

Un focus sui dati Italiani dell'indagine TIMSS & PIRLS 2011

Leonardo Grilli[^], Fulvia Pennoni^{*}, Carla Rampichini[^], Isabella Romeo^{*}

[^]Dipartimento di Statistica, Informatica, Applicazioni 'G. Parenti', Università degli Studi di Firenze
grilli@disia.unifi.it, rampichini@disia.unifi.it

^{*}Dipartimento di Statistica e Metodi Quantitativi, Università degli Studi di Milano-Bicocca, Milano
fulvia.pennoni@unimib.it, isabella.romeo@unimib.it

1. Introduzione

Nelle ultime due decadi, tra le varie indagini internazionali mirate ad acquisire una maggiore comprensione dei livelli degli apprendimenti educativi degli studenti, due delle più diffuse su larga scala sono le indagini *Trends in International Mathematics and Science Study* (TIMSS) e *Progress in International Reading Literacy Study* (PIRLS), realizzate dall'ente *International Association for the Evaluation of the Educational Achievement* (IEA). I dati ottenuti attraverso queste rilevazioni rientrano tra quelli che nella terminologia anglosassone sono definiti *International Large-Scale Assessment Data*.

TIMSS 2011 è l'indagine volta a rilevare le competenze degli studenti in matematica e scienze del quarto anno della scuola primaria e di terza media. Questa indagine viene effettuata ogni quattro anni a partire dal 1995. PIRLS 2011 è l'indagine volta a rilevare le competenze della lingua madre degli studenti del quarto anno di scuola primaria. A partire dal 2001 viene effettuata ogni 5 anni. I livelli di apprendimento stimati a livello paese e area geografica si basano su metodi statistici che consentono di ottenere delle stime corrette e dei valori consistenti per la variabilità delle stime stesse (Wu, 2005).

Nel 2011, per la prima volta, queste due indagini sono state svolte sul medesimo campione di studenti, permettendo di analizzare in modo congiunto i risultati in italiano, matematica e scienze degli studenti di quarta. Oltre ai test volti a valutare le competenze degli studenti, l'indagine prevede la somministrazione di un questionario agli stessi, ai genitori, agli insegnanti ed ai dirigenti scolastici della scuola. La base di dati

combinata (Foy, 2013) fornisce informazioni su studenti di 34 paesi e 3 entità amministrative e contiene informazioni relative a 185.475 studenti, 171.098 genitori, 14.258 insegnanti e 6.469 dirigenti scolastici. Dal 2012 questi dati sono resi disponibili per quella che in gergo tecnico viene definita *secondary analysis* (Rutkowski, Gonzalez, Joncas, von Davier, 2010).

I dati riguardanti il campione di scuole, classi e studenti relativi all'Italia comprendono: 202 scuole, 239 classi e 4.125 studenti. Il campionamento a grappoli a due stadi è stato effettuato, come per gli altri paesi partecipanti all'indagine, considerando le scuole unità di primo stadio e le classi unità di secondo stadio. Tutti gli studenti delle classi selezionate sono inclusi nel campione. Si precisa che le scuole sono stratificate in base all'area geografica ed in base ad altre caratteristiche rilevanti e selezionate in relazione al numero di studenti iscritti (Martin & Mullis, 2012).

Per quanto riguarda gli insegnanti, il numero totale è pari a 483. Nel sistema educativo attuale la scelta e l'attribuzione delle discipline insegnate dipendono dalla progettazione del collegio docenti, ogni classe può avere da uno a tre insegnanti che insegna una combinazione diversa di materie. Nel campione osservato ci sono 2 insegnanti nella maggior parte delle classi. I dati sono oggetto di una ricerca in corso nell'ambito del progetto Futuro in Ricerca finanziato dal Governo Italiano dal titolo: "*Modelli mistura e a variabili latenti per l'inferenza causale e l'analisi di dati socio-economici*" (Grilli, Pennoni, Rampichini, Romeo, 2014).

2. Descrizione del campione italiano

Le caratteristiche dello studente considerate nell'analisi sono: il genere, la lingua parlata a casa, la frequentazione della scuola dell'infanzia, le risorse culturali possedute nell'abitazione¹ e le conoscenze preliminari del bambino in italiano e matematica. Il 69% degli studenti intervistati ha frequentato per almeno tre anni la scuola dell'infanzia e il 22% non parla Italiano a casa. Tra le caratteristiche dell'insegnante sono state considerate: il genere, gli anni di insegnamento ed il possesso della laurea. La maggior parte degli insegnanti è donna con un'età compresa tra i 40 e i 59 anni. Il 35% degli insegnanti possiede un diploma di laurea. Mediamente, ogni insegnante ha 23 anni di esperienza nell'insegnamento. A livello di classe sono state aggregate le variabili sopra elencate per lo studente e considerate anche altre variabili dichiarate dal dirigente scolastico della scuola come ad esempio la sicurezza della scuola e l'adeguatezza delle risorse. Infine, come proxy dello stato socio-economico della zona dove opera la scuola è stato considerato il reddito lordo pro capite a livello provinciale relativo al 2010. Questa variabile non è presente nella banca dati ed è stata ricavata da un'altra banca dati fornita dall'Istituto Tagliacarne di Unioncamere.

Tra le scuole campionate, il 34% si trova in una grande città e il 53% ha una distribuzione dell'orario scolastico su 5 giorni settimanali. Nel 65% delle scuole selezionate gli studenti italiani sono più del 90%, mentre le scuole con una più bassa percentuale di studenti svantaggiati dal punto di vista socio-economico rappresentano il 39% del campione. Occorre notare che il valore medio dell'Italia di alcuni indicatori riferiti al background dello studente e alle risorse della scuola è inferiore al valore medio dei paesi OCSE.

3. Risultati preliminari

L'analisi è stata condotta solo sugli studenti che non presentano valori mancanti per nessuna delle variabili considerate: 3.841 studenti e 237 classi. La Tabella 1 riporta i punteggi medi per le tre materie in base ad alcune delle variabili

qualitative presenti nella base di dati, mentre la Tabella 2 illustra il punteggio medio per il primo e ultimo quartile della distribuzione di alcune delle variabili continue selezionate dalla base di dati.

Osserviamo che, come noto in letteratura, le bambine hanno un punteggio medio peggiore in matematica e scienze rispetto ai coetanei maschi. I bambini che hanno frequentato la scuola dell'infanzia risultano avere punteggi più alti nelle tre materie rispetto a quelli che non l'hanno frequentata. I bambini italiani ottengono risultati migliori, rispetto ai loro pari la cui lingua parlata a casa non è l'italiano. Quando l'insegnante di italiano è di genere femminile si nota un innalzamento dei punteggi per gli studenti in questa materia. I punteggi sono più alti in lettura se l'insegnante ha meno di 29 anni e in scienze se l'insegnante ha meno di 45 anni. Si osservano punteggi più alti per tutte le materie, ed in particolare per matematica e scienze, per gli studenti che frequentano scuole con una più bassa percentuale di studenti svantaggiati dal punto di vista socio-economico, scuole che si trovano in aree con più di 50.000 abitanti e scuole con una programmazione scolastica su sei giorni settimanali.

Considerando la Tabella 2 si evince che le variabili relative alle conoscenze preliminari di italiano e matematica e alle risorse culturali disponibili nell'abitazione sono associate positivamente ai risultati sulle tre materie a tutti i livelli considerati: studente, classe e scuola. Inoltre, a livello classe, altre variabili rilevanti sono la percentuale di studenti che ha frequentato la scuola dell'infanzia e la percentuale di studenti che non parla italiano a casa. A livello scuola, tutte le variabili considerate risultano essere importanti. Il valore aggiunto pro capite, misurato a livello provinciale, è associato positivamente ai punteggi nelle tre materie. Si osserva un'alta correlazione tra i punteggi ottenuti dagli studenti nelle diverse materie. Poiché il livello di apprendimento raggiunto dipende sia dalle caratteristiche dello studente che dall'organizzazione scolastica, dagli insegnanti e dalle interazioni con i compagni, ci si attende che i risultati di studenti appartenenti alla stessa classe siano più simili tra loro rispetto a quelli di studenti appartenenti a classi diverse. Per valutare questo aspetto, abbiamo scomposto la variabilità complessiva nelle due quote tra ed entro le classi. In base alla Tabella 3 osserviamo che la correlazione tra materie è più alta tra classi piuttosto che entro classi. Inoltre, italiano mostra

¹ Questo indicatore è stato ricavato considerando il numero di libri, i supporti disponibili per lo studio e il livello educativo e lavorativo dei genitori (Martin & Mullis, 2013).

la più bassa variabilità a livello classe, probabilmente perché rappresenta la materia più influenzata dalle caratteristiche degli studenti. Infine, matematica e scienze, mostrano una più bassa variabilità entro classi, probabilmente perché tali materie sono di solito insegnate dal medesimo insegnante.

Queste prime analisi descrittive evidenziano l'importanza del background dello studente e delle risorse della scuola nel determinare il livello di apprendimento in italiano, matematica e scienze. Tuttavia, l'analisi descrittiva fin qui condotta deve essere approfondita, al fine di considerare congiuntamente i fattori esplicativi del livello di apprendimento e la particolare natura dei dati esaminati. La ricerca pertanto è incentrata sulla specificazione e sulla stima di un modello statistico che tenga conto in maniera adeguata: della struttura gerarchica dei dati (studenti in classi), della natura multivariata della risposta (apprendimento dello studente in italiano, matematica e scienze) e della variabilità insita nel processo di raccolta dei dati e di misurazione dell'apprendimento.

Riferimenti bibliografici

- Foy, P. (2013). *TIMSS & PIRLS 2011 user guide for the fourth grade combined international database*, Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Grilli, L., Pennoni, F., Rampichini, C., Romeo, I. (2014). Multivariate multilevel modelling of student achievement data, *Proceedings SIS 2014 47th Scientific meeting of the Italian Statistical Society*, 11-13 Giugno 2014, Cagliari, 1-6.
- Martin, M.O. & Mullis, I.V.S (Eds.) (2013). *TIMSS & PIRLS 2011: Relationships Among Reading, Mathematics, and Science Achievement at the Fourth Grade Implications for Early Learning*, Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Martin, M.O. & Mullis, I.V.S. (Eds.) (2012). *Methods and procedures in TIMSS and PIRLS 2011*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Rutkowski, L., Gonzalez, E., Joncas, M. von Davier, M. (2010). International Large-Scale Assessment Data: Issues in Secondary Analysis and Reporting. *Educational Research*, 39, 142–151.
- Wu, M. (2005). The role of plausible values in large-scale surveys, *Studies in Educational Evaluation*, 31, 114-128.

Tabella 1 Punteggi medi ottenuti in Italiano, Matematica e Scienze e numero complessivo per studenti, insegnanti e scuole, in base alle variabili qualitative presenti nella base di dati (media internazionale pari a 500).

	Italiano	Matematica	Scienze	N	
	<i>Genere</i>				
	Maschio	522,20	507,54	522,87	2.040
	Femmina	528,48	497,06	515,23	2.085
	<i>Scuola dell'infanzia</i>				
Studente	No	503,71	483,53	496,73	1.265
	Si	534,96	510,52	528,87	2.860
	<i>Lingua parlata a casa</i>				
	Italiano	534,59	509,04	528,30	3.230
	Altre lingue	492,11	477,71	485,49	895
	<i>Genere</i>				
	Maschio	505,48	500,64	517,37	215
	Femmina	526,47	502,45	519,23	3.910
	<i>Età</i>				
Insegnante	<25	539,06	506,22	542,61	53
	25-29	533,86	493,38	514,97	66
	30-39	522,36	505,38	528,87	345
	40-49	524,39	501,71	513,09	1.615
	50-59	527,43	502,43	520,37	1.583
	>60	526,41	504,92	530,44	290
	<i>Livello educativo</i>				
	Non laureato	524,44	502,07	517,68	3.146
	Laureato	528,38	502,82	522,73	979
	<i>Percentuale di studenti italiani</i>				
	< 90%	526,66	501,97	518,43	1.455
	> 90%	524,13	501,33	518,42	2.586
	<i>Percentuale di studenti svantaggiati</i>				
	>10%	521,44	499,45	514,31	2.340
	<10%	529,84	504,45	523,64	1.590
	<i>Area in cui è collocata la scuola</i>				
Scuola	Piccola città	523,31	500,20	517,64	2.544
	Grande città	528,67	505,12	520,88	1.527
	<i>Reddito medio nell'area</i>				
	Alto	533,29	515,09	530,92	296
	Medio	528,57	504,47	522,38	3.107
	Basso	506,67	485,00	497,11	668
	<i>Numero di abitanti nell'area</i>				
	< 50.000	523,01	500,82	517,34	2.782
	> 50.000	529,64	504,39	521,75	1.242
	<i>Numero di giorni scolastici</i>				
	5 giorni	522,84	499,65	515,54	2.180
	6 giorni	528,04	504,62	522,55	1.885

Tabella 2 Punteggi medi ottenuti in Italiano, Matematica e Scienze per il primo ultimo ed quartile delle variabili continue presenti nella base di dati (media internazionale pari a 500).

	Italiano		Matematica		Scienze	
	Primo quartile	Ultimo quartile	Primo quartile	Ultimo quartile	Primo quartile	Ultimo quartile
<i>Variabili relativa alla Classe</i>						
Anni di insegnamento dell'insegnante	528,22	531,47	507,67	510,55	527,83	532,52
Conoscenze preliminari medie nella classe	526,48	519,64	502,77	500,15	521,94	512,44
Percentuale di femmine nella classe	515,11	523,20	492,86	495,43	507,58	514,03
Percentuale di studenti che hanno frequentato la scuola dell'infanzia	510,23	536,61	487,88	513,85	501,11	533,79
Percentuale di studenti che non parlano italiano a casa	544,96	506,48	520,84	483,74	541,18	496,80
Media di classe delle risorse culturali disponibili nell'abitazione	500,49	542,52	480,48	518,70	490,33	537,46
<i>Variabili relative alla Scuola</i>						
Adeguatezza delle risorse e dell'ambiente	519,62	534,22	492,58	513,47	510,70	529,32
Media di scuola delle risorse possedute a casa	502,69	540,72	482,45	517,96	492,65	536,78
Conoscenze preliminari medie nella scuola	528,66	518,43	503,77	497,93	523,22	510,24
Sicurezza e disciplina nella scuola	518,21	532,50	496,30	505,66	509,94	524,10
Misura del valore aggiunto pro capite	505,68	533,68	482,84	508,91	494,50	527,42

Tabella 3 Matrice di correlazione dei punteggi in Italiano, Matematica e Scienze.

	Matrice di correlazione tra alunni della stessa classe			Matrice di correlazione tra classi			Matrice di correlazione tra risposte (totale)			% VarCov a livello di classe		
	Ita.	Mat.	Sci.	Ita.	Mat.	Sci.	Ita.	Mat.	Sci.	Ita.	Mat.	Sci.
Ita.	1,00			1,00			1,00			17,27		
Mat.	0,71	1,00		0,92	1,00		0,75	1,00		25,78	25,47	
Sci.	0,81	0,74	1,00	0,97	0,98	1,00	0,84	0,80	1,00	24,70	31,74	26,71