

Automated feedback delivery: a structured model for computer-based adaptive testing

Restituzione automatica del feedback: un modello strutturato per le prove adattative computer based

Emanuela Botta

Sapienza University of Rome, Dept. of developmental psychology and educational research, Rome (Italy)

OPEN ACCESS

Double blind peer review

Citation: Botta, E. (2023). Automated feedback delivery: a structured model for computer-based adaptive testing. *Italian Journal of Educational Research*, 30, 73-84.
<https://doi.org/10.7346/sird-012023-p73>

Corresponding Author: Emanuela Botta
Email: emanuela.botta@uniroma1.it

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. IJEDuR is the official journal of Italian Society of Educational Research (www.sird.it).

Received: February 8, 2023

Accepted: April 20, 2023

Published: June 30, 2023

Pensa MultiMedia / ISSN 2038-9744
<https://doi.org/10.7346/sird-012023-p73>

Abstract

This paper will present an integration of an adaptive model for the estimation of ability in mathematics implemented as part of a consortium research project of Sapienza University. In addition to recalling the characteristics of this type of test and the advantages it has over linear tests, such as the greater precision of the estimation for each student regardless of his or her ability level, an automatic feedback model connected to the test itself will be illustrated, as well as the results from the literature that shows how similar models are appreciated by students, in formative and summative assessment. The model is structured in accordance with the main research results on feedback and it could be an essential schools' tool for formative assessment, the detection of misconceptions or errors of individuals and groups of students. The use of adaptive models is connected to the focus on equity in the system.

Keywords: Assessment; Adaptive Testing; Feedback; Skills in Mathematics; CBT.

Riassunto

In questo lavoro si presenta un'integrazione di un modello adattativo per la stima delle abilità in matematica su larga scala implementato nell'ambito di una ricerca della Sapienza. Oltre a ricordare le caratteristiche di questo tipo di prova e i vantaggi che presenta rispetto a quelle lineari, come la maggiore precisione della stima per ogni studente indipendentemente dal suo livello di abilità, si illustra un modello automatico di feedback connesso alla prova stessa e i risultati di letteratura che mostrano come modelli simili siano notevolmente apprezzati dagli studenti, nell'ambito sia della valutazione formativa sia di quella sommativa. Il modello è strutturato in accordo con i principali risultati di ricerca sul feedback e potrebbe essere per i docenti uno strumento essenziale per la valutazione formativa, l'individuazione di misconcezioni o errori sia dei singoli sia della classe. Merita osservare che l'uso di modelli adattativi è generalmente connesso all'attenzione all'equità del sistema.

Parole chiave: Valutazione; Prove adattative; Feedback; Abilità in matematica; CBT.

1. Introduzione

La questione della valutazione, delle sue molteplici funzioni e delle modalità con cui viene praticata nelle scuole e a livello di sistema è oggetto di continuo dibattito. In questo breve articolo non esploreremo tutte le possibili sfumature del problema ma, vista l'ampia diffusione di strumenti o applicazioni basati su complessi algoritmi matematici e sull'intelligenza artificiale per la stima delle abilità degli studenti in vari campi e la rapidità con cui sta aumentando la facilità di utilizzo degli stessi, ci concentreremo su come sia possibile progettare un modello di restituzione dei risultati che abbia una funzione educativa e non solo descrittiva o addirittura comparativa. Due elementi vanno tenuti infatti in debita considerazione, da una parte occorre constatare che a livello di sistema, per la valutazione dei sistemi scolastici o per esami di ammissione a master e corsi di laurea, le prove sono sempre più spesso orientate su modelli adattativi, dall'altra si deve prendere atto del fatto che ad esse non è correlata una restituzione individuale dei risultati che vada oltre l'assegnazione di un livello o di una posizione in graduatoria. A titolo di esempio citiamo il *Graduate Management Admission Test* (GMAT) per la pre-valutazione dei candidati all'ammissione a diversi Master in *Business Administration* (MBA) o ad altri corsi di management post-laurea, il *Graduate Record Examinations* (GRE) per l'ammissione a varie scuole di specializzazione, *Business School* e *Law School* nel mondo e infine, a livello di scuola primaria e secondaria, il primo esperimento di prova adattativa messa direttamente a disposizione delle scuole, il PAT (Mathematics Adaptive Test), sviluppato dall'NZCER¹ per la Nuova Zelanda, il quale però, nonostante l'intento dichiaratamente educativo si limita anch'esso, al momento, a restituire risultati comparativi rispetto agli standard nazionali. A livello globale anche L'OCSE ha cominciato a realizzare i test PISA sulla base di un modello adattativo² multilivello molto simile a quello realizzato nell'ambito di questa ricerca.

Se ci soffermiamo a riflettere sulla circostanza che la spinta all'adozione di prove adattative è dovuta al fatto che rispetto alle prove lineari esse sono sempre significativamente più informative, a livello individuale, per studenti con qualsiasi livello di abilità, anche agli estremi dell'intervallo assunto come riferimento (Botta, 2021a; Botta, 2021b; Lord, 1980; Wainer et al., 2000; Weiss & Kingsbury, 1984; Weiss, 1985), allora ciò che sembra mancare, e su cui la letteratura comincia ad indagare, è l'utilizzo delle migliori proprietà misuratorie delle prove adattative ai fini della raccolta e restituzione di informazioni che possano essere usate per una valutazione formativa.

Merita altresì ricordare che in Italia, con il passaggio delle prove Invalsi alla somministrazione computer based, dal grado 8 al grado 13, alle scuole sono venuti a mancare strumenti essenziali sui quali lavorare proprio in termini di valutazione formativa, individuazione di misconcezioni o errori diffusi nel gruppo classe e carenze specifiche dei singoli alunni.

Scopo di questo lavoro è dunque quello di illustrare i primi risultati di ricerca in questo campo e di proporre, in associazione alla prova adattativa multilivello effettivamente realizzata e somministrata sul campo a un campione di oltre 4000 studenti (Botta, 2021a), un modello di restituzione dei risultati, individualizzato e orientato alla valutazione per l'apprendimento.

2. Le prove adattative e la restituzione del feedback

2.1 Il contesto di riferimento

La letteratura in merito all'utilizzo di questo tipo di prove sia ai fini della valutazione formativa, per l'apprendimento, sia di quella sommativa è piuttosto ampia ma esistono relativamente pochi studi (Barker, 2010; Lilley & Barker, 2007; Roos, Plake & Wise, 1992; Lilley, Barker & Britton, 2005) che mostrano come sia possibile implementare modelli automatici di feedback connessi alle prove stesse e quanto questo sia notevolmente apprezzato dagli studenti. In questi studi è stato assunto che lo scopo primario di una valutazione formativa sia informare gli studenti sui loro punti di forza e di debolezza, in accordo con Morgan et al. (2004) e Brown et al. (1997). La valutazione formativa è dunque intesa come mirata a restituire

1 <https://www.nzcer.org.nz/tests/pat-mathematics-adaptive>, ultimo accesso il 27/02/2023.

2 https://www.oecd-ilibrary.org/education/introduction-of-multistage-adaptive-testing-design-in-pisa-2018_b9435d4b-en

agli studenti un feedback utile a migliorare nell'apprendimento e nel rendimento (performance). I risultati positivi di questo approccio sono stati ampiamente riportati da Charman (2002), Sly & Rennie (2002) e Steven & Hesketh (2002). I successivi studi di Lilley e Barker si sono focalizzati sul miglioramento del modello di feedback, correlandolo ad esempio alla tassonomia delle abilità cognitive di Bloom (Bloom, 1956), cercando di superare l'idea di confrontare in ogni caso i risultati del singolo con quelli di una media di riferimento, e arricchendo il modello con una serie di proposte di esercizi di rinforzo con un livello di difficoltà adeguato al loro livello di abilità stimato. Come verificato anche nei nostri studi precedenti (Botta, 2021a; Botta, 2021b), infatti, mettere gli studenti di fronte ad esercizi per loro troppo difficili, sembra innescare processi di demotivazione, sforzo cognitivo e conseguente deplezione dell'io, che influiscono negativamente sulla capacità dello studente di applicarsi in seguito anche su prove che sarebbero per lui perfettamente adeguate. Risulta interessante rilevare che da questi studi emerge anche quanto gli studenti apprezzino questo tipo di restituzione, in particolare i suggerimenti per il miglioramento dei propri apprendimenti e l'individuazione di possibili punti di forza e debolezza, e come questi non siano positivamente correlati, in modo statisticamente significativo, con il rendimento nella prova (Lilley & Barker, 2007).

2.2 La prova adattativa utilizzata per la ricerca

È evidente che per la progettazione del modello di restituzione dei risultati agli studenti occorre tenere nella dovuta considerazione sia la prova, con le sue caratteristiche tecniche e misuratorie, sia il costrutto della matematica a cui essa fa riferimento.

La prova utilizzata per la ricerca è una prova adattativa multilivello per la stima delle abilità in matematica degli studenti del grado 10 (Botta, 2021a). Essa è stata strutturata secondo un modello a 3 livelli, del tipo 1-3-3, comprendente sette possibili percorsi, ciascuno composto da 46 item (Figura 1). Il primo livello è costituito da un unico modulo, composto da 16 item, e centrato attorno all'abilità media stimata, con il modello IRT a un parametro, sul campione utilizzato nelle fasi di progettazione e pretest (circa 10000 studenti), mentre ciascuno dei livelli successivi comprende tre moduli distinti, ciascuno centrato attorno ad un diverso livello di abilità (bassa, media e alta), e composto da 18 item a livello 2 e da 12 item a livello 3. Ogni studente svolge il modulo iniziale, uno solo dei tre moduli del livello 1 e uno solo dei tre moduli del livello 3.

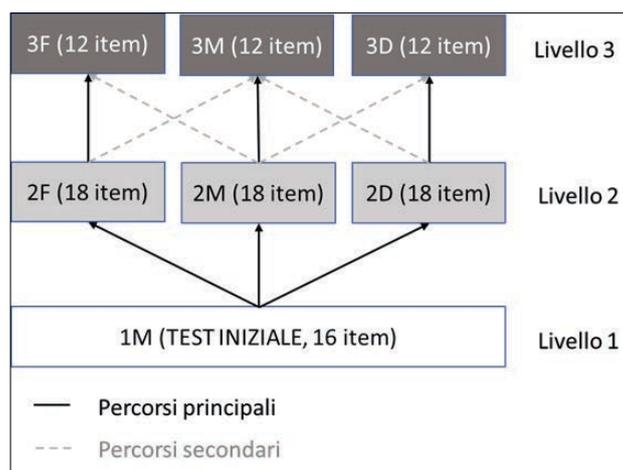


Figura 1: Modello MST della prova

Le frecce in figura indicano i sette possibili percorsi all'interno della prova. La stima dell'abilità dello studente viene calcolata al termine di ogni livello considerando sempre tutti gli item a cui ha risposto e sulla base di tale stima si determina quale dovrà essere il modulo somministrato allo studente nel livello successivo. I percorsi risultano dunque suddivisi in due tipologie: i percorsi principali, lungo i quali l'abilità stimata dello studente si mantiene all'interno dello stesso intervallo di riferimento, e quelli secondari,

lungo i quali l'abilità stimata dello studente varia, passando ad esempio da bassa a media. Tutti i moduli di un livello sono costituiti dallo stesso numero item e bilanciati in relazione alle dimensioni e ai contenuti che caratterizzano il costrutto dell'abilità che si vuole misurare.

Il costrutto di riferimento della prova tiene in debita considerazione sia i riferimenti normativi principali, Indicazioni nazionali per i Licei (MIUR, 2010a) e Linee guida per gli Istituti tecnici (MIUR, 2010b) e professionali (MIUR, 2019), sia il Quadro di riferimento dell'Invalsi per le prove di matematica (Invalsi, 2018), sia la necessità di selezionare e definire obiettivi specifici di apprendimento la cui acquisizione sia effettivamente verificabile con una prova adattativa somministrata in formato computer based. Esso risulta dunque articolato in quattro ambiti di contenuto, Numeri (NU), Dati e previsioni (DP), Relazioni e funzioni (RF) e Spazio e figure (SF) e in tre dimensioni, Conoscere, Risolvere problemi e Argomentare. "Le dimensioni costituiscono un raggruppamento degli obiettivi di apprendimento, fondato sull'idea che le attività matematiche si riferiscano essenzialmente o all'argomentare o al risolvere problemi e che queste due dimensioni non siano indipendenti l'una dall'altra e richiedano, per essere attuate, conoscenze su concetti, linguaggio formale e procedure. La dimensione semiotica della rappresentazione è considerata trasversale alle altre e assume in ciascuna di esse aspetti diversi" (Botta, 2021a). Gli ambiti costituiscono un altro raggruppamento degli obiettivi di apprendimento in relazione ai contenuti specifici della matematica.

Percorso	Moduli	DP	NU	RF	SF	Totale per percorso
<i>Percorso 1</i>	1M+2F+3F	12	12	11	11	46
<i>Percorso 2</i>	1M+2F+3M	12	12	11	11	46
<i>Percorso 3</i>	1M+2M+3F	12	11	12	11	46
<i>Percorso 4</i>	1M+2M+3M	12	11	12	11	46
<i>Percorso 5</i>	1M+2M+3D	12	11	12	11	46
<i>Percorso 6</i>	1M+2D+3M	11	12	11	12	46
<i>Percorso 7</i>	1M+2D+3D	11	12	11	12	46

Tabella 1: Numero di item per ambito in ogni percorso

Percorso	Moduli	Conoscere	Risolvere problemi	Argomentare	Totale per percorso
<i>Percorso 1</i>	1M+2F+3F	26	16	4	46
<i>Percorso 2</i>	1M+2F+3M	27	14	5	46
<i>Percorso 3</i>	1M+2M+3F	24	17	5	46
<i>Percorso 4</i>	1M+2M+3M	25	15	6	46
<i>Percorso 5</i>	1M+2M+3D	23	17	6	46
<i>Percorso 6</i>	1M+2D+3M	23	16	7	46
<i>Percorso 7</i>	1M+2D+3D	21	18	7	46

Tabella 2: Numero di item per dimensione in ogni percorso

Questa prova misura dunque essenzialmente conoscenze e abilità, anche se comprende quesiti adatti a misurare in parte gli aspetti relativi alle competenze nel risolvere problemi e nell'argomentare. Come osservato anche in Botta (2021a) "[...] Tali aspetti sono comunque misurabili parzialmente; risulta infatti difficile costruire e valutare organicamente, con ridotti margini di errore, le abilità nel sostenere argomentazioni e nell'ideare dimostrazioni complesse, così come quelle necessarie a porsi e risolvere problemi di una certa difficoltà e complessità, o ancora quelle relative alla costruzione di modelli matematici per situazioni complesse. Generalmente si richiede quindi allo studente di scegliere l'argomentazione più adeguata a sostenere una determinata tesi, di confermare con un controesempio la falsità di un'affermazione, di stabilire il valore di verità di una proposizione o di completare una dimostrazione inserendo alcune parole chiave. Nell'ambito della formalizzazione si richiede di individuare la modellizzazione più adatta a rappresentare una situazione problematica e non si richiede di riportare i passaggi effettuati nella risoluzione di un problema ma solamente la soluzione trovata."

3. Il modello di restituzione

3.1 Linee guida per la progettazione del modello

In accordo con i risultati delle ricerche sulla valutazione formativa (Black & William, 2009; Hattie & Timperley, 2007; Hattie, 2016; Bennett, 2011) il modello di feedback proposto fa riferimento agli obiettivi di apprendimento verificati nella prova e previsti dal costrutto e non solo ai singoli errori commessi. Inoltre, cerca di fornire spunti di riflessione pensati come feedback sui processi utilizzati per svolgere il compito. Il completamento della prova con uno strumento automatico di restituzione dei risultati finalizzato alla valutazione formativa è fortemente connesso alla funzione del feedback di livello 1, sul compito, e di livello 2, sui processi, e alla necessità che esso sia fornito in tempi rapidi e sulla base di elementi concreti. Come osservato da Hattie (2016) infatti, “il feedback a livello di compito e prodotto è efficace soprattutto se mira a fornire un riscontro oggettivo (ad esempio giusto o sbagliato), se porta ad acquisire informazioni ulteriori o diverse, e se sviluppa una maggiore conoscenza superficiale. [...] Questo feedback sul prodotto compito è essenziale e getta le basi sulle quali possono poi essere costruite efficacemente l’elaborazione (livello 2) e l’autoregolazione (livello 3). [...] Al secondo livello il feedback è diretto ai processi [...]. Con questo tipo di feedback è possibile fornire modalità alternative di elaborazione, ridurre il carico cognitivo, aiutare a sviluppare strategie di apprendimento e rilevazione degli errori e suggerire di fare una nuova e più efficace ricerca di informazioni, di cogliere i collegamenti fra le idee e di utilizzare strategie per svolgere il compito”. Gli spunti di riflessione proposti allo studente possono inoltre influire sul livello dell’autoregolazione, migliorando le capacità di autovalutazione degli studenti e aiutandoli ad accettare e comprendere il feedback ricevuto.

In quest’ottica l’idea di valutazione formativa che si è assunta come riferimento è sostanzialmente quella di Black e William (2009), secondo cui un’esperienza educativa è formativa nella misura in cui si ricercano, si interpretano e si utilizzano le evidenze sugli obiettivi di apprendimento raggiunti e queste attività forniscono elementi essenziali sulla definizione dei successivi passi da compiere nel processo educativo al fine di garantire che l’esito sia migliore di quello che si sarebbe avuto in assenza di tali elementi.

Le caratteristiche del modello proposto e illustrato nel paragrafo che segue sono dunque:

- la correlazione dei risultati riportati con il profilo in uscita da ogni percorso della prova adattativa, in termini di traguardi e obiettivi specifici di apprendimento, definito in forma analitica per ciascun percorso nella banca e richiamato nel modello stesso;
- la puntuale individuazione ed esplicitazione del raggiungimento dei diversi obiettivi di apprendimento;
- l’assenza di confronto dei risultati ottenuti con la media di popolazione o con altre medie, sempre al fine di correlare il feedback all’obiettivo di apprendimento più che alla comparazione con i pari o con teorici livelli di riferimento;
- la presenza di suggerimenti e spunti di riflessione basati sull’analisi delle risposte fornite, sia errate sia corrette, da tutti gli studenti tramite categorizzazione e analisi dal punto di vista specifico della disciplina oggetto della prova, nel nostro caso la matematica. Questi vengono aggiunti al catalogo della banca di item ex-post e periodicamente aggiornati in occasione di nuove somministrazioni di ogni singolo item.

3.2 Illustrazione del modello

Per l’illustrazione del modello si è scelto di selezionare tre studenti con abilità media provenienti da tre diversi percorsi, nello specifico il 2, il 4 e il 6, tutti aventi come esito una stima dell’abilità ricadente nell’intervallo dell’abilità media. Alla base di questa scelta vi sono due obiettivi: il primo è per l’appunto illustrare il modello, utilizzando l’esito ottenuto per uno di questi studenti, l’altro è quello di mostrare alcune caratteristiche peculiari delle prove adattative indispensabili per comprendere perchè esse si prestino più di una qualsiasi prova lineare ad essere utilizzate per la raccolta di elementi validi per la valutazione del singolo studente.

Il modello di restituzione individuale è dunque costituito da 2 parti:

- il risultato generale e
- il risultato per ambito di contenuto con i relativi suggerimenti.

A titolo di esempio si riportano le schede del modello relative allo studente che ha affrontato il percorso 2. La parte relativa al risultato generale, Figura 2, riporta i risultati ottenuti in termini quantitativi, articolati sia per Dimensione sia per Ambito di contenuto. Questa prima parte consente allo studente di individuare, anche graficamente, non solo il rendimento generale nella prova, ma anche le eventuali aree di forza o debolezza in relazione al costruito. Per questo studente è evidente che in termini di dimensioni le difficoltà si concentrano nel Risolvere problemi, mentre mostra un buon livello di abilità nella dimensione dell'Argomentare. Tenendo conto del percorso affrontato, che nella parte centrale è costituito da domande di difficoltà bassa anziché media, il docente potrebbe ragionevolmente ritenere di dover proporre allo studente attività e quesiti più sfidanti in questa dimensione, ad esempio facendolo passare dall'individuazione della giustificazione corretta per una data affermazione fra alcune proposte ad attività di argomentazione autonoma o guidata o di dimostrazione. Nel farlo potrà tenere conto anche dei risultati dello studente in relazione all'ambito di contenuto facendolo lavorare prima su Numeri e Dati e previsioni e solo in seguito, dopo adeguate attività di rinforzo, su Spazio e figure e Relazioni e funzioni.

La parte relativa al risultato per ambito di contenuto viene elaborata per ciascuno dei quattro ambiti previsti dal costruito ma qui presentiamo solamente quella relativa all'ambito Numeri, riportata nelle figure 3 e 4.

Nella prima parte di questa scheda, Figura 3, vengono riportati la descrizione dell'ambito Numeri, il risultato ottenuto nella prova relativamente a questo singolo ambito e la tabella analitica di tutti gli obiettivi di apprendimento connessi alle domande a cui lo studente ha risposto. Inoltre per ciascuna domanda è riportata una sintetica descrizione di quanto la domanda richiedeva, per consentire un'agevole connessione con la domanda stessa. In questo caso occorre infatti ricordare che nel caso in cui la prova sia utilizzata per la valutazione di sistema non è possibile rendere pubbliche le domande (come nel nostro caso), pertanto l'inserimento della descrizione della richiesta è un'accortezza necessaria a consentire allo studente e al docente di comprendere il feedback restituito e di poterlo conservare. Infine, in corrispondenza di ogni domanda è riportato un simbolo atto a individuare immediatamente se la risposta fornita dallo studente è corretta o errata. La lettura di questa parte della scheda permette dunque di avere un'idea chiara e completa degli obiettivi di apprendimento di cui è stato verificato il raggiungimento e anche di delineare, rispetto all'ambito di contenuto in oggetto, il profilo dello studente. Fra i vari elementi che si possono trarre vale la pena di soffermarsi in particolare sul caso del calcolo della percentuale. Come si può ben osservare le domande che richiedono la conoscenza e l'uso della percentuale sono diverse, ma lo studente si è trovato in difficoltà solamente sull'ultima, l'unica che ricade nella dimensione Risolvere problemi. Questa informazione assume maggiore rilievo non appena viene connessa a quelle ricavabili dalla scheda del Risultato generale, che evidenziavano un problema diffuso a tutti gli ambiti proprio nella dimensione Risolvere problemi.

Nella seconda parte di questa scheda, Figura 4, per ciascuna domanda a cui è stata fornita una risposta errata vengono riportate le possibili difficoltà che lo studente può avere incontrato e alcune domande che svolgono la funzione di indirizzare lo studente sia verso l'individuazione della specifica difficoltà incontrata sia verso la riflessione, da una parte rispetto al possesso delle conoscenze e delle abilità necessarie e dall'altra rispetto ai processi messi in atto per lo svolgimento del compito. La disamina delle possibili difficoltà e dei suggerimenti emerge dall'analisi delle risposte fornite da tutti gli studenti e non solo da quella dello studente che sta ricevendo il feedback. In generale dunque essa non sarà esaustiva e tenderà ad arricchirsi all'aumentare del numero di soggetti a cui la prova sarà somministrata. I suggerimenti di revisione e gli spunti di riflessione devono quindi essere integrati, per ciascun item, all'interno del catalogo della banca di item utilizzata per la prova, in modo da essere automaticamente reperibili in corrispondenza di ogni risposta errata fornita dallo studente. Come osservato da Lilley et al. (2005), sebbene essi siano intrinsecamente connessi con la domanda a cui lo studente ha fornito una risposta errata non devono replicare la domanda stessa o fornire la risposta corretta, perché questo non favorirebbe l'attivarsi dello studente in un processo di riflessione e ricerca teso a migliorare la sua comprensione e i suoi apprendimenti.

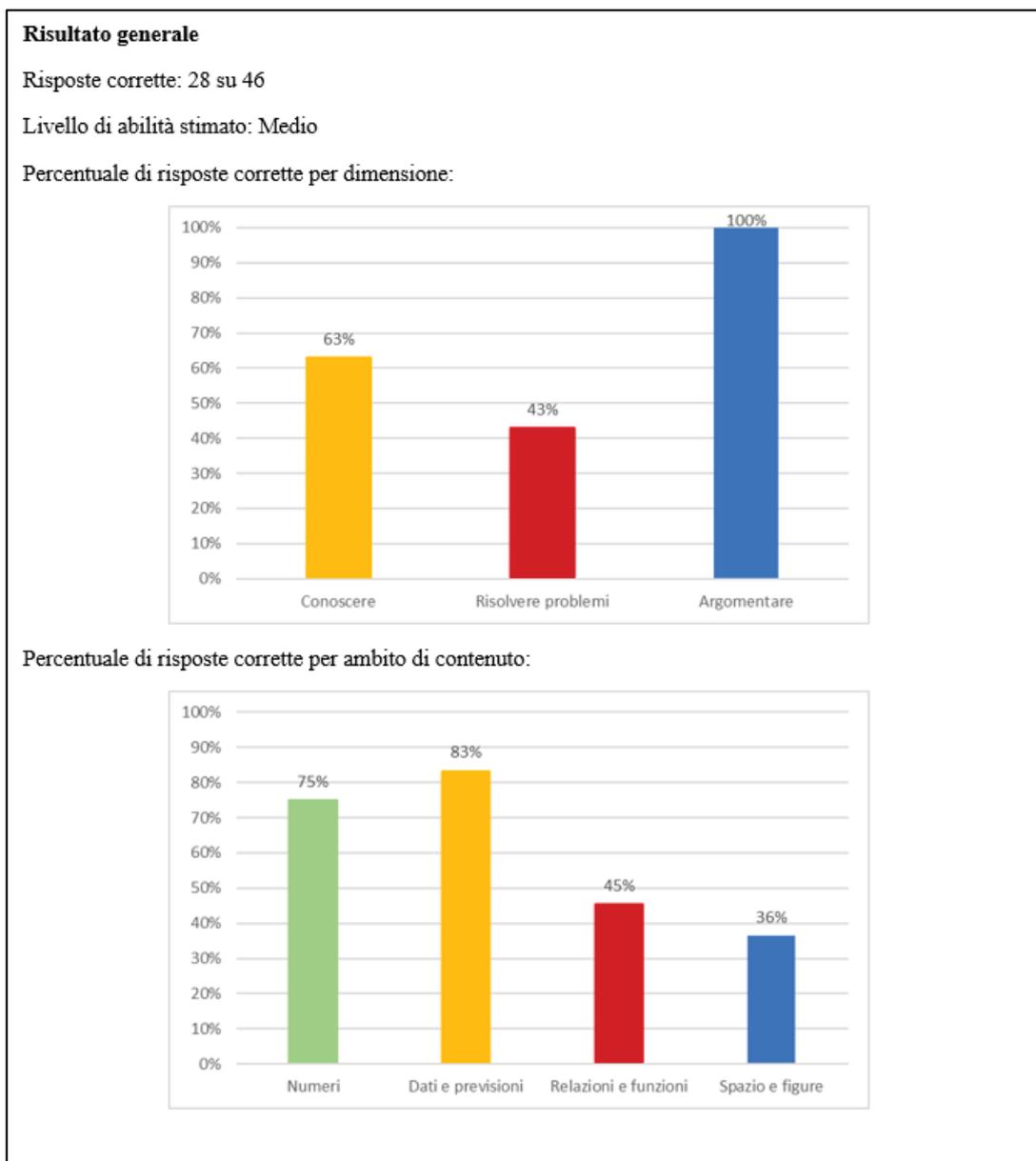


Figura 2: Modello di restituzione – Risultato generale

Risultato per ambito di contenuto

Ambito: Numeri

L'ambito Numeri è focalizzato sulle conoscenze e sulle abilità degli studenti nel rappresentare numeri, ordinarli, eseguire calcoli, fare stime appropriate in base al contesto, usare i rapporti e il ragionamento proporzionale, conoscere e applicare le proprietà dei numeri e delle operazioni anche per risolvere problemi e argomentare, sia nel contesto interno alla matematica sia in altri contesti. A questo ambito di contenuto afferiscono anche le abilità relative al calcolo letterale.

Risposte corrette: 9 su 12

La tabella che segue riporta il risultato che hai ottenuto nella prova relativamente a questo ambito, specificando a quali obiettivi di apprendimento sono correlate le domande che ti sono state poste:

Obiettivo di apprendimento	Descrizione dell'item	Dimensione	Risposta
1. Trasformare semplici formule o espressioni algebriche contenenti le quattro operazioni e l'elevamento a potenza fra monomi e polinomi.	Svolgere il prodotto di due binomi di primo grado.	Conoscere	✓
2. Risolvere semplici disequazioni o riconoscerne l'insieme delle soluzioni.	Riconoscere l'insieme delle soluzioni di una disequazione di secondo grado.	Conoscere	✗
3. Ordinare numeri reali e utilizzare con disinvoltura i simboli =, <, >, ... e la retta dei numeri. Fare stime di numeri, quantità o lunghezze e approssimare a un dato numero di cifre dopo la virgola.	Riconoscere la posizione sulla retta dei numeri di un numero dato, espresso in lettere, confrontandolo con un numero in notazione scientifica.	Conoscere	✗
4. Riconoscere proprietà dei numeri attraverso espressioni letterali che li rappresentano.	Riconoscere fra diverse espressioni letterali di primo grado l'unica che può rappresentare un numero dispari per qualsiasi valore attribuito alla variabile nell'insieme dei numeri naturali.	Argomentare	✓
5. Trasformare semplici formule o espressioni algebriche contenenti le quattro operazioni e l'elevamento a potenza fra monomi e polinomi	Svolgere il prodotto di un monomio per un binomio di primo grado.	Conoscere	✓
6. Determinare la verità o falsità di affermazioni sui numeri e sul loro ordinamento	Riconoscere fra diverse argomentazioni proposte quella che giustifica correttamente la verità di un'affermazione relativa ai numeri pari e dispari, rappresentati con un'espressione letterale. In particolare è necessario riconoscere che la veridicità dell'affermazione per un singolo valore attribuito alla variabile nell'insieme dei numeri naturali non implica che essa sia vera sempre.	Argomentare	✓
7. Conoscere e utilizzare il concetto di rapporto e quello di proporzionalità diretta e inversa. Dividere una quantità in relazione a un rapporto dato.	Comprendere che la situazione problematica è riconducibile a una proporzionalità diretta, individuare dalla lettura di un testo complesso contenente anche dati e tabelle i due rapporti da porre in relazione di uguaglianza e scegliere, fra diversi valori dati, l'unico che la rende vera.	Conoscere	✓
8. Calcolare una data percentuale di una quantità.	Calcolare in modo diretto una certa percentuale di una quantità data.	Conoscere	✓
9. Tradurre le informazioni fornite in un testo (registro verbale) in una rappresentazione numerica (tabella di valori), simbolica (formula) o grafica.	Scegliere, fra quelle fornite, la corretta rappresentazione simbolica della descrizione verbale di un'uguaglianza fra espressioni letterali.	Risolvere problemi	✓
10. Risolvere problemi diretti e inversi relativi al calcolo di percentuali.	Risolvere un problema sulla determinazione del prezzo scontato di un prodotto conoscendo il prezzo iniziale, non scontato, e le percentuali di sconto da applicare.	Risolvere problemi	✗
11. Conoscere il ruolo dello zero e dell'uno nelle operazioni nell'insieme dei numeri reali. Conoscere e utilizzare la legge di annullamento del prodotto	Dato il valore del prodotto di due numeri, espressi simbolicamente in forma letterale, selezionare fra le diverse proposte l'unica condizione necessaria che deve essere soddisfatta dai due numeri dati affinché l'uguaglianza risulti vera.	Conoscere	✓
12. Determinare la verità o falsità di affermazioni sui numeri e sul loro ordinamento.	Individuare fra quelle date l'unica giustificazione corretta ad un'affermazione sul valore del massimo comune divisore fra tre numeri naturali noti.	Argomentare	✓

Legenda: risposte corrette, ✓, o errate, ✗

Figura 3: Modello di restituzione – Risultato per ambito Numeri, prima parte

La tabella che segue riporta la descrizione delle difficoltà che potresti aver incontrato per rispondere ad una certa domanda e alcuni suggerimenti e spunti di riflessione per aiutarti a comprendere cosa potresti fare per migliorare.

Item	Possibili difficoltà incontrate	Suggerimenti
2	La difficoltà potrebbe stare nel riconoscere che la somma di un numero ignoto elevato alla seconda e di un numero positivo dato è positiva per qualsiasi valore attribuito all'incognita. Inoltre non è indispensabile risolvere la disequazione utilizzando uno dei procedimenti che hai studiato. In questo caso la scelta di risolvere formalmente la disequazione potrebbe rivelarsi poco efficace, rendendo più difficile individuare l'insieme delle soluzioni.	<ul style="list-style-type: none"> - Hai cercato di utilizzare un procedimento risolutivo che conoscevi? - Hai cercato di procedere come se si trattasse di una disequazione di primo grado "spostando" i termini numerici a destra del segno di disequazione? - Hai ritenuto che potesse essere corretto estrarre la radice n-esima del primo e del secondo membro della disequazione? Questo è sempre possibile? - Hai provato a chiederti che segno ha un qualsiasi numero reale elevato a potenza con esponente pari? - Sapresti dire che segno ha la somma di numeri positivi o quella di numero positivo più lo zero?
3	Le possibili difficoltà risiedono nel comprendere che l'ordine di grandezza dei due numeri, dati con rappresentazioni diverse, è lo stesso e nell'individuare il passo utilizzato sulla retta dei numeri, che non necessariamente corrisponde sempre ad una singola unità di misura.	<ul style="list-style-type: none"> - Hai familiarità con l'uso della retta dei numeri? - Pensi di essere in grado di individuare su una retta dei numeri un passo non corrispondente ad una unità di misura? (Ad esempio in una retta su cui sono riportati solo numeri pari o dispari) - Sai individuare l'ordine di grandezza delle potenze di dieci rappresentate in lettere, come mille o un miliardo? - Sai individuare l'ordine di grandezza delle potenze di dieci espresse in forma numerica come 10^3 o 10^7? - Sapresti rappresentare un numero come ad esempio 3520 in notazione scientifica? - Pensi che per confrontare due numeri molto grandi o molto piccoli possa essere utile rappresentarli entrambi nello stesso modo?
10	Le difficoltà aggiuntive, rispetto ad esempio all'item relativo all'obiettivo 8, risiedono nel fatto che viene richiesto di applicare due sconti, e che la determinazione del prezzo scontato richiede l'ulteriore passaggio logico di sottrarre al prezzo iniziale il valore dello sconto individuato con il calcolo della relativa percentuale.	<ul style="list-style-type: none"> - Hai verificato di saper individuare correttamente una certa percentuale di un numero dato, come ad esempio il 20% di 400? (Come hai risposto alle domande 7 e 8?) - Sapresti indicare come calcolare il prezzo scontato del 20% di un articolo che costa inizialmente 400 euro? - Se dovessi applicare un ulteriore sconto di quale prezzo calcoleresti la nuova percentuale, sempre di 400 euro?

Figura 4: Modello di restituzione – Risultato per ambito Numeri, seconda parte

3.3 Prove adattative e stima dell'abilità dello studente

Ci soffermiamo ora ad analizzare alcuni elementi che potrebbero rivelarsi utili al docente nel momento in cui decidesse di procedere ad una disamina dei risultati di vari gruppi di studenti. Affinché si possa attribuire a tali informazioni il corretto significato è infatti indispensabile sapere che mentre in una prova lineare, o in un insieme di prove lineari parallele, si può assumere che le prove abbiano tutte mediamente la stessa difficoltà e dunque che ci sia, seppure con un certo margine di errore, una certa corrispondenza fra il punteggio ottenuto da diversi studenti e la stima della loro abilità, in una prova adattativa questo non avviene.

A titolo di esempio consideriamo tre studenti la cui abilità è stata stimata al termine della prova come media, provenienti rispettivamente dai percorsi 2, 4 e 6. La stima è stata effettuata con un modello IRT a un parametro e quindi tutti gli studenti, pur avendo effettuato percorsi diversi, e dunque prove diverse, sono posti su un'unica scala delle abilità e i risultati dell'uno sono direttamente confrontabili con quelli di un altro.

Come si può osservare in Figura 5 i tre percorsi scelti hanno in comune il primo e l'ultimo modulo, rispettivamente 1M e 3M, mentre si differenziano per il modulo di livello 2, che per il percorso 2 è 2F, e contiene domande con un indice di difficoltà b basso, per il percorso 4 è 2M, che contiene domande con un indice di difficoltà medio, e per il percorso 6 è 2D, che contiene domande con un indice di difficoltà alto. Si consideri che la difficoltà media è $b = -0,406$ nella scala del modello IRT.

I punteggi grezzi degli studenti sulla prova nel suo complesso e nell'ambito Numeri sono riportati nella Tabella 3, e come si può osservare variano, sull'intera prova da 18 a 28, e nell'ambito numeri da 6 a 9.

Studente	Punteggio su Numeri	Punteggio totale
Percorso 2	9	28
Percorso 4	6	23
Percorso 6	6	18

Tabella 3: Punteggi grezzi di studenti di percorsi diversi

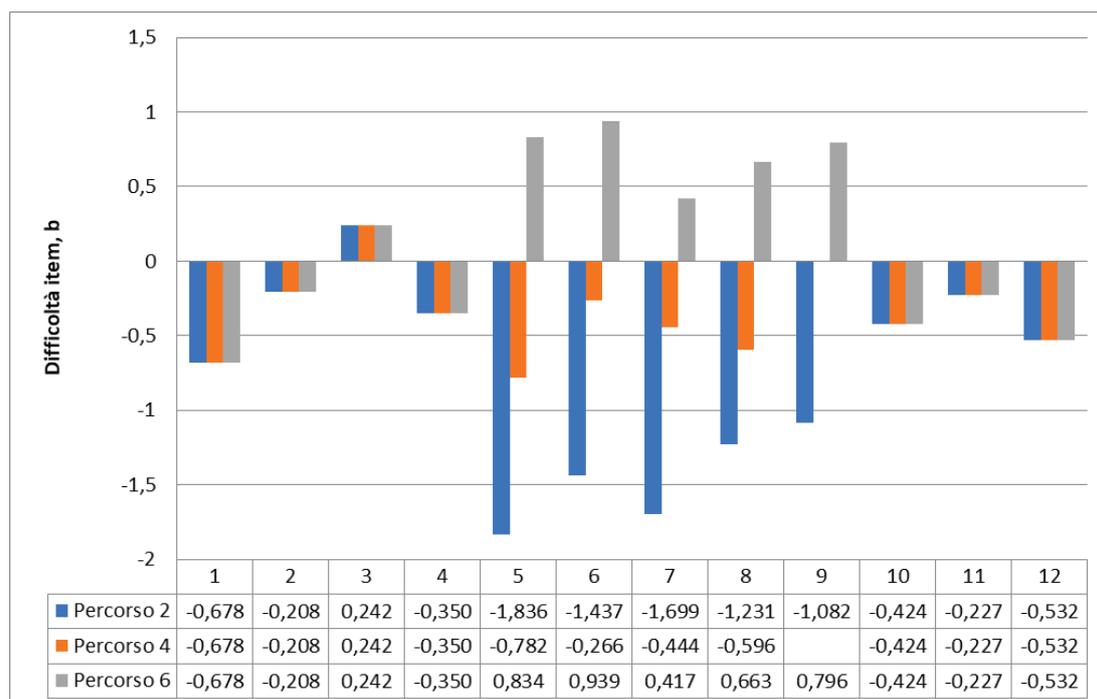


Figura 5: Difficoltà delle domande nei tre percorsi 2, 4 e 6

Se osserviamo in Tabella 4 il tracciato delle risposte dei tre studenti alle domande dell'ambito Numeri, dove 1 indica una risposta corretta, 9 una risposta non data e 0 una risposta errata, possiamo comprendere ancora meglio come in una prova adattativa la stima dell'abilità dello studente dipenda non solo dal numero di risposte corrette che ha dato ma anche dalla difficoltà delle domande e dal percorso che ha svolto.

Studente	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12	Tot. Risposte corrette
Percorso 2	1	0	0	1	1	1	1	1	1	0	1	1	9
Percorso 4	1	1	1	0	0	1	9	0		1	0	1	6
Percorso 6	1	1	1	0	0	0	0	0	1	1	9	1	6

Tabella 4: Tracciati nell'ambito Numeri di studenti di percorsi diversi

Se ci limitassimo a prendere in considerazione, ai fini della valutazione dei tre studenti, solamente i punteggi ottenuti, trascurando tutti gli elementi ricavabili dalla restituzione del feedback illustrati nei paragrafi precedenti potremmo giungere all'errata conclusione che lo studente del percorso 2 ha raggiunto più obiettivi di apprendimento degli altri ed ha rispetto a loro un'abilità maggiore. Solo la lettura analitica dei risultati ci è dunque utile ai fini della valutazione.

4. Discussione e conclusioni

Il presente lavoro descrive la progettazione di un modello di feedback automatico connesso all'utilizzo di una prova adattativa computer based in un contesto di valutazione formativa. Il lavoro è un approfondimento di una ricerca precedente mirata a progettare, costruire e somministrare una prova adattativa multilivello e a verificarne le migliori proprietà misuratorie rispetto alle classiche prove lineari (Botta, 2021a). Nonostante non sia stato ancora possibile testare sul campo l'efficacia del modello di feedback automatico proposto si ritiene che l'analisi delle sue caratteristiche e potenzialità possano già offrire una nuova prospettiva sull'uso delle prove adattative anche come strumenti per la valutazione formativa oltre che a garanzia di una maggiore equità della valutazione di sistema (Botta, 2019). L'uso di prove computer based e online è in aumento sia nell'istruzione secondaria sia nelle università. Il feedback di questi test usualmente

si limita a fornire le risposte corrette alle domande, e più di rado offre alcuni esempi di domande simili da utilizzare per esercitarsi, come proposto anche nella piattaforma Invalsi Open³. Ciò che caratterizza il modello presentato è invece la combinazione di prove adattative e feedback formativo, studiato per una restituzione individuale dei risultati che sia strettamente connessa al livello di preparazione dello studente nel momento in cui svolge la prova. Come noto, infatti (Weiss, 1985; Botta, 2021a), la somministrazione a uno studente di domande per lui troppo facili o troppo difficili è poco informativa e un feedback basato su di esse sarebbe in ogni caso di scarsa utilità. Nel primo caso infatti gli forniremmo informazioni su contenuti e abilità già stabilmente acquisiti e nel secondo su altri ampiamente al di fuori della sua zona di sviluppo prossimale (Vygotskij, 1987). Nelle prove adattative gli studenti vengono messi alla prova al confine tra ciò che hanno già fatto proprio e ciò che sono pronti ad approssiare. Si tratta di un confine importante, poiché a questo livello gli studenti sono ben motivati e non si lasciano scoraggiare da domande sufficientemente difficili da essere sfidanti o demotivare da domande che trovano facilmente accessibili. È a questo livello dunque che possiamo intervenire per migliorare la loro comprensione di un argomento.

A livello macroscopico inoltre, nell'ipotesi di un uso "libero" della banca di item da parte delle scuole per la predisposizione di prove adattative di vario grado, da somministrare in base alle loro esigenze, come uno strumento fra tanti altri a disposizione dei docenti, l'associazione ad ogni singolo item dell'ambito di contenuto, della dimensione e dell'obiettivo specifico consentirà loro di selezionare le domande non solo in base alla difficoltà ma anche in relazione al curriculum di istituto. Viceversa, in un uso della prova a livello di valutazione di sistema, questo approccio permetterà, ad esempio, di identificare quegli obiettivi che non sono stati raggiunti perché non previsti all'interno del curriculum, magari perché procrastinati nel tempo a vantaggio di altri obiettivi o esigenze emergenti dal contesto. Questo renderà accettabile, trasparente e chiaro ciò che diversamente potrebbe apparire come un fallimento, del singolo, della classe o della scuola. Inoltre una scelta di questo tipo potrebbe restituire alle scuole quel patrimonio informativo perso con il passaggio dal cartaceo al computer based, consentendo comunque di preservare la riservatezza delle specifiche domande.

In prospettiva, sempre nel contesto dei progetti della Sapienza, si sta lavorando alla realizzazione di una prova adattativa computer based per l'ultima classe della scuola primaria e la prima classe della secondaria di I grado, con l'obiettivo di farne uno strumento liberamente utilizzabile dalle scuole. In quest'ambito si prevede di testare anche l'efficacia e il gradimento del modello di restituzione individuale proposto in questo articolo.

Conflict of interests

The author declares no conflict of interest.

References

- Barker, T. (2010). An automated feedback system based on adaptive testing: Extending the model. *International Journal of Emerging Technologies in Learning (iJET)*, 5(2), 11-14. <https://doi.org/10.3991/ijet.v5i2.1235>
- Bennett, R. E. (2011) Formative assessment: a critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), 5-25. <https://doi.org/10.1080/0969594X.2010.513678>
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)*, 21(1), 5-31. <https://doi.org/10.1007/s11092-008-9068-5>
- Bloom, B. S. (1956). *Taxonomy of educational objectives. Handbook 1, Cognitive domain: the classification of educational goals*. London: Longman.
- Botta, E. (2021a). *Sperimentazione di un modello adattativo multilivello per la stima delle abilità in matematica nelle rilevazioni su larga scala*. Roma: Nuova Cultura.

3 https://invalsi-areaprove.cineca.it/index.php?get=static&pag=esempi_prove_grado_10, ultimo accesso 28/2/2023

- Botta, E. (2021b). Percorsi secondari di una prova adattativa multilivello e valutazione formativa. *Excellence and Innovation in Learning and Teaching-Open Access*, 6(2), 58-73. <https://doi.org/10.3280/exioa2-2021oa13019>
- Botta, E. (2019). Prove standardizzate ed equità. In G. Benvenuto, P. Sposetti, G. Szpunar, *Tutti i bisogni educativi sono speciali. Riflessioni, ricerche, esperienze didattiche* (n. 14, pp. 163-174). Roma: Nuova Cultura.
- Brown, G., Bull, J., & Pendlebury, M. (1997). *Assessing student learning in Higher Education*. London: Routledge Falmer.
- Charman, D. (2002). Issues and impacts of using computer-based assessments (CBAs) for formative assessment. In S. Brown, P. Race and J. Bull, *Computer-assisted assessment in Higher Education*, 85-93. London: Kogan Page.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>
- Hattie, J. (2016). *Apprendimento visibile, insegnamento efficace: Metodi e strategie di successo dalla ricerca evidence-based*. Trento: Erickson.
- INVALSI (2018). *Il Quadro di riferimento delle prove di Matematica del Sistema nazionale di valutazione*, URL: https://invalsi-areaprove.cineca.it/docs/file/QdR_MATEMATICA.pdf, data di consultazione 28/02/2023.
- Lilley, M., & Barker, T. (2007). Students' perceived usefulness of formative feedback for a computer-adaptive test. In *Proceedings of the 5th European Conference on elearning: ECEL* (p. 208). Academic Conferences Limited.
- Lilley, M., Barker, T., & Britton, C. (2005). *Automated Feedback for a Computer-Adaptive Test: A Case Study*. URL: <https://uhra.herts.ac.uk/handle/2299/2019>.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- MIUR (2010a). *Indicazioni nazionali riguardanti gli obiettivi specifici di apprendimento concernenti le attività e gli insegnamenti compresi nei piani degli studi previsti per i percorsi liceali*, Roma. URL: <https://www.gazzettaufficiale.it/gunewsletter/dettaglio.jsp?service=1&datagu=2010-12-14&task=dettaglio&numgu=291&redaz=-010G0232&tmstp=1292405356450>
- MIUR (2010b). *Istituti tecnici: linee guida per il passaggio al nuovo ordinamento*, Roma. URL: <https://www.gazzettaufficiale.it/eli/id/2010/09/22/10A11375/sg>
- MIUR (2019). *Linee guida per favorire e sostenere l'adozione del nuovo assetto didattico e organizzativo dei percorsi di istruzione professionale*. URL: <https://www.miur.gov.it/-/linee-guida-per-favorire-e-sostenere-l-adozione-del-nuovo-assetto-didattico-e-organizzativo-dei-percorsi-di-istruzione-professionale>
- Morgan, C., Dunn, L., Parry, S., & O'Reilly, M. (2004). *The student assessment handbook*. London: Routledge Falmer.
- Sly, L., & Rennie, L. J. (2002). Computer managed learning as an aid to formative assessment in higher education software. In S. Brown, P. Race & J. Bull, *Computer-assisted assessment in Higher Education* (pp. 113-120). London: Kogan Page.
- Steven, C., & Hesketh, I., (2002). Increasing learner responsibility and support with the aid of adaptive formative assessment using QM designer software. In S. Brown, P. Race & J. Bull, *Computer-assisted assessment in Higher Education*, 103-112. London: Kogan Page.
- Roos, L. L., Plake, B. S., & Wise, S. L. (1992). *The Effects of Feedback in Computerized Adaptive and Self-Adapted Tests*.
- Vygotskij, L. S. (1987). *Il processo cognitivo*. Torino: Universale Scientifica Boringhieri.
- Wainer, H., Dorans, N. J., Flaugher, R., Green, B. F., & Mislevy, R. J. (2000). *Computerized adaptive testing: A primer*. New York: Routledge.
- Weiss, D. J., & Kingsbury, G. G. (1984). Application of computerized adaptive testing to educational problems. *Journal of Educational Measurement*, 21(4), 361-375. <https://doi.org/10.1111/j.1745-3984.1984.tb01040.x>
- Weiss, D. J. (1985). Adaptive testing by computer. *Journal of consulting and clinical psychology*, 53(6), 774-789. <https://doi.org/10.1037/0022-006X.53.6.774>