

PP3: risultati della rilevazione degli apprendimenti degli studenti

1. Premessa

Scopo della rilevazione degli apprendimenti è quello di fornire alle scuole un riferimento per valutare il grado di preparazione che ognuna di esse fornisce ai propri allievi.

Anche il Progetto Pilota per la Valutazione dell'Istruzione relativo all'anno scolastico 2003-2004 (PP3) ha conservato la volontarietà della partecipazione da parte delle istituzioni scolastiche. Ne è conseguita (come per il precedente) l'esigenza di affiancare a quella volontaria una rilevazione su un campione probabilistico nazionale, in modo da offrire alle scuole un riferimento rispetto al quale misurare i propri esiti ed individuare i propri punti di forza e di debolezza. Per il campione probabilistico si è utilizzato un piano di campionamento a due stadi con stratificazione delle scuole¹.

Ai fini dell'indagine le scuole sono state pertanto articolate in due gruppi: il campione probabilistico (589 scuole, 589 classi e 21.138 studenti) e il gruppo di scuole che ha aderito volontariamente (8.461 scuole, 68.031 classi e 1.388.390 studenti) con somministrazione di prove cartacee o informatiche. Ne discende che l'indagine ha due valenze: una formativa (indagine volontaria) per le scuole in quanto possono verificare in piena autonomia il percorso didattico scelto e una di sistema (indagine campionaria) in quanto dà indicazioni sui punti di forza e di debolezza del sistema istruzione.

Su indicazione del Ministro:

- gli ambiti disciplinari sono stati: l'apprendimento della lingua italiana, della matematica e delle scienze, perché queste discipline di base sono presenti in tutti i livelli scolastici ed in tutti gli indirizzi di studio ed implicano alcune conoscenze ed abilità fondamentali che sono tra quelle indicate come competenze chiave a livello internazionale;
- i livelli scolastici sono stati: le classi II e IV della scuola primaria, la classe I della secondaria di I grado e le classi I e III della secondaria di II grado.

2. Partecipazione

La partecipazione al PP3 delle scuole (volontarie e campione) è riportata nella tabella che segue per modalità di somministrazione: cartacea (in tutte le classi prove somministrate su supporto cartaceo, 94,7%), informatica (in tutte le classi prove somministrate su supporto informatico, 1,7%), mista (in alcune classi prove cartacee e nelle rimanenti informatiche, 3,6%).

Tipo di somministrazione	Primaria	Secondaria I grado	Secondaria II grado	Totale Scuole	Totale Istituzioni scolastiche
Cartacea	3668	2965	1934	8567	6468
Informatica	22	39	91	152	137
Mista	88	83	160	331	282
Totale	3778	3087	2185	9050	6887

Le scuole statali rappresentano 89,8% del totale e le paritarie il 10,2%.

Nella tabella seguente è riportata la partecipazione in termini di classi, studenti ed insegnanti.

¹ Per scuola si intende l'istituzione scolastica contata per ogni ordine con cui partecipa.

Livello	Classi	Studenti	Insegnanti	Scuole
II primaria	15.335	282.517	49.783	3.736
IV primaria	15.368	282.119	46.104	3.747
I secondaria I grado	14.888	321.963	47.751	3.087
I secondaria II grado	13.173	312.357	41.704	2.173
III secondaria II grado	11.655	248.955	34.966	2.176
Totale	70.419	1.447.911	220.308	9.050*

*Il totale delle scuole non è la somma delle scuole aderenti per livello.

Hanno partecipato alla somministrazione cartacea 68.620 classi (97,4%) e alla somministrazione informatica 1799 (2,6%); in termini di studenti il 97,3% (1.409.528 studenti) ha utilizzato il mezzo cartaceo e il 2,6% (38.383 studenti) la modalità informatica.

3. Prove

Per una descrizione delle prove puntuale ed approfondita si veda il Rapporto finale sui Progetti Pilota pubblicato nel sito dell'INValSI². In questo paragrafo si riportano solo alcune informazioni utili alla lettura dei dati.

I Progetti Pilota hanno utilizzato per la raccolta dati, nei tre anni di sperimentazione, prove oggettive con quesiti a scelta multipla per le aree disciplinari oggetto d'indagine. Le prove sono state costruite da gruppi di esperti delle singole discipline provenienti dal mondo accademico e da quello della scuola, coordinati dal GdL sulla valutazione e con la collaborazione dell'INValSI per gli aspetti docimologici. Gli esperti hanno individuato nuclei disciplinari essenziali e, grazie alla presenza di insegnanti in servizio potenziata nel PP3, anche dei "programmi" effettivamente e generalmente sviluppati nelle classi.

I criteri che hanno guidato la costruzione delle prove del PP3 si possono così riassumere:

- uniformità della struttura: tutte le prove sono state costituite da quesiti a scelta multipla con quattro risposte (tre per la II elementare) di cui una sola corretta; questa struttura è stata scelta per consentire la lettura ottica delle prove per l'elaborazione dei dati al fine di restituire i risultati alle scuole in tempi brevi (entro il 31/07/2004);
- durata di ogni singola prova non superiore ai 60 minuti (30 per la II classe della scuola elementare; le prove sono state leggermente sovradimensionate rispetto al tempo);
- costruzione degli item rispetto a un numero limitato di conoscenze ed abilità ritenute irrinunciabili, acquisite nel biennio precedente, tenendo conto sia delle indicazioni nazionali (ancorché sperimentali) per la scuola primaria e per la scuola secondaria di primo grado sia dei programmi in vigore solitamente svolti; alcune conoscenze ed abilità sono state provate in modo ricorsivo e gradualmente più complesso; i quesiti sono stati strutturati in modo da sollecitare una riflessione piuttosto che risposte meccaniche;
- predisposizione di una prova per disciplina e per livello; si è scelto di non differenziare le prove disciplinari per la classe terza superiore perché le prove tendono a verificare conoscenze ed abilità ritenute essenziali, qualunque sia il tipo di Istituto e che devono essere acquisite all'inizio del nuovo biennio;
- taratura dei quesiti nel pretesting (prova su un numero limitato di classi che precede la somministrazione vera e propria).

² <http://www.invalsi.it>

4. Risultati

Si presentano i principali risultati per livello. Si utilizzano i punteggi normalizzati a 100 (il punteggio coincide con la % media di risposte corrette, 100 punteggio massimo e 0 punteggio minimo).

Sebbene i risultati del PP3 presentino, come si vedrà nel seguito, alcune analogie con i risultati del PP2, non è possibile fare comparazioni diacroniche in quanto le prove non sono costituite da item “equivalenti”. Infatti l’Italia non ha ancora sviluppato, a differenza di altri paesi che hanno una storia nel campo della valutazione, banche di item comparabili nel tempo (“equivalenti”): i Progetti Pilota hanno iniziato a lavorare in questo senso. È stato costituito un gruppo di esperti disciplinari che nell’arco dei due anni ha prodotto un considerevole numero di item (circa 250). Tutti questi item sono stati provati e i dati raccolti consentiranno nei prossimi anni di selezionare item equivalenti che ci permetteranno confronti diacronici anche nel campo degli apprendimenti.

Emerge, comunque, dai risultati e dalle osservazioni dei docenti un’attenzione particolare degli insegnanti di matematica per la geometria che ha determinato un riequilibrio delle prestazioni nelle aree di contenuti della matematica.

Per tutti i livelli, i risultati medi sono riassunti in una tabella non per comparazioni fra discipline, prive di significato, ma per comodità di lettura.

II Classe Primaria.

Area geografica³	Italiano (12 quesiti – 30 min.)	Matematica (16 quesiti – 30 min.)	Scienze (7 quesiti – 25 min.)
Nord Ovest	70	74	78
Nord Est	72	77	78
Centro	73	77	81
Sud	74	81	86
Sud e Isole	78	83	84
Italia	73	79	82

Considerato il numero esiguo di quesiti per le prove di italiano e scienze, non si sono elaborati i dati per categorie di conoscenza. Solo per matematica i quesiti sono stati suddivisi in *numero* (lettura e scrittura di numeri naturali, esecuzione d’addizioni e sottrazioni, ecc.) e *geometria* (identificare e denominare correttamente forme geometriche, attuazione percorsi e riconoscimento caratteristiche). La prestazione risulta sia complessivamente che per area geografica, per sesso e per regolarità degli studi, migliore in geometria (Italia = 81) rispetto a numero (Italia = 77).

I risultati per regolarità degli studi sono riportati di seguito.

³ Nord ovest: Val D’Aosta; Piemonte, Liguria, Lombardia; Nord est: Trentino Alto Adige, Veneto, Friuli Venezia Giulia, Emilia Romagna; Centro: Toscana, Umbria, Marche, Lazio; Sud: Abruzzo, Molise, Campania, Puglia; Sud e isole: Basilicata, Calabria, Sicilia, Sardegna.

Regolarità negli studi⁴	Italiano (12 quesiti – 30 min.)	Matematica (16 quesiti – 30 min.)	Scienze (7 quesiti – 25 min.)	Popolazione stimata %
>1 anno di ritardo	63	82	76	0*
1 anno di ritardo	67	73	82	1
Regolare	73	79	82	95
Anticipo	74	80	82	4
Italia	73	79	82	100

*Popolazione stimata = 0,2.%

Dai dati emerge quanto segue:

- gli alunni della II classe della scuola primaria sono stati in grado di rispondere alle prove; le prove probabilmente non hanno tenuto in debito conto che la somministrazione è avvenuta a marzo e cinque mesi incidono molto sui livelli di maturazione degli alunni;
- non vi sono differenze significative nei risultati fra maschi (52% della popolazione) e femmine (48% della popolazione);
- gli alunni in anticipo (4% della popolazione) e in ritardo (1% della popolazione) non determinano il punteggio medio (contano molto poco); il punteggio medio è determinato dagli studenti regolari (95%);
- gli esiti non sono uniformi rispetto al territorio, ma evidenziano differenze che dovranno essere confermate anche nelle prossime rilevazioni e studiate anche con altri strumenti; dai risultati emerge una differenza statisticamente significativa fra nord ovest e sud isole per italiano, fra centro nord e sud in matematica e scienze;
- in matematica la prestazione risulta migliore (con differenza statisticamente significativa) in geometria che in numero.

IV Classe Primaria.

Area geografica	Italiano (28 quesiti – 45 min.)	Matematica (28 quesiti – 45 min.)	Scienze (25 quesiti – 45 min.)
Nord Ovest	72	70	73
Nord Est	72	70	74
Centro	71	71	74
Sud	75	74	76
Sud e Isole	77	76	80
Italia	73	72	75

Le differenze che compaiono nella tabella non sono statisticamente significative.

I risultati per regolarità degli studi sono riportati di seguito.

⁴ Anticipo: tutti gli studenti con età inferiore a quella prevista per l'anno di corso frequentata; regolare: tutti gli studenti con età corrispondente all'anno di corso frequentato.

Regolarità negli studi	Italiano (28 quesiti – 45 min.)	Matematica (28 quesiti – 45 min.)	Scienze (25 quesiti – 45 min.)	Popolazione stimata %
>1 anno di ritardo	78	73	79	0*
1 anno di ritardo	65	65	66	1
Regolare	73	72	75	96
Anticipo	83	78	81	3
Italia	73	72	75	100

*Popolazione stimata = 0,3.

Dai dati emerge quanto segue:

- anche per questo livello le prove sono probabilmente sottodimensionate rispetto alle conoscenze ed abilità degli alunni: i risultati ci suggeriscono che si può osare di più; questa sottostima è dovuta all'impossibilità di eseguire il pretesting negli stessi livelli e nello stesso periodo dell'indagine: la rilevazione sugli apprendimenti ha scadenza annuale; nel PP3 il pretesting è stato condotto a fine ottobre su un campione di giudizio di V classi e sono stati scelti in prevalenza quesiti con indice di difficoltà medio facile;
- non vi sono differenze significative fra maschi (51% della popolazione) e femmine (49% della popolazione):
- in italiano, non ci sono differenze significative nelle prestazioni fra il testo espositivo e narrativo (il testo espositivo precedeva il narrativo); le prestazioni poi risultano statisticamente più basse nelle conoscenze morfosintattiche che nella comprensione degli aspetti pragmatici e semantici del testo; per quanto riguarda le conoscenze lessicali si nota nel PP3 una certa povertà, ma questa categoria è rappresentata solo da 3 quesiti e quindi poco significativa;
- in matematica, non vi sono differenze significative nelle prestazioni fra le conoscenze in geometria e in numero; qualche differenza significativa rimane nella categoria misure fra Nord ovest e il Sud e Isole;
- in scienze, si nota una carenza, peraltro già messa in evidenza dai risultati di analoghe indagini nazionali ed internazionali, nell'area tematica "uomo/ambiente", mentre le prestazioni migliori si hanno nell'area tematica "metodo sperimentale";
- le prestazioni per area geografica sono uniformi in tutte e tre le discipline; gli errori standard sono tali da non rendere significative le differenze.

I Classe Secondaria I Grado.

Area geografica	Italiano (30 quesiti – 45 min.)	Matematica (28 quesiti – 60 min.)	Scienze (25 quesiti – 40 min.)
Nord Ovest	55	59	66
Nord Est	53	58	66
Centro	54	57	64
Sud	50	54	62
Sud e Isole	52	54	64
Italia	53	56	64

L'unica differenza statisticamente significativa per le prestazioni a livello territoriale è in Italiano fra Nord ovest e Sud (differenza = $4,9 \pm 1,5$). Tale differenza dovrà essere confermata dai dati della prossima rilevazione.

I risultati per regolarità degli studi sono riportati in tabella. Il ritardo è salito dall'1% della scuola primaria al 6%, mentre l'anticipo si mantiene costante. Gli studenti in anticipo non

costituiscono un gruppo le cui prestazioni sono statisticamente differenti dagli studenti regolari.

Regolarità negli studi	Italiano (30 quesiti – 45 min.)	Matematica (28 quesiti – 60 min.)	Scienze (25 quesiti – 40 min.)	Popolazione stimata %
>1 anno di ritardo	39	41	46	1
1 anno di ritardo	41	40	49	5
Regolare	53	57	65	90
Anticipo	57	60	68	4
Italia	53	56	64	100

Dai dati emerge quanto segue:

- le prove risultano più equilibrate di quelle della scuola primaria: i risultati rispecchiano il carattere di transizione del livello scolastico;
- vi è una differenza significativa fra femmine (47% della popolazione) e maschi (53% della popolazione) nelle prestazioni in italiano (differenza = $4,2 \pm 1,0$), riconfermando quando era già emerso lo scorso anno;
- in italiano, vi sono differenze significative nelle prestazioni fra il testo espositivo e il testo narrativo; le prestazioni risultano migliori nel testo espositivo (differenza = $12,1 \pm 1,0$), che è posto al primo posto; è da notare che molti somministratori hanno notato che il testo narrativo non era giudicato interessante dagli studenti; anche per questo livello le prestazioni migliori sono nella comprensione testuale, le peggiori nelle conoscenze lessicali e nelle conoscenze morfosintattiche e semantiche;
- in matematica, non vi sono differenze significative nelle prestazioni fra le conoscenze in geometria e in numero; qualche differenza significativa rimane in geometria fra Nord ovest e il Sud e Isole;
- in scienze, come per la classe IV della primaria, le migliori prestazioni si hanno nell'area tematica "elementi di metodo sperimentale" mentre le peggiori nell'area tematica "trasformazioni", non presente in IV primaria;
- le prestazioni per area geografica è abbastanza uniforme: la differenza nelle prestazioni fra nord ovest e sud e isole va confermata dalla prossima rilevazione.

I Classe Secondaria II Grado.

Tipologia d'istruzione⁵	Italiano (30 quesiti – 45 min.)	Matematica (28 quesiti – 60 min.)	Scienze (25 quesiti – 40 min.)
Istruzione Classica	59	65	59
Istruzione Professionale	37	39	38
Istruzione Artistica	47	47	52
Istruzione Tecnica	46	54	53
Istituti Superiori	49	54	52
Italia	49	55	52

Si riconfermano, anche se le misure sono eseguite con strumenti diversi, i dati dello scorso anno per tipologia d'istruzione; per disciplina si nota che le prestazioni in scienze sono più

⁵ Classica = tutti i tipi di liceo; Professionale = istituti professionale; Artistica = liceo artistico e istituti d'arte; Tecnico = istituti tecnici; Istituti superiori = istituti di istruzione secondaria di II grado di diverso ordine e tipo.

vicine alle prestazioni delle altre discipline rispetto alla classe I della scuola secondaria di I grado, dipendendo dalla diversa filosofia con cui sono state costruite le prove di scienze: infatti, le prove di scienze per la secondaria superiore di II grado sono caratterizzate dalla presenza di quesiti per ambiti disciplinari (Fisica, Chimica, ecc.); sono meno significativi rispetto al I ciclo i dati aggregati per Italia, ad esempio la regolarità negli studi per disciplina.

Regolarità negli studi	Italiano (30 quesiti – 45 min.)	Matematica (28 quesiti – 60 min.)	Scienze (25 quesiti – 40 min.)	Popolazione stimata %
>1 anno di ritardo	41	45	47	4
1 anno di ritardo	41	44	43	10
Regolare	51	56	54	82
Anticipo	56	60	54	4
Italia	49	55	52	100

Dai dati emerge quanto segue:

- l’istruzione classica ha le prestazioni migliori (tutte le differenze sono statisticamente significative) nelle discipline oggetto d’indagine; ha però la più bassa presenza di ritardi (6% della sua popolazione) e la più alta presenza di anticipi (9% della sua popolazione). È da notare che il 30% degli studenti frequenta un liceo. Di contro l’istruzione professionale, che ha i risultati più bassi, ha il 30% dei suoi studenti in ritardo e solo l’1% in anticipo (solo il 69% è regolare);
- vi è una differenza significativa statisticamente fra femmine (51% della popolazione) e maschi (49% della popolazione) nelle prestazioni in italiano e scienze (le femmine hanno prestazioni migliori);
- in italiano, vi sono differenze significative nelle prestazioni fra il testo argomentativo e il testo letterario. Le prestazioni risultano migliori per il testo letterario (differenza = $3,0 \pm 0,8$) in tutti i tipi di scuole ad esclusione dei licei dove non vi sono differenze significative; le prestazioni migliori si hanno nelle conoscenze lessicali e morfosintattiche a differenza del I ciclo;
- in matematica le prestazioni migliori in relazioni e funzioni e in dati e previsioni in tutti i tipi d’istruzione mentre non risultano particolarmente brillanti neppure nell’istruzione classica le prestazioni in numeri e geometria (molto simili);
- in scienze, i migliori risultati si hanno per l’area tematica “elementi di metodo sperimentale” (solo 4 quesiti) mentre i risultati meno buoni in Scienze della Terra (4 quesiti). L’andamento è molto simile nei vari tipi d’istruzione anche se i risultati dei licei sono i più alti.

III Classe Secondaria II Grado.

Tipologia d'istruzione⁶	Italiano (38 quesiti – 60 min.)	Matematica (30 quesiti – 60 min.)	Scienze (25 quesiti – 40 min.)
Istruzione Classica	67	60	66
Istruzione Professionale	45	37	44
Istruzione Artistica	50	37	52
Istruzione Tecnica	53	47	56
Istituti Superiori	54	47	56
Italia	56	49	57

Dai dati emerge quanto segue:

- l'istruzione classica ha le prestazioni migliori (tutte le differenze sono statisticamente significative) nelle discipline oggetto d'indagine; ha però la più bassa presenza di ritardi (9% della sua popolazione) e la più alta presenza di anticipi (10% della sua popolazione); di contro l'istruzione professionale, che ha i risultati più bassi, ha il 39% dei suoi studenti in ritardo e solo il 2% in anticipo (solo il 59% è regolare);
- vi è una differenza significativa statisticamente fra femmine (48% della popolazione) e maschi (52% della popolazione) nelle prestazioni in italiano; nei licei in matematica le prestazioni dei maschi sono migliori (differenze statisticamente significative);
- in italiano le prestazioni più basse si sono avute nel testo poetico e nelle conoscenze morfosintattiche;
- in matematica, le prestazioni migliori in numero e dati e previsioni;
- in scienze, le prestazioni migliori in biologia (6 quesiti) e le peggiori in elementi di metodo sperimentale.

⁶ Classica = tutti i tipi di liceo; Professionale = istituti professionale; Artistica = liceo artistico e istituti d'arte; Tecnico = istituti tecnici; Istituti superiori = istituti di istruzione secondaria di II grado di diverso ordine e tipo.