione e di intervento possibili, consentendo alle scuole di scegliere tra essi e di adattarli alle specificità contestuali nel corso del processo.

- Valorizzare le scuole attraverso l'offerta di prove on demand.

Fondamentale è ricorrere a prove formative on demand, anche di livello complesso, tratte da una banca dati nazionale; le prove dovrebbero essere mirate a verificare il raggiungimento di obiettivi standard, ma anche a valorizzare le differenze sia nel senso della semplificazione sia nel senso dell'eccellenza, con libertà alle scuole di gestire tempi e modi di applicarle. Vanno decisamente valorizzate le scuole che accettano prove sfidanti ed audit.

- Predisporre modelli di intervento didattico basati su evidenze, per dimostrare come si superano praticamente le difficoltà.

Alla base di tutto, c'è il quesito cruciale su che cosa si voglia/possa fare dei dati, una volta acquisite le conoscenze sullo stato degli apprendimenti. Si tratta di un nesso fondamentale, senza il quale l'intero sistema di valutazione appare vacuo. Come sottolinea anche l'OECD (2013), un punto fondamentale è legare la valutazione al miglioramento delle pratiche didattiche, basate su evidenze di efficacia.

Qui merita ricordare che la ricerca didattica ha fatto rilevanti avanzamenti negli ultimi anni e che i centri legati alle reti internazionali che si rifanno EBE forniscono ormai dati affidabili su metodologie, tecniche e modelli didattici risultati efficaci (Calvani, 2012; Hattie & Anderman, 2013; Mitchell, 2014; Slavin, 2002)¹⁵.

¹⁵ Per un primo approfondimento si consigliano i numeri 2, 3 e 4 di "Form@re" nell'anno 2014, dedicati all'EBE ed il lavoro di Vivanet (2014a).

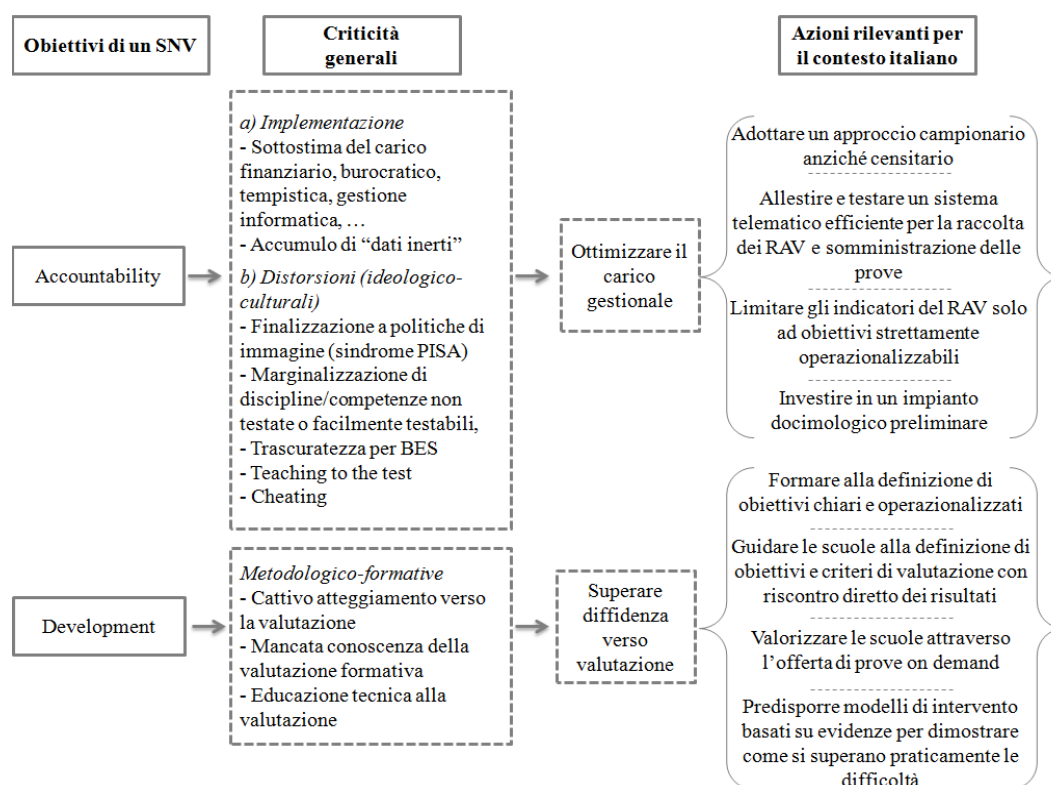


Figura 1. Criticità generali e azioni rilevanti per il contesto italiano.

4. Un esempio

Concludiamo aggiungendo un esempio su come il sistema potrebbe operare. Un buon modello tra i sistemi esistenti volti a consentire processi di valutazione formativa online può essere considerato il sistema neozelandese e-asTTle (<http://e-asttle.tki.org.nz/>), già illustrato (Calvani & Vivanet, 2014)¹⁶. Tale modello, avvalendosi di un sistema informatico, consente di fornire in tempo reale un feedback allo studente distinguendo gli obiettivi raggiunti (risultati in linea coi traguardi nazionali), le lacune (obiettivi che avrebbero dovuto essere raggiunti ma che non lo sono), i punti di forza (aspetti in cui si è andati oltre gli standard nazionali), ed obiettivi da raggiungere (obiettivi non ancora raggiunti ma in cui si è in pari con gli andamenti nazionali).

Nella scheda di Figura 2 abbiamo adattato il modello suddetto ai traguardi per lo sviluppo delle competenze al termine della scuola primaria per la storia, come indicati nelle Indicazioni Nazionali (MIUR, 2012, p. 53)¹⁷. Per ciascuno dei traguardi il sistema può consentire l'applicazione di prove di valutazione applicabili in itinere o alla fine,

¹⁶ Altri sistemi sono indicati nel lavoro di Vivanet (2014b).

¹⁷ Nella parte alta abbiamo riportato alcuni dei traguardi indicati nelle Indicazioni Nazionali (MIUR, 2012), facendo corrispondere esempi di prove (i titoli ai traguardi sono stati aggiunti dallo scrivente per comodità). Nella parte sottostante è esemplificato un esempio ipotetico di feedback che l'alunno potrebbe ricevere, differenziato per punti di forza, obiettivi raggiunti, obiettivi da raggiungere, lacune.

utilizzabili in formato diverso in quanto basate su un repository ampio di item. L'applicazione consente la valutazione con feedback immediato (Figura 3) ed è riproponibile più volte durante il processo. I dati acquisiti dal sistema servono a loro volta a parametrizzare i valori di riferimento¹⁸.

Scheda 1. Esempio di impiego di un sistema informatizzato volto a fornire un feedback in itinere per lo sviluppo delle competenze per la storia nella scuola primaria

TRAGUARDI

Riconoscimento tracce storiche (traguardi 1-2-3)

1. L'alunno riconosce elementi significativi del passato nel suo ambiente di vita.
2. Riconosce ed esplora in modo via via più approfondito le tracce storiche presenti nel territorio e comprende l'importanza del patrimonio artistico e culturale.
3. Usa la fascia del tempo per organizzare informazioni, conoscenze, periodi e individuare successioni, contemporaneità, durate, periodizzazioni.

Prova: il repository dispone di alcune centinaia di immagini (reperti tipici delle varie culture e fasi storiche: vengono pescate randomicamente una decina di immagini e l'alunno deve saperle collocare in modo pur approssimativo su una fascia del tempo.

Relazioni ambiente-uomo (traguardo 4)

4. Individua relazioni tra gruppi umani e contesti spaziali.

Prova: il repository dispone di diverse decine di quesiti a scelta multipla o con associazioni o diagrammi da riempire che chiedono di individuare possibili rapporti causali (del tipo agricoltura e abbandono del nomadismo; carestia e popolazione; aumento produttività e aumento di popolazione; agricoltura e scrittura; città e professioni specializzate). Per il test vengono pescati randomicamente alcuni quesiti.

Tipologie di conoscenze storiche (traguardo 5)

5. Organizza le informazioni e le conoscenze tematizzando ed utilizzando le concettualizzazioni pertinenti (tipologia di conoscenza).

Prova: il repository si avvale di alcune decine di quesiti a scelta multipla in cui si chiede di riconoscere la natura dell'informazione (ad esempio come appartenenza all'ambito economico/culturale/sociale/politico...) oppure selezionando un titolo adeguato ad un breve brano). Per il test vengono pescati randomicamente alcuni quesiti.

Figura 2. Esempio di scheda per una valutazione formativa online.

¹⁸ Un'esperienza di questo tipo, cioè la somministrazione di un test sulla competenza digitale, applicato su un campione nazionale di scuole, è stato condotto dall'Università di Firenze avvalendosi delle funzionalità di sondaggio presenti nell'ambiente Moodle (Calvani, Fini, Ranieri & Picci, 2012). Il testing, svolto in aula informatica, mentre raccoglie i dati utili per una valutazione a livello nazionale, fornisce allo stesso tempo un immediato feedback con suggerimenti operativi alle classi che vi si sottopongono.

Punti di forza Traguardo 1 – Riconoscimento tracce storiche <i>Per questo traguardo i punteggi sono andati oltre l'obiettivo da raggiungere ☺</i>	Obiettivi da raggiungere Traguardo 5 – Tipologie di conoscenze <i>Occorre ancora lavorare per questo traguardo. Si suggeriscono attività di riflessione sulle tipologie di informazione storica (vedi esempi all'URL...).</i>
Obiettivi raggiunti Traguardi 2 e 3 – Riconoscimento tracce storiche <i>Per questi traguardi si sono già raggiunti gli obiettivi ☺</i>	Lacune Traguardo 4 – Relazioni ambiente-uomo <i>Si suggerisce di dedicare particolare attenzione a questo traguardo e alle sue implicazioni concettuali: ci si può avvalere di giochi e simulazioni relativi alla comprensione di meccanismi di causa effetto (vedi esempi all'URL...). Si suggerisce una nuova prova tra 2 mesi.</i>

Figura 3. Esempio di feedback che il sistema informatico potrebbe fornire in tempo reale ad alunni di una scuola primaria, i quali abbiano partecipato ad una valutazione in itinere in rapporto ai traguardi delle competenze nella storia (modello adattato da e-asTTle)¹⁹.

Costruire un kit articolato di prove di valutazione, come quelle a cui qui abbiamo accennato per i traguardi della scuola primaria, rimane alla portata di un lavoro ragionevolmente conseguibile da parte di un team qualificato, e comunque rappresenterebbe un'infinitesima parte del dispendio di energia che gli insegnanti dedicano ogni anno, con esiti per lo più incerti, all'interno della programmazione didattica ordinaria²⁰.

5. Conclusione

La valutazione scolastica rappresenta uno dei temi attualmente più dibattuti nelle politiche europee. In questo lavoro ci siamo chiesti, distinguendo gli obiettivi in un'ottica

¹⁹ Si ipotizza l'applicazione alla fine del quarto anno, in previsione del conseguimento di tutti i traguardi per il quinto anno.

²⁰ Non entriamo qui nel merito di un esame articolato delle diverse tipologie di prove, alcune delle quali non possono essere affrontate con un testing automatico, in particolare in quell'ambito che va sotto il nome di "competenze". Su questo aspetto, e su come altri Paesi si stiano attrezzando cfr. Vivanet (2014b). Laddove non è possibile un feedback automatico si tratta di fornire alle scuole facsimili di prove, e modelli di rubriche; ad es. per un traguardo come "Racconta i fatti studiati e sa produrre semplici testi storici" all'alunno si potrebbe chiedere di ricostruire un episodio accaduto basandosi su indizi e testimonianze, valutando criticamente la loro affidabilità. Il sistema può fornire all'insegnante un sistema di valutazione a gradi, con un repertorio di esempi (del tipo PISA).

di accountability e development, quali sono le criticità principali in cui un SNV si imbatte, e quali azioni sono da raccomandare per il contesto italiano che ha avviato di recente un complesso programma in tal senso. Per quanto riguarda l'ambito della accountability spiccano le problematiche gestionali e finanziarie, con il distacco tra i tempi di rilevazione, restituzione dei dati, ed il rischio di una raccolta di dati inerti. A ciò si legano possibili "distorsioni", da quelle più legate a questioni di difesa dell'immagine, alla messa in secondo piano di discipline meno facilmente testabili, alla trascuratezza verso le problematiche dell'inclusione e delle diversità (Bisogni Educativi Speciali), a fenomeni indotti come il cheating e il teaching to the test. Per quanto riguarda il development, in primo piano possiamo mettere le criticità relative agli atteggiamenti verso la valutazione, alla scarsa conoscenza della sua natura dinamica e formativa.

Nel contesto italiano diventa ancora più cruciale allestire sistemi facilmente gestibili, capaci di mostrare subito di essere efficaci, garantendo una restituzione utile dei dati in tempo reale. Una particolare criticità risiede nella ormai radicata diffidenza verso la valutazione, percepita come una operazione estranea all'apprendimento ed invasiva rispetto alla propria autonomia. La strada da perseguire a questo riguardo è quella di mostrare, con esempi convincenti l'effettiva utilità della valutazione, facendo vedere come possa tradursi in concrete indicazioni di miglioramento.

In questo quadro va messo al centro un piano di formazione dei dirigenti scolastici centrato sulla valutazione formativa e sulle recenti acquisizioni di didattica efficace in ambito evidence based, affiancato da una progettazione ed implementazione di un repertorio di strumenti e modelli per la valutazione da fornire alle scuole. L'Invalsi dovrebbe preoccuparsi di allestire, in raccordo con le Indicazioni Nazionali (MIUR, 2012), un menu di strumenti di varia complessità, per la valutazione di obiettivi/performance da conseguire, che le scuole possano accogliere in toto o da cui possano desumere varianti in funzione delle peculiarità contestuali. La disponibilità di un sistema di interazione informatica in questo contesto gioca un ruolo chiave per dare flessibilità al sistema consentendo un'offerta articolata con frequenti momenti di valutazione formativa.

È anche opportuno lasciare gran parte della valutazione on demand con sistemi di ricompense per scuole che accettano prove sfidanti ed audit. Al momento in cui la valutazione è restituita al soggetto, mostrandogli concretamente cosa può fare e cosa può ricavarne, la diffidenza è destinata a cadere e la valutazione, nella sua accezione formativa, può essere riconosciuta per quello che è, il fattore più importante per favorire l'apprendimento.

Al momento l'Invalsi ha messo in moto un'imponente macchina di rilevazione di dati. Ma è ben chiaro cosa occorre perché non rimangano dati inerti?

Bibliografia

- Barzanò, G., & Grimaldi, E. (2013). «Policy» valutative e contesti di applicazione. Caratteristiche procedurali. *ECPS - Journal of Educational, Cultural and Psychological Studies*, 3(6), 159–189.
- Calvani, A. (2012). *Per un'istruzione evidence based. Analisi teorico-metodologica internazionale sulle didattiche efficaci e inclusive*. Trento: Erickson.

- Calvani, A., Fini, A., Ranieri, M., & Picci, P. (2012). Are Young Generations In Secondary School Digitally Competent? A Study On Italian Teenagers. *Computers and Education*, 58(2), 797–807.
- Calvani, A., & Vivanet, G. (2014). Evidence Based Education e modelli di valutazione formativa per le scuole. *ECPS - Journal of Educational, Cultural and Psychological Studies*, 1(9), 127–146.
- e-asTTle. Online assessment tool. <http://e-asttle.tki.org.nz/> (ver. 30.12.2014).
- EPPI Centre. Evidence for Policy and Practice Information and Co-ordinating Centre (2002). *A Systematic Review of the Impact of Summative Assessment and Tests on Students' Motivation for Learning*. Review conducted by the Assessment and Learning Synthesis Group, EPPI, London.
- Eurydice Network (2009). *National Testing of Pupils in Europe: Objectives, Organisation and Use of Results*. http://eacea.ec.europa.eu/education/eurydice/documents/thematic_reports/109EN.pdf (ver. 30.12.2014).
- Fondazione Giovanni Agnelli (2014). *La valutazione della scuola. A che cosa serve e perché è necessaria all'Italia*. Roma-Bari: Laterza.
- Greaney, V., & Kellaghan, T. (1995). *Equity Issues in Public Examinations in Developing Countries*. Washington, DC: World Bank.
- Greaney, V., & Kellaghan, T. (2008). *Assessing National Achievement Levels in Education*. Washington, DC: World Bank.
- Guilfoyle, C. (2006). NCLB: Is There Life Beyond Testing?. *Educational Leadership*, 64(3), 8–13.
- Hattie, J. (2011). Feedback in schools. In R. Sutton, M.J. Hornsey, K.M. Douglas (Eds.). *Feedback: The communication of praise, criticism, and advice*. New York: Peter Lang Publishing.
- Hattie, J. (2013). What is the nature of evidence that makes a difference to learning?. *Form@re - Open Journal per la formazione in rete*, 13(2), 6–21.
- Hattie, J., & Anderman, E.M. (Eds.) (2013). *International Guide to Student Achievement*. London: Routledge.
- Hopkins, D., Pennock, D., Ritzen, J., Ahtaridou, E., & Zimmer, K. (2008). *External evaluation of the policy impact of PISA*. OECD document EDU/PISA/GB(2008)35/REV1. Paris: OECD Publishing.
- Hout, M., & Elliott, S. (Eds.) (2011). *Incentives and Test-Based Accountability in Education*. Washington, DC: The National Academies Press.
- Invalsi. Archivio prove. www.invalsi.it (ver. 30.12.2014).
- Koretz, D. (2010). *Implications of Current Policy for Educational Measurement*. Education Testing Service (ETS).
- McNeil, L. (2000). *Contradictions of school reform: Educational costs of standardized testing*. New York: Routledge.

- Milanowski, A., & Heneman, H. (2001). Assessment of teacher reactions to a standards based teacher evaluation system: A pilot study. *Journal of Personnel Evaluation in Education*, 15(3), 193–212.
- Mitchell, D. (2014). *What Really Works in Special and Inclusive Education: Using evidence-based teaching strategies* (2nd ed.). New York: Routledge.
- MIUR. Ministero dell'Istruzione, dell'Università e della Ricerca (2012). Indicazioni nazionali per il curriculum della scuola dell'infanzia e del primo ciclo d'istruzione. *Annali della Pubblica Istruzione*. No. Speciale. http://www.annaliistruzione.it/var/ezflow_site/storage/original/application/55f6425315450eb079ff3e4da917750c.pdf (ver. 30.12.2014).
- MIUR. Ministero dell'Istruzione, dell'Università e della Ricerca (2014). Direttiva del 18 settembre 2014, n. 11. *Priorità strategiche del Sistema Nazionale di Valutazione per gli anni scolastici 2014/15, 2015/16 e 2016/17*. http://www.istruzione.it/allegati/2014/DIRETTIVA_SISTEMA_NAZIONALE_DI_VALUTAZIONE.pdf (ver. 30.12.2014).
- OECD. Organization of Economic Co-operation and Development (2012). *Equity and Quality in Education. Supporting Disadvantaged Students and Schools*. http://www.oecd-ilibrary.org/education/equity-and-quality-in-education_9789264130852-en (ver. 30.12.2014).
- OECD. Organization of Economic Co-operation and Development (2013). *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*. <http://www.oecd.org/edu/school/synergies-for-better-learning.htm> (ver. 30.12.2014).
- Olsen, J.P. (1989). *Modernization Programs in Perspective*. Institutional Analysis of Organizational Change, Notat 89/46, Bergen: LOS-Senter.
- Phelps, R.P. (2009). Educational achievement testing: critiques and rebuttals. In R.P. Phelps (Ed.), *Correcting Fallacies about Educational and Psychological Testing* (pp. 89–146). Washington, DC: American Psychological Association.
- Popham, W.J. (2001). Teaching to the Test?. *Educational Leadership*, 58(6), 16–21.
- Popham, W.J. (2008). Anchoring down the data. *Educational Leadership*, 66(4), 85–86.
- Popham, W.J. (2011). *Transformative assessment in action: An inside look at applying the process*. Alexandria: ASCD.
- Reeves, D. (2008). Looking deeper into the data. *Educational Leadership*, 66(4), 89–90.
- Scriven, M. (1967). The methodology of evaluation. In R.W. Tyler, R.M. Gagné, & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (pp. 39–83). Chicago, IL: Rand McNally.
- Slavin, R.E. (2002). Evidence-based education policies: Transforming educational practice and research. *Educational Researcher*, 31(7), 15–21.
- Toch, T. (2006). Margins of Error: The Education Testing Industry in the No Child Left Behind Era. *Education Sector Reports*. www.educationsector.org/publications/margins-error-testing-industry-no-child-left-behind-era (ver. 30.12.2014).

- Trinchero, R. (2014). Il Servizio Nazionale di Valutazione e le prove Invalsi. Stato dell'arte e proposte per una valutazione come agente di cambiamento. *Form@re - Open Journal per la formazione in rete*, 14(4).
- Vertecchi, B. (2014). Invalsi, risparmiamo con test a campione. *L'Unità*, 7/5/2014. <http://cerca.unita.it/ARCHIVE/xml/2635000/2632482.xml?key=Benedetto+Vertecchi&first=1&orderby=1&f=fir> (ver. 30.12.2014).
- Vivanet, G. (2014a). *Che cos'è l'evidence based education*. Roma: Carocci.
- Vivanet, G. (2014b). La valutazione degli apprendimenti scolastici. Un quadro internazionale. *Form@re - Open Journal per la formazione in rete*, 14(4).
- Wang, L., Beckett, G., & Brown, L. (2006). Controversies of Standardized Assessment in School Accountability Reform: A Critical Synthesis of Multidisciplinary Research Evidence. *Applied Measurement in Education*. 19(4), 305–328.