Form@re
Open Journal per la formazione in rete

articoli

The influence of the time variable on Invalsi Tests of math

L'influenza della variabile tempo nelle Prove Invalsi di matematica

Stefania Pinnelli[a], Marina De Nunzio[b,1]

[a] *Università del Salento, stefania.pinnelli@unisalento.it*
[b] *Pedagogista, marinadenunzio@libero.it*

Abstract

The essay presents the results of a survey aimed at evaluating how the *time* variable influences the performance of Invalsi math tests in the Italian secondary school. It is worth to say that these tests are intended to detect not specific knowledge or skills but how students can use what he/she has learnt in order to solve new problems. The survey wants to find out in which terms the *time available* variable influences a different expression of competences. Indeed it offers a pedagogical-didactic interpretation space that gives qualitative and not just quantitative meaning to results.

Keywords: Invalsi; time available; math; secondary school; standardized test.

Abstract

Il contributo espone i risultati di un'indagine volta a valutare l'influenza della variabile *tempo* nelle performance alle Prove Invalsi di matematica nella scuola secondaria di primo grado. Premesso che tali prove non hanno l'obiettivo di rilevare specifiche conoscenze o abilità, ma come l'allievo riesce ad usare ciò che ha appreso per risolvere problemi nuovi, lo scopo dell'indagine è rilevare in che termini la *variabile tempo* influenzi una differente espressione delle competenze e, pertanto, offrire uno spazio di interpretazione pedagogico-didattica che dia senso qualitativo e non solo quantitativo ai risultati.

Parole chiave: Invalsi; tempo; matematica; scuola secondaria; performance.

[1] L'indagine è frutto di un lavoro congiunto tra gli autori, tuttavia Pinnelli ha curato in particolare i paragrafi 1, 5 e 6 e De Nunzio i paragrafi 2, 3 e 4. Si ringraziano le scuole che hanno aderito al progetto.

© 2018 Author(s). Open Access article distributed under CC BY-NC-ND 4.0

1. Introduzione

La valutazione rappresenta sempre uno dei principali temi di dibattito all'interno della riflessione pedagogica e nel contesto scolastico: la valutazione degli alunni e la valutazione del sistema scolastico sono due temi che nell'attuale sistema organizzativo ministeriale sono correlati. Le Prove dell'Istituto Nazionale per la Valutazione del Sistema di Istruzione e Formazione (Invalsi) e il Rapporto di Autovalutazione delle scuole, espresso nel RAV e monitorato dal Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR), costituiscono alcuni esempi.

L'Invalsi nasce nel 1999 come trasformazione del Centro Europeo dell'Educazione (CEDE) proprio con l'intento di costruire un sistema di valutazione nazionale, finalizzato al miglioramento della scuola stessa.

Tra le altre l'Invalsi costruisce annualmente la prova da somministrare durante gli Esami di Stato alla fine della scuola secondaria di primo grado: si tratta di "una prova scritta, a carattere nazionale, volta a verificare i livelli generali e specifici di apprendimento conseguiti dagli studenti" (L. n. 176/2007, art 4-ter).

Tale prova, denominata Prova Nazionale, è entrata ufficialmente in vigore dall'anno scolastico 2008-2009 e sino all'anno 2016-2017 ha concorso nella valutazione finale dell'esame del primo ciclo di istruzione, rappresentando "uno strumento per la certificazione degli apprendimenti degli studenti" (Fondazione Giovanni Agnelli, 2014, p. 68). A partire dall'anno scolastico 2017-2018, a seguito della Riforma della legge n. 107/2015, i risultati di tale prova costituiscono un prerequisito per l'ammissione agli Esami di Stato (D.M. n. 741/2017, art. 2).

Considerato il valore in termini di impatto sulla lettura delle competenze degli alunni e considerato lo stress emotivo e cognitivo che tali prove provocano in alunni, scuole e famiglie, la presente ricerca ha inteso indagare quanto la fotografia rimandata dalla Prova Nazionale di matematica possa rappresentare fedelmente gli apprendimenti degli studenti negli ambiti di indagine, essendo la stessa una prova a tempo.

L'obiettivo della ricerca, in particolare, ha inteso verificare se il tempo concesso ai ragazzi per lo svolgimento della Prova comporti un cambiamento in relazione alle performance. L'ipotesi è che le competenze, seppure riconducibili a comportamenti competenti traducibili in situazioni di valutazioni, siano soggette a tempi di realizzazione personali e a processi di elaborazione difficilmente riconducibili a standard.

L'obiettivo della ricerca, pertanto, non è stato interrogarsi sulla validità di tali prove, né sulla loro funzionalità nei processi di miglioramento, ma verificare se *la misura dell'apprendimento* dell'allievo ottenuta con questi strumenti sia *parziale o totale* nell'ottica della certificazione delle competenze e delle conoscenze prevista dall'Esame di Stato.

Uno studio di Chamberlain, Daly & Spalding del 2011, teso a esplorare le cause scatenanti l'ansia da esame, ha evidenziato che tra di esse si colloca anche *il tempo disponibile per completare la prova*. In particolare, avere tempo contingentato per effettuare la prova può avere un effetto positivo, perché stimola lo studente a non disperdersi, ma anche uno negativo, perché lo studente può vivere situazioni di panico e confusione.

Il tempo assegnato per effettuare una prova o, comunque, la percezione che di esso hanno gli studenti può quindi essere un elemento da tenere in considerazione perché potrebbe alterare la misurazione degli apprendimenti.

Ricerche internazionali (Lesaux, Pearson & Siegel, 2006; Sireci, Li & Scarpati, 2003; Zuriff, 2000) hanno inoltre dimostrato che avere più tempo per lo svolgimento di una prova d'esame permette agli esaminati, con e senza difficoltà, di raggiungere punteggi maggiori ai test.

Inoltre le Prove Invalsi di matematica hanno lo scopo di valutare "non tanto e non solo la conoscenza teorica di alcuni concetti, ma la capacità di utilizzarli nell'esempio concreto" (Invalsi, 2014, p. 15). In particolare, esse intendono indagare gli ambiti e i processi contenuti nelle *Indicazioni per il curricolo per la scuola dell'infanzia e per il primo ciclo d'istruzione* (Ministero della Pubblica Istruzione, 2007).

Focalizzando l'attenzione sui processi indagati dalle prove (Figura 1), è facile rendersi conto che tra di essi alcuni prevedono non un'applicazione algoritmica (meccanica) di formule, ma attività che possono caratterizzarsi come un processo costruttivo di ricerca della soluzione.

Codice	**Processi**
1	Conoscere e padroneggiare i contenuti specifici della matematica (oggetti matematici, proprietà, strutture...).
2	Conoscere e padroneggiare algoritmi e procedure (in ambito aritmetico, geometrico...)
3	Conoscere e padroneggiare diverse forme di rappresentazione e sapere passare da una all'altra (verbale, scritta, simbolica, grafica, ...).
4	Sapere risolvere problemi utilizzando gli strumenti della matematica (individuare e collegare le informazioni utili, confrontare strategie di soluzione, individuare schemi risolutivi di problemi come ad esempio sequenza di operazioni, esporre il procedimento risolutivo...).
5	Sapere riconoscere in contesti diversi il carattere misurabile di oggetti e fenomeni e saper utilizzare strumenti di misura (saper individuare l'unità o lo strumento di misura più adatto in un dato contesto, saper stimare una misura,...).
6	Acquisire progressivamente forme tipiche del pensiero matematico (congetturare, verificare, giustificare, definire, generalizzare, ...).
7	Utilizzare la matematica appresa per il trattamento quantitativo dell'informazione in ambito scientifico, tecnologico, economico e sociale (descrivere un fenomeno in termini quantitativi, interpretare una descrizione di un fenomeno in termini quantitativi con strumenti statistici o funzioni, utilizzare modelli matematici per descrivere e interpretare situazioni e fenomeni, ...).
8	Saper riconoscere le forme nello spazio (riconoscere forme in diverse rappresentazioni, individuare relazioni tra forme, immagini o rappresentazioni visive, visualizzare oggetti tridimensionali a partire da una rappresentazione bidimensionale e, viceversa, rappresentare sul piano una figura solida, saper cogliere le proprietà degli oggetti e le loro relative posizioni, ...).

Figura 1. Processi attivati nella risoluzione dei quesiti (Invalsi, 2012, p. 40).

Ricordiamo, infatti, che i quesiti Invalsi "non puntano a rilevare specifiche conoscenze o abilità apprese a scuola ma a rilevare come l'allievo riesce ad usare ciò che ha appreso a scuola per risolvere problemi nuovi e non routinari (concetto noto come competenza). […] Questa caratteristica è indispensabile per rilevare non il semplice ricordo mnemonico di nozioni o la ripetizione meccanica di procedure, ma come l'allievo usa le procedure e nozioni apprese quando si trova a dover risolvere un problema. […] I quesiti richiedono quindi la messa in campo di capacità di ragionamento più o meno complesso" (Trinchero, 2014, p. 37).

Anche all'interno delle prove usate per la presente ricerca ricorrono quesiti di differente grado di complessità a seconda che rimandino a conoscenze scolastiche o che richiedano procedure di ragionamento[2].

Quanto tempo sia necessario per individuare/elaborare una soluzione di una situazione problematica non è semplice da quantificare. Le Prove Invalsi, avendo come obiettivo la comparazione, devono necessariamente stabilire un tempo standard partendo da indagini statistiche. Ma se l'obiettivo dell'Invalsi è misurare il livello di apprendimento di un allievo, il limite temporale imposto può rappresentare un ostacolo all'estrinsecazione della sua competenza?

Al di là del tempo medio utilizzato da un campione di alunni per la risoluzione di un quesito, siamo sicuri che una fascia di alunni non riuscirebbe ad ottenere risultati adeguati avendo a disposizione un tempo maggiore? E, soprattutto, è corretto ritenere questi risultati meno validi di quelli raggiunti con il tempo dato?

2. Disegno della ricerca

L'indagine empirica è stata articolata secondo il seguente disegno di ricerca.

2.1. Selezione classi campione

Per la ricerca sono stati coinvolti alunni delle classi terze di scuole secondarie di primo grado. La proposta di partecipare al progetto è stata inviata a sette scuole del territorio (Lecce e provincia); tre di quelle scuole hanno dato disponibilità a partecipare al progetto. Delle tre scuole aderenti al progetto sono state coinvolte otto classi terze selezionate dalle stesse scuole. I criteri suggeriti ai docenti per l'individuazione delle classi sono stati: classi eterogenee per performance evitando pertanto classi eccellenti o particolarmente deboli; eterogeneità per genere; classi che non avessero usato già le Prove Invalsi selezionate per la ricerca. Complessivamente, hanno partecipato al progetto otto classi, per un totale di 188 studenti.

	MASCHI	FEMMINE	TOT
SOTTOGRUPPO A	31	32	63
SOTTOGRUPPO B	34	40	74
TOT	65	72	137

Figura 2. Distribuzione degli allievi nei sottogruppi.

In fase di analisi dei risultati, il campione ha subìto una riduzione. Infatti il campione non probabilistico finale è formato da 137 studenti, nello specifico 72 femmine e 65 maschi. Sono stati esclusi dal gruppo iniziale gli alunni che non hanno partecipato a tutte le somministrazioni o alunni le cui prove erano incomplete o non valutabili, alunni in situazione di disabilità. Ciascuna delle 8 classi, per un totale di 137 alunni, è stata distinta in due sottogruppi, anche essi assortiti in modo da distribuire equamente le performance

[2] Si rimanda, a titolo d'esempio, ai quesiti D9 e D22 della Prova Invalsi del 2010/11: il quesito D22 richiede di identificare la corretta applicazione di una formula, mentre il quesito D9 rimanda ad un ragionamento molto più articolato.

scolastiche degli studenti[3]. Si sono così costruiti il Sottogruppo A e il Sottogruppo B (Figura 2).

2.2. Strumenti utilizzati

Per la valutazione delle performance matematiche sono state selezionate due[4] Prove Invalsi degli anni precedenti. In particolare si è scelto di usare quella relativa all'anno scolastico 2010/2011 (denominata *prova α*) e quella relativa all'anno scolastico 2011/2012 (denominata *prova β*). Si sono altresì utilizzate le relative griglie di correzione e di attribuzione dei punteggi.

2.3. Indagine sul campo

Nel mese di Maggio 2017 ai due sottogruppi sono state proposte due somministrazioni di Prove Invalsi (Prova α e Prova β) in due giorni distinti, ma ravvicinati, affinché tra le due somministrazioni non si potesse verificare un *significativo* aumento delle competenze matematiche degli allievi. Ogni gruppo, pertanto, ha svolto entrambe le prove avendo a disposizione lassi di tempo differenti. In particolare, in Figura 3 è riportata l'organizzazione delle somministrazioni.

	1° SOMMINISTRAZIONE **Tempo: 120 minuti**	**2° SOMMINISTRAZIONE** **Tempo: 75 minuti**
SOTTOGRUPPO A	Prova α	Prova β
SOTTOGRUPPO B	Prova β	Prova α

Figura 3. Organizzazione delle somministrazioni.

La somministrazione è avvenuta all'interno delle aule dei ragazzi, alla presenza di un docente della classe (non necessariamente quello di matematica) e di almeno un ricercatore. I banchi dei ragazzi sono stati separati e posizionati per file; i ragazzi sono stati fatti sedere in modo da non avere vicino compagni che stavano effettuando la stessa prova (infatti nell'aula c'erano ragazzi che svolgevano la Prova α e ragazzi che svolgevano la Prova β).

Prima di affrontare la prova, ai ragazzi è stato spiegato che stavano per affrontare una simulazione di Prova Invalsi come quella che avrebbero affrontato durante gli esami. Si è spiegato che il motivo della somministrazione aveva fini di ricerca (ma si è taciuto il motivo della stessa per non influenzare i ragazzi), tuttavia per evitare che i ragazzi prendessero sotto gamba l'attività, si è spiegato loro che i ricercatori avrebbero avuto i risultati di prove in modalità anonima, ma che i docenti avrebbero avuto i risultati delle prove.

Si è quindi provveduto ad assegnare ad ogni ragazzo un codice identificativo.

[3] La composizione dei sottogruppi ha seguito il criterio di distribuzione omogena delle differenti fasce di voto scolastico degli studenti in modo da rendere i sottogruppi ugualmente eterogenei.

[4] Si è scelto di selezionare due prove per poter tenere sotto controllo il fattore *prova*, ovvero si voleva controllare che gli eventuali effetti del tempo disponibile nelle risposte ai quesiti non fossero legati ad una specifica prova ma fossero generalizzabili ad una qualsiasi.

3. Analisi dei risultati

Il primo tipo di analisi svolto ha preso in considerazione le medie dei punteggi conseguiti dai ragazzi ad entrambe le prove. La Figura 4 riporta le medie delle prove.

	1° SOMMINISTRAZIONE (120 minuti)	**2° SOMMINISTRAZIONE (75 minuti)**
Sottogruppo A (63 alunni)	Prova α 36,84	Prova β 25,9
Sottogruppo B (74 alunni)	Prova β 26.76	Prova α 34.64

Figura 4. Media dei punteggi delle due prove nelle due somministrazioni.

Ad un primo sguardo si nota immediatamente che nel Sottogruppo A la performance svolta in più tempo ha fatto registrare punteggi più alti in maniera significativa rispetto alla prova svolta in 75 minuti. Questo, pertanto, sembrerebbe confermare che ad una diminuzione di tempo a disposizione per svolgere la prova corrisponda una minore *possibilità* di dimostrare i propri apprendimenti.

Se però si osservano le medie delle due prove ottenute dal Sottogruppo B, possiamo notare che si verifica la condizione inversa, ovvero che al diminuire del tempo le performance sono migliorate. Cosa può aver provocato dei risultati così contraddittori?

Per poter comprendere a cosa sia dovuta questa differente situazione occorre confrontare le due prove, in particolare i punteggi medi della Prova α in 120 minuti con quelli della Prova β svolti nello stesso tempo.

Come indicato nella Figura 4, i punteggi medi della Prova α, nelle due somministrazioni sono generalmente più alti dei punteggi medi della Prova β. Un t-test rileva differenze significative tra le due prove in entrambe le somministrazioni. Nella prima somministrazione (120 minuti), i punteggi medi sono pari a 36,84 e 26,76 rispettivamente per α e β, con una differenza significativamente diversa da zero (t=6,968; p=0,000). Allo stesso modo, nella seconda somministrazione (75 minuti) lo scarto tra i punteggi medi di α e β denota una differenza significativa tra i due test (t=-6,062; p=0,000). Pertanto la Prova α e la Prova β sono significativamente diverse tra loro e la Prova α è più facile (poiché i risultati in entrambe le modalità di somministrazione sono significativamente più alti) rispetto alla β.

Tale primo dato, seppure interessante perché evidenzia il fattore *fortuna* nella valutazione conclusiva degli studenti, non permette di verificare se il tempo della prova abbia influito sui risultati della performance per lo stesso sottogruppo di soggetti. È pertanto necessario confrontare i risultati alla stessa prova svolta in tempi differenti dai due sottogruppi di studenti, facendo affidamento sul fatto che i campioni di partenza siano stati costruiti distribuendo in modo omogeneo i ragazzi per livello di apprendimento, per genere e per esperienza sulle prove.

In Figura 5 sono rappresentate le distribuzioni dei punteggi conseguiti alla Prova α ottenuti nella somministrazione a 120 e a 75 minuti. Dai boxplot si evince che la prova effettuata in minor tempo ha fatto registrare punteggi mediamente inferiori. Analizzando le medie dei risultati, possiamo notare che tra di esse esiste uno scarto, ma un t-test ha evidenziato scarsa significatività di tale differenza (t=1,35; p=0,17).

Prendendo in considerazione la Prova β svolta in 120 minuti dal Sottogruppo B e in 75 dal Sottogruppo A (la distribuzione dei punteggi ottenuti in questa prova viene rappresentata

nella Figura 6), possiamo rilevare che la differenza tra i punteggi medi delle due prove è ancora più piccola. Anche in questo caso un t-test ha dimostrato la non significatività di tale scarto (t=0,67; p=0,49).

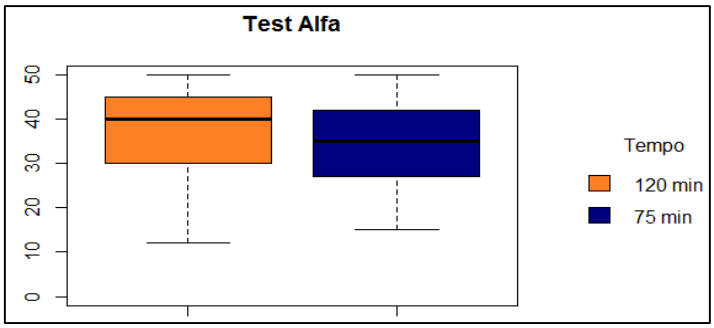

Figura 5. Distribuzione dei punteggi della Prova α nelle somministrazioni in tempi differenti.

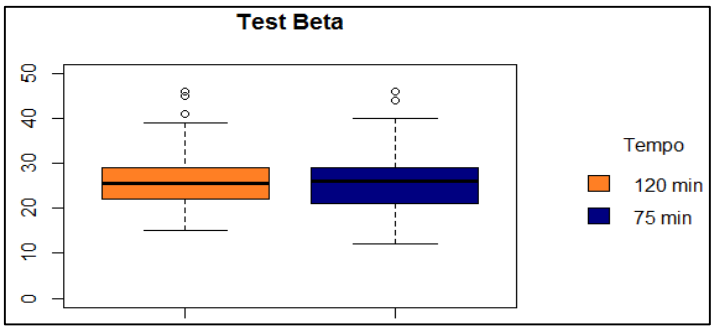

Figura 6. Distribuzione dei punteggi della Prova β nelle somministrazioni in tempi differenti.

SOTTOGRUPPO A - PROVA α (120 minuti)

Voti	Conteggio di frequenza
4	2
5	11
6	9
7	6
8	9
9	20
10	6
Totale complessivo	**63**

SOTTOGRUPPO A - PROVA β (75 minuti)

Voti	Conteggio di frequenza
4	16
5	24
6	14
7	3
8	4
9	1
10	1
Totale complessivo	**63**

SOTTOGRUPPO B - PROVA β (120 minuti)

Voti	Conteggio di frequenza
4	12
5	39
6	8
7	5
8	8
9	1
10	1
Totale complessivo	**74**

SOTTOGRUPPO B - PROVA α (75 minuti)

Voti	Conteggio di frequenza
4	7
5	13
6	12
7	8
8	15
9	12
10	7
Totale complessivo	**74**

Figura 7. Frequenza dei voti in entrambe le prove nelle due somministrazioni.

In sintesi, quindi, attraverso l'analisi delle medie è emerso che il tempo assegnato per effettuare una Prova Invalsi non sembra influire sulle performance degli allievi.

Una seconda analisi ha preso in considerazioni i *voti* finali conseguiti dai ragazzi alle prove attribuiti sulla base delle tabelle fornite dall'Invalsi. Più specificamente, è interessante osservare le frequenze dei voti associati alle due somministrazioni distinte per le due prove (la Figura 7 riporta le specifiche), in particolare è utile andare a osservare come varia il numero di studenti con voti sufficienti (dal 6 in su) nelle quattro combinazioni di situazioni.

I ragazzi del Sottogruppo *A* che hanno raggiunto punteggi sufficienti nella prima somministrazione (Prova α) sono 50, mentre quelli che hanno raggiunto almeno la sufficienza nella seconda somministrazione (Prova β) sono stati 23. L'immagine sottostante (Figura 8) mostra le percentuali di prestazioni sufficienti per il Sottogruppo A nelle due prove.

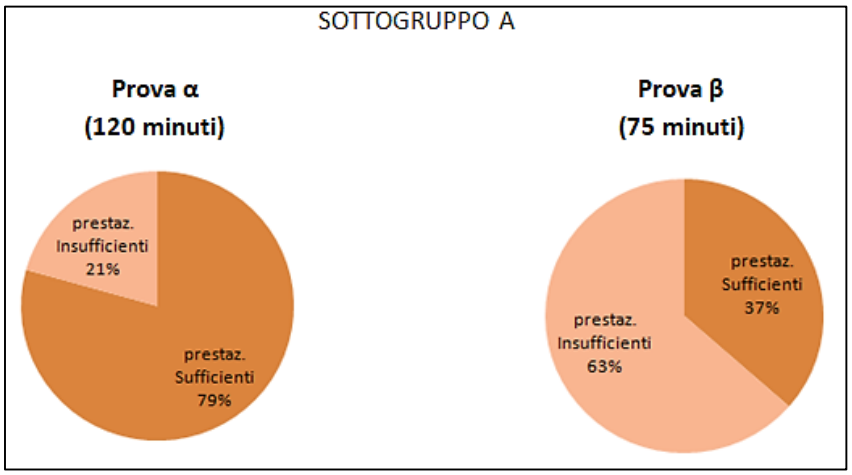

Figura 8. Percentuali prestazioni sufficienti alle prove svolte dal Sottogruppo A.

Per il Sottogruppo B il numero degli studenti con sufficienza è passato da 23 in prima somministrazione (Prova β) a 54 in seconda somministrazione (Prova α). La Figura 9 mostra le percentuali di prestazioni sufficienti per il Sottogruppo B nelle due prove.

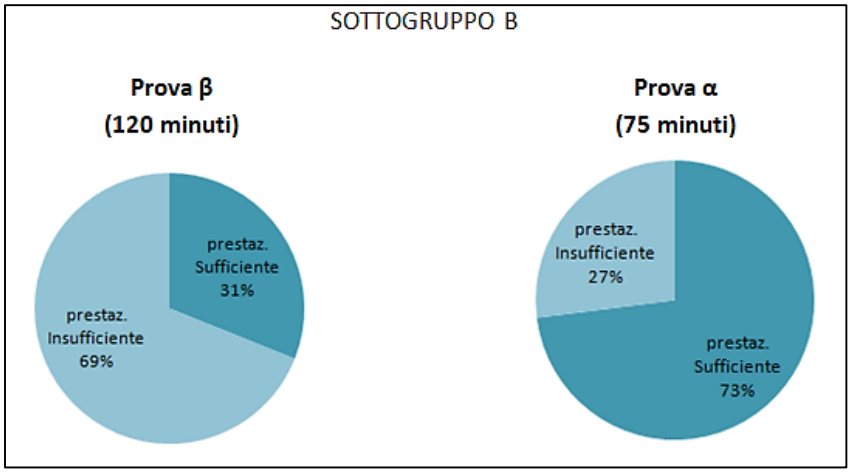

Figura 9. Percentuali prestazioni sufficienti alle prove svolte dal Sottogruppo B.

Anche con questo tipo di confronto si evidenzia la discrepanza tra le due prove: la Prova β ha evidentemente messo in difficoltà gli studenti.

Se invece confrontiamo i dati relativi alla stessa prova ma svolta in tempi differenti, possiamo osservare che non si evidenziano cambiamenti sostanziali, quindi la variabile *tempo* non sembra aver influito sui voti delle performance.

Si è tuttavia condotta un'ulteriore analisi (sia per la Prova α, che per la Prova β) che ha preso in considerazione i *coefficienti di difficoltà*[5] dei singoli item delle prove. In particolare l'obiettivo dell'indagine era controllare se questi coefficienti subissero delle variazioni al variare del tempo assegnato per lo svolgimento della prova. Si è pertanto proceduto a calcolare distintamente i parametri di difficoltà dei vari item, sia per la Prova α che per la β, a seconda che queste fossero state svolte in 120 o 75 minuti. Tali coefficienti sono stati ricavati attraverso un'analisi delle performance dei ragazzi elaborata attraverso il Modello di Rasch. In Figura 10 vengono mostrati i parametri calcolati con la Rasch Analysis e quelli stimati direttamente da Invalsi, indicati nei rapporti tecnici relativi alle specifiche prove; inoltre sono inserite le medie di tali coefficienti.

PROVA α

Item	Prova 120'	Prova 75'	Stime Invalsi
MEDIA	**-0,59**	**-0,42**	**-0,63**

PROVA β

Item	Prova 120'	Prova 75'	Stime Invalsi
MEDIA	**0,02**	**0,15**	**-0,07**

Figura 10. Medie dei parametri di difficoltà distinti per ogni prova nelle due soluzioni temporali, valutati secondo il Modello di Rasch e stime riportate dai documenti Invalsi.

In relazione ad entrambe le prove, confrontando le media dei tassi di difficoltà degli item a 120 e 75 minuti si evidenzia che la difficoltà aumenta quando il tempo per espletare la prova diminuisce. Inoltre la distribuzione della differenza dei valori non ha media nulla (p= 0,012), il che significa che la variazione non è attribuibile al caso, quindi la differenza è significativa (distribuzioni non gaussiane in base al test di Kolmogorov-Smirnov; test adottato *sign test* per campioni appaiati) (Figura 11).

PROVA α	PROVA β
• confronto Alfa 120 vs Invalsi: (p=0,73) ipotesi nulla non rigettabile; • confronto Alfa 75 vs Invalsi: (p=0,04) ipotesi nulla rigettata cioè differenza significativa.	• confronto Beta 120 vs Invalsi: (p=0,27) ipotesi nulla non rigettabile; • confronto Beta 75 vs Invalsi: (p=0,0002) ipotesi nulla rigettata cioè differenza significativa.

Figura 11. Report risultati del confronto tra le medie dei parametri di difficoltà delle prove svolte in 120 o 75 minuti con le stime Invalsi.

[5] Nella costruzione delle Prove Invalsi, dopo una prima valutazione sintetica sull'attendibilità generale di ogni Prova, l'analisi viene approfondita applicando i metodi della Rasch Analysis (RA). "La RA consente inoltre di focalizzare l'attenzione non solo sull'intera prova, ma anche sulle singole domande, poiché per ogni quesito viene stimato il livello di difficoltà sulla stessa scala sulla quale è misurato il livello di performance di ogni allievo" (Invalsi, 2011, p. 14).

Dal confronto tra i coefficienti di difficoltà delle prove a 120 e 75 minuti con le stime Invalsi (analisi di significatività statistica effettuata con *sign test)* è emerso che le stime di difficoltà interna degli item Invalsi si sovrappongono maggiormente a quelle stimate per la prova svolta in 120 minuti, invece si discostano significativamente da quelle svolte in 75 almeno per il campione di alunni preso in esame.

Ciò significa che, per converso, con più tempo gli studenti esaminati possono raggiungere i risultati attesi per un campione standard Invalsi in entrambe le prove.

4. Discussione

Come già ampiamente descritto nel paragrafo precedente, l'analisi delle medie dei punteggi alle prove ha fatto emergere dati contrastanti all'interno dei due sottogruppi campione. In particolare, infatti, se si considerasse esclusivamente ciò che è avvenuto nelle performance del Sottogruppo A si sarebbe tentati di confermare quanto ipotizzato all'avvio della ricerca, ovvero che il tempo concesso per effettuare una Prova Invalsi influisce sull'esito della performance stessa. Ma i risultati emersi nel Sottogruppo B contrastano con ciò perché manifestano una situazione opposta. Cosa dire quindi sulla relazione tra tempo e performance? Come è emerso attraverso le analisi effettuate con il t-test, non è possibile comparare le performance degli studenti di uno stesso sottogruppo nelle due prove perché le due prove sono significativamente diverse. L'analisi effettuata confrontando su ciascuna prova i risultati dei due sottogruppi al variare del tempo ha dimostrato che le variazioni tra i risultati dei ragazzi esistono (più tempo ha prodotto punteggi/voti maggiori) ma non sono significative.

Andando invece a considerare le componenti interne delle prove, ossia l'indice di difficoltà degli item, le cose cambiano. L'analisi condotta sui coefficienti di difficoltà degli item delle prove ha fatto emergere che al diminuire del tempo i coefficienti di difficoltà delle prove aumentano, anche in maniera significativa. Inoltre i coefficienti di difficoltà degli item stimati dall'Invalsi (che sono tarati su prove da svolgere in 75 minuti) sono sovrapponibili a quelli calcolati tramite la Rasch Analysis a partire dai risultati delle prove svolte in 120 minuti dagli studenti presi in esame nella presente ricerca. Questo può significare che gli studenti delle classi in esame nel presente studio non raggiungono gli standard Invalsi in 75 minuti (quindi manifestano prestazioni minori rispetto a quelle delle classi campione), ma riescono a farlo avendo a disposizione tempi maggiori. Il che sembrerebbe confermare l'ipotesi di partenza ovvero che il tempo della prova è una variabile significativa nell'ottica dell'estrinsecazione delle proprie abilità in ambito matematico.

5. Conclusioni

Dalle analisi svolte prendendo in esame solo i punteggi finali o i voti raggiunti dai ragazzi nelle Prove Invalsi di matematica, al variare del tempo sembrerebbe non esserci una significativa influenza della variabile tempo sui risultati. Tuttavia l'analisi dei coefficienti di difficoltà delle prove evidenzia l'esistenza di una correlazione inversa tra tempo assegnato per lo svolgimento della prova e difficoltà della stessa.

Probabilmente non è possibile evincere tale correlazione solo dall'analisi dei punteggi o dei voti raggiunti dagli allievi perché il metodo, usato da Invalsi, di attribuzione degli stessi, per *inscatolamenti* successivi, riduce l'evidenza di questo dato. Ogni prova, infatti, è

costituita da diverse domande alcune delle quali suddivise in più item. Le domande o gli item sono classificate in diversi blocchi (A, B, C), in relazione al loro grado di difficoltà. Il punteggio finale deriva dalla somma dei sotto-punteggi. Tali sotto-punteggi vengono attribuiti ai blocchi in base non al numero di risposte corrette date ai vari item che compongono ciascun blocco, ma in un *range* di risposte per le quali rispondendo ad un minimo o al massimo degli item il punteggio non cambia. Si consideri, ad esempio, l'attribuzione dei punteggi per il blocco C della Prova α: esso è composto da quattro domande o item e il punteggio massimo assegnabile è cinque. I cinque punti però sì raggiungono anche semplicemente rispondendo correttamente a due item o domande, come previsto dalla tabella Invalsi. Risulta chiaro quindi che il ragazzo che ha svolto tutti e quattro i quesiti del blocco C in realtà ha perso tempo perché sarebbe stato sufficiente svolgere solo due di quei quesiti per ottenere il massimo punteggio, ma ovviamente i ragazzi non conoscono la suddivisione in blocchi e i relativi item che ne fanno parte.

Per tale meccanismo l'analisi globale dei risultati al test non fa emergere una correlazione significativa tra punteggi e tempo. Andando invece a considerare l'indice di difficoltà di ciascun item delle prove tale conclusione è smentita dalla correlazione tra indici di difficoltà delle prove al variare del tempo.

Questo risultato non mette in discussione la validità delle Prove Invalsi che, ripetiamo, hanno scopo censuario, ma fa sorgere qualche perplessità nell'ottica della valutazione delle abilità matematiche.

A tal proposito occorre fare un'ulteriore constatazione: l'analisi svolta ha evidenziato che le prove scelte per questa ricerca hanno una oggettiva differenza in relazione al loro coefficiente di difficoltà. E questo risulta ancora più evidente se si nota che lo stesso sottogruppo di ragazzi manifesta prestazioni estremamente differenti nelle due prove: nel Sottogruppo A le prestazioni sufficienti o più che sufficienti (dal 6 in su) passano dal 79% (Prova α) al 37% (Prova β) del campione e per il Sottogruppo B dal 31% (Prova β) al 73% (Prova α).

Tale constatazione pone un interrogativo: partendo dal presupposto che queste Prove siano Prove Nazionali ovvero sono state strumento di esame rispettivamente negli anni 2011 e 2012 e che "la Prova Invalsi svolta a conclusione dell'Esame di Stato della scuola secondaria di primo grado serve anche a certificare gli apprendimenti degli studenti" (Fondazione Giovanni Agnelli, 2014, p. 68), quale dei valori (punteggi/voti) raggiunti dagli studenti può essere considerato il valore degli apprendimenti degli studenti? Ricordiamo, infatti, che il voto ottenuto dagli allievi alle Prova Nazionale fino al 2017 ha concorso alla valutazione finale dell'esame del primo ciclo di istruzione ed oggi è altrettanto significativa poiché rappresenta un prerequisito per l'ammissione agli Esami di Stato. Inoltre un'analisi dei coefficienti di difficoltà delle Prove Nazionali svolte dal 2010 al 2017 (tali dati sono inseriti nei Rapporti tecnici pubblicati dagli Invalsi annualmente) mostra una notevole variabilità tra le prove, pertanto uno studente con competenze medie[6] può ottenere risultati estremamente diversi pur lavorando al massimo; infatti, oltre alle sue competenze a seconda della prova che è chiamato a svolgere, il tempo a disposizione e il fattore fortuna

[6] Per uno studente poco competente anche la prova più semplice può essere irrisolvibile, mentre per un ragazzo particolarmente abile o plus dotato non sarà certo una prova più complessa a bloccare l'estrinsecazione della sua competenza. Basti considerare i ragazzi che nella nostra analisi non hanno raggiunto la sufficienza nella prova più semplice o quelli che hanno raggiunto il punteggio massimo nella prova più difficile (nove insufficienti nella Prova α e quattro ragazzi con punteggi ottimali nella Prova β).

hanno un ruolo considerevole. Si evidenzia come la gestione del tempo sia un elemento critico per l'emersione delle competenze proprio in quei ragazzi con voti medi che costituiscono la maggior parte degli alunni. Come è noto dall'anno scolastico 2017/2018, a seguito del D.M n. 741/2017, per le prove viene concesso più tempo, passando da 75 a 90 minuti, ma saranno prove da svolgere al computer. Questa novità nella somministrazione, se apparentemente sembra aver colto il problema della *variabile tempo* disponibile per lo svolgimento delle prove e quindi rispondere al bisogno di maggiore calma e riflessione, dall'altra rischierà di vanificare tale dato dovendo fare i conti con lo strumento informatico, con la connessione della rete scolastica oltre alla familiarità stessa dello strumento nella risoluzione di un compito di matematica.

In conclusione, considerato il rapporto che la variabile tempo ha nell'emersione delle competenze matematiche in rapporto alle difficoltà delle prove, considerato che tale elemento incide sulla fascia media dei ragazzi, considerato inoltre che l'elaborazione di processi risolutivi può seguire processi creativi e complessi che fanno parte della specifica attuazione di competenze, si ritiene che affrancare le Prove Invalsi da margini temporali così rigorosi, oltre a contribuire ad abbassare i livelli di ansia individuale, possa contribuire ad una più realistica espressione dei processi di apprendimento.

Bibliografia

Chamberlain, S., Daly, A.S., & Spalding, V. (2011). Assessment and qualifications alliance. *Pastoral care in education*, *29*, 193−205.

Decreto Ministeriale 3 ottobre 2017, n. 741. *Esami di stato conclusivo del primo ciclo di istruzione*.

Fondazione Giovanni Agnelli. (2014). *La valutazione della scuola. A che cosa serve e perché è necessaria all'Italia*. Bari: Laterza.

Invalsi. Istituto Nazionale per la Valutazione del Sistema di Istruzione e Formazione (2011). *Le prove INVALSI 2011, Rapporto tecnico sulle caratteristiche delle prove INVALSI 2011*, http://www.invalsi.it/esamidistato1011/documenti/Rapporto_tecnico_prove_invalsi_2011.pdf (ver. 15.07.2018).

Invalsi. Istituto Nazionale per la Valutazione del Sistema di Istruzione e Formazione (2012). *Rilevazioni Nazionali sugli apprendimenti 2011/12, Il quadro di sistema*. http://www.invalsi.it/snv2012/documenti/Rapporti/Rapporto_rilevazione_apprendimenti_2012.pdf (ver. 15.07.2018).

Invalsi. Istituto Nazionale per la Valutazione del Sistema di Istruzione e Formazione (2014). *Il Decennale delle Prove INVALSI. Esiti, strumenti e riflessioni verso il Sistema Nazionale di Valutazione. Roma, 4 e 5 dicembre 2014.* http://www.invalsi.it/invalsi/doc_eventi/12-2014/4/Documento_DecennaleProveInvalsi.pdf (ver. 15.07.2018).

Legge 25 ottobre 2007, n. 176. *Conversione in legge, con modificazioni, del decreto-legge 7 settembre 2007, n. 147, recante disposizioni urgenti per assicurare l'ordinato avvio dell'anno scolastico 2007-2008 ed in materia di concorsi per ricercatori universitari* (GU Serie Generale n.250 del 26-10-2007).

Legge 13 luglio 2015, n. 107, *Riforma del sistema nazionale di istruzione e formazione e delega per il riordino delle disposizioni legislative vigenti*. (GU Serie Generale n.162 del 15-07-2015).

Lesaux, N.K., Pearson, M.R., & Siegel, L.S. (2006). The effects of time and untimed testing conditions on the reading comprehension performance of adults with reading disabilities. *Reading and Writingvol*, *19*(1), 21−48. https://link.springer.com/article/10.1007/s11145-005-4714-5 (ver. 15.07.2018).

Ministero della Pubblica Istruzione (2007). *Indicazioni per il curricolo per la scuola dell'infanzia e per il primo ciclo d'istruzione*, https://archivio.pubblica.istruzione.it/normativa/2007/allegati/dir_310707.pdf (ver. 15.07.2018).

Runyan, M.K. (1991). The effect of extra time on reading comprehension scores for university students with and without learning disabilities. *Journal of Learning Disabilities*, *24*(2), 104−108.

Sireci, S.G., Li, S., & Scarpati, S. (2003). The effects of test accommodation on test performance: a review of the literature. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.586.2376 (ver. 15.07.2018).

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*, *4*(14). http://www.fupress.net/index.php/formare/article/view/15794 (ver. 15.07.2018).

Zuriff, G.E. (2000). Extra examination time for students with learning disabilities: an examination of the maximum potential thesis. *Applied Measurement in Education*, *3*, 99−117.