

LE PROVE INVALSI 2011

Relazione tecnica sulle caratteristiche psicometrico-misuratorie delle prove INVALSI 2011 sostenute dagli studenti delle classi II e V della scuola primaria, I e III della secondaria di primo grado e della classe II della scuola secondaria di secondo grado

*Rapporto tecnico
sulle caratteristiche
delle prove
INVALSI 2011*

Il gruppo di lavoro assegnato al Servizio Nazionale di Valutazione ed ai Progetti di ricerca finanziati con i fondi straordinari allo stesso correlati è composto da: Roberto Ricci (responsabile), Monica Amici, Andrea Biggera, Emanuela Cuzzucoli, Alessandra Fazio, Elisabetta Figura, Teresa Fiorino, Cristina Lasorsa, Caterina Ponzio.

Hanno collaborato alla redazione del presente rapporto:

Monica Amici, Emiliano Campodifiori, Michele Cardone, Patrizia Falzetti, Elisabetta Figura, Francesca Fortini, Cristina Lasorsa, Letizia Marcucci, Angela Martini, Antonella Mastrogiovanni, Alessia Mattei, Rita Marzoli, Monica Papini, Roberto Ricci, Maddalena Tozzi.

Si ringrazia il Servizio statistico dell'INVALSI che ha garantito in tempi rapidi la predisposizione di un *dataset* con elevati standard qualitativi ed effettuato le analisi proposte nel presente rapporto.

Si ringrazia, inoltre, il *Servizio Hardware e Reti SHR* (Carlo Di Giovamberardino e Antonio Severoni) e il *Servizio Web* (Alessandro Borsella e Stefano Famiglietti) dell'INVALSI per il fondamentale supporto informatico.

Si ringraziano gli Uffici Scolastici regionali e provinciali, i Referenti regionali e provinciali, i Dirigenti scolastici, gli insegnanti, gli studenti, le segreterie delle scuole campione e gli osservatori esterni per il loro impegno e contributo all'indagine.

LE PROVE INVALSI 2011

RAPPORTO TECNICO SULLE CARATTERISTICHE DELLE PROVE INVALSI 2011

| | |
|---|----|
| CAPITOLO I: La costruzione delle prove standardizzate..... | 1 |
| 1. Premessa..... | 1 |
| 2. La costruzione delle prove del Servizio nazionale di valutazione..... | 5 |
| 2.1 I quadri di riferimento per la valutazione | 5 |
| 2.2 Gli autori delle prove..... | 6 |
| 2.3 La costruzione delle prove..... | 6 |
| 2.3.1. L'analisi qualitativa ex ante..... | 7 |
| 2.3.2. Il pre-test..... | 7 |
| 2.3.3. L'analisi quanti-qualitativa ex post: la calibrazione delle domande..... | 9 |
| 2.3.4. La composizione dei fascicoli definitivi..... | 9 |
| CAPITOLO II: La valutazione delle prove standardizzate | 11 |
| 1. La valutazione delle prove del Servizio nazionale di valutazione..... | 11 |
| 2. L'indice α di Cronbach..... | 12 |
| 3. La Rasch Analysis per la stima della difficoltà delle domande..... | 14 |
| 3.1. La scuola primaria | 14 |
| 3.2. La scuola secondaria di primo grado | 19 |
| 3.2.1. La Prova nazionale | 21 |
| 3.3. La scuola secondaria di secondo grado | 23 |
| 4. Il placement relativo delle domande rispetto al punteggio di Rasch dei rispondenti | 26 |
| 5. L'indice di discriminazione statistica | 38 |
| Riferimenti bibliografici..... | 50 |

CAPITOLO I: La costruzione delle prove standardizzate

1. Premessa

La costruzione di una prova standardizzata è il frutto di un lungo e articolato processo tecnico-scientifico, il cui rispetto è fondamentale se si vuole giungere alla definizione di strumenti di misurazione, pur nei limiti insiti negli strumenti stessi, in grado di fornire informazioni solide e significative sui livelli di apprendimento conseguiti dagli allievi ai quali le prove sono state sottoposte.

Ancora prima di entrare nel merito delle prove è opportuno fare preliminarmente chiarezza su alcuni termini che consentono più facilmente di comprendere le finalità delle prove standardizzate, almeno nelle modalità di utilizzo previste dall'INVALSI sulla base del mandato ricevuto dal legislatore. Le prove INVALSI hanno lo scopo principale di misurare i livelli di apprendimento raggiunti dagli studenti italiani relativamente ad alcuni aspetti di base di due ambiti fondamentali: la comprensione della lettura e la matematica. In termini ancora più espliciti, mediante le prove del Servizio nazionale di valutazione (SNV) si vuole giungere a una misura dei livelli di apprendimento nei suddetti ambiti, comunemente ritenuti condizione necessaria per un accesso consapevole alla cittadinanza attiva. Da ciò segue immediatamente che gli ambiti oggetto di misurazione delle prove INVALSI non esauriscono di certo i saperi e le competenze prodotte dalla scuola. Infine, ma non da ultimo, è importante sottolineare l'intenzionalità con la quale si è utilizzato il termine misurare e non valutare, proprio per evitare l'equivoco che potrebbe sorgere da quest'ultimo termine. Infatti, è noto che in campo scolastico la valutazione è qualcosa di più ampio della misurazione, anche se quest'ultima è un elemento imprescindibile per la prima. Le prove INVALSI si pongono quindi come obiettivo di fornire alla scuola nel suo complesso misure solide e robuste e, soprattutto, utilizzabili in un'ottica comparativa.

In primo luogo è bene mettere opportunamente a fuoco gli elementi che caratterizzano una prova standardizzata, somministrata a diverse centinaia di migliaia di studenti, e gli aspetti che la differenziano dalle prove, strutturate o meno, predisposte comunemente dai docenti per la valutazione dei loro allievi, sia *in itinere* che alla fine di un determinato percorso didattico o formativo.

Tipicamente, una prova standardizzata somministrata su larga scala si pone come obiettivo quello di misurare, mediante una stima di tipo statistico, i livelli di apprendimento raggiunti dagli allievi

lungo una scala in grado di rappresentare adeguatamente tutti i risultati degli studenti, da quelli più bassi a quelli più elevati. Da ciò ne discende in modo diretto una prima grossa differenza con le prove che comunemente sono predisposte dai docenti per i loro allievi. Nelle prove standardizzate, come quelle del SNV, è normale che vi siano domande complesse alle quali è in grado di rispondere solo una piccola o anche piccolissima minoranza degli allievi e quesiti molto semplici, accessibili alla quasi totalità degli studenti. Una prova standardizzata deve, nel limite del possibile, prevedere domande distribuite lungo tutta la scala di misura oggetto di interesse. In una prova di classe, invece, non è normale e nemmeno auspicabile che vi siano quesiti potenzialmente non accessibili alla totalità della classe, proprio perché la prova è costruita per quella determinata classe. Le prove standardizzate, al contrario, sono costruite avendo a riferimento l'intera popolazione scolastica (ad esempio, tutti gli allievi della classe seconda della scuola primaria) e quindi è assolutamente comune che vi siano alcuni particolari quesiti di difficoltà troppo elevata per la quasi totalità degli allievi di una determinata classe poiché non è detto che particolari livelli molto elevati di competenza siano necessariamente presenti in tutte le classi, mentre lo sono certamente a livello di popolazione.

Ma le differenze tra una prova standardizzata somministrata su larga scala e le prove di scuola non si limitano a quelle che riguardano gli aspetti misuratori, ma riguardano anche altri elementi legati propriamente agli assunti alla base delle prove stesse.

In primo luogo, le prove standardizzate utilizzate nelle rilevazioni su larga scala si basano su un quadro di riferimento pubblico (<http://www.invalsi.it/snv1011/>), a disposizione di tutti coloro che siano interessati a conoscere la cornice all'interno della quale vengono costruite le prove stesse. Come avviene in tutte le ricerche nazionali e internazionali, il quadro di riferimento riveste un ruolo cruciale, anche se sovente sottovalutato, specie nel contesto italiano, poiché mediante tale documento sono resi espliciti i contenuti delle prove, le tipologie di domande, i processi cognitivi indagati e i compiti richiesti agli allievi. Ma c'è ancora di più. All'interno dei quadri di riferimento sono esplicitati anche i criteri di scelta dei quesiti, dei testi delle prove di comprensione della lettura e, anche se in modo sintetico, delle modalità in funzione delle quali le singole domande vengono selezionate.

La conoscenza dei quadri di riferimento è fondamentale, inoltre, per comprendere il reale valore informativo delle prove proposte agli studenti, ma anche i limiti delle prove stesse. È alquanto frequente, non solo nel contesto italiano, che le prove standardizzate vengano criticate, paventando

un'attribuzione eccessiva di significato alle prove stesse, quando invece all'interno dei quadri di riferimento è già esplicitamente dichiarato che, per loro natura, le prove standardizzate non possono, né devono, essere intese come uno strumento utile per tutti gli scopi di misurazione e di valutazione.

È assolutamente opportuno e auspicabile che all'interno del Paese si sviluppi un ampio dibattito sulle prove che permetta di dare spazio a diversi punti di vista, sia sul piano culturale sia su quello più propriamente tecnico. Tuttavia, affinché la discussione, anche accesa, si traduca in una vera occasione di miglioramento, è necessario che siano noti alcuni aspetti di fondo di tipo psicometrico che sono alla base delle prove stesse. In questa direzione la conoscenza approfondita dei quadri di riferimento per la valutazione sembra condizione necessaria per poter accedere a un livello di discussione e confronto in grado di fornire un apporto conoscitivo positivo e costruttivo. Solo in questo modo è possibile giungere alla costruzione di prove sempre più adeguate all'esigenza di mettere a disposizione del sistema scolastico uno strumento, fra gli altri, per conoscere in un'ottica comparativa i risultati raggiunti dagli studenti che frequentano le scuole italiane.

Se la conoscenza dei quadri di riferimento è un elemento essenziale per comprendere gli assunti sui quali si basano le prove, anche alcune conoscenze di base di tipo psicometrico sono altrettanto importanti per valutare la qualità delle prove. Specie per le prove di comprensione della lettura, talvolta vengono sollevate delle critiche che lasciano trasparire una conoscenza non del tutto adeguata delle metodologie, consolidate anche a livello internazionale, secondo le quali sono costruite le prove.

Una obiezione tipica che viene sollevata alle prove di comprensione della lettura riguarda la scelta del testo, lamentando spesso un disegno più o meno esplicito di natura ideologica nella scelta dei brani. È invece necessario sottolineare che nella scelta dei testi non intervengono in alcun modo considerazioni di carattere ideologico, fatta eccezione, se si vuole, per la norma di carattere prudenziale di evitare testi che possano “urtare in modo evidente la sensibilità di una parte degli allievi” (come si legge a pag. 8 del quadro di riferimento per l'italiano).

Ma esiste un'obiezione ancora più frequente e importante, vale a dire l'indicazione della risposta “giusta”, che secondo una critica ricorrente o “non c'è nel testo” o non sarebbe distinguibile come tale dalle altre alternative di risposta errate (*distrattori*), in una parola sarebbe frutto di “personali” preferenze dei costruttori del test.

Per quanto riguarda il primo punto (la risposta non c'è nel testo), è importante osservare che l'idea di comprensione del testo su cui le prove INVALSI sono basate è molto più ampia di quella di una semplice individuazione di informazioni esplicitamente date nel testo e dunque in alcuni casi la risposta giusta deve essere individuata attraverso processi di ragionamento che vanno al di là di quanto il testo dice alla lettera e che possono anche implicare, oltre che l'enciclopedia personale dello studente-lettore, una comparazione delle alternative proposte fino a identificare quella più corretta. Questo perché i *distrattori* per “funzionare” – ma su questo si tornerà nelle parti successive del rapporto – debbono esser abbastanza plausibili da attrarre le scelte di una parte degli alunni, quelli, in pratica, che padroneggiano di meno l'abilità¹ (o costrutto latente) che la prova intende misurare. Se la risposta giusta venisse, a qualunque livello di abilità, sempre preferita rispetto ai distrattori, questa sarebbe un'indicazione molto forte – secondo le regole che presiedono alla costruzione di domande a scelta multipla – che i distrattori non funzionano e che vanno riformulati o sostituiti.

Nello specifico, nelle critiche secondo cui l'indicazione della risposta giusta risponderebbe a soggettive preferenze o interpretazioni dell'INVALSI, e dunque non sarebbe giustificata, quello che solitamente sfugge è che tutte le prove del Servizio nazionale di valutazione, prima di arrivare alla loro stesura definitiva, sono preliminarmente testate, vale a dire sottoposte a campioni nazionali di alunni dell'età per cui il test viene predisposto e che le risposte degli studenti alle domande sono analizzate statisticamente secondo due approcci: quello della teoria cosiddetta “classica” dei test (TCT) e quello della “risposta all'item” (IRT). Ciò significa che per ciascun testo nel suo insieme e per ognuna delle domande del pre-test, le analisi forniscono una serie di dati che permettono di intervenire a ragion veduta sugli item e sulle alternative di risposta (risposta corretta e distrattori), scartando, correggendo o modificando là dove questo, in base ai risultati delle analisi, sia opportuno.

Considerazioni analoghe valgono anche per le prove di matematica, anche se le problematiche suddette, per quanto presenti, risultano meno rilevanti. Infatti, la mancanza di un testo di riferimento al quale sono legate diverse domande semplifica in parte la costruzione della prova. Inoltre, le caratteristiche della disciplina e le competenze ad essa collegate sono per loro natura, in una qualche misura, meno difficili da testare mediante domande standardizzate.

¹ In ambito psicometrico il termine *abilità* è inteso in un'accezione più generale e meno tecnica di quanto avvenga nelle scienze dell'educazione. Infatti, in questo contesto per abilità si intende semplicemente il costrutto latente che la prova intende misurare.

2. La costruzione delle prove del Servizio nazionale di valutazione

In questo paragrafo sono sinteticamente illustrate le fasi che caratterizzano la costruzione delle prove del Servizio nazionale di valutazione, incluse quelle somministrate all'interno dell'esame di Stato conclusivo del primo ciclo di istruzione, la cosiddetta Prova nazionale.

La costruzione di una prova standardizzata è il frutto di un lungo e articolato processo tecnico-scientifico, il cui rispetto è fondamentale se si vuole giungere alla definizione di strumenti in grado di misurare adeguatamente il livello di preparazione raggiunto dagli studenti in un determinato ambito di osservazione. Infatti, non sempre è noto che la costruzione di una prova standardizzata richiede grandi sforzi e tempi piuttosto lunghi, mai inferiori ai 15-18 mesi, e il rispetto di una procedura molto articolata e complessa.

2.1 I quadri di riferimento per la valutazione

L'INVALSI, come altri organismi che in ambito internazionale si occupano di ricerche sui livelli di apprendimento², è partito dalla definizione dei quadri di riferimento per la valutazione (INVALSI, 2011a, 2011b). Essi rivestono un ruolo fondamentale, spesso sottovalutato, quando non addirittura ignorato, per la costruzione delle prove standardizzate da somministrare agli allievi. Mediante il quadro di riferimento per la valutazione (QdR) vengono definiti gli ambiti, i processi cognitivi e i compiti oggetto di rilevazione, delimitando quindi il campo rispetto al quale sono costruite le prove. Inoltre, aspetto molto importante, il QdR permette di definire e circoscrivere il valore informativo delle prove che in base ad esso vengono costruite, chiarendone la portata e i limiti.

Il QdR elaborato dall'INVALSI è il frutto di uno studio articolato su più anni in cui sono stati valutati, anche in chiave comparativa, sia la normativa attualmente vigente in Italia circa i contenuti dell'insegnamento primario e secondario sia i quadri di riferimento prodotti negli ultimi anni dalla IEA (AA.VV, 2005) e dall'OCSE (OCSE, 2007) e, non da ultimo, la prassi didattica. Il QdR è uno strumento fondamentale per tutti coloro che, a vario titolo, sono coinvolti nel processo di costruzione e utilizzo delle prove oggettive standardizzate. Il QdR costituisce, infatti, il documento fondamentale per gli autori delle prove, per gli esperti che ne curano la revisione sia sotto il profilo dei contenuti che sotto quello misuratorio, per i docenti che sono chiamati a interpretare i risultati dei loro allievi e per i cosiddetti *stakeholder* che utilizzano i risultati delle rilevazioni standardizzate nazionali per valutare i livelli di apprendimento garantiti dal sistema educativo nel suo complesso.

² Ad esempio l'OCSE (Organizzazione per la Cooperazione e lo Sviluppo Economico) per la ricerca PISA e la IEA (International Association for the Evaluation of Educational Achievement) per le ricerche TIMSS e PIRLS.

2.2 Gli autori delle prove

La costruzione delle prove oggettive standardizzate richiede un grosso sforzo per individuare le domande più adeguate sia in termini di rispondenza al QdR sia rispetto alla formulazione e alla capacità misuratoria dal punto di vista tecnico-statistico. Da questo segue che l'INVALSI ha bisogno, per ogni livello scolastico interessato dal Servizio nazionale di valutazione (SNV), di un numero di domande molto elevato tra le quali poter scegliere le migliori e le più adeguate. Di norma, per la costruzione di una prova serve un numero di quesiti tre o quattro volte superiore a quello che effettivamente compare nella prova stessa somministrata agli allievi. Inoltre, l'INVALSI ha bisogno che le domande siano formulate in modo pienamente accessibile agli studenti, oltre che, ovviamente, in maniera ineccepibile sotto il profilo contenutistico-disciplinare.

Per realizzare il difficile compito di reperire un grosso numero di domande di elevata qualità, l'INVALSI si avvale della fondamentale collaborazione di oltre 200 docenti ed esperti provenienti dal mondo della scuola e dell'università. Essi, sulla base di indicazioni fornite dall'istituto e in relazione al QdR, producono proposte di quesiti o di intere prove da sottoporre al vaglio di una commissione di esperti nazionali e internazionali che opera presso l'INVALSI.

La decisione dell'INVALSI di rivolgersi ad una così ampia platea di docenti ed esperti si è rivelata strategica poiché si è dimostrato che solo chi opera direttamente nel mondo delle scuole e dell'università è in grado di formulare domande adeguate e significative. La conoscenza approfondita dei programmi, delle prassi, dei processi cognitivi e delle difficoltà degli studenti è fondamentale e irrinunciabile per costruire delle prove adeguate e in grado di fornire informazioni ricche e approfondite per tutti i soggetti interessati. Infine, ma non da ultimo, il consistente numero di autori rende possibile disporre di un'ampia varietà di quesiti sia rispetto alla modalità di formulazione sia rispetto ai contenuti. Ciò è fondamentale, poiché se si vuole evitare che si inducano nella scuola fenomeni non desiderabili di addestramento alle prove standardizzate è necessario che queste siano molto varie da un anno all'altro, sia rispetto ai contenuti sia alle modalità con le quali i quesiti sono formulati.

2.3 La costruzione delle prove

Dopo che gli autori hanno inviato all'INVALSI le loro proposte di quesiti, comincia il processo di costruzione vero e proprio delle prove. Esso si articola in diverse fasi, ognuna delle quali caratterizzata da aspetti differenti.

2.3.1. L'analisi qualitativa ex ante

Terminata la raccolta e la classificazione delle domande inviate dagli autori, un gruppo di esperti effettua una prima valutazione qualitativa in funzione della rispondenza delle domande al QdR, al livello scolastico per il quale devono essere proposte le prove e alla composizione del fascicolo (*booklet*) che dovrà essere inviato al pre-test. Questa prima fase richiede il lavoro di alcuni mesi di una commissione, articolata per ambiti disciplinari, composta sia da ricercatori dell'INVALSI che da esperti nazionali ed internazionali nell'ambito della costruzione di prove oggettive e delle analisi statistico-psicometriche.

L'analisi qualitativa *ex ante* è molto importante perché permette di escludere le domande non coerenti con le finalità del SNV e di effettuare un primo adattamento di quelle ritenute idonee per essere inviate al pre-test. Infatti, molto spesso accade che gli autori proponano delle domande molto interessanti e potenzialmente appropriate, ma non prive di ambiguità o di elementi che potrebbero mettere in difficoltà il rispondente. A differenza dei quesiti di una prova somministrata direttamente dall'autore in cui possono essere fornite agli esaminati delle spiegazioni o dei chiarimenti, nel caso di una prova standardizzata rivolta a tutti gli studenti di un certo livello scolastico ciò non è possibile ed è quindi necessario che tutti gli elementi di ambiguità o incertezza siano rimossi a priori.

Un altro controllo molto importante che viene effettuato in questa fase è una prima verifica dell'indipendenza delle domande. Questa analisi permette di verificare se vi siano quesiti le cui risposte suggeriscano la risposta ad altre domande. Tale controllo è importante soprattutto per le prove di comprensione della lettura in cui è facile che, in una prima stesura, vi siano delle opzioni di risposta che in un qualche modo forniscono dei suggerimenti per altre domande.

2.3.2. Il pre-test

La fase illustrata precedentemente si conclude con la composizione dei cosiddetti fascicoli, solitamente in numero almeno doppio rispetto al necessario, la cui adeguatezza viene valutata mediante il pre-test somministrato ad allievi con caratteristiche analoghe a quelle di coloro che dovranno svolgere le prove SNV. Poiché non è possibile che gli allievi coinvolti nel pre-test (*field trial*) siano gli stessi che parteciperanno alla rilevazione cosiddetta principale, è necessario che il *field trial* sia realizzato durante l'anno scolastico successivo a quello dello studio principale. È evidente che ciò pone dei forti vincoli di natura operativa poiché costringe l'INVALSI a operare

con un ampio anticipo, determinando così i tempi complessivi richiamati nell'introduzione per la costruzione delle prove.

La realizzazione del pre-test si basa sull'estrazione di un campione casuale di classi con rappresentatività nazionale. Per le prove del SNV 2010-11 sono stati coinvolti oltre 5000 studenti in tutto il Paese cercando di rappresentare adeguatamente tutte le principali tipologie di scuole e di studenti. Inoltre, per garantire il rispetto di un protocollo di somministrazione uniforme in tutte le scuole, le prove sono state svolte dagli studenti alla presenza di un somministratore dell'INVALSI e corrette direttamente dall'INVALSI. Terminata la somministrazione e l'acquisizione dei dati, inizia una fase di natura molto tecnica che si conclude con la costruzione dei *dataset* sui quali verranno effettuate le analisi statistiche. La predisposizione dei *dataset*, realizzata dal Servizio statistico dell'INVALSI, è molto delicata e impone una serie di verifiche incrociate assolutamente necessarie per poter effettuare analisi solide e robuste.

2.3.3. L'analisi quanti-qualitativa ex post: la calibrazione delle domande

I dati raccolti mediante il pre-test vengono poi analizzati mediante l'applicazione di appropriati modelli statistico-psicometrici sostanzialmente ascrivibili alla cosiddetta teoria classica dei test (*TCT*) e alla teoria della risposta (*Item Response Theory*) (Barbaranelli e Natali, 2005).

La realizzazione delle analisi quantitative dei dati del pre-test, la cosiddetta calibrazione, richiede competenze statistiche molto specifiche per le quali l'INVALSI si avvale anche della collaborazione delle università italiane maggiormente specializzate in questo particolare ambito di ricerca³. La capacità misuratoria di ogni domanda viene analizzata mediante modelli statistici in grado di stabilire la coerenza di ciascuna opzione di risposta rispetto al costrutto oggetto di valutazione, rispetto al livello di abilità/competenza del rispondente e rispetto alla difficoltà specifica della domanda stessa. I richiamati modelli statistici permettono, inoltre, di valutare il cosiddetto potere discriminante di ciascuna domanda, ovvero la capacità di ogni quesito di distinguere adeguatamente gli allievi in termini di abilità/competenza in funzione della risposta fornita (Falocci *et al.*, 2010).

Sulla base delle informazioni tratte dalle analisi quantitative dei dati del pre-test la commissione degli esperti effettua un'ulteriore analisi qualitativa delle domande testate. Durante questa fase, che richiede alcuni mesi, vengono selezionate in via definitiva le domande che saranno inserite nelle prove somministrate a tutti gli studenti. Solo quei quesiti che mostrano adeguati requisiti di chiarezza, affidabilità e validità possono entrare a far parte dei cosiddetti fascicoli definitivi. In genere, solo un terzo delle domande che vengono selezionate durante la prima fase, l'analisi *ex ante*, superano anche quelle successive e sono quindi ritenute idonee per i fascicoli definitivi.

2.3.4. La composizione dei fascicoli definitivi

Dopo che ciascuna domanda è stata analizzata sotto il profilo quali-quantitativo, la commissione degli esperti compone i fascicoli che saranno somministrati durante la rilevazione principale. Anche questa fase è molto importante e richiede la valutazione di diversi aspetti. In primo luogo, i fascicoli devono essere composti in funzione dei tempi di compilazione, stimati sulla base dei risultati del pre-test, e del livello complessivo di difficoltà, dati i vincoli di composizione definiti dai quadri di riferimento per la valutazione. Riguardo a quest'ultimo aspetto, non sempre è noto che la

³ Dipartimento di Scienze statistiche "P. Fortunati" dell'Università di Bologna e Dipartimento di Economia, Finanza e Statistica dell'Università di Perugia.

composizione di una prova standardizzata rivolta all'accertamento su scala nazionale dei livelli di apprendimento non risponde agli stessi criteri che guidano la costruzione delle verifiche di classe. Una prova standardizzata nazionale deve essere in grado di misurare i risultati degli studenti all'interno di una scala di abilità/competenza molto lunga, dai livelli più bassi a quelli di eccellenza. È quindi normale che all'interno di una prova di questo genere vi siano anche dei quesiti molto difficili ai quali solo una piccola percentuale di allievi è in grado di rispondere. Talvolta ciò genera un certo smarrimento negli insegnanti, specie della scuola primaria, ma non deve essere motivo di allarme se vi sono alcune domande alle quali solo una piccola percentuale di studenti è in grado di rispondere, poiché esse sono finalizzate alla misurazione di livelli di apprendimento molto elevati, quindi percentualmente meno rappresentati.

CAPITOLO II: La valutazione delle prove standardizzate

1. La valutazione delle prove del Servizio nazionale di valutazione

La valutazione delle caratteristiche di una prova standardizzata può avvenire su diversi piani, alcuni dei quali riguardano specificamente le proprietà misuratorie della prova stessa. Sovente le prove standardizzate, non solo quelle proposte dall'INVALSI, sono oggetto di discussione, specie relativamente alla loro difficoltà, più o meno accentuata, chiarezza, eventuale ambiguità e così via. È certamente fuor di dubbio che la discussione sulle prove è molto importante poiché da essa possono scaturire proposte per migliorare le prove stesse e per costruirne di sempre più adeguate per la misurazione dei livelli di apprendimento che si producono all'interno del sistema scolastico formale. Tuttavia, quando l'oggetto d'interesse è la prova intesa come strumento di misura, allora esistono metodi statistico-psicometrici che permettono di valutare in modo scientificamente robusto le caratteristiche misuratorie delle prove.

La valutazione sulla difficoltà di una prova è sovente effettuata utilizzando come metro l'esperienza individuale e un limitato numero di casi osservati. Ovviamente questa modalità di valutazione, utile in prima approssimazione, non può essere presa a riferimento per un giudizio tecnico e scientificamente fondato circa il livello di difficoltà poiché non permette di comparare un elevato numero di risposte e quindi una varietà di comportamenti sufficientemente ampia.

La teoria dei test, specie nella sua forma più moderna che va sotto il nome di *Item Response Theory* e di *Rasch Analysis*⁴, permette di effettuare solide e robuste valutazioni sulla difficoltà complessiva di una prova, sulla sua coerenza generale e sull'adeguatezza delle domande che la compongono.

Analogamente a quanto avviene in altre rilevazioni quantitative sugli apprendimenti, sia a livello nazionale sia a livello internazionale, l'INVALSI utilizza metodi e tecniche la cui validità trova riscontro in tutta la letteratura di settore. In particolare, per ogni parte delle prove SNV⁵ (Italiano o Matematica), sono state effettuate diverse analisi basate sui dati forniti da campioni di non meno di 1300 classi, come nel caso della Prova nazionale, fino a raggiungere le 2300 classi, come è avvenuto per la scuola secondaria di secondo grado.

⁴ Esula dallo scopo del presente lavoro qualsiasi approfondimento circa le differenze metodologiche e teoriche tra *Item Response Theory* e *Rasch Analysis*. A tal riguardo si veda, ad esempio, BARBARANELLI C., NATALI E. (2005), *I test psicologici: teorie e modelli psicometrici*, Carrocci Editore, Roma.

⁵ Inclusa la Prova nazionale

Gli strumenti tecnico-statistici principali con i quali comunemente vengono valutate le caratteristiche misuratorie di una prova standardizzata sono:

1. valutazione generale di affidabilità della prova mediante l'indice α di *Cronbach*;
2. la *Rasch Analysis* per la stima della difficoltà delle domande;
3. *placement* relativo delle domande rispetto al punteggio di Rasch assegnato a ciascun allievo;
4. l'indice di discriminazione statistica.

2. L'indice α di Cronbach

Il primo indice, cioè l' α di *Cronbach*, permette di esprimere una valutazione sintetica e generale sulla coerenza complessiva della prova e quindi sull'attendibilità delle informazioni che da essa si possono desumere. Più precisamente, l' α di *Cronbach* consente di comprendere se l'oggetto della valutazione di una prova ha caratteristiche prevalenti di unitarietà o meno (Cronbach, 1951), ovvero, in termini più espliciti, se le domande che compongono la prova sono tra loro coerenti e volte alla misurazione dello stesso oggetto. Come è noto in letteratura, l' α di *Cronbach* non è sufficiente per poter esprimere un giudizio sicuro e fondato sull'aspetto predetto, ma consente di effettuare un'analisi preliminare che deve essere ulteriormente approfondita con altri strumenti. A tal fine, la *Rasch Analysis* consente di effettuare gli approfondimenti necessari per valutare se e in che misura la prova proposta agli studenti ha sufficienti caratteristiche di affidabilità per poter formulare misurazioni robuste e informative sui livelli di apprendimento raggiunti.

L' α di *Cronbach* è un indice che ha come estremi 0 e 1 e tanto più si avvicina ad 1, tanto maggiore è la coerenza complessiva della prova oggetto d'interesse. La letteratura psicometrica fornisce valori di riferimento che definiscono degli standard generalmente accettati per esprimere un primo giudizio orientativo sulla coerenza generale di una prova.

Tavola II.1. Soglie di α per prove su larga scala

| Valori di α | Attendibilità della prova |
|--------------------|---------------------------|
| 0,80 o superiore | Molto buona |
| da 0,70 a 0,80 | Buona |
| da 0,50 a 0,70 | Modesta |
| inferiore a 0,50 | Inadeguata |

I valori soglia riportati nella tavola II.1 hanno naturalmente una valenza orientativa e dipendono dalla tipologia di prova presa in esame. Se, come nel caso delle prove SNV, inclusa la Prova nazionale (PN), i rispondenti sono tutti gli allievi che frequentano una data classe, allora possono essere prese in considerazione delle prove che presentino valori di α almeno superiori a 0,70.

La tavola II.2 riporta i valori dell'indice α di *Cronbach* per i fascicoli di Italiano e Matematica delle prove SNV 2010-2011 e della Prova nazionale 2011.

Tavola II.2. Valori di α delle prove SNV-PN

| Livello | Fascicolo | Valori di α |
|--------------------------|------------|--------------------|
| II primaria | Italiano | 0,88 |
| | Matematica | 0,81 |
| V primaria | Italiano | 0,87 |
| | Matematica | 0,88 |
| I sec. di I gr. | Italiano | 0,93 |
| | Matematica | 0,88 |
| III sec. di I gr. | Italiano | 0,86 |
| | Matematica | 0,83 |
| II sec. di II gr. | Italiano | 0,91 |
| | Matematica | 0,89 |

Il confronto dei dati delle tavole II.1 e II.2 permette di esprimere un primo giudizio provvisorio molto positivo circa la coerenza complessiva di ciascuno dei fascicoli che compongono le prove SNV e della Prova nazionale. In termini ancora più espliciti, gli elevati valori degli indici riportati nella tavola II.2 consentono di ritenere che le due parti delle prove (Italiano e Matematica) siano coerenti dal punto di vista misuratorio, ovvero siano costruite in modo che l'oggetto di misurazione abbia una sua unitarietà prevalente. Ciò significa che i fascicoli permettono di esprimere misure affidabili del livello di preparazione degli allievi ai quali è stata presentata la prova. Anche se, come sarà mostrato in seguito, l' α di Cronbach è solo un primo strumento di valutazione, i valori della tavola II.2 fanno pensare che la prova, complessivamente intesa, non presenta problemi di ambiguità macroscopiche o incoerenza tali da mettere in difficoltà i rispondenti.

3. La Rasch Analysis per la stima della difficoltà delle domande

Dopo una prima valutazione sintetica sull'attendibilità generale delle prove SNV e della Prova nazionale, l'analisi è stata approfondita applicando i metodi della cosiddetta *Rasch Analysis* (RA). In termini assolutamente generali, la principale caratteristica della RA è quella di ottenere sulla stessa scala la stima della difficoltà di ciascuna domanda e la misura del livello di prestazione del rispondente⁶. Quest'ultimo aspetto è molto importante dal punto di vista interpretativo poiché consente di comprendere a quale livello di difficoltà si colloca ciascuna domanda e, allo stesso tempo, di comprendere quali e quanti sono gli allievi che mostrano livelli di apprendimento superiori o inferiori alla difficoltà di una determinata domanda o di un insieme di quesiti.

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. Quest'ultimo aspetto è di fondamentale importanza non solo dal punto di vista misuratorio, ma anche da quello più propriamente interpretativo e sostantivo. Infatti, in questo modo è possibile comprendere ciò che l'allievo è in grado di fare o che conosce. Risulta quindi evidente l'enorme potenzialità di questa metodologia anche dal punto di vista didattico e per la promozione di azioni volte al miglioramento che possono essere realizzate dalle singole scuole.

Le tavole del presente paragrafo riportano i cosiddetti *parametri di difficoltà* di ciascuna domanda delle prove (fascicolo d'Italiano e di Matematica) SNV⁷ per la classe II della scuola primaria. Tali parametri sono definiti su una scala convenzionale centrata sul valore 0 e compresa tra i valori -4 e +4. Ciò significa che domande con parametri di difficoltà inferiori a 0 sono facili o molto facili se tali parametri assumono valori vicino a -4, mentre quelle con parametri positivi sono più difficili fino a divenire molto difficili tanto più ci si avvicina a +4.

3.1. La scuola primaria

Le tavole II.3 e II.4 riportano i parametri di difficoltà delle prove di Italiano e Matematica della classe II primaria.

⁶ Si veda nota 1 di pag. 4.

⁷ I risultati della prova preliminare di lettura della II primaria sarà oggetto, insieme ai risultati del questionario studente somministrato in V primaria e nella I secondaria di primo grado, di un apposito approfondimento che l'INVALSI pubblicherà nel corso dell'anno scolastico 2011-12.

Tavola II.3. Parametri di difficoltà delle domande del fascicolo di Italiano (II primaria)

| COMPRESIONE DELLA LETTURA | | | ESERCIZIO 1 ⁸ | | | ESERCIZIO 2 ⁹ | | |
|---------------------------|------|------------|--------------------------|------|------------|--------------------------|------|------------|
| N. item | Item | Difficoltà | N. item | Item | Difficoltà | N. item | Item | Difficoltà |
| 1 | A1 | -0,9 | 19 | B1 | -1,5 | 34 | C1 | -2,3 |
| 2 | A2 | -1,1 | 20 | B2 | -2,4 | 35 | C2 | -1,0 |
| 3 | A3 | -0,8 | 21 | B3 | -1,6 | 36 | C3 | -1,1 |
| 4 | A4 | -0,5 | 22 | B4 | -0,5 | 37 | C4 | -2,7 |
| 5 | A5 | -0,7 | 23 | B5 | -2,0 | 38 | C5 | -1,7 |
| 6 | A6 | 0,0 | 24 | B6 | -2,2 | - | - | - |
| 7 | A7 | -1,6 | 25 | B7 | -1,7 | - | - | - |
| 8 | A8 | -2,5 | 26 | B8 | -1,8 | - | - | - |
| 9 | A9 | -0,3 | 27 | B9 | -1,7 | - | - | - |
| 10 | A10 | -1,3 | 28 | B10 | -1,4 | - | - | - |
| 11 | A11 | -0,8 | 29 | B11 | -0,3 | - | - | - |
| 12 | A12 | -0,5 | 30 | B12 | -1,7 | - | - | - |
| 13 | A13 | 1,1 | 31 | B13 | -0,6 | - | - | - |
| 14 | A14 | -0,3 | 32 | B14 | -1,1 | - | - | - |
| 15 | A15 | 0,0 | 33 | B15 | -0,2 | - | - | - |
| 16 | A16 | -0,6 | - | - | - | - | - | - |
| 17 | A17 | 0,0 | - | - | - | - | - | - |
| 18 | A18 | -2,0 | - | - | - | - | - | - |

Tavola II.4. Parametri di difficoltà delle domande del fascicolo di Matematica (II primaria)

| N. item | Item | Difficoltà | Ambito | N. item | Item | Difficoltà | Ambito |
|---------|------|------------|-------------------|---------|-------|------------|-----------------|
| 1 | D1 a | -0,8 | Dati e previsioni | 15 | D9 a | -2,1 | Numeri |
| 2 | D1 b | -1,8 | Dati e previsioni | 16 | D9 b | -0,7 | Numeri |
| 3 | D2 a | -0,6 | Numeri | 17 | D9 c | -0,7 | Numeri |
| 4 | D2 b | -1,5 | Numeri | 18 | D9 d | -1,1 | Numeri |
| 5 | D3 a | -0,9 | Numeri | 19 | D10 a | -0,1 | Spazio e figure |
| 6 | D3 b | -0,4 | Numeri | 20 | D10 b | 0,4 | Spazio e figure |
| 7 | D4 | -0,2 | Spazio e figure | 21 | D11 | 1,3 | Numeri |
| 8 | D5 a | -2,2 | Numeri | 22 | D12 | 0,5 | Numeri |
| 9 | D5 b | -2,6 | Numeri | 23 | D13 | -1,1 | Spazio e figure |
| 10 | D5 c | 0,1 | Numeri | 24 | D14 | 0,7 | Numeri |
| 11 | D6 | -2,1 | Numeri | 25 | D15 | 1,3 | Numeri |
| 12 | D7 a | -0,8 | Spazio e figure | 26 | D16 | -1,3 | Spazio e figure |
| 13 | D7 b | 1,0 | Spazio e figure | 27 | D17 | -0,3 | Numeri |
| 14 | D8 | 0,4 | Numeri | 28 | D18 | -1,3 | Spazio e figure |

La lettura dei dati delle tavole II.3 mette in luce come che la prova di Italiano della seconda primaria possa considerarsi, in base alle risposte effettivamente espresse dagli allievi, medio-facile, confermando quindi tutte le ipotesi emerse durante la fase di pre-test delle prove medesime. La difficoltà contenuta delle prove della seconda primaria è coerente con le scelte effettuate

⁸ L'esercizio richiede all'allievo di indicare se quindici coppie di parole hanno significato uguale o contrario.

⁹ L'esercizio richiede il collegamento tra le due parti in cui sono suddivise cinque frasi.

dall'INVALSI in sede di progettazione degli strumenti di misurazione per bambini di 7 anni al termine del loro secondo anno di scuola. E' importante, come indicato dalla letteratura del settore, che bambini così piccoli affrontino prove standardizzate che non creino loro senso di frustrazione per l'eccesso di difficoltà, ma che, al contempo, contengano domande sufficientemente stimolanti per evitare fenomeni di disinteresse e di calo di attenzione durante lo svolgimento della prova stessa. Tuttavia, l'analisi congiunta dei risultati dell'analisi delle difficoltà delle domande delle prove dei due anni scolastici 2009-10 e 2010-11 indica chiaramente la necessità di inserire nelle prove proposte dall'INVALSI anche domande in grado di stimare più precisamente la parte più alta della scala di abilità. Naturalmente ciò deve avvenire nel pieno rispetto della peculiarità delle prove proposte a bambini di soli 7 anni, senza creare quindi nei rispondenti la sensazione di richieste troppo complesse. Le analisi proposte nella tavola II.3 rispondono in modo incontrovertibile a molte critiche rivolte alle prove INVALSI della II primaria. Sovente, docenti o analisti hanno affermato che le prove proposte ai bambini sono troppo complesse per alunni di soli 7 anni. Come spesso accade, le valutazioni puramente qualitative, per quanto importanti, non sono sufficienti per esprimere un giudizio sulla difficoltà di una prova, specie se basate sull'osservazione di poche decine di casi. Differenti, invece, sono i risultati che emergono quando si analizzano le risposte effettivamente fornite da diverse centinaia di migliaia di allievi, ossia tutti i bambini che frequentano la seconda classe della scuola primaria. Infatti, accade frequentemente che le capacità di comprensione della lettura degli alunni venga sottostimata, con la conseguenza che non vengono proposti loro testi ritenuti complessi, ma che in realtà non lo sono realmente per i bambini stessi. Anche in questo senso le prove standardizzate proposte a tutti gli alunni della classe seconda primaria sono molto utili poiché consentono di fornire una misura solida e comparativamente informativa sulla reale comprensione della lettura dei bambini di 7 anni che frequentano le scuole italiane, statali e paritarie.

La tavola II.4 dà ragione di una prova di livello di difficoltà medio-basso, ma con la presenza anche di alcune domande un po' più impegnative. Come accade sovente, la prova di Matematica risulta maggiormente in grado di rappresentare anche i livelli più alti della scala di abilità dei rispondenti, anche se, come già messo in luce per quella di Italiano, anche per la Matematica è necessario prevedere per il futuro alcuni quesiti in grado di permettere una stima più accurata dei livelli più alti delle abilità dei rispondenti. Infine, per concludere sulla comparazione dei livelli di difficoltà dei quesiti delle due prove rivolte alla classe seconda della scuola primaria, la prova di Matematica pare meglio articolata in termini di distribuzione delle difficoltà delle domande, ma è necessario

rimarcare la maggiore semplicità di realizzare ciò con i quesiti di matematica anziché con quelli di italiano. Infatti, questi ultimi sono legati a un testo e quindi i margini di libertà nella scelta delle domande sono decisamente minori. Tuttavia, proprio in questa direzione si concretizza una sfida per l'INVALSI, ossia quella di trovare testi adatti a bambini di 7 anni, ma sufficientemente ricchi e densi per formulare domande di diversi livelli di difficoltà, in grado, quindi, di stimare accuratamente tutti i livelli di abilità dei rispondenti, da quelli più modesti a quelli più elevati.

Nell'Appendice 1 sono riportate le distribuzioni di frequenza delle risposte fornite dagli allievi a ciascuna domanda. Esse possono essere utili per approfondire a livello di singolo quesito quanto illustrato in termini complessivi a commento delle tavole II.3 e II.4.

Le tavole II.5 e II.6 riportano i parametri di difficoltà delle prove di Italiano e Matematica della classe V primaria.

Tavola II.5. Parametri di difficoltà delle domande del fascicolo di Italiano (V primaria)

| TESTO NARRATIVO | | | TESTO ESPOSITIVO | | | GRAMMATICA | | |
|-----------------|------|------------|------------------|------|------------|------------|-------|------------|
| N. item | Item | Difficoltà | N. item | Item | Difficoltà | N. item | Item | Difficoltà |
| 1 | A1 | -0,6 | 18 | B1 | -0,6 | 33 | C1_a1 | -1,9 |
| 2 | A2 | -0,5 | 19 | B2 | -1,8 | 34 | C1_a2 | -2,7 |
| 3 | A3 | -1,0 | 20 | B3 | -1,0 | 35 | C1_b1 | -3,5 |
| 4 | A4 | -1,7 | 21 | B4 | -1,6 | 36 | C1_b2 | -2,6 |
| 5 | A5 | -1,0 | 22 | B5 | -1,1 | 37 | C1_b3 | -2,8 |
| 6 | A6 | -2,1 | 23 | B6 | 0,1 | 38 | C2 | -0,5 |
| 7 | A7 | -2,7 | 24 | B7 | -0,2 | 39 | C3_a | -4,3 |
| 8 | A8 | -1,6 | 25 | B8 | -0,2 | 40 | C3_b | -4,4 |
| 9 | A9 | -0,9 | 26 | B9 | 0,4 | 41 | C3_c | -1,7 |
| 10 | A10 | -1,8 | 27 | B10 | 0,2 | 42 | C3_d | -3,3 |
| 11 | A11 | -0,1 | 28 | B11 | 0,0 | 43 | C3_e | -3,2 |
| 12 | A12 | -1,6 | 29 | B12 | -1,6 | 44 | C3_f | -1,5 |
| 13 | A13 | -0,8 | 30 | B13 | 0,2 | 45 | C3_g | -2,4 |
| 14 | A14 | -0,4 | 31 | B14 | -0,4 | 46 | C3_h | -1,6 |
| 15 | A15 | -1,1 | 32 | B15 | 0,2 | 47 | C3_i | -3,8 |
| 16 | A16 | -1,8 | - | - | - | 48 | C3_l | -3,2 |
| 17 | A17 | -1,1 | - | - | - | 49 | C3_m | -2,8 |
| - | - | - | - | - | - | 50 | C3_n | -1,2 |
| - | - | - | - | - | - | 51 | C3_o | -3,3 |
| - | - | - | - | - | - | 52 | C3_p | -3,5 |
| - | - | - | - | - | - | 53 | C3_q | -3,7 |
| - | - | - | - | - | - | 54 | C3_r | -3,7 |
| - | - | - | - | - | - | 55 | C4 | -1,6 |
| - | - | - | - | - | - | 56 | C5 | 1,1 |
| - | - | - | - | - | - | 57 | C6 | -0,9 |
| - | - | - | - | - | - | 58 | C7 | -0,9 |
| - | - | - | - | - | - | 59 | C8 | -0,9 |
| - | - | - | - | - | - | 60 | C9 | 0,4 |
| - | - | - | - | - | - | 61 | C10 | 1,0 |

Tavola II.6. Parametri di difficoltà delle domande del fascicolo di Matematica (V primaria)

| N. item | Item | Difficoltà | Ambito | N. item | Item | Difficoltà | Ambito |
|---------|-------|------------|----------------------|---------|-------|------------|----------------------|
| 1 | D1_a | -1,9 | Dati e previsioni | 26 | D18 | 0,1 | Numeri |
| 2 | D1_b | -2,1 | Dati e previsioni | 27 | D19 | -1,1 | Numeri |
| 3 | D1_c | -2,2 | Dati e previsioni | 28 | D20 | -2,4 | Spazio e figure |
| 4 | D1_d | -1,6 | Dati e previsioni | 29 | D21_a | -3,1 | Dati e previsioni |
| 5 | D2 | -1,5 | Numeri | 30 | D21_b | 0,0 | Dati e previsioni |
| 6 | D3 | -2,0 | Numeri | 31 | D22 | -0,7 | Numeri |
| 7 | D4_a | -2,7 | Relazioni e funzioni | 32 | D23_a | -1,5 | Spazio e figure |
| 8 | D4_b | -2,2 | Relazioni e funzioni | 33 | D23_b | 0,4 | Spazio e figure |
| 9 | D5 | -1,1 | Numeri | 34 | D24_a | -2,0 | Relazioni e funzioni |
| 10 | D6 | -0,6 | Relazioni e funzioni | 35 | D24_b | -0,4 | Relazioni e funzioni |
| 11 | D7 | 1,6 | Relazioni e funzioni | 36 | D24_c | 0,7 | Relazioni e funzioni |
| 12 | D8 | -1,5 | Numeri | 37 | D25 | -0,7 | Dati e previsioni |
| 13 | D9 | -1,1 | Spazio e figure | 38 | D26 | -0,8 | Spazio e figure |
| 14 | D10 | 0,4 | Spazio e figure | 39 | D27 | 0,3 | Spazio e figure |
| 15 | D11 | -0,6 | Numeri | 40 | D28_a | -3,0 | Relazioni e funzioni |
| 16 | D12 | -1,4 | Numeri | 41 | D28_b | -2,5 | Relazioni e funzioni |
| 17 | D13 | 0,5 | Relazioni e funzioni | 42 | D28_c | -2,4 | Relazioni e funzioni |
| 18 | D14 | -0,4 | Numeri | 43 | D29_a | -1,6 | Numeri |
| 19 | D15 | 0,4 | Spazio e figure | 44 | D29_b | -1,9 | Numeri |
| 20 | D16_a | -0,6 | Spazio e figure | 45 | D29_c | -0,5 | Numeri |
| 21 | D16_b | 0,6 | Spazio e figure | 46 | D29_d | -0,6 | Numeri |
| 22 | D17_a | -3,3 | Dati e previsioni | 47 | D30 | -1,7 | Spazio e figure |
| 23 | D17_b | -1,5 | Dati e previsioni | - | - | - | - |
| 24 | D17_c | -1,8 | Dati e previsioni | - | - | - | - |
| 25 | D17_d | -1,0 | Dati e previsioni | - | - | - | - |

Anche per le prove delle V primaria si confermano i risultati dei pre-test e le ipotesi formulate in quella sede escono corroborate dai risultati della somministrazione a tutti gli allievi della quinta classe della scuola primaria.

I valori dei parametri di difficoltà della prova di Italiano riportati nella tavola II.5 mostrano che essa è formata da quesiti adeguati per gli allievi della V primaria poiché i livelli di difficoltà delle domande, pur prevalendo quelle medio-facili, si distribuiscono maggiormente lungo la scala delle abilità. Inoltre, è importante osservare che la maggior parte di domande molto facili sono, in realtà, i 16 item che compongono la domanda C3 della parte riservata alla grammatica. Ciò significa che i 16 item di questa domanda devono essere considerati nel loro complesso come afferenti ai livelli bassi della scala di abilità, ossia essi formano una domanda semplice, volta alla misurazione di livelli di apprendimento molto basilari. Come avviene spesso, se si confronta la prova di Italiano di V primaria con quella della classe II primaria, si osserva che nella prima sono presenti più quesiti di maggiore difficoltà, anche se non particolarmente elevata. Tale fenomeno è ancora più evidente se,

come si vedrà in seguito, il confronto è effettuato con la prova di Italiano della I secondaria di primo grado. Ciò, almeno in parte, si spiega con la maggiore articolazione della prova. Infatti, non bisogna dimenticare che, a differenza della prova di Italiano della II primaria, le altre prove di Italiano sono composte da due testi, uno narrativo-letterario e l'altro espositivo-argomentativo. Tuttavia, è indubbio che anche la prova di II primaria, *mutatis mutandis*, debba in futuro comprendere alcuni quesiti in grado di descrivere i livelli più alti della scala di abilità.

La tavola II.6 riporta i parametri di difficoltà della prova di Matematica. Anche in questo caso si riscontra ciò che è già stato evidenziato per la prova di Matematica della II primaria, ossia una sostanziale adeguatezza in termini di distribuzione delle difficoltà, anche se mancano ancora alcuni quesiti in grado di stimare con maggiore accuratezza i livelli alti e molto alti della scala di abilità.

Nell'Appendice 2 sono riportate le distribuzioni di frequenza delle risposte fornite dagli allievi a ciascuna domanda della prova di Italiano e Matematica. Esse possono essere utili per approfondire a livello di singola domanda quanto esposto in termini complessivi a commento delle tavole II.5 e II.6.

3.2. La scuola secondaria di primo grado

Le tavole II.7 e II.8 riportano i parametri di difficoltà delle prove di Italiano e Matematica della classe I della scuola secondaria di primo grado.

Tavola II.7. Parametri di difficoltà delle domande del fascicolo di Italiano (I sec. I gr.)

| TESTO NARRATIVO | | | TESTO ESPOSITIVO | | | GRAMMATICA | | |
|-----------------|------|------------|------------------|------|------------|------------|-------|------------|
| N. item | Item | Difficoltà | N. item | Item | Difficoltà | N. item | Item | Difficoltà |
| 1 | A1 | -0,7 | 21 | B1 | -0,5 | 38 | C1_1 | -1,3 |
| 2 | A2 | 0,6 | 22 | B2 | -0,4 | 39 | C1_2 | -1,8 |
| 3 | A3 | -1,7 | 23 | B3 | 0,0 | 40 | C1_3 | -1,9 |
| 4 | A4 | -0,5 | 24 | B4 | -0,8 | 41 | C2 | -0,8 |
| 5 | A5 | 0,3 | 25 | B5 | -0,2 | 42 | C3_a | -1,7 |
| 6 | A6 | 0,3 | 26 | B6_a | -0,9 | 43 | C3_b | -2,8 |
| 7 | A7 | 0,6 | 27 | B6_b | -1,5 | 44 | C3_c | -2,5 |
| 8 | A8 | 0,3 | 28 | B6_c | -1,2 | 45 | C3_d | -0,9 |
| 9 | A9 | 0,9 | 29 | B6_d | -1,5 | 46 | C3_e | -2,5 |
| 10 | A10 | -0,8 | 30 | B6_e | -0,5 | 47 | C3_f | -0,9 |
| 11 | A11 | -0,3 | 31 | B7 | 1,0 | 48 | C4_a | -1,8 |
| 12 | A12 | 2,1 | 32 | B8 | 0,1 | 49 | C4_b | -1,3 |
| 13 | A13 | 0,1 | 33 | B9 | -0,2 | 50 | C4_c | -1,1 |
| 14 | A14 | -1,8 | 34 | B10 | 0,8 | 51 | C5_a | 0,1 |
| 15 | A15 | -0,6 | 35 | B11 | -0,7 | 52 | C5_b | 0,3 |
| 16 | A16 | -0,3 | 36 | B12 | -0,7 | 53 | C5_c | -0,7 |
| 17 | A17 | -1,0 | 37 | B13 | -0,9 | 54 | C5_d | 1,1 |
| 18 | A18 | 1,4 | - | - | - | 55 | C6 | -0,4 |
| 19 | A19 | 0,5 | - | - | - | 56 | C7 | -1,9 |
| 20 | A20 | -0,3 | - | - | - | 57 | C8 | 0,6 |
| - | - | - | - | - | - | 58 | C9_aa | -1,0 |
| - | - | - | - | - | - | 59 | C9_ab | -0,1 |
| - | - | - | - | - | - | 60 | C9_ac | -1,7 |
| - | - | - | - | - | - | 61 | C9_ad | -0,5 |
| - | - | - | - | - | - | 62 | C9_ae | -1,0 |
| - | - | - | - | - | - | 63 | C9_af | -0,9 |
| - | - | - | - | - | - | 64 | C9_ag | 0,0 |
| - | - | - | - | - | - | 65 | C9_ah | -0,5 |
| - | - | - | - | - | - | 66 | C9_ai | -0,4 |
| - | - | - | - | - | - | 67 | C9_al | 0,0 |
| - | - | - | - | - | - | 68 | C9_am | -0,5 |
| - | - | - | - | - | - | 69 | C9_an | 0,5 |
| - | - | - | - | - | - | 70 | C9_ba | -1,8 |
| - | - | - | - | - | - | 71 | C9_bb | -1,4 |
| - | - | - | - | - | - | 72 | C9_bc | -1,6 |
| - | - | - | - | - | - | 73 | C9_bd | -1,5 |
| - | - | - | - | - | - | 74 | C9_be | -1,4 |
| - | - | - | - | - | - | 75 | C9_bf | -1,6 |
| - | - | - | - | - | - | 76 | C9_bg | -1,4 |
| - | - | - | - | - | - | 77 | C9_bh | -1,5 |
| - | - | - | - | - | - | 78 | C9_bi | -1,6 |
| - | - | - | - | - | - | 79 | C9_bl | -1,8 |
| - | - | - | - | - | - | 80 | C9_bm | -1,5 |
| - | - | - | - | - | - | 81 | C9_bn | -1,9 |

Tavola II.8. Parametri di difficoltà delle domande del fascicolo di Matematica (I sec. I gr.)

| N. item | Item | Difficoltà | Ambito | N. item | Item | Difficoltà | Ambito |
|---------|-------|------------|----------------------|---------|-------|------------|----------------------|
| 1 | D1 | -0,8 | Numeri | 23 | D15 | 0,5 | Dati e previsioni |
| 2 | D2 | -0,3 | Spazio e figure | 24 | D16_a | -1,6 | Spazio e figure |
| 3 | D3_a | -2,8 | Relazioni e funzioni | 25 | D16_b | 0,2 | Spazio e figure |
| 4 | D3_b | -0,5 | Relazioni e funzioni | 26 | D17_a | 0,4 | Spazio e figure |
| 5 | D3_c | 0,0 | Relazioni e funzioni | 27 | D17_b | 0,9 | Spazio e figure |
| 6 | D3_d | -0,3 | Relazioni e funzioni | 28 | D18 | 0,7 | Numeri |
| 7 | D4 | -1,3 | Dati e previsioni | 29 | D19 | 2,7 | Relazioni e funzioni |
| 8 | D5 | -0,8 | Spazio e figure | 30 | D20 | 1,0 | Relazioni e funzioni |
| 9 | D6 | -2,4 | Spazio e figure | 31 | D21 | 0,2 | Dati e previsioni |
| 10 | D7 | 0,7 | Spazio e figure | 32 | D22_a | 0,5 | Numeri |
| 11 | D8 | 2,4 | Numeri | 33 | D22_b | 1,1 | Numeri |
| 12 | D9 | 0,5 | Spazio e figure | 34 | D22_c | -0,1 | Numeri |
| 13 | D10_a | -0,4 | Relazioni e funzioni | 35 | D23 | -0,2 | Numeri |
| 14 | D10_b | 0,6 | Relazioni e funzioni | 36 | D24 | 0,5 | Numeri |
| 15 | D10_c | 0,5 | Relazioni e funzioni | 37 | D25_a | -0,2 | Relazioni e funzioni |
| 16 | D11 | 1,1 | Numeri | 38 | D25_b | 0,7 | Relazioni e funzioni |
| 17 | D12 | 0,8 | Spazio e figure | 39 | D26 | -0,4 | Spazio e figure |
| 18 | D13_a | -1,0 | Dati e previsioni | 40 | D27_a | -1,5 | Dati e previsioni |
| 19 | D13_b | -1,0 | Dati e previsioni | 41 | D27_b | -0,3 | Dati e previsioni |
| 20 | D13_c | 0,2 | Dati e previsioni | 42 | D28 | 0,4 | Numeri |
| 21 | D14_a | 1,1 | Numeri | 43 | D29 | 0,3 | Spazio e figure |
| 22 | D14_b | 1,7 | Numeri | - | - | - | - |

I valori riportati nelle tavole II.7 e II.8 confermano la tendenza emersa nell'analisi comparativa delle prove della scuola primaria. Senza creare eccessi o sproporzioni tra i diversi livelli di difficoltà delle domande, le prove rivolte alla prima classe della scuola secondaria di primo grado, quindi ad allievi che di norma hanno 12 anni, si caratterizzano per la distribuzione dei quesiti lungo tutta la scala di abilità, sia per Italiano sia per Matematica. Tale distribuzione lungo l'intera scala è molto importante poiché ciò significa che le prove sono in grado di rappresentare con precisione i livelli di preparazione raggiunti da tutti gli allievi, da quelli che hanno già raggiunto traguardi formativi elevati a quelli che si trovano in condizioni di minore vantaggio.

3.2.1. La Prova nazionale

A partire dal 2008, al termine della scuola secondaria di primo grado e all'interno dell'esame di Stato, si svolge la cosiddetta Prova nazionale. Essa risponde a finalità in parte diverse rispetto a quelle del Servizio nazionale di valutazione, tuttavia, dal punto di vista della costruzione delle prove non vi sono delle differenze sostanziali, sia dal punto di vista strettamente tecnico sia per quanto riguarda il quadro di riferimento. In ogni caso, poiché la Prova nazionale contribuisce, anche se solo nella misura di un settimo rispetto al totale, al voto finale di licenza, una particolare attenzione è

rivolta alla scelta della difficoltà dei quesiti, proprio per non alterare lo spirito generale che ispira l'esame di Stato conclusivo del primo ciclo di istruzione.

Le tavole II.9 e II.10 riportano i parametri di difficoltà delle prove di Italiano (denominata in sede di esame fascicolo 2) e Matematica (denominata in sede di esame fascicolo 1) della classe III della scuola secondaria di primo grado.

Tavola II.9. Parametri di difficoltà delle domande del fascicolo di Italiano (III sec. I gr.)

| TESTO LETTERARIO | | | TESTO ESPOSITIVO-ARGOMENTATIVO | | | GRAMMATICA | | |
|------------------|------|------------|--------------------------------|------|------------|------------|------|------------|
| N. item | Item | Difficoltà | N. item | Item | Difficoltà | N. item | Item | Difficoltà |
| 1 | A1 | 0,4 | 20 | B1 | -1,1 | 39 | C1 | -1,3 |
| 2 | A2 | -1,2 | 21 | B2 | -1,1 | 40 | C2 | -0,5 |
| 3 | A3 | -3,1 | 22 | B3 | 0,2 | 41 | C3 | -0,8 |
| 4 | A4 | -1,3 | 23 | B4 | 0,1 | 42 | C4 | -1,9 |
| 5 | A5 | -1,9 | 24 | B5 | -1,4 | 43 | C5 | -2,8 |
| 6 | A6 | -1,5 | 25 | B6 | 0,1 | 44 | C6 | -1,6 |
| 7 | A7 | -1,7 | 26 | B7 | -1,7 | 45 | C7 | -2,2 |
| 8 | A8 | -0,4 | 27 | B8 | -0,2 | 46 | C8 | -0,1 |
| 9 | A9 | -0,2 | 28 | B9 | -2,0 | 47 | C9 | 0,3 |
| 10 | A10 | -1,0 | 29 | B10 | -1,3 | 48 | C10 | -0,9 |
| 11 | A11 | -1,2 | 30 | B11 | -1,9 | 49 | C11 | -1,0 |
| 12 | A12 | -1,9 | 31 | B12 | -1,4 | - | - | - |
| 13 | A13 | -1,3 | 32 | B13 | -2,0 | - | - | - |
| 14 | A14 | -2,4 | 33 | B14 | -2,3 | - | - | - |
| 15 | A15 | -1,4 | 34 | B15 | -0,1 | - | - | - |
| 16 | A16 | -0,2 | 35 | B16 | -1,5 | - | - | - |
| 17 | A17 | -0,8 | 36 | B17 | -1,0 | - | - | - |
| 18 | A18 | -1,3 | 37 | B18 | -0,1 | - | - | - |
| 19 | A19 | -0,1 | 38 | B19 | -0,3 | - | - | - |

Tavola II.10. Parametri di difficoltà delle domande del fascicolo di Matematica (III sec. I gr.)

| N. item | Item | Difficoltà | Ambito | N. item | Item | Difficoltà | Ambito |
|---------|-------|------------|----------------------|---------|-------|------------|----------------------|
| 1 | D1 a | -2,2 | Dati e previsioni | 18 | D12 | 0,8 | Spazio e figure |
| 2 | D1 b | -2,2 | Dati e previsioni | 19 | D13 | -1,3 | Numeri |
| 3 | D2 | -0,8 | Numeri | 20 | D14 | 0,2 | Numeri |
| 4 | D3 a | -1,7 | Relazioni e funzioni | 21 | D15 | -1,0 | Dati e previsioni |
| 5 | D3 b | -1,8 | Relazioni e funzioni | 22 | D16 | -2,0 | Dati e previsioni |
| 6 | D3 c | -0,3 | Relazioni e funzioni | 23 | D17 | -0,4 | Relazioni e funzioni |
| 7 | D4 | -2,0 | Spazio e figure | 24 | D18 | -0,7 | Numeri |
| 8 | D5 | -1,0 | Relazioni e funzioni | 25 | D19 | -0,9 | Numeri |
| 9 | D6 a | 1,1 | Spazio e figure | 26 | D20 | 0,0 | Spazio e figure |
| 10 | D6 b | 1,3 | Spazio e figure | 27 | D21 a | -1,1 | Dati e previsioni |
| 11 | D7 | -1,8 | Numeri | 28 | D21 b | 0,2 | Dati e previsioni |
| 12 | D8 a | -0,8 | Numeri | 29 | D22 | -0,2 | Spazio e figure |
| 13 | D8 b | 0,9 | Numeri | 30 | D23 | -0,9 | Relazioni e funzioni |
| 14 | D9 | -0,4 | Spazio e figure | 31 | D24 | -1,5 | Relazioni e funzioni |
| 15 | D10 a | -0,5 | Relazioni e funzioni | 32 | D25 | -0,3 | Numeri |
| 16 | D10 b | 0,7 | Relazioni e funzioni | 33 | D26 | -2,0 | Spazio e figure |
| 17 | D11 | 2,1 | Dati e previsioni | - | - | - | - |

I dati delle tavole II.9 e II.10 confermano, soprattutto per la prova di Italiano, una certa prevalenza di domande a un livello medio e medio-basso di difficoltà, mentre per la prova di Matematica sono presenti anche quesiti con una difficoltà maggiore.

3.3. La scuola secondaria di secondo grado

In attuazione delle direttive ministeriali 74/2008 e 67/2010 le prove del SNV in quest'anno scolastico hanno coinvolto anche il secondo anno della scuola secondaria di secondo grado.

Prima di riproporre le analisi mostrate in precedenza per le altre classi interessate dalla rilevazione del SNV, sono necessarie alcune considerazioni di carattere generale. In primo luogo, le rilevazioni del SNV si collocano alla fine dell'obbligo di istruzione, quando gli studenti hanno, di norma, frequentato dieci anni di istruzione.

L'estensione delle prove del SNV anche al secondo anno della scuola secondaria di secondo grado ha reso necessarie alcune scelte; bisogna infatti tener conto del fatto che queste prove sono proposte a studenti che si trovano al termine dell'obbligo di istruzione, indipendentemente dal tipo di scuola e dall'indirizzo di studio frequentato e che devono quindi possedere competenze di base, sia se continueranno i loro studi sia se si inseriranno nel mondo del lavoro. In questa fase di avvio della rilevazione si è pertanto ritenuto opportuno predisporre prove uniche per tutti i tipi di scuola (licei, istituti tecnici e professionali), volte alla misurazione degli apprendimenti di base relativi a due ambiti fondamentali: 1) la comprensione di testi scritti di varia natura e la riflessione sulla lingua e 2) la matematica. La scelta di sottoporre agli studenti prove non differenziate rispetto agli indirizzi di studio implica necessariamente che esse facciano riferimento a contenuti, abilità e competenze comuni a tutti i tipi di scuola e che i quesiti proposti siano in grado di coprire tutta un'ampia gamma di livelli di difficoltà.

Effettuate le dovute precisazioni preliminari, anche per le prove della scuola secondaria di secondo grado è fondamentale prendere in considerazione gli stessi parametri proposti per le scuole del primo ciclo di istruzione.

Le tavole II.11 e II.12 riportano i parametri di difficoltà delle prove di Italiano e Matematica della classe II della scuola secondaria di secondo grado.

Tavola II.11. Parametri di difficoltà delle domande del fascicolo di Italiano (II sec. II gr.)

| TESTO 1 e 2 | | | TESTO 3 e 4 | | | GRAMMATICA | | |
|-------------|-------------------|------------|-------------|--------------------|------------|------------|-------|------------|
| N. item | Item | Difficoltà | N. item | Item | Difficoltà | N. item | Item | Difficoltà |
| 1 | A1 | -1,5 | 21 | C1 | -2,6 | 45 | E1 | -1,1 |
| 2 | A2 | -1,7 | 22 | C2 | -2,1 | 46 | E2 | -0,9 |
| 3 | B1 | -0,2 | 23 | C3 | -2,2 | 47 | E3 | -0,5 |
| 4 | B2 | 0,3 | 24 | C4 | 0,0 | 48 | E4 | -1,8 |
| 5 | B3 | -0,5 | 25 | D1 ¹⁰ | -0,5 | 49 | E5 | -0,5 |
| 6 | B4 | -2,3 | 26 | D2 | -2,8 | 50 | E6 | -0,3 |
| 7 | B5 | -2,0 | 27 | D3 | -0,8 | 51 | E7 | 0,3 |
| 8 | B6 ¹¹ | 1,1 | 28 | D4 | -0,7 | 52 | E8 a | 0,6 |
| 9 | B7 | -0,5 | 29 | D5 a | -1,8 | 53 | E8 b | 0,3 |
| 10 | B8 | 0,2 | 30 | D5 b | -1,7 | 54 | E9 | -1,8 |
| 11 | B9 a | -0,8 | 31 | D6 | -0,1 | 55 | E10 a | -1,2 |
| 12 | B9 b | -0,6 | 32 | D7 | -0,6 | 56 | E10 b | -2,1 |
| 13 | B10 | -1,5 | 33 | D8 | -2,1 | 57 | E10 c | -1,6 |
| 14 | B11 | -0,2 | 34 | D9 a | 0,1 | 58 | E10 d | -0,3 |
| 15 | B12 | -1,2 | 35 | D9 b ¹² | -0,7 | 59 | E11 a | -0,1 |
| 16 | B13 | 0,2 | 36 | D10 | -2,4 | 60 | E11 b | -0,4 |
| 17 | B14 | -0,2 | 37 | D11 | -0,1 | 61 | E11 c | 0,1 |
| 18 | B15 ¹³ | -1,3 | 38 | D12 | -0,1 | 62 | E12 a | -2,0 |
| 19 | B16 | -0,5 | 39 | D13 | -1,1 | 63 | E12 b | -1,9 |
| 20 | B17 | -0,8 | 40 | D14 | -1,5 | - | - | - |
| - | - | - | 41 | D15 a1 | -1,5 | - | - | - |
| - | - | - | 42 | D15 a2 | -1,7 | - | - | - |
| - | - | - | 43 | D15 a3 | -1,7 | - | - | - |
| - | - | - | 44 | D15 b | 0,0 | - | - | - |

¹⁰ La domanda D1 è a credito parziale, il passaggio dalla prima alla seconda categoria si ha con il valore -1,4.

¹¹ La domanda B6 è a credito parziale, il passaggio dalla prima alla seconda categoria si ha con il valore 0,2.

¹² La domanda D9_b è a credito parziale, il passaggio dalla prima alla seconda categoria si ha con il valore -1,7.

¹³ La domanda B15 è a credito parziale, il passaggio dalla prima alla seconda categoria si ha con il valore -1,2.

Tavola II.12. Parametri di difficoltà delle domande del fascicolo di Matematica (II sec. II gr.)

| N. item | Item | Difficoltà | Ambito | N. item | Item | Difficoltà | Ambito |
|---------|--------|-------------------|----------------------|---------|--------|------------|----------------------|
| 1 | MD1_a | -2,4 | Dati e Previsioni | 25 | MD13_a | -2,6 | Relazioni e Funzioni |
| 2 | MD1_b | -1,4 | Dati e Previsioni | 26 | MD13_b | 2,9 | Relazioni e Funzioni |
| 3 | MD1_c | -2,2 | Dati e Previsioni | 27 | MD13_c | 2,5 | Relazioni e Funzioni |
| 4 | MD2 | 0,2 | Dati e Previsioni | 28 | MD14 | 2,1 | Relazioni e Funzioni |
| 5 | MD3 | 0,0 | Spazio e Figure | 29 | MD15 | 1,3 | Numeri |
| 6 | MD4 | 0,3 | Numeri | 30 | MD16 | 1,5 | Numeri |
| 7 | MD5 | 2,6 | Numeri | 31 | MD17 | 0,3 | Spazio e Figure |
| 8 | MD6 | -0,4 | Dati e Previsioni | 32 | MD18 | 1,0 | Spazio e Figure |
| 9 | MD7_a | 1,0 | Relazioni e Funzioni | 33 | MD19 | -0,5 | Dati e Previsioni |
| 10 | MD7_b | 1,3 | Relazioni e Funzioni | 34 | MD20_a | -0,9 | Numeri |
| 11 | MD8_a | 0,3 | Spazio e Figure | 35 | MD20_b | 0,0 | Numeri |
| 12 | MD8_b | 0,3 | Spazio e Figure | 36 | MD20_c | -0,3 | Numeri |
| 13 | MD9_a | 0,3 ¹⁴ | Spazio e Figure | 37 | MD20_d | 0,6 | Numeri |
| 14 | MD9_b | 1,5 | Spazio e Figure | 38 | MD21 | 1,9 | Relazioni e Funzioni |
| 15 | MD10 | 2,4 | Numeri | 39 | MD22 | 1,6 | Numeri |
| 16 | MD11_a | -0,6 | Relazioni e Funzioni | 40 | MD23_a | 0,1 | Numeri |
| 17 | MD11_b | 0,9 | Relazioni e Funzioni | 41 | MD23_b | 0,7 | Numeri |
| 18 | MD11_c | 0,6 | Relazioni e Funzioni | 42 | MD24 | 0,5 | Relazioni e Funzioni |
| 19 | MD12_a | -2,2 | Dati e Previsioni | 43 | MD25 | 2,4 | Relazioni e Funzioni |
| 20 | MD12_b | -1,8 | Dati e Previsioni | 44 | MD26 | -0,2 | Relazioni e Funzioni |
| 21 | MD12_c | -0,5 | Dati e Previsioni | 45 | MD27 | 0,0 | Relazioni e Funzioni |
| 22 | MD12_d | -1,5 | Dati e Previsioni | 46 | MD28 | -0,6 | Numeri |
| 23 | MD12_e | 0,2 | Dati e Previsioni | 47 | MD29 | -0,1 | Numeri |
| 24 | MD12_f | -1,9 | Dati e Previsioni | 48 | MD30 | -0,4 | Spazio e Figure |

I dati della tavola II.11 mostrano che la prova di Italiano della seconda classe della scuola secondaria di secondo grado si caratterizza per la prevalenza di domande con un livello di difficoltà medio e medio-basso. Ciò è coerente con le decisioni assunte in fase di progettazione della prova poiché si è ritenuto opportuno, in una prima fase di avvio, predisporre una prova, specie per la comprensione della lettura e della grammatica, non troppo complessa, proprio per favorire un inserimento graduale delle prove standardizzate del SNV nella scuola secondaria di secondo grado.

Per la prova di Matematica si sono prese decisioni in parte diverse. Infatti, la sezione di matematica si caratterizza per la netta prevalenza di domande singole, non legate da un macro-stimolo unico, come nel caso dei testi (narrativo, espositivo, ecc.) della prova di comprensione della lettura. In questo caso è stato più semplice introdurre domande di complessità maggiore in grado di misurare con più precisione tutti i livelli di abilità, compresi quelli più elevati. Infatti, come si evince chiaramente dalla tavola II.12, le domande della prova di Matematica si distribuiscono in modo pressoché uniforme lungo tutta la scala di abilità.

¹⁴La domanda D9_a è a credito parziale, il passaggio dalla prima alla seconda categoria si ha con il valore -0,7.

4. Il placement relativo delle domande rispetto al punteggio di Rasch dei rispondenti

L'ultimo passo per la valutazione complessiva di una prova è lo studio del cosiddetto *placement* relativo tra difficoltà delle domande e livelli di apprendimento raggiunti dai rispondenti. Dal punto di vista teorico, il *placement* ideale è dato da una distribuzione della difficoltà delle domande in grado di coprire tutto l'intervallo della scala su cui insistono i livelli di apprendimento degli allievi. Nelle situazioni concrete ciò non è perfettamente realizzabile, specie nelle prove come quelle di comprensione della lettura caratterizzate da un insieme di domande legate a uno o più testi. È però importante che il *placement* sia soddisfacente, ovvero che non vi siano intervalli, soprattutto quelli intermedi, totalmente privi di domande o di allievi con livelli di prestazione che si collocano proprio in quegli intervalli. Inoltre, come già accennato nel commento alle tavole del paragrafo precedente, la scelta della proporzione di domande lungo la scala delle abilità deve tenere conto dell'età degli allievi ai quali la prova è presentata. In altre parole, quando ci si rivolge a bambini molto piccoli è opportuno che le domande tendano a popolare maggiormente i livelli medi e medio-bassi della scala di difficoltà, proprio per creare nel rispondente un senso di adesione alla verifica, tanto più importante, tanto più quest'ultimo si trova a frequentare i primi anni della scuola primaria.

Le figure da II.1 a II.10 mostrano il posizionamento relativo delle domande delle prove SNV e della Prova nazionale lungo la stessa scala sulla quale si rappresentano i risultati conseguiti dagli allievi.

Figura II.1. *Placement* della prova di Italiano (II primaria)

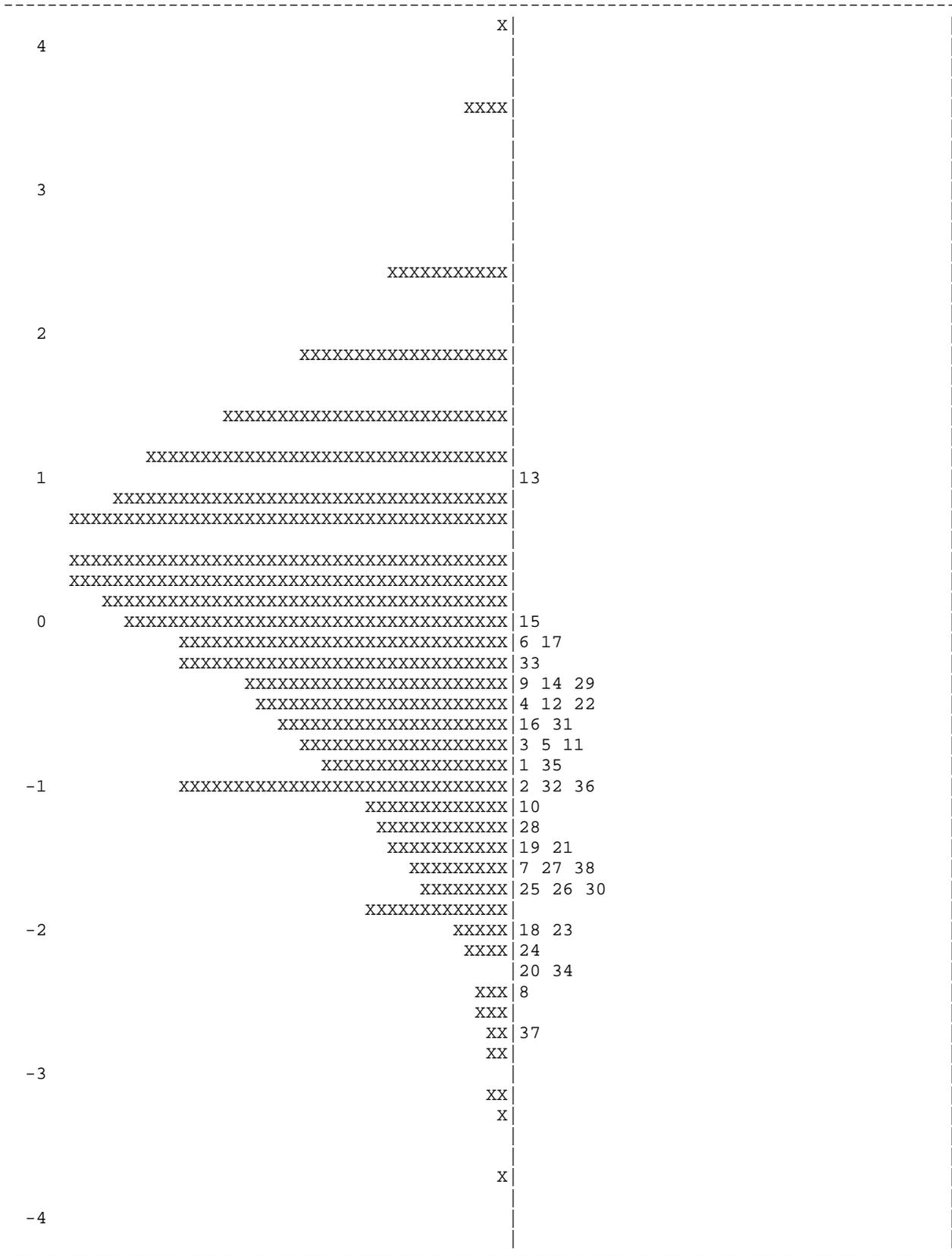


Figura II.2. *Placement* della prova di Matematica (II primaria)



I grafici delle figure II.1 e II.2 confermano i dati contenuti nelle tavole II.3 e II.4. Nella parte sinistra dei grafici si trova la distribuzione dei risultati degli allievi che hanno partecipato alla prova SNV (dati campionari) della seconda classe della scuola primaria, mentre nella parte destra la collocazione sulla stessa scala (asse verticale tratteggiato) della difficoltà delle domande¹⁵. Più precisamente, tutti gli studenti rappresentati dalle crocette che si trovano in una posizione inferiore a quella occupata da una data domanda hanno una scarsa probabilità (inferiore al 50 per cento) di rispondere correttamente alla domanda medesima, mentre coloro che sono posizionati rispetto all'asse verticale al di sopra di una domanda hanno una probabilità alta (superiore al 50 per cento) di rispondere correttamente e tale probabilità aumenta all'aumentare della distanza studente-domanda.

In questa sede è bene rimarcare la differenza tra una prova tradizionale di scuola e una prova standardizzata a carattere nazionale. Quest'ultima si pone come obiettivo principale quello di misurare i livelli di apprendimento raggiunti dagli studenti che frequentano le scuole italiane, ovvero si propone di misurare i risultati raggiunti in un determinato ambito da allievi con diversi livelli di abilità potenzialmente posizionati su ogni punto dell'asse verticale, ad esempio, delle figure II.1 e II.2. Ciò significa che in una prova standardizzata nazionale o somministrata su larga scala è assolutamente normale che vi siano alcuni quesiti ai quali solo una piccola minoranza di studenti è in grado di rispondere correttamente, senza che da questo si debba dedurre che la preparazione dei rispondenti sia inadeguata. Infatti, queste particolari domande sono finalizzate all'individuazione di livelli di apprendimento particolarmente elevati. In assenza di queste domande più difficili la prova non sarebbe in grado di dare indicazioni circa i livelli di apprendimento degli studenti più abili.

Le figure da II.3 a II.10 illustrano il *placement* relativo delle prove di Italiano e Matematica della V primaria, della I e III secondaria di primo grado e della II secondaria di secondo grado.

¹⁵ Le domande sono etichettate in base al numero progressivo come riportato nelle tavole II.3-II.12.

Figura II.3. *Placement* della prova di Italiano (V primaria)

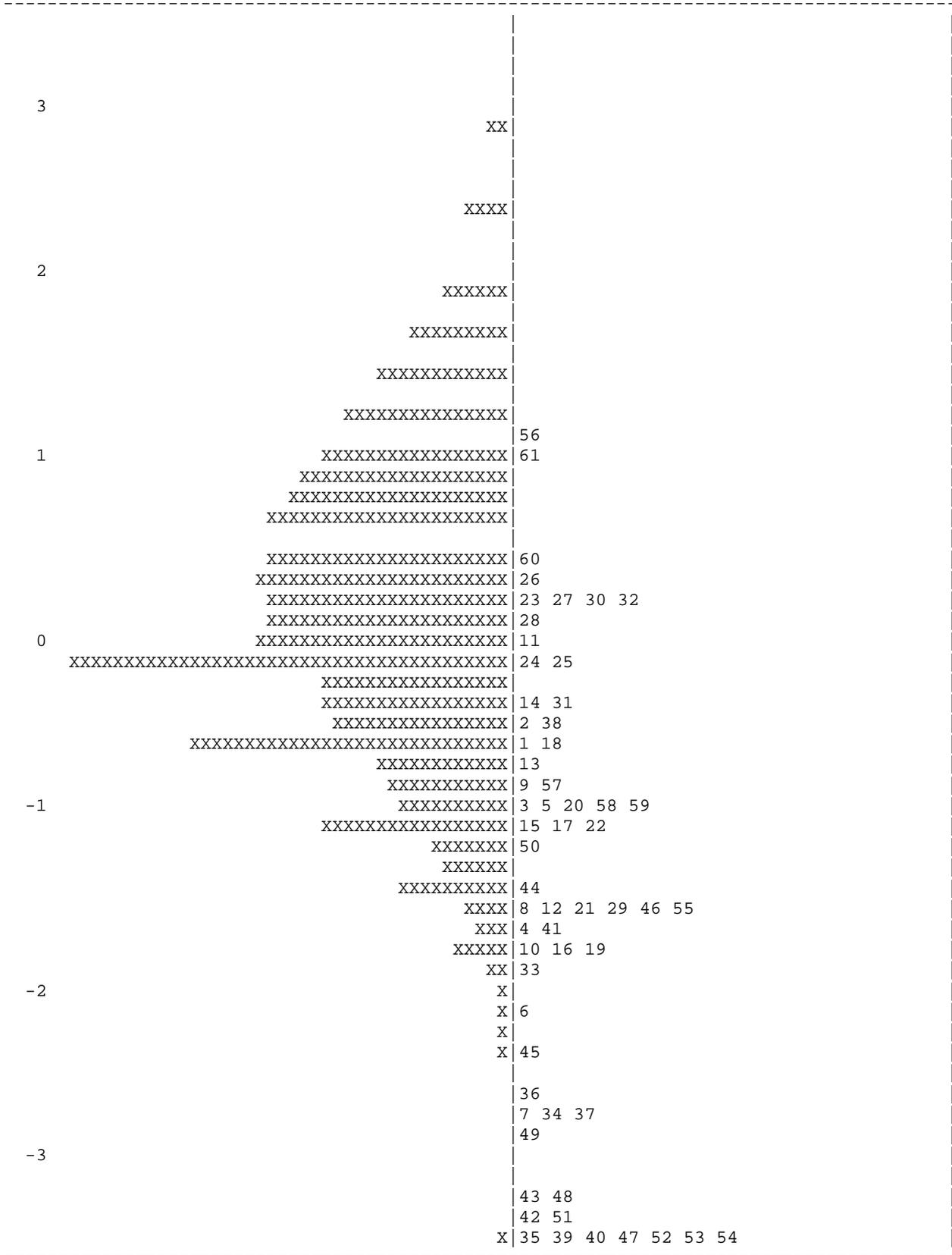


Figura II.4. *Placement* della prova di Matematica (V primaria)

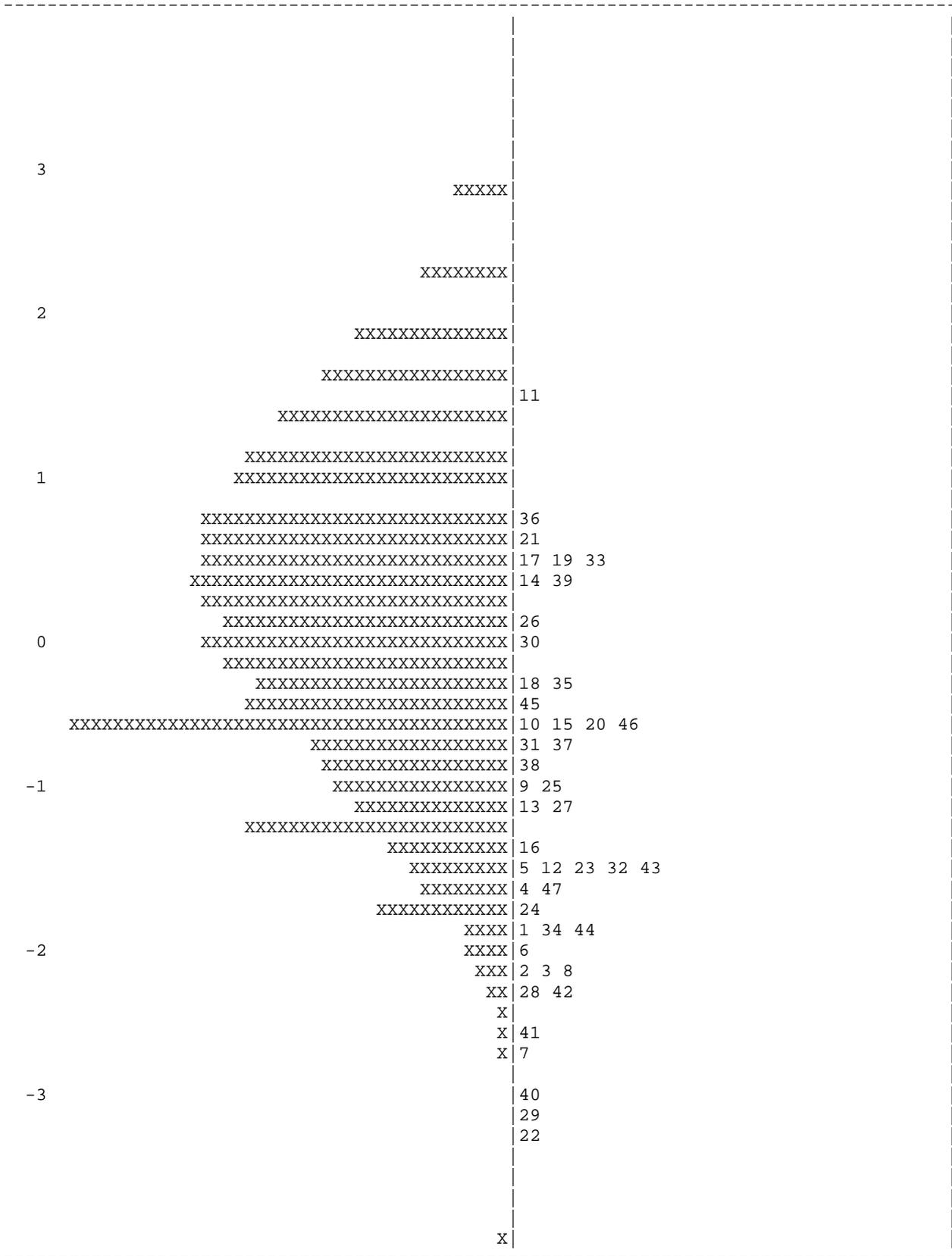


Figura II.6. *Placement* della prova di Matematica (I sec. I gr.)

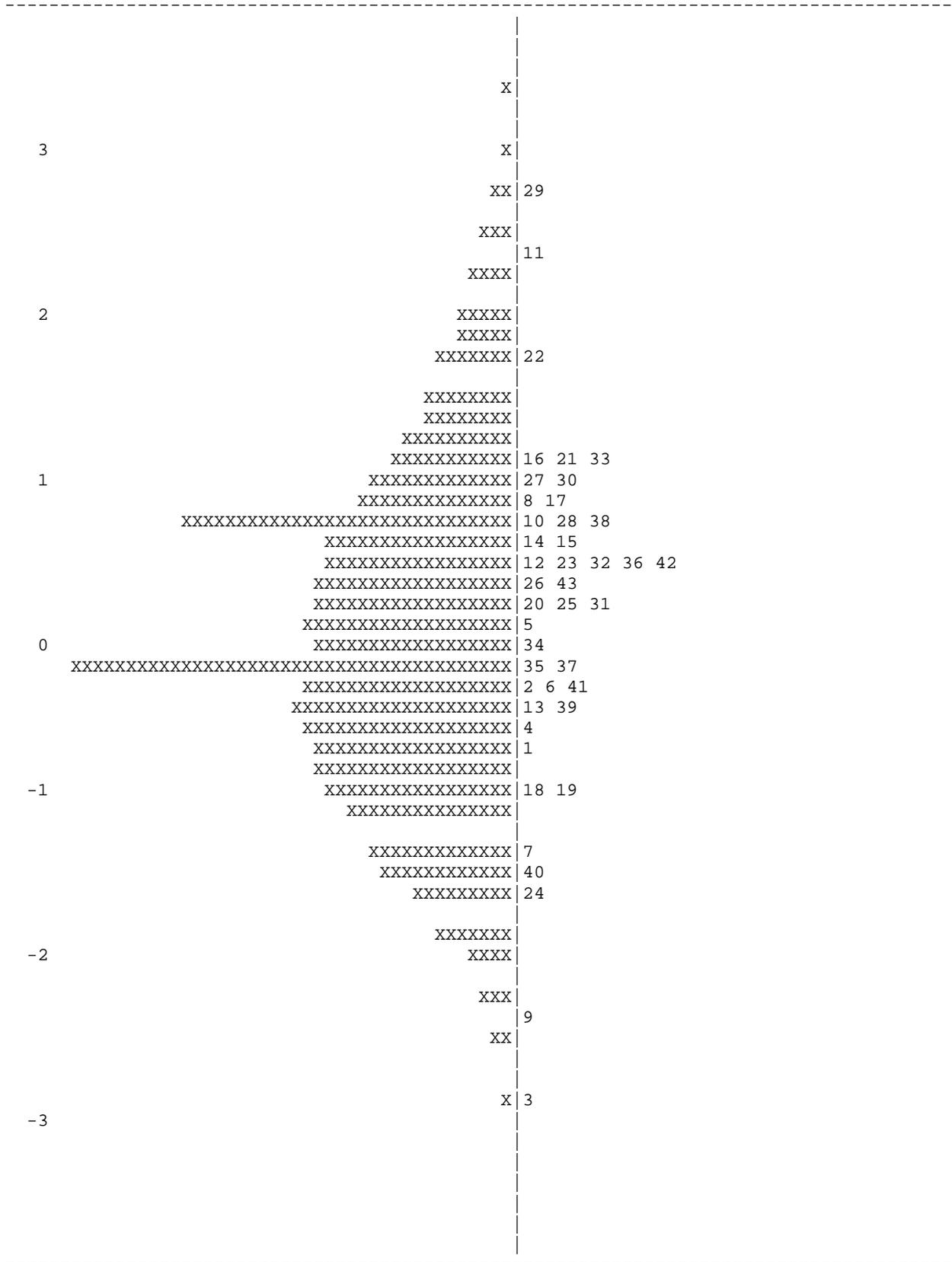


Figura II.7. *Placement* della prova di Italiano (III sec. I gr.)

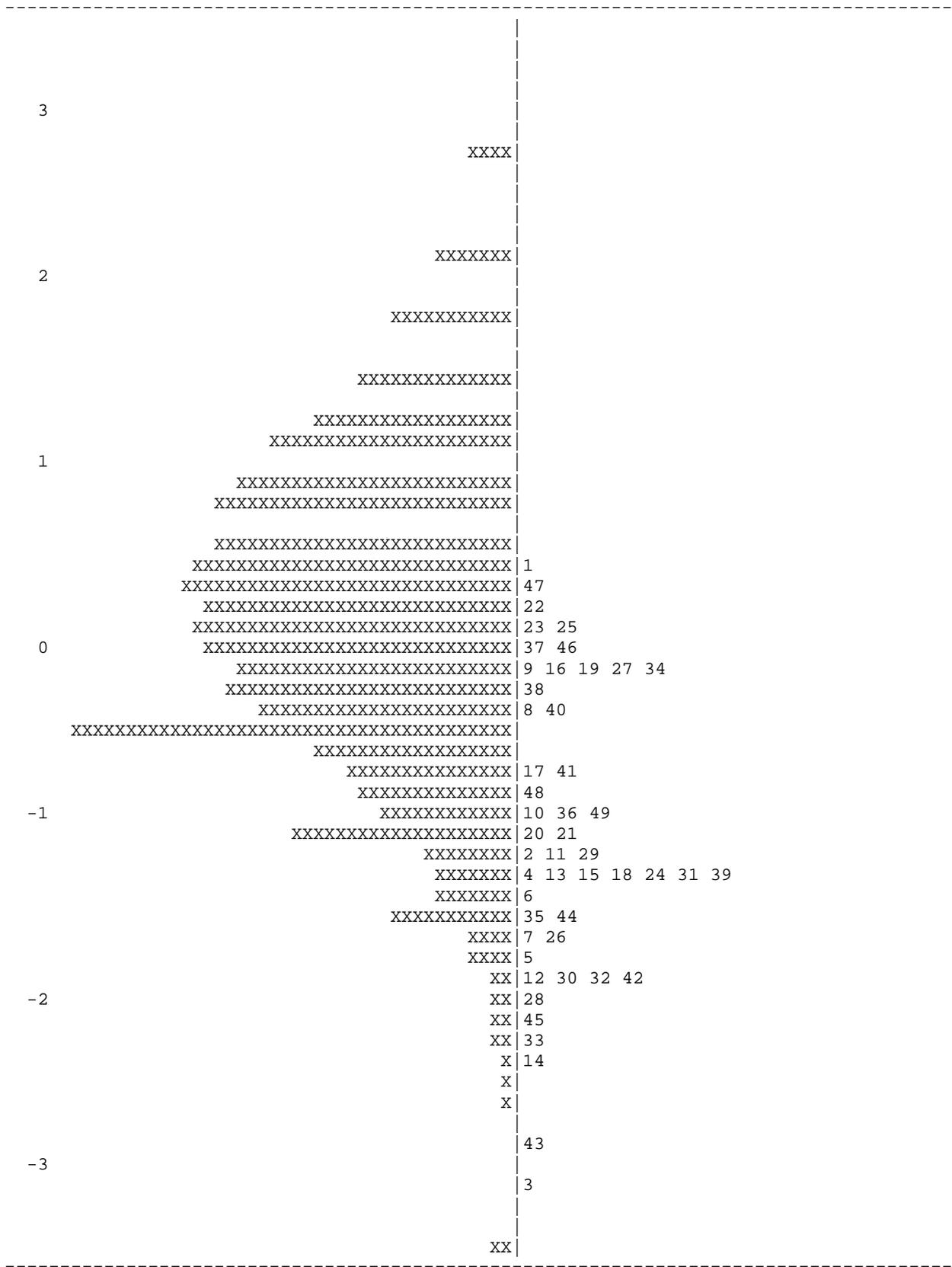


Figura II.8. *Placement* della prova di Matematica (III sec. I gr.)

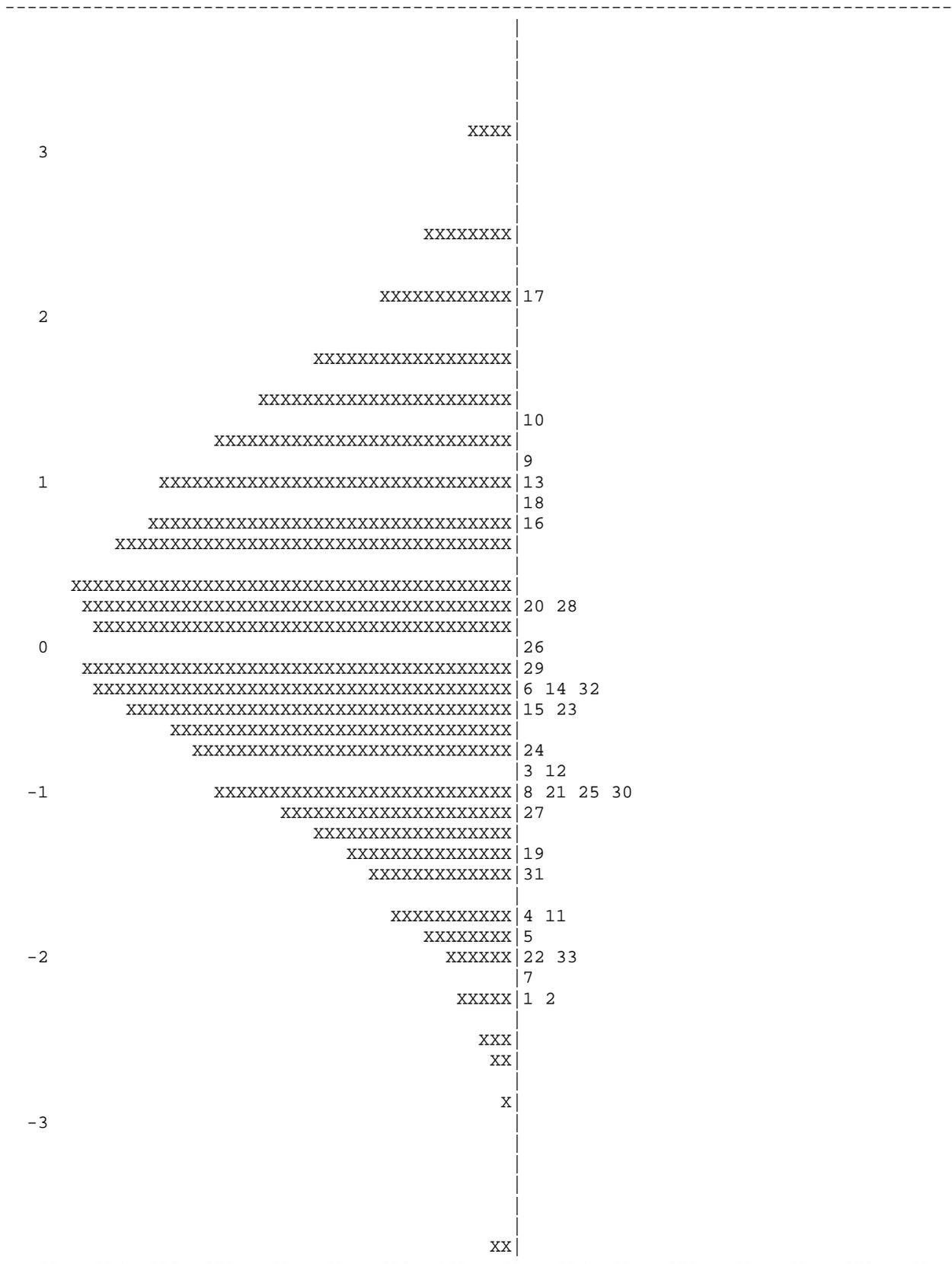


Figura II.9. *Placement* della prova di Italiano (II sec. II gr.)

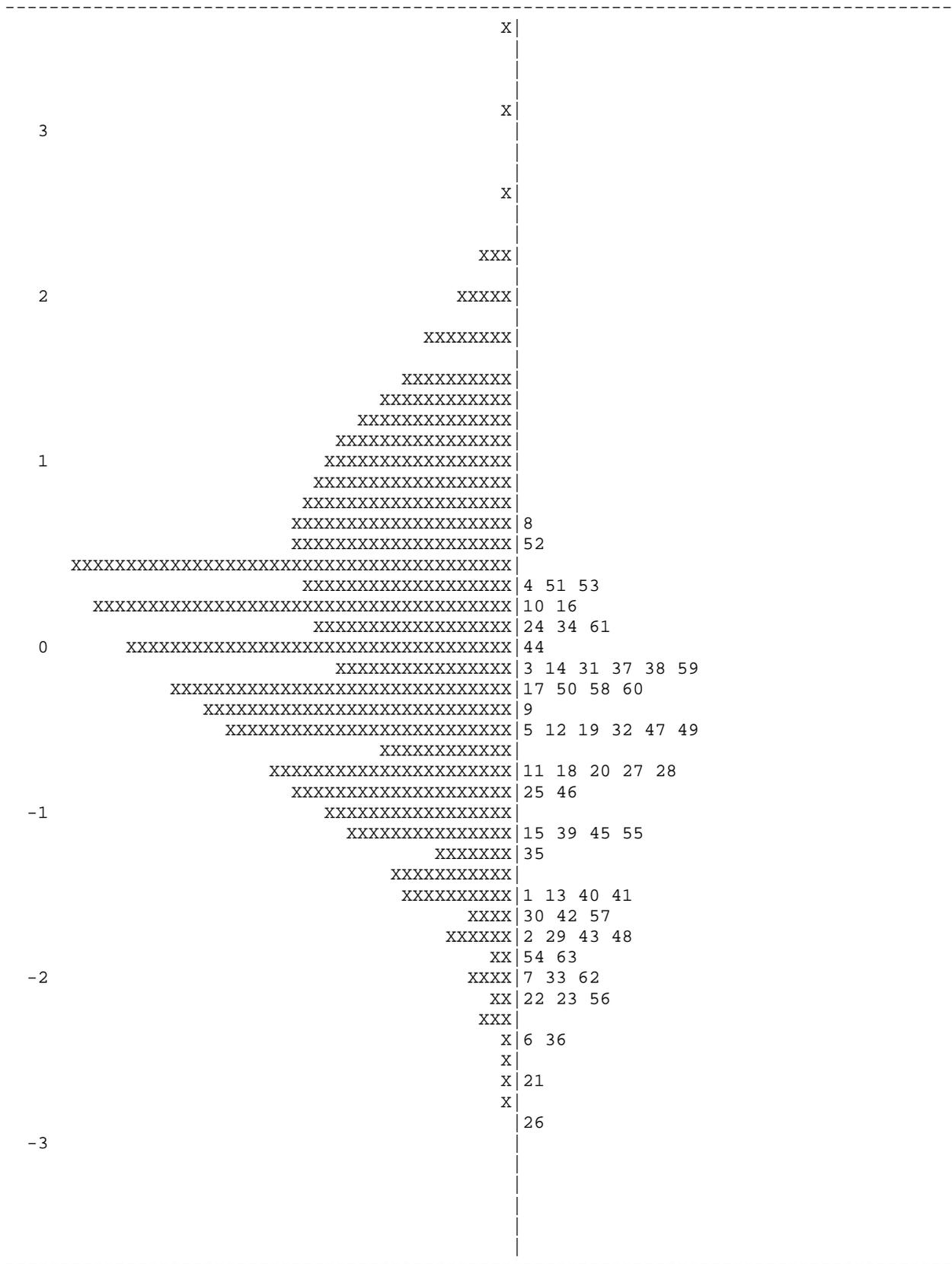
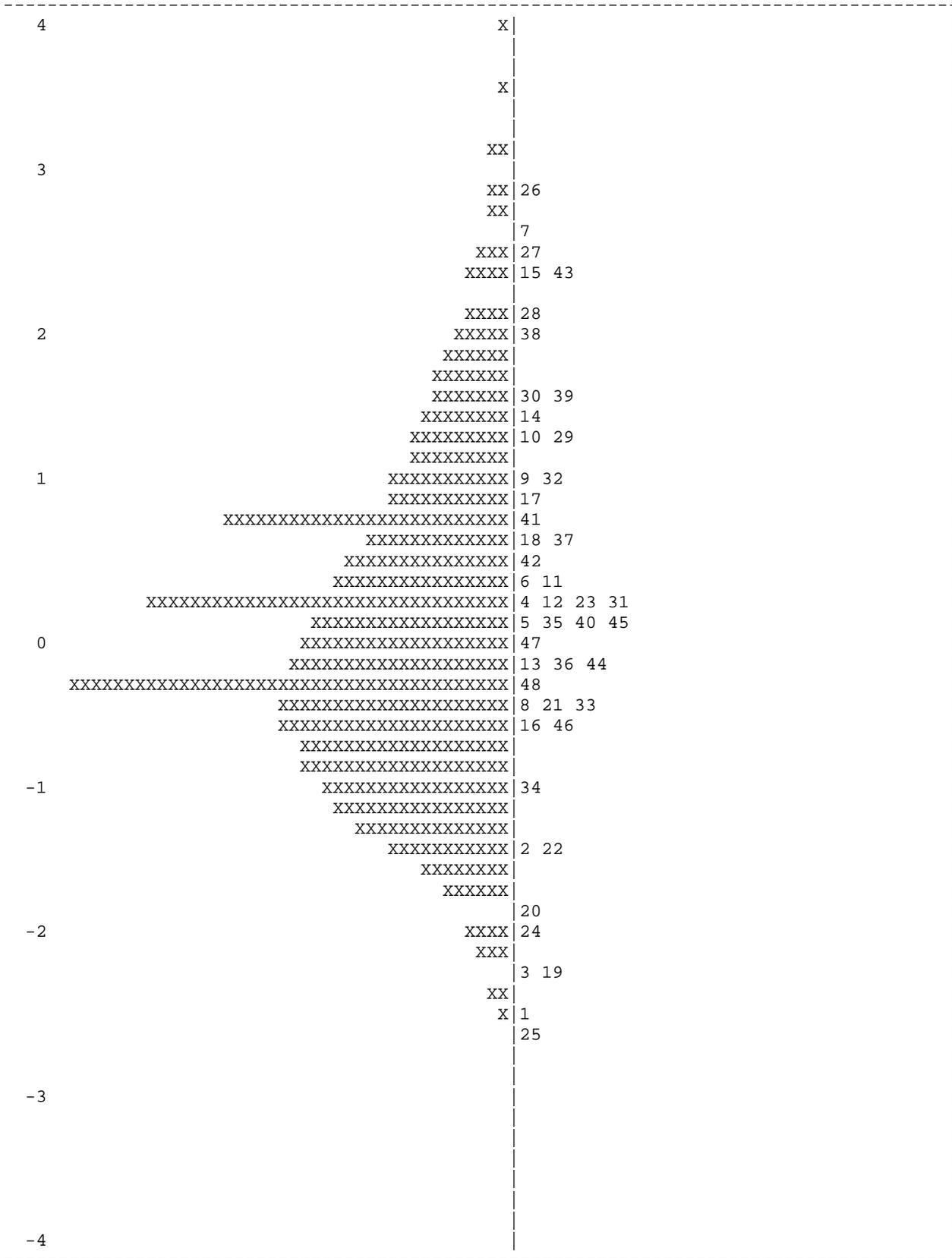


Figura II.10. *Placement* della prova di Matematica (II sec. II gr.)



Se si osservano in sequenza le figure II.1-II.10 si può avere un'idea abbastanza precisa delle modalità con le quali sono state costruite le prove SNV per le diverse classi e quelle della Prova nazionale e, conseguentemente, delle scelte operative che hanno permesso di realizzare quanto suggerito dalla letteratura specifica nel campo della costruzione delle prove standardizzate. In particolare, si può notare come sia necessario trovare soluzioni convincenti e adatte all'età degli alunni per inserire, specie nelle prove di Italiano, un numero più elevato di quesiti in grado di fornire una misura più precisa e particolareggiata dei livelli più alti della scala di abilità.

5. L'indice di discriminazione statistica

Le analisi fin qui proposte mancano ancora di un elemento per esprimere un giudizio complessivo adeguatamente approfondito sulle prove e su ciascuna domanda che le compongono. È necessario valutare se le diverse domande contengano o meno delle ambiguità che possano aver tratto in inganno i rispondenti. Per ogni domanda è stata valutata sia in sede di pre-test sia *ex post*, ovvero sui risultati della rilevazione vera e propria, la capacità misuratoria di ciascuna opzione. In termini più espliciti, in una domanda formulata in modo adeguato e priva di ambiguità solo l'opzione corretta si deve associare positivamente con il risultato complessivo della prova, nel senso che coloro che rispondono correttamente a una data domanda devono, in media, ottenere un risultato complessivo nella prova migliore di coloro che rispondono in modo errato o non rispondono affatto.

La tavola II.13 mostra, a titolo esemplificativo, le analisi che devono essere condotte su ciascuna domanda per verificarne l'adeguatezza misuratoria.

Tavola II.13. Dati generali di una domanda a risposta multipla con quattro opzioni (C = risp. corretta)

| | | Ricostruire il significato globale del testo, integrando più informazioni e concetti | | | | |
|-------------------------|-----------|--|------------------------------|-------------------------------|---------------|---------------------|
| <i>Processi:</i> | | <i>Indice di discriminazione:</i> | 0,28 | <i>Livello di difficoltà:</i> | -0,14 | |
| | Frequenza | Frequenza percentuale | Correlazione punto biseriale | t(p) | Abilità media | Deviazione standard |
| | (2) | (3) | (4) | (5) | (6) | (7) |
| A | 86852 | 15,0 | -0,14 | -22,67(,000) | -0,25 | 0,69 |
| B | 121568 | 21,0 | -0,09 | -14,94(,000) | -0,13 | 0,69 |
| C | 306743 | 53,1 | 0,28 | 46,33(,000) | 0,19 | 0,77 |
| D | 55225 | 9,6 | -0,13 | -21,79(,000) | -0,3 | 0,71 |
| Risposta non valida (7) | 2886 | 0,5 | -0,05 | -8,08(,000) | -0,45 | 0,73 |
| Item non raggiunto (8) | 114 | 0,0 | -0,06 | -10,37(,000) | -2,78 | 0,02 |
| Mancata risposta (9) | 4528 | 0,8 | -0,08 | -12,35(,000) | -0,54 | 0,8 |

La prima indicazione (processi) consente di collegare la specifica domanda al quadro di riferimento mediante l'individuazione del processo cui si riferisce il quesito oggetto d'interesse. L'utilizzatore di queste analisi può pertanto esaminare tutte le domande di un determinato livello e quindi comprendere, anche in chiave di processi indagati, quali siano le caratteristiche di queste domande e, di conseguenza, verificare se un determinato gruppo di allievi abbia raggiunto o meno quel dato livello.

L'indice di discriminazione consente di valutare in termini quantitativi la capacità di una domanda di individuare gruppi di rispondenti in funzione del loro livello di preparazione. In altri termini, tanto più una domanda è *discriminativa*, tanto più essa è in grado di misurare la variazione di probabilità di fornire la risposta corretta anche per piccole variazioni di abilità del rispondente. Gli standard di letteratura indicano che l'indice di discriminazione deve raggiungere almeno il valore 0,20 e può considerarsi buono quando supera i valori di 0,25-0,28¹⁶.

I dati contenuti nella tavola II.13 si riferiscono a un'ipotetica domanda di Italiano con risposta a scelta multipla in cui l'opzione "C" è quella corretta. La prima valutazione molto semplice è effettuata sulla base dei dati delle colonne (2) e (3) che riportano le frequenze assolute e percentuali delle risposte. Esse indicano, con tutta evidenza, che la domanda oggetto d'interesse è abbastanza semplice. Naturalmente, questa è un'informazione del tutto insufficiente per valutarne

¹⁶ Si veda *Pisa Technical Report* (2003) pag. 123.

l'adeguatezza della domanda, sia in relazione al potere misuratorio dell'opzione di risposta corretta che alle altre opzioni. I dati della colonna (4) sono in grado di fornire un'informazione molto importante per ogni opzione di risposta. La *correlazione punto-biserial* rappresenta la correlazione tra la probabilità di scegliere una data opzione e l'abilità complessiva dell'allievo. Tale legame deve essere negativo, ovvero inverso, per le opzioni di risposta non corrette e positivo solo per quella esatta. Più precisamente, una domanda a risposta multipla è ben formulata se, in media, coloro che rispondono correttamente a quella domanda lo fanno anche in buona parte delle altre. Con tutta evidenza, ciò vale a condizione che la prova sia coerente e quindi i risultati forniti siano attendibili¹⁷. Per esempio, la domanda in esame nella tavola II.13 mette in evidenza che tendono a fornire la risposta corretta gli studenti che rispondono bene alla maggior parte delle domande (la correlazione punto-biserial è +0,28). Al contrario, scelgono le opzioni non corrette gli allievi con punteggi complessivi più bassi (le correlazioni punto-biseriali sono tutte negative). I dati della colonna (5) consentono di valutare la significatività della correlazione punto-biserial. Nel caso in esame, tutti i valori della colonna (4) sono significativi in senso propriamente statistico¹⁸. Infine, la colonna (6) permette di valutare su una scala di Rasch il livello medio di *performance* degli allievi che scelgono una determinata risposta, mentre la colonna (7) fornisce una misura della variabilità di ciascuno dei risultati riportati in colonna (6). È importante notare che, in media, il livello di preparazione degli allievi che scelgono l'opzione corretta di risposta è più elevato di quello di coloro che scelgono le altre opzioni e ciò fornisce un'ulteriore indicazione positiva circa l'adeguatezza della domanda presa in esame.

Nelle Appendici 1 - 5 sono riportate le stesse analisi della tavola II.13 per ciascuna domanda delle prove SNV e della Prova nazionale. Inoltre, le tavole II.14.-II.23. riportano i valori del coefficiente di correlazione punto-biserial per le opzioni di risposta corrette di tutte le domande delle prove SNV e della Prova nazionale.

¹⁷ La valutazione dell'attendibilità e coerenza complessiva della prova è effettuata mediante l' α di Cronbach.

¹⁸ Dal punto di vista propriamente statistico un determinato valore di sintesi, in questo caso il coefficiente di correlazione punto-biserial, è statisticamente significativo se il valore $t(p)$ ha una probabilità, indicata tra parentesi, di prodursi per solo caso inferiore a 0,05 (5%). Nel caso in esame, tutti i valori di $t(p)$ mostrano valori di probabilità sostanzialmente nulli e, pertanto, tutti i coefficienti di correlazione punto-biserial sono statisticamente significativi.

Tavola II.14. Coefficienti di correlazione punto-biseriali delle opzioni corrette di risposte
(Italiano - II primaria)

| COMPRESIONE DELLA LETTURA | | ESERCIZIO 1 ¹⁹ | | ESERCIZIO 2 ²⁰ | |
|---------------------------|------------------------------|---------------------------|------------------------------|---------------------------|------------------------------|
| Item | Correlazione punto biseriale | Item | Correlazione punto biseriale | Item | Correlazione punto biseriale |
| A1 | 0,39 | B1 | 0,48 | C1 | 0,42 |
| A2 | 0,36 | B2 | 0,53 | C2 | 0,50 |
| A3 | 0,42 | B3 | 0,61 | C3 | 0,48 |
| A4 | 0,30 | B4 | 0,47 | C4 | 0,46 |
| A5 | 0,44 | B5 | 0,50 | C5 | 0,48 |
| A6 | 0,38 | B6 | 0,49 | - | - |
| A7 | 0,35 | B7 | 0,63 | - | - |
| A8 | 0,40 | B8 | 0,61 | - | - |
| A9 | 0,40 | B9 | 0,57 | - | - |
| A10 | 0,44 | B10 | 0,48 | - | - |
| A11 | 0,24 | B11 | 0,42 | - | - |
| A12 | 0,34 | B12 | 0,61 | - | - |
| A13 | 0,23 | B13 | 0,45 | - | - |
| A14 | 0,39 | B14 | 0,51 | - | - |
| A15 | 0,36 | B15 | 0,35 | - | - |
| A16 | 0,48 | - | - | - | - |
| A17 | 0,31 | - | - | - | - |
| A18 | 0,44 | - | - | - | - |

¹⁹ L'esercizio richiede all'allievo di indicare se quindici coppie di parole hanno significato uguale o contrario.

²⁰ L'esercizio richiede il collegamento tra le due parti in cui sono suddivise cinque frasi.

Tavola II.15. Coefficienti di correlazione punto-biseriali delle opzioni corrette di risposte
(Matematica-II primaria)

| Item | Correlazione punto biseriale | Ambito |
|-------|------------------------------|-------------------|
| D1_a | 0,40 | Dati e previsioni |
| D1_b | 0,43 | Dati e previsioni |
| D2_a | 0,55 | Numeri |
| D2_b | 0,52 | Numeri |
| D3_a | 0,56 | Numeri |
| D3_b | 0,52 | Numeri |
| D4 | 0,11 | Spazio e figure |
| D5_a | 0,37 | Numeri |
| D5_b | 0,35 | Numeri |
| D5_c | 0,49 | Numeri |
| D6 | 0,31 | Numeri |
| D7_a | 0,29 | Spazio e figure |
| D7_b | 0,47 | Spazio e figure |
| D8 | 0,38 | Numeri |
| D9_a | 0,47 | Numeri |
| D9_b | 0,46 | Numeri |
| D9_c | 0,51 | Numeri |
| D9_d | 0,44 | Numeri |
| D10_a | 0,38 | Spazio e figure |
| D10_b | 0,29 | Spazio e figure |
| D11 | 0,49 | Numeri |
| D12 | 0,47 | Numeri |
| D13 | 0,37 | Spazio e figure |
| D14 | 0,29 | Numeri |
| D15 | 0,38 | Numeri |
| D16 | 0,30 | Spazio e figure |
| D17 | 0,31 | Numeri |
| D18 | 0,38 | Spazio e figure |

Tavola II.16. Coefficienti di correlazione punto-biseriali delle opzioni corrette di risposte
(Italiano-V primaria)

| TESTO NARRATIVO | | TESTO ESPOSITIVO | | GRAMMATICA | |
|-----------------|------------------------------|------------------|------------------------------|------------|------------------------------|
| Item | Correlazione punto biseriale | Item | Correlazione punto biseriale | Item | Correlazione punto biseriale |
| A1 | 0,25 | B1 | 0,46 | C1_a1 | 0,28 |
| A2 | 0,46 | B2 | 0,49 | C1_a2 | 0,33 |
| A3 | 0,47 | B3 | 0,31 | C1_b1 | 0,26 |
| A4 | 0,43 | B4 | 0,35 | C1_b2 | 0,32 |
| A5 | 0,38 | B5 | 0,37 | C1_b3 | 0,32 |
| A6 | 0,47 | B6 | 0,28 | C2 | 0,24 |
| A7 | 0,31 | B7 | 0,45 | C3_a | 0,22 |
| A8 | 0,41 | B8 | 0,39 | C3_b | 0,20 |
| A9 | 0,27 | B9 | 0,44 | C3_c | 0,38 |
| A10 | 0,47 | B10 | 0,31 | C3_d | 0,32 |
| A11 | 0,22 | B11 | 0,23 | C3_e | 0,19 |
| A12 | 0,44 | B12 | 0,34 | C3_f | 0,34 |
| A13 | 0,42 | B13 | 0,35 | C3_g | 0,34 |
| A14 | 0,44 | B14 | 0,42 | C3_h | 0,37 |
| A15 | 0,39 | B15 | 0,41 | C3_i | 0,23 |
| A16 | 0,35 | - | - | C3_l | 0,34 |
| A17 | 0,19 | - | - | C3_m | 0,34 |
| - | - | - | - | C3_n | 0,27 |
| - | - | - | - | C3_o | 0,27 |
| - | - | - | - | C3_p | 0,23 |
| - | - | - | - | C3_q | 0,22 |
| - | - | - | - | C3_r | 0,21 |
| - | - | - | - | C4 | 0,40 |
| - | - | - | - | C5 | 0,37 |
| - | - | - | - | C6 | 0,29 |
| - | - | - | - | C7 | 0,38 |
| - | - | - | - | C8 | 0,43 |
| - | - | - | - | C9 | 0,35 |
| - | - | - | - | C10 | 0,35 |

Tavola II.17. Coefficienti di correlazione punto-biseriali delle opzioni corrette di risposte
(Matematica-V primaria)

| Item | Correlazione punto biseriale | Ambito | Item | Correlazione punto biseriale | Ambito |
|-------|------------------------------|----------------------|-------|------------------------------|----------------------|
| D1_a | 0,35 | Dati e previsioni | D18 | 0,54 | Numeri |
| D1_b | 0,33 | Dati e previsioni | D19 | 0,54 | Numeri |
| D1_c | 0,37 | Dati e previsioni | D20 | 0,23 | Spazio e figure |
| D1_d | 0,38 | Dati e previsioni | D21_a | 0,27 | Dati e previsioni |
| D2 | 0,46 | Numeri | D21_b | 0,43 | Dati e previsioni |
| D3 | 0,33 | Numeri | D22 | 0,48 | Numeri |
| D4_a | 0,30 | Relazioni e funzioni | D23_a | 0,46 | Spazio e figure |
| D4_b | 0,35 | Relazioni e funzioni | D23_b | 0,40 | Spazio e figure |
| D5 | 0,46 | Numeri | D24_a | 0,45 | Relazioni e funzioni |
| D6 | 0,41 | Relazioni e funzioni | D24_b | 0,57 | Relazioni e funzioni |
| D7 | 0,15 | Relazioni e funzioni | D24_c | 0,51 | Relazioni e funzioni |
| D8 | 0,47 | Numeri | D25 | 0,39 | Dati e previsioni |
| D9 | 0,44 | Spazio e figure | D26 | 0,32 | Spazio e figure |
| D10 | 0,36 | Spazio e figure | D27 | 0,28 | Spazio e figure |
| D11 | 0,44 | Numeri | D28_a | 0,27 | Relazioni e funzioni |
| D12 | 0,35 | Numeri | D28_b | 0,23 | Relazioni e funzioni |
| D13 | 0,39 | Relazioni e funzioni | D28_c | 0,28 | Relazioni e funzioni |
| D14 | 0,43 | Numeri | D29_a | 0,45 | Numeri |
| D15 | 0,40 | Spazio e figure | D29_b | 0,35 | Numeri |
| D16_a | 0,48 | Spazio e figure | D29_c | 0,54 | Numeri |
| D16_b | 0,46 | Spazio e figure | D29_d | 0,48 | Numeri |
| D17_a | 0,25 | Dati e previsioni | D30 | 0,40 | Spazio e figure |
| D17_b | 0,23 | Dati e previsioni | - | - | - |
| D17_c | 0,31 | Dati e previsioni | - | - | - |
| D17_d | 0,24 | Dati e previsioni | - | - | - |

Tavola II.18. Coefficienti di correlazione punto-biseriali delle opzioni corrette di risposte
(Italiano-I sec. di I gr.)

| TESTO NARRATIVO | | TESTO ESPOSITIVO | | GRAMMATICA | |
|-----------------|------------------------------|------------------|------------------------------|------------|------------------------------|
| Item | Correlazione punto biseriale | Item | Correlazione punto biseriale | Item | Correlazione punto biseriale |
| A1 | 0,39 | B1 | 0,44 | C1_1 | 0,45 |
| A2 | 0,29 | B2 | 0,37 | C1_2 | 0,37 |
| A3 | 0,35 | B3 | 0,24 | C1_3 | 0,42 |
| A4 | 0,39 | B4 | 0,37 | C2 | 0,43 |
| A5 | 0,22 | B5 | 0,35 | C3_a | 0,29 |
| A6 | 0,37 | B6_a | 0,35 | C3_b | 0,28 |
| A7 | 0,28 | B6_b | 0,29 | C3_c | 0,28 |
| A8 | 0,34 | B6_c | 0,35 | C3_d | 0,24 |
| A9 | 0,26 | B6_d | 0,30 | C3_e | 0,34 |
| A10 | 0,39 | B6_e | 0,23 | C3_f | 0,29 |
| A11 | 0,40 | B7 | 0,44 | C4_a | 0,45 |
| A12 | 0,21 | B8 | 0,40 | C4_b | 0,41 |
| A13 | 0,20 | B9 | 0,35 | C4_c | 0,42 |
| A14 | 0,37 | B10 | 0,29 | C5_a | 0,23 |
| A15 | 0,35 | B11 | 0,31 | C5_b | 0,26 |
| A16 | 0,21 | B12 | 0,40 | C5_c | 0,35 |
| A17 | 0,42 | B13 | 0,32 | C5_d | 0,04 |
| A18 | 0,06 | - | - | C6 | 0,38 |
| A19 | 0,33 | - | - | C7 | 0,35 |
| A20 | 0,27 | - | - | C8 | 0,39 |
| - | - | - | - | C9_aa | 0,52 |
| - | - | - | - | C9_ab | 0,49 |
| - | - | - | - | C9_ac | 0,48 |
| - | - | - | - | C9_ad | 0,54 |
| - | - | - | - | C9_ae | 0,52 |
| - | - | - | - | C9_af | 0,52 |
| - | - | - | - | C9_ag | 0,33 |
| - | - | - | - | C9_ah | 0,53 |
| - | - | - | - | C9_ai | 0,50 |
| - | - | - | - | C9_al | 0,47 |
| - | - | - | - | C9_am | 0,54 |
| - | - | - | - | C9_an | 0,32 |
| - | - | - | - | C9_ba | 0,57 |
| - | - | - | - | C9_bb | 0,56 |
| - | - | - | - | C9_bc | 0,55 |
| - | - | - | - | C9_bd | 0,60 |
| - | - | - | - | C9_be | 0,55 |
| - | - | - | - | C9_bf | 0,57 |
| - | - | - | - | C9_bg | 0,55 |
| - | - | - | - | C9_bh | 0,60 |
| - | - | - | - | C9_bi | 0,55 |
| - | - | - | - | C9_bl | 0,59 |
| - | - | - | - | C9_bm | 0,59 |
| - | - | - | - | C9_bn | 0,57 |

Tavola II.19. Coefficienti di correlazione punto-biseriali delle opzioni corrette di risposte
(Matematica- I sec. di I gr.)

| Item | Correlazione punto biseriale | Ambito | Item | Correlazione punto biseriale | Ambito |
|-------|------------------------------|----------------------|-------|------------------------------|----------------------|
| D1 | 0,37 | Numeri | D15 | 0,53 | Dati e previsioni |
| D2 | 0,52 | Spazio e figure | D16_a | 0,44 | Spazio e figure |
| D3_a | 0,17 | Relazioni e funzioni | D16_b | 0,42 | Spazio e figure |
| D3_b | 0,46 | Relazioni e funzioni | D17_a | 0,43 | Spazio e figure |
| D3_c | 0,50 | Relazioni e funzioni | D17_b | 0,39 | Spazio e figure |
| D3_d | 0,47 | Relazioni e funzioni | D18 | 0,41 | Numeri |
| D4 | 0,33 | Dati e previsioni | D19 | 0,32 | Relazioni e funzioni |
| D5 | 0,46 | Spazio e figure | D20 | 0,42 | Relazioni e funzioni |
| D6 | 0,25 | Spazio e figure | D21 | 0,52 | Dati e previsioni |
| D7 | 0,43 | Spazio e figure | D22_a | 0,28 | Numeri |
| D8 | 0,39 | Numeri | D22_b | 0,32 | Numeri |
| D9 | 0,21 | Spazio e figure | D22_c | 0,38 | Numeri |
| D10_a | 0,37 | Relazioni e funzioni | D23 | 0,43 | Numeri |
| D10_b | 0,51 | Relazioni e funzioni | D24 | 0,31 | Numeri |
| D10_c | 0,39 | Relazioni e funzioni | D25_a | 0,48 | Relazioni e funzioni |
| D11 | 0,47 | Numeri | D25_b | 0,60 | Relazioni e funzioni |
| D12 | 0,30 | Spazio e figure | D26 | 0,35 | Spazio e figure |
| D13_a | 0,40 | Dati e previsioni | D27_a | 0,30 | Dati e previsioni |
| D13_b | 0,46 | Dati e previsioni | D27_b | 0,48 | Dati e previsioni |
| D13_c | 0,53 | Dati e previsioni | D28 | 0,40 | Numeri |
| D14_a | 0,40 | Numeri | D29 | 0,23 | Spazio e figure |
| D14_b | 0,46 | Numeri | - | - | - |

Tavola II.20. Coefficienti di correlazione punto-biseriali delle opzioni corrette di risposte
(Italiano-III sec. di I gr.)

| TESTO LETTERARIO | | TESTO ESPOSITIVO-ARGUMENTATIVO | | GRAMMATICA | |
|------------------|------------------------------|--------------------------------|------------------------------|------------|------------------------------|
| Item | Correlazione punto biseriale | Item | Correlazione punto biseriale | Item | Correlazione punto biseriale |
| A1 | 0,32 | B1 | 0,27 | C1 | 0,35 |
| A2 | 0,32 | B2 | 0,47 | C2 | 0,36 |
| A3 | 0,24 | B3 | 0,24 | C3 | 0,42 |
| A4 | 0,41 | B4 | 0,41 | C4 | 0,42 |
| A5 | 0,35 | B5 | 0,37 | C5 | 0,23 |
| A6 | 0,31 | B6 | 0,39 | C6 | 0,42 |
| A7 | 0,40 | B7 | 0,36 | C7 | 0,35 |
| A8 | 0,46 | B8 | 0,22 | C8 | 0,38 |
| A9 | 0,39 | B9 | 0,34 | C9 | 0,53 |
| A10 | 0,30 | B10 | 0,50 | C10 | 0,44 |
| A11 | 0,21 | B11 | 0,45 | C11 | 0,44 |
| A12 | 0,34 | B12 | 0,48 | - | - |
| A13 | 0,41 | B13 | 0,42 | - | - |
| A14 | 0,31 | B14 | 0,34 | - | - |
| A15 | 0,29 | B15 | 0,42 | - | - |
| A16 | 0,29 | B16 | 0,47 | - | - |
| A17 | 0,29 | B17 | 0,38 | - | - |
| A18 | 0,29 | B18 | 0,40 | - | - |
| A19 | 0,30 | B19 | 0,39 | - | - |

Tavola II.21. Coefficienti di correlazione punto-biseriali delle opzioni corrette di risposte
(Matematica- III sec. di I gr.)

| Item | Correlazione punto biseriale | Ambito | Item | Correlazione punto biseriale | Ambito |
|-------|------------------------------|----------------------|-------|------------------------------|----------------------|
| D1_a | 0,33 | Dati e previsioni | D12 | 0,22 | Spazio e figure |
| D1_b | 0,35 | Dati e previsioni | D13 | 0,54 | Numeri |
| D2 | 0,34 | Numeri | D14 | 0,48 | Numeri |
| D3_a | 0,45 | Relazioni e funzioni | D15 | 0,46 | Dati e previsioni |
| D3_b | 0,37 | Relazioni e funzioni | D16 | 0,34 | Dati e previsioni |
| D3_c | 0,44 | Relazioni e funzioni | D17 | 0,43 | Relazioni e funzioni |
| D4 | 0,31 | Spazio e figure | D18 | 0,40 | Numeri |
| D5 | 0,10 | Relazioni e funzioni | D19 | 0,52 | Numeri |
| D6_a | 0,44 | Spazio e figure | D20 | 0,39 | Spazio e figure |
| D6_b | 0,45 | Spazio e figure | D21_a | 0,43 | Dati e previsioni |
| D7 | 0,36 | Numeri | D21_b | 0,46 | Dati e previsioni |
| D8_a | 0,43 | Numeri | D22 | 0,36 | Spazio e figure |
| D8_b | 0,46 | Numeri | D23 | 0,50 | Relazioni e funzioni |
| D9 | 0,36 | Spazio e figure | D24 | 0,42 | Relazioni e funzioni |
| D10_a | 0,44 | Relazioni e funzioni | D25 | 0,44 | Numeri |
| D10_b | 0,48 | Relazioni e funzioni | D26 | 0,34 | Spazio e figure |
| D11 | 0,16 | Dati e previsioni | - | - | - |

Tavola II.22. Coefficienti di correlazione punto-biseriali delle opzioni corrette di risposte
(Italiano-II sec. di II gr.)

| TESTO 1 e 2 | | TESTO 3 e 4 | | GRAMMATICA | |
|-------------|------------------------------|-------------|------------------------------|------------|------------------------------|
| Item | Correlazione punto biseriale | Item | Correlazione punto biseriale | Item | Correlazione punto biseriale |
| A1 | 0,38 | C1 | 0,18 | E1 | 0,43 |
| A2 | 0,45 | C2 | 0,33 | E2 | 0,57 |
| B1 | 0,25 | C3 | 0,27 | E3 | 0,41 |
| B2 | 0,22 | C4 | 0,53 | E4 | 0,41 |
| B3 | 0,23 | D1 | 0,31 | E5 | 0,39 |
| B4 | 0,28 | D2 | 0,36 | E6 | 0,62 |
| B5 | 0,36 | D3 | 0,38 | E7 | 0,26 |
| B6 | 0,20 | D4 | 0,42 | E8_a | 0,48 |
| B7 | 0,34 | D5_a | 0,52 | E8_b | 0,56 |
| B8 | 0,25 | D5_b | 0,49 | E9 | 0,36 |
| B9_a | 0,37 | D6 | 0,42 | E10_a | 0,43 |
| B9_b | 0,48 | D7 | 0,38 | E10_b | 0,45 |
| B10 | 0,41 | D8 | 0,35 | E10_c | 0,50 |
| B11 | 0,32 | D9_a | 0,28 | E10_d | 0,48 |
| B12 | 0,32 | D9_b | 0,55 | E11_a | 0,51 |
| B13 | 0,51 | D10 | 0,43 | E11_b | 0,49 |
| B14 | 0,55 | D11 | 0,37 | E11_c | 0,52 |
| B15 | 0,47 | D12 | 0,43 | E12_a | 0,45 |
| B16 | 0,27 | D13 | 0,26 | E12_b | 0,46 |
| B17 | 0,23 | D14 | 0,36 | - | - |
| - | - | D15_a1 | 0,43 | - | - |
| - | - | D15_a2 | 0,47 | - | - |
| - | - | D15_a3 | 0,47 | - | - |
| - | - | D15_b | 0,54 | - | - |

Tavola II.23. Coefficienti di correlazione punto-biseriali delle opzioni corrette di risposte
(Matematica- II sec. di II gr.)

| Item | Correlazione punto biseriale | Ambito | Item | Correlazione punto biseriale | Ambito |
|-------|------------------------------|----------------------|-------|------------------------------|----------------------|
| D1_a | 0,23 | Dati e Previsioni | D13_a | 0,34 | Relazioni e Funzioni |
| D1_b | 0,41 | Dati e Previsioni | D13_b | 0,38 | Relazioni e Funzioni |
| D1_c | 0,26 | Dati e Previsioni | D13_c | 0,36 | Relazioni e Funzioni |
| D2 | 0,44 | Dati e Previsioni | D14 | 0,26 | Relazioni e Funzioni |
| D3 | 0,36 | Spazio e Figure | D15 | 0,34 | Numeri |
| D4 | 0,57 | Numeri | D16 | 0,52 | Numeri |
| D5 | 0,32 | Numeri | D17 | 0,42 | Spazio e Figure |
| D6 | 0,49 | Dati e Previsioni | D18 | 0,51 | Spazio e Figure |
| D7_a | 0,30 | Relazioni e Funzioni | D19 | 0,39 | Dati e Previsioni |
| D7_b | 0,08 | Relazioni e Funzioni | D20_a | 0,29 | Numeri |
| D8_a | 0,63 | Spazio e Figure | D20_b | 0,43 | Numeri |
| D8_b | 0,49 | Spazio e Figure | D20_c | 0,51 | Numeri |
| D9_a | 0,38 | Spazio e Figure | D20_d | 0,51 | Numeri |
| D9_b | 0,60 | Spazio e Figure | D21 | 0,20 | Relazioni e Funzioni |
| D10 | 0,43 | Numeri | D22 | 0,44 | Numeri |
| D11_a | 0,44 | Relazioni e Funzioni | D23_a | 0,42 | Numeri |
| D11_b | 0,60 | Relazioni e Funzioni | D23_b | 0,56 | Numeri |
| D11_c | 0,33 | Relazioni e Funzioni | D24 | 0,39 | Relazioni e Funzioni |
| D12_a | 0,37 | Dati e Previsioni | D25 | 0,38 | Relazioni e Funzioni |
| D12_b | 0,39 | Dati e Previsioni | D26 | 0,48 | Relazioni e Funzioni |
| D12_c | 0,27 | Dati e Previsioni | D27 | 0,54 | Relazioni e Funzioni |
| D12_d | 0,40 | Dati e Previsioni | D28 | 0,44 | Numeri |
| D12_e | 0,39 | Dati e Previsioni | D29 | 0,45 | Numeri |
| D12_f | 0,35 | Dati e Previsioni | D30 | 0,46 | Spazio e Figure |

È importante notare che tutti i quesiti della prova mostrano le stesse caratteristiche individuate nella tavola II.13, confermando così la buona qualità della prova come strumento di misurazione e, soprattutto, corroborando le ipotesi effettuate in sede di pre-test. In termini ancora più espliciti, tutti i quesiti delle prove SNV mostrano una correlazione punto-biseriali positiva con l'esito complessivo della prova medesima. Come già evidenziato nel commento della tavola II.13 questo è un aspetto molto importante sotto il profilo della qualità in senso statistico-misuratorio della prova e solo in pochissimi casi il coefficiente di correlazione punto-biseriali assume valori inferiori a 0,20. Infine, i dati delle tavole II.14- II.23 permettono di eliminare, su una base empirica e razionalmente fondata, la preoccupazione che siano presenti ambiguità o passaggi poco chiari nelle domande che costituiscono le prove SNV.

Riferimenti bibliografici

Mullis, I.V.S., Martin, M.O., Ruddock, G.J., O'Sullivan, C.Y., Arora, A., & Erberber, E. (2005). *TIMSS 2007 assessment frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Barbaranelli, C. & Natali, E. (2005). *I test psicologici: teorie e modelli psicometrici*. Roma: Carrocci Editore.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.

Falocci N., Gnaldi, M., Matteucci, M. & Mignani, S. (2010). *La validazione statistica di test standardizzati di profitto: principali aspetti di metodo e due casi di studio sulla valutazione degli apprendimenti nella scuola primaria*. Collana Working Paper INVALSI, 9. http://www.invalsi.it/download/wp/wp09_Falocci.pdf

INVALSI (2011a). *Quadro di riferimento di Italiano*. (<http://www.invalsi.it/snv1011/>)

INVALSI (2011b). *Quadro di riferimento di Matematica*. (<http://www.invalsi.it/snv1011/>)

OCSE (2007). *Valutare le competenze in scienze, lettura e matematica*. Roma: Armando Editore.