

10. Effetti di variabili individuali e di variabili scolastiche sulla comprensione della lettura: analisi multilivello dei dati PISA 2009 dell'Alto Adige

10.1 Introduzione: il modello statistico

Nello studio di fenomeni complessi è importante cercare di stabilire delle relazioni tra variabili in un qualche modo legate all'oggetto della ricerca. Nelle indagini sui sistemi educativi, quest'ultimo aspetto assume un ruolo particolarmente rilevante ed è quindi cruciale ricorrere a modelli statistici che siano in grado di rappresentare adeguatamente i dati rilevati.

Negli ultimi anni, anche negli studi sulla scuola, hanno trovato una certa diffusione metodi di analisi di regressione a più livelli, in grado di dar conto della struttura gerarchica che tipicamente caratterizza i dati sulla scuola. Nel contesto educativo, infatti, la popolazione oggetto di indagine presenta una struttura "annidata" (*nested*), in cui sono individuabili distinti livelli di appartenenza, di ordine via via più elevato, per ciascuna unità di osservazione. Le unità di livello più basso, vale a dire gli studenti, non s'incontrano da soli, ma raggruppati in unità di ordine superiore, le classi, le quali a loro volta sono riunite in scuole, le scuole nelle province, ecc. Tranne casi particolari ed infrequenti¹, la scelta di un modello che non tenga in adeguato conto la struttura gerarchica insita nei dati e la correlazione tra le osservazioni appartenenti ad uno stesso gruppo ad ogni livello² determina una rilevante perdita di informazione, portando a stime distorte dei parametri oggetto d'interesse e in particolare dei loro errori standard (Bryk e Raudenbush, 2002). Inoltre, in molte ricerche condotte sulla scuola, il piano di campionamento riproduce anch'esso la struttura gerarchica della popolazione studiata e quindi a maggior ragione è opportuno ricorrere a metodi di analisi che siano in grado di render conto di tale organizzazione dei dati. L'indagine PISA, per tornare al caso che qui c'interessa, si basa anch'essa su un campionamento a due stadi (*two-stage sampling*), in cui, al primo stadio, all'interno di ciascun paese, vengono selezionate le scuole con probabilità proporzionale alla loro dimensione, mentre al secondo stadio vengono estratti 35 studenti, fra tutti gli alunni quindicenni iscritti, con probabilità uguale in ciascuna scuola campionata³.

Per comprendere la logica dei modelli multilivello, si farà riferimento ad un semplice esempio, considerando, in un primo momento, il consueto metodo di analisi di regressione ordinaria che non tiene conto della struttura gerarchica dei dati, per poi introdurre un modello lineare gerarchico (o *multilevel*) ed esaminarne i vantaggi sia sotto il profilo interpretativo sia sotto quello più propriamente statistico-inferenziale.

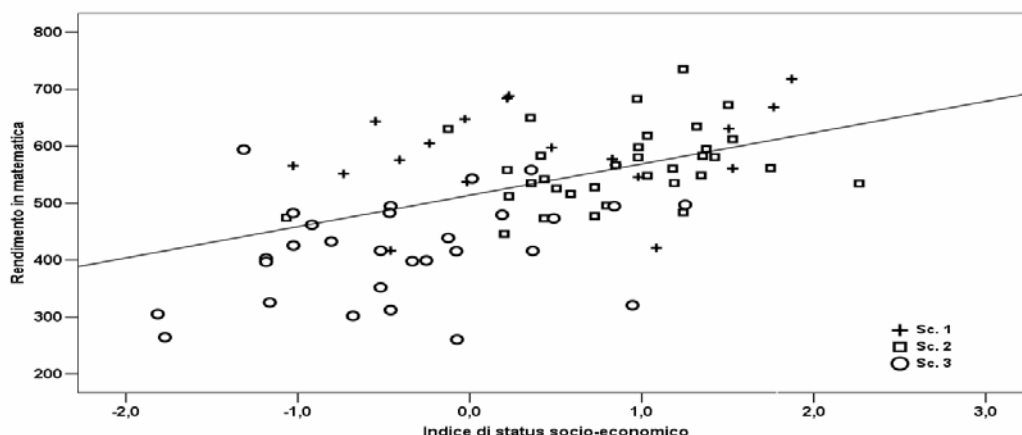
Si supponga di aver effettuato un'indagine sugli studenti di tre scuole e che per ciascun alunno siano state rilevate due variabili: lo status socio-economico, misurato mediante un indicatore sintetico i cui valori sono indicati sull'asse orizzontale del grafico in figura 1, ed il punteggio conseguito in una prova di matematica, i cui valori sono riportati sull'asse verticale.

¹ Se la relazione tra le variabili oggetto d'interesse è la stessa all'interno di ciascun gruppo o se i gruppi sono formati al loro interno da unità omogenee, allora la struttura gerarchica dei dati può essere ignorata.

² Una delle assunzioni su cui si basa la regressione ordinaria è invece l'indipendenza delle osservazioni. Nella realtà le unità appartenenti allo stesso gruppo sono verosimilmente più omogenee tra loro delle unità appartenenti a gruppi diversi e mostrano dunque un certo grado di correlazione.

³ Nel caso di scuole con un numero di quindicenni inferiore a 35 vengono selezionati tutti gli studenti (OCSE, 2007).

Figura 1. Relazione tra rendimento in matematica e status socio-economico



Come si può vedere, l'andamento complessivo dei punti nel grafico – ciascuno dei quali corrisponde ad uno studente – mostra un'associazione concorde e positiva tra il rendimento in matematica (Y) ed i valori dell'indice di status socio-economico (X). In termini formali, l'associazione tra le misure delle due variabili rilevate sugli studenti – senza tener conto per ora del loro raggruppamento nelle tre scuole – può essere rappresentata dalla retta della figura 1, detta “retta di regressione ordinaria” ed espressa dalla seguente equazione⁴:

$$Y_i = \alpha + \beta X_i + e_i \quad (1)$$

dove Y_i (con $i = 1, 2, \dots, n$) è il rendimento in matematica dello studente i -mo, il coefficiente α è il valore medio assunto da Y quando la variabile X – vale a dire lo status socio-economico dello studente – è pari a 0⁵, il coefficiente β , comunemente denominato “coefficiente di regressione lineare”, indica l'aumento medio di Y in seguito ad un incremento unitario di X e, infine, il termine e_i corrisponde all'errore della stima, cioè alla differenza tra il valore di Y effettivamente osservato sullo studente i -mo ed il valore stimato in base alla (1). Tale differenza rappresenta, quindi, quella parte di variabilità dei punteggi in matematica che non viene spiegata dalla condizione socio-economica dello studente.

Tuttavia, il modello (1) ora considerato non tiene alcun conto dei tre gruppi (le scuole) in cui gli alunni sono suddivisi. Specie nelle ricerche sui sistemi d'istruzione, l'appartenenza ad un gruppo (la classe, la scuola, ecc.) può avere invece un ruolo importante. Accade spesso, infatti, che gli appartenenti a uno stesso gruppo siano più simili fra loro dei membri di gruppi diversi, e ciò a causa di processi di selezione o autoselezione nella formazione dei gruppi o dell'esposizione a fattori ambientali comuni una volta costituiti. Risulta, dunque, cruciale per l'analisi dei dati la scelta di un modello statistico che sia in grado di tener conto di questa loro caratteristica⁶.

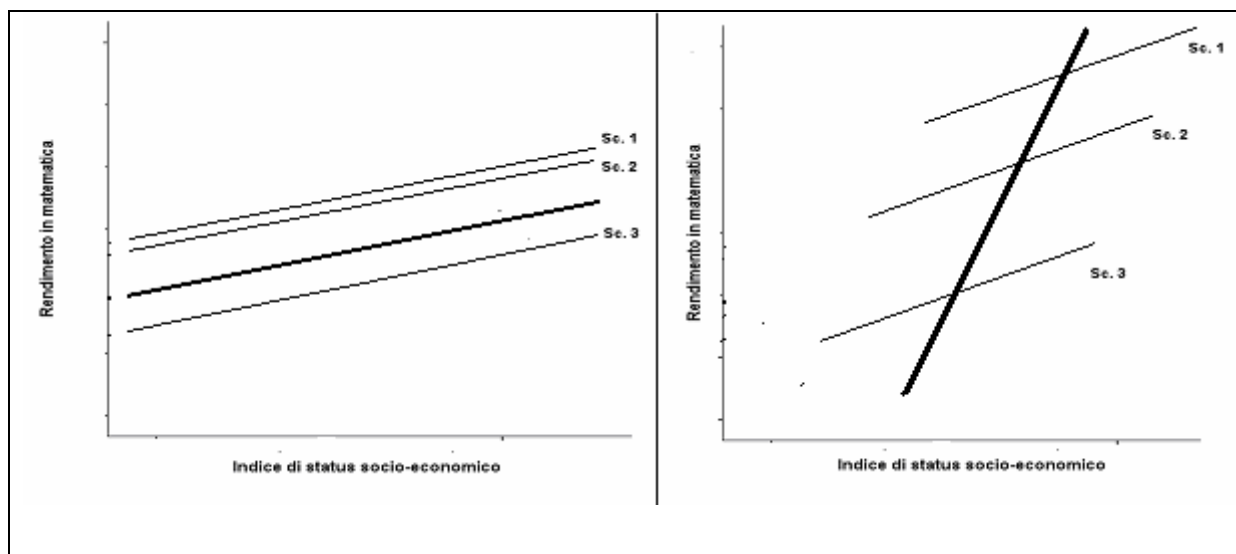
⁴ Il modello (1) rappresenta il caso più semplice con una sola variabile esplicativa X . Con ovvio significato dei simboli, esso può essere esteso al caso con k variabili esplicative: $Y_i = \alpha + \beta_1 X_{1i} + \dots + \beta_k X_{ki} + e_i$.

⁵ L'interpretazione dei risultati sovente è facilitata se la variabile X , anziché essere espressa nella sua metrica naturale, viene riscalata mediante la sottrazione del suo valore medio. In questo modo il valore di α può essere inteso come il valore medio assunto dalla Y quando la X ha un valore eguale alla sua media. Nell'esempio in esame, α indicherebbe il valore medio del rendimento in matematica per uno studente con un indice di status socio-economico pari alla media dell'indice calcolata sull'insieme di tutti gli n studenti considerati.

⁶ Nel caso della ricerca PISA i dati relativi ai singoli stati, o alle entità sub-nazionali, sono organizzati secondo una gerarchia a due livelli: studenti e scuole, non essendo previsto nel campionamento il livello della classe, e dunque ci limitiamo qui ad illustrare questo caso. I modelli multilivello possono tuttavia tener conto, se i dati lo consentono, anche di tre (o più) livelli: ad es., studente, classe e scuola, oppure: studente, scuola, paese (o regione).

Se si prende in considerazione la suddivisione degli studenti nelle tre scuole, la situazione rappresentata dalla figura 1 può infatti cambiare radicalmente ed assumere configurazioni molto diverse.

Figura 2. Relazione tra rendimento in matematica e status s.e. in generale e per scuola



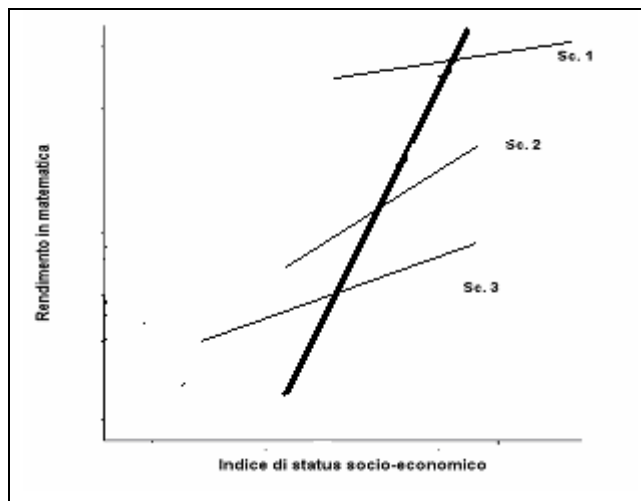
La figura 2 esemplifica due situazioni opposte che potrebbero emergere quando si tenga conto del raggruppamento degli studenti. Il grafico nella parte sinistra della figura 2 rappresenta una realtà in cui tutta la gamma di condizioni socio-economiche dell'intera popolazione di studenti è presente in ognuna delle scuole e l'effetto dello status sul rendimento in matematica è lo stesso in ciascuna di esse: le rette di regressione dei singoli istituti hanno infatti la medesima inclinazione una rispetto all'altra e rispetto alla retta di regressione generale (in grassetto), cioè la retta che si ottiene senza tener conto dei gruppi. In ogni scuola, inoltre, vi sono alunni con bassi, medi e alti punteggi in matematica. Solo le altezze dei gradienti⁷ variano, denotando una modesta oscillazione del rendimento medio delle scuole intorno alla media complessiva. Nel caso in esame, il ricorso ad un modello di regressione ordinaria non produrrebbe risultati molto differenti rispetto ad un modello multilivello.

Assai diversa è la situazione raffigurata nel grafico di destra. Qui, in primo luogo, si può notare come le scuole accolgano alunni con un ventaglio di condizioni socio-economiche che si diversifica da un istituto all'altro, globalmente più elevate nel caso della scuola 1, più disagiate nel caso della scuola 3 e intermedie nel caso della scuola 2. Le rette di regressione dei tre istituti sono ancora parallele fra loro ma mostrano un'inclinazione notevolmente minore di quella della retta di regressione generale. Ciò riflette la maggiore omogeneità dell'origine sociale degli studenti all'interno dei singoli istituti che ne rende meno influente il peso sul rendimento. Per conseguenza, data la forte associazione fra le due variabili, anche la gamma e il livello medio dei punteggi in matematica delle tre scuole risultano alquanto differenti fra loro e rispetto alla media generale. In questo secondo caso, i risultati che si otterrebbero da un modello di regressione ordinaria darebbero un'immagine distorta della situazione reale e solo un modello gerarchico è in grado di rappresentarla adeguatamente poiché capace di tenere adeguato conto della struttura dei dati.

La figura 3, infine, rappresenta un caso in cui la differenza tra le scuole è ancora più considerevole.

⁷ Gradiente sta qui, di fatto, per "retta di regressione". L'altezza del gradiente è l'ordinata (asse verticale) del punto, su ciascuna retta, in corrispondenza dell'ascissa (asse orizzontale) uguale al valore medio dell'indice di status socio-economico.

Figura 3. Relazione tra rendimento in matematica e status s.e. in generale e per scuola



In quest'ultimo esempio, non solo la posizione delle rette delle singole scuole rispetto all'asse X denota come esse reclutino alunni con condizioni socio-economiche differenti una dall'altra, ma in più l'effetto dell'origine sociale sul rendimento in matematica appare anch'esso differente da un istituto all'altro. In questo caso le rette di regressione delle tre scuole hanno, infatti, una pendenza che cambia sia fra l'una e l'altra sia rispetto all'inclinazione della retta generale. Pertanto, per una variazione unitaria della variabile X si producono incrementi di grandezza differente sulla variabile Y all'interno di ciascuna scuola e dunque non solo le medie in matematica dei tre istituti saranno fra loro diverse, ma anche la relazione fra status sociale e rendimento sarà modulata in maniera differente, ora con una accentuazione, ora con una moderazione del suo effetto. Di conseguenza, a maggior ragione l'uso di un modello di regressione ordinaria produrrebbe in questo caso risultati sensibilmente diversi rispetto ad un modello di regressione multilivello.

Sulla base di questa breve esposizione⁸, possiamo ora estendere il modello di regressione della (1) al caso più generale⁹ in cui i dati sono organizzati secondo una struttura gerarchica. Al fine di poter distinguere le unità di secondo livello (le scuole), in cui sono raggruppate le unità di primo livello (gli studenti), è necessario introdurre un indice di scuola $j = 1, 2, \dots, J$. La formulazione multilivello della (1) diviene pertanto:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + e_{ij} \quad (2)$$

La (2) permette di esprimere il rendimento in matematica Y_{ij} dell'alunno i -mo della scuola j -ma in funzione del rendimento medio della scuola (β_{0j}) e dell'effetto medio (β_{1j}) dello status socio-economico sul rendimento in matematica, sempre all'interno della scuola j -ma. Anche la (2) prevede una componente di errore, con lo stesso significato illustrato per il modello (1).

A loro volta, i coefficienti β_{0j} e β_{1j} possono essere espressi in funzione di una media di gruppo, γ_{00} e γ_{10} , di una qualche caratteristica di gruppo Z_j (ad es., lo status sociale medio della scuola, o il suo statuto pubblico o privato) più una componente di errore, u_{0j} e u_{1j} , associata al particolare istituto (figura 3).

In formule si ha:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + u_{0j} \quad \text{con } u_{0j} \sim N(0, \sigma_{u0}^2) \quad (3)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}Z_j + u_{1j} \quad \text{con } u_{1j} \sim N(0, \sigma_{u1}^2) \quad (4)$$

⁸ Per una esposizione più esaustiva della modellistica in questione e dei nodi ad essa collegati, vedi Bryk/Raudenbush (2002), Hox (2002), Paccagnella (2006).

La sostituzione della (3) e della (4) nella (2) fornisce il modello “completo” o “combinato”, che facilmente può essere esteso al caso in cui le variabili esplicative (predittori) di primo e secondo livello siano più di una.

Nelle analisi multilivello, il primo passo consiste nello stimare un modello cosiddetto “nullo” o “vuoto”¹⁰, in cui non vi sono predittori né a livello 1 né a livello 2 e si assume che il valore di Y_{ij} dipenda solo dalla media dell'intera popolazione (γ_{00}), dalla componente di errore associata al gruppo di appartenenza (u_{0j}) e da una componente di errore individuale (e_{ij}). La variabilità di Y_{ij} viene in tal modo ad esser scomposta nella somma di due componenti: la varianza “tra” i gruppi (σ_{uo}^2) e la varianza “entro” i gruppi (σ^2). È così possibile calcolare il “coefficiente di correlazione intraclasse” (ρ), che permette di valutare il grado di omogeneità tra le osservazioni appartenenti allo stesso gruppo. In formula si ha:

$$\rho = \frac{\sigma_{uo}^2}{\sigma_{uo}^2 + \sigma^2} \in [0; 1] \quad (5).$$

Il coefficiente ρ assume valori compresi tra 0 e 1. Esso esprime dunque la proporzione tra la varianza di secondo livello e la somma delle varianze dei due livelli, o varianza totale. Quanto più ρ si avvicina ad 1, tanto più forte è l'effetto del raggruppamento e quindi maggiore l'utilità di un'analisi multilivello. Introducendo nella stima dei modelli successivi al modello base una o più variabili esplicative, di primo e secondo livello, le componenti di varianza “tra” ed “entro” si riducono e questo consente di valutare, per differenza rispetto alla varianza iniziale, il contributo di ciascun predittore, o gruppo di predittori, alla spiegazione della variabilità dei dati.

10.2 Descrizione della procedura d'analisi e delle variabili considerate

Stando a quanto accennato nella chiusa del paragrafo precedente, l'analisi condotta sui dati relativi alle prestazioni in comprensione della lettura degli studenti quindicenni della provincia di Bolzano¹¹ ha preso avvio con il calcolo di un primo modello (denominato “0”) che non contiene variabili esplicative né di primo né di secondo livello e si limita a scomporre la varianza complessiva dei punteggi nella componente tra le scuole e in quella tra gli alunni all'interno delle scuole. Come si può vedere dalle tabelle da 1 a 3 alle pagine seguenti¹², la varianza dovuta alle differenze tra le scuole¹³ rappresenta quasi il 47% della varianza totale ed è dunque, come era d'altronde da attendersi data la natura a filiere dell'organizzazione scolastica nel grado secondario superiore, piuttosto elevata, giustificando pienamente l'opportunità di un'analisi a più livelli. La restante varianza, corrispondente al 53% circa, è invece da attribuire alle differenze tra gli alunni all'interno delle scuole.

Al primo livello, o livello studente, è stata presa in considerazione, oltre alle caratteristiche socio-demografiche degli studenti, una nutrita serie di altre variabili individuali potenzialmente in grado di incidere sulla prestazione in lettura.

⁹ Più generale perché la regressione ordinaria può esser considerata come corrispondente al caso particolare (visto nel primo esempio di figura 2) in cui la popolazione di studenti di ciascuna scuola riproduce sostanzialmente al proprio interno la situazione riscontrabile nella popolazione complessiva per quanto riguarda l'andamento delle variabili oggetto d'interesse.

¹⁰ Equivalente all'analisi della varianza ad una via con effetti casuali.

¹¹ Va detto che dall'analisi sono stati esclusi gli studenti della scuola media, in quanto essendo essi costituiti dai soli studenti che a 15 anni si trovano ancora in questo livello scolastico, non possono esser considerati rappresentativi degli istituti secondari di primo grado. Il campione considerato nell'analisi qui illustrata è così risultato composto da 2129 studenti, suddivisi in 81 scuole secondarie di secondo grado (27 licei, 21 istituti tecnici, 11 istituti professionali e 22 scuole di formazione professionale).

¹² In ognuna di esse, i risultati del modello “0” sono riportati in prima colonna.

¹³ Essa equivale di fatto al coefficiente intraclasse, o coefficiente ρ , moltiplicato per 100.

Per valutare l'effetto sul rendimento delle caratteristiche degli alunni, i relativi indicatori sono stati inseriti nei modelli stimati per blocchi di variabili affini. In una prima fase, i singoli blocchi sono stati introdotti a turno uno alla volta e si sono così calcolati sei distinti modelli, uno per ciascun gruppo di variabili. In un secondo momento, le variabili di ciascun gruppo che avevano mostrato un effetto significativo sul risultato in lettura, a parità delle altre appartenenti allo stesso blocco, sono state tutte insieme inserite in un ulteriore modello, per stimare di nuovo il contributo di ognuna alla variabilità dei risultati e identificare quelle che mantenevano un effetto significativo "al netto" delle altre (vedi modello 1_7 in tabella 1).

Un'analoga procedura, ma un po' più articolata, è stata seguita nella stima degli effetti delle variabili di livello 2, anch'esse riunite in gruppi. In un primo tempo, ciascun gruppo di variabili affini è stato esaminato separatamente mediante una serie di modelli in cui sono stati inseriti solo predittori di secondo livello allo scopo di valutarne l'influenza sul punteggio medio in lettura delle scuole (*means as outcomes regression*). Successivamente, le variabili esplicative sopravvissute a questa prima fase per aver mostrato un effetto significativo, sono state introdotte in un successivo modello (2_7 in tabella 2) e ne è stato ricalcolato l'effetto.

Da ultimo, è stata stimata una serie di modelli finali con le sole caratteristiche degli studenti e delle scuole che avevano superato il vaglio delle fasi precedenti. Le variabili degli alunni sono state inserite simultaneamente, mentre le variabili di scuola sono state aggiunte una dopo l'altra, così da valutarne l'effetto sotto il controllo delle variabili individuali e arrivare progressivamente alla stima di un modello conclusivo con tutte le variabili rilevanti di primo e secondo livello (vedi tabella 3).

Per maggiore chiarezza, diamo di seguito l'elenco delle variabili a livello studente e a livello scuola, suddivise nei vari blocchi¹⁴. Accanto alla descrizione sintetica di ciascuna compare la denominazione con cui la variabile è indicata nelle tre tabelle dove sono riportati i risultati dell'analisi, mentre, per le variabili categoriali, sono indicate tra parentesi le categorie, di cui la prima è quella assunta come base di riferimento¹⁵.

A. Variabili di livello 1 (studente):

A.1) Variabili socio-demografiche:

- Indice di status socio-economico-culturale, semplice e al quadrato: ESCS, ESCS²
- Genere (maschio; femmina): FEMMINA
- Origine (autoctona; immigrata): IMMIG
- Lingua parlata a casa (lingua del test; un'altra lingua): LINGCASA
- Numero di figli (fino a tre; superiore a tre): NUM_FIGLI

A.2) Variabili scolastiche:

- Frequenza della scuola dell'infanzia (frequenza per più di un anno; nessuna frequenza o per un anno al massimo): FREQ_SI
- Regolarità del percorso precedente (alunno regolare o in anticipo; in ritardo): RITARDO
- Voto riportato a scuola nella lingua del test: VOTO_LT
- Titolo di studio più alto che lo studente si aspetta di arrivare a conseguire: TITOLO

A.3) Atteggiamento verso la scuola e strategie di apprendimento

- Atteggiamento verso la scuola: ATSCHL
- Uso di strategie metacognitive basate sul comprendere e il cercar di ricordare: UNDREM
- Uso di strategie metacognitive basate sul riassunto dei contenuti: METASUM
- Uso di strategie basate sull'apprendimento mnemonico: MEMOR

¹⁴ Per una descrizione più dettagliata di tutte le variabili considerate nelle analisi si rinvia all'Appendice, dove sono date anche le principali statistiche ad esse relative e le percentuali di valori mancanti.

¹⁵ Nel caso di variabili politomiche, sono state create delle variabili *dummy*, pari al numero di categorie diminuito di 1.

- Uso di strategie basate sull'elaborazione delle informazioni: ELAB
- Uso di strategie di controllo: CSTRAT

A.4) Variabili relative alla disponibilità e all'uso di Tecnologie dell'Informazione e della Comunicazione (TIC):

- Disponibilità di risorse tecnologiche: ICTRES
- Disponibilità delle TIC a casa: ICTHOME
- Uso delle TIC per fare i compiti di scuola: HOMSCH
- Fiducia nella propria capacità di svolgere compiti d'alto livello al computer: HIGHCONF
- Uso del computer per passatempo: ENTUSE
- Atteggiamento verso il computer: ATTCOMP

A.5) Variabili relative alle abitudini di lettura:

- Tempo dedicato alla lettura per proprio piacere (non legge; fino a 30' al giorno; più di 30' al giorno): TEMPO_LET1, TEMPO_LET2
- Frequentazione biblioteche interne ed esterne alla scuola: LIBUSE
- Grado di piacere personale ricavato dalla lettura: JOYREAD
- Grado di varietà dei materiali di lettura affrontati: DIVREAD
- Lettura on-line: ONLNREAD

A.6) Variabili relative all'influenza dei genitori:

- Sostegno ai figli nelle prime fasi di apprendimento della lettura: PRESUPP
- Motivazione personale dei genitori alla lettura: MOTREAD
- Risorse (libri, giornali, riviste, ecc.) presenti in casa e a disposizione dei figli: READRES
- Sostegno attuale all'impegno nella lettura dei figli: CURSUPP

B. Variabili di livello 2 (scuola):

B.1) Variabili di contesto esterno:

- Localizzazione dell'istituto (in un centro con non più di 15.000 abitanti; in un centro con più di 15.000 abitanti): UBIC
- Esposizione alla competizione con altre scuole (con nessun'altra scuola; con almeno un'altra scuola): COMPET

B.2) Tipologia d'istituto:

- Gruppo linguistico di appartenenza della scuola (tedesco/ladino; italiano): LINGSC¹⁶
- Tipo di indirizzo (Liceo; Istituto Tecnico; Istituto Professionale; scuola o centro di formazione professionale): TECNICO, PROFESS, CFP

B.3) Composizione del corpo studentesco:

- Media dell'indice di status socio-economico-culturale degli studenti: SCHESCS
- Dimensione della scuola, semplice e al quadrato: SCHSIZE, SCHSIZE²

¹⁶ Dato il numero esiguo, gli studenti delle scuole di lingua ladina sono stati accorpati a quelli delle scuole di lingua tedesca.

- Percentuale di ragazze: PCGIRLS
- Percentuale di studenti d'origine immigrata: PCIMMIG

B.4) *Variabili relative alle risorse umane e materiali:*

- Numero di studenti per insegnante: STRATIO
- Carenza quantitativa e qualitativa degli insegnanti di lingua, matematica e scienze: TCSHORT
- Proporzione di computer a uso didattico rispetto agli studenti della classe modale: IRATCOMP
- Qualità delle attrezzature per l'insegnamento: SCMATEDU
- Disponibilità delle TIC a scuola: ICTSCH
- Uso a scuola delle TIC: USESCH

B.5) *Variabili relative ai processi:*

- Partecipazione dei docenti: TCHPARTI
- Leadership del capo d'istituto: LDRSHP
- Offerta di attività extra-curricolari: EXCURACT
- Comportamenti degli insegnanti: TEACBEHA
- Comportamenti degli studenti: STUDEBEHA
- Relazioni tra docenti e studenti: STUDREL
- Clima disciplinare: DISCLIMA
- Pressione dei genitori per alti livelli di risultato (da nessuno; da alcuni o molti): PRESSGE
- Percezione da parte dei genitori della qualità della scuola: PQSCHOOL
- Coinvolgimento dei genitori nella vita della scuola: PARINVOL

B.6) *Variabili relative all'insegnamento della lingua del test:*

- Media ore settimanali curricolari di lingua (LORESET)
- Stimolo da parte degli insegnanti all'impegno nel leggere: STIMREAD
- Qualità del lavoro in classe sulla lettura di testi: STRSTRAT
- Lettura per la scuola di testi narrativi: RFSINTRP
- Lettura per la scuola di testi non continui: RFSNCONT
- Lettura per la scuola di testi letterari o di storia della letteratura: RFSTRLIT
- Lettura per la scuola di testi funzionali: RFSFUMAT

È opportuno rilevare che mentre le variabili di primo livello sono indicatori costruiti sulla base delle risposte al questionario-studente o, nel caso delle variabili del gruppo A.6, al questionario genitori, le variabili di secondo livello sono ottenute in parte dal questionario-scuola, compilato dal dirigente o da un suo delegato, in parte aggregando a livello di scuola le risposte a domande del questionario-studente o del questionario-genitori. È questo il caso, ad esempio, dell'indice medio di status sociale dell'istituto o del numero medio di ore curricolari di insegnamento della lingua del test per settimana, ma anche degli indici sul clima relazionale e disciplinare, o di percezione della qualità della scuola, che sono basati sulla media delle risposte degli alunni o dei genitori di uno stesso istituto.

Prima di chiudere, può esser utile fornire alcune ulteriori precisazioni sulle scelte fatte in relazione al problema dei valori mancanti (*missing values*) e alla specificazione dei modelli stimati. Per quanto riguarda il primo punto, i valori mancanti di tutte le variabili continue sono stati imputati con la media di scuola se si trattava di variabili di primo livello, o con la media dell'indirizzo di appartenenza, quando le variabili erano indicatori di secondo livello. Nel caso di variabili dicotomiche (*dummy*), i valori mancanti sono stati assimilati ad una delle due categorie in gioco, in base a considerazioni legate alle particolarità di ciascuna e della situazione esaminata. Al fine di tener sotto controllo le eventuali distorsioni che potevano conseguire dal procedimento d'imputazione, per ogni variabile di livello 1 con

valori mancanti è stata costruita una nuova variabile di tipo dicotomico (*missing dummy*), codificata con il valore “0” se il valore della variabile ad essa associata era presente e con il valore “1” se il valore risultava invece mancante. Quando, per ognuna delle variabili, la percentuale di valori mancanti raggiungeva o superava il 5%, la corrispondente *missing dummy* è stata inserita nei calcoli per valutarne l’eventuale effetto sul rendimento in lettura¹⁷.

In riferimento al secondo e più rilevante punto, in tutti i modelli stimati le pendenze delle rette di regressione sono state costrette sul loro valore medio e solo le intercette sono state lasciate libere di variare casualmente. Tutte le variabili individuali di tipo continuo, inoltre, sono state centrate sulla media generale (vale a dire la media degli studenti della provincia di Bolzano), mentre le variabili dicotomiche non sono state centrate¹⁸.

10.3 Gli effetti delle caratteristiche individuali degli studenti

La tabella 1 riporta i risultati della stima dei modelli contenenti solamente le variabili esplicative di primo livello¹⁹.

Tabella 1: Effetti delle variabili degli studenti sui risultati in lettura PISA 2009

	Modello 00	Modello 1_1	Modello 1_2	Modello 1_3	Modello 1_4	Modello 1_5	Modello 1_6	Modello 1_7
Intercetta	489,7	481,5	501,1	490,7	490,1	482,7	489,8	489,8
Livello Studente								
ESCS		5,5**						0,1
ESCS ²		-3,7*						-2,4
FEMMINA		31,2***						10,6**
IMMIG		-59,4***						-41,9***
LINGCASA		-14,8***						-8,4**
NUM_FIGLI		-3,4						-
FREQ_SI			-7,4					-
RITARDO			-43,3***					-30,1***
VOTO_LT			16,2***					9,3***
TITOLO			6,4***					4,5***
ATSCHL				2,1				-
UNDREM				8,0***				5,3***
METASUM				19,8***				10,5***
MEMOR				-10,8***				-7,2***
ELAB				-1,7				-
CSTRAT				7,3***				2,5
ICTRES					9,4***			4,5**
ICTHOME					-0,6			-

¹⁷ Di fatto, nessuna delle *missing dummy* si è rivelata significativa, per cui esse non appaiono nelle tabelle sui risultati dell’analisi.

¹⁸ Per il problema della centratura delle variabili nei modelli *multilevel* si veda: Paccagnella, 2006.

¹⁹ Tutti i modelli commentati in questo paragrafo e nei successivi sono stati calcolati utilizzando il software HLM 6.06 (Raudenbush *et al.*, 2008). Le equazioni sottese ai vari modelli stimati sono presentate nell’Appendice 2.

	Modello 00	Modello 1_1	Modello 1_2	Modello 1_3	Modello 1_4	Modello 1_5	Modello 1_6	Modello 1_7
HOMSCH					-11,5***			-11,4***
HIGHCON					8,2***			4,3**
ENTUSE					-4,0*			-0,2
ATTCOMP					4,7**			3,1*
TEMPO_LET1						16,9***		12,3**
TEMPO_LET2						17,3***		17,7***
LIBUSE						-15,4***		-9,1***
JOYREAD						19,6***		10,9***
DIVREAD						5,7*		5,6**
ONLNREAD						4,2*		3,3
PRESUPP							2,5	-
MOTREAD							7,5***	3,1*
READRES							5,0**	3,0*
CURSUPP							-6,8**	-6,7***
Componenti casuali								
Varianza Livello1	4416,1	4001,4	3690,4	3899,6	4214,5	3743,3	4341,6	2797,4
Varianza livello 2	3886,5	3072,1	2328,8	2707,6	3716,0	2826,9	3615,6	1407,7
Varianza between (%)	46,8							
Riduz. varianza within (%)		9,4	16,4	11,7	4,6	15,2	1,7	36,7
Riduz. varianza between (%)		21,0	40,1	30,3	4,4	27,3	7,0	63,8

* 0.05 < p-value ≤ 0.10; ** 0.01 < p-value ≤ 0.05; *** p-value ≤ 0.01

Nel leggere i dati presentati nella tabella, è opportuno tener presente che in ciascuno dei modelli da 1_1 a 1_7, i valori delle intercette rappresentano il punteggio stimato in lettura, all'interno delle scuole, di uno studente tipico che presenta valori eguali alla media²⁰ nel caso di variabili continue e appartiene alla categoria assunta come base di riferimento nel caso di variabili categoriali. I valori registrati nelle colonne corrispondenti ad ognuno dei sette modelli indicano la variazione del punteggio individuale in lettura prodotta, rispetto all'intercetta, da una variazione unitaria²¹ dei predittori inseriti nel calcolo.

Come sovente si riscontra nelle ricerche sulla scuola, le variabili socio-demografiche (modello 1_1) mostrano un'associazione con i risultati conseguiti dagli allievi. L'indice di status socio-economico-culturale (ESCS) della famiglia di provenienza²² ha, a livello individuale, un'incidenza non forte ma tuttavia significativa, che tende però progressivamente a diminuire fino ad annullarsi, come indica il segno negativo dell'effetto, anch'esso significativo, dell'ESCS quadratico; ciò denota una relazione di tipo curvilineo tra la variabile in esame e il risultato in lettura.

²⁰ Si ricordi quanto detto alla nota 5.

²¹ Tale variazione corrisponde alla deviazione standard per gli indicatori continui standardizzati (come ad es. l'ESCS), al passaggio dalla categoria di riferimento alla categoria opposta per le variabili categoriali dicotomiche e all'incremento di una unità per le variabili quantitative non standardizzate (ad es. le ore curriculari di scienze). In alcuni casi, come la dimensione della scuola e la percentuale di ragazze, l'unità equivale rispettivamente a 100 studenti e 10 punti percentuali.

²² Questo indicatore – il cui ruolo è particolarmente rilevante, come si vedrà, al secondo livello - misura la qualità dell'ambiente familiare di uno studente mediante l'integrazione di tre dimensioni: a) il livello di istruzione dei genitori, b) il prestigio sociale del lavoro che essi svolgono, c) la presenza in casa di risorse educative (ad es. un dizionario, ecc.) e di beni di tipo materiale e culturale.

L'essere di genere femminile comporta un aumento consistente (31 punti), a parità delle altre condizioni, del punteggio in lettura. Il risultato conferma quanto già era emerso dalle precedenti rilevazioni PISA, in cui la superiorità delle femmine nella comprensione di testi si è rivelata una costante in tutti i paesi che hanno partecipato alle diverse tornate dell'indagine.

L'origine immigrata è associata a un significativo effetto negativo di notevole entità, comportando quasi 60 punti di svantaggio per uno studente straniero rispetto a un alunno autoctono. Anche il parlare a casa una lingua diversa da quella nella quale è stata svolta la prova comporta una diminuzione del punteggio di una quindicina di punti.

Passando al gruppo delle variabili relative al percorso e alla carriera scolastica degli alunni (modello 1_2), possiamo constatare che, come era da attendersi, l'essere in ritardo negli studi si traduce in un effetto sensibilmente negativo sul punteggio in lettura. L'aver riportato in Italiano un voto²³ più alto di una unità di deviazione standard rispetto alla media provinciale è invece associato a un incremento del punteggio, e così pure le aspettative degli studenti relative al titolo di studio che prevedono di conseguire in futuro. Quanto più elevato è il livello del diploma che gli alunni pensano di poter arrivare a ottenere²⁴, tanto migliore è il punteggio in lettura. L'unica variabile di questo gruppo che non discrimina fra gli studenti è la frequenza della scuola dell'infanzia: il punteggio degli alunni che non l'hanno frequentata, o che l'hanno fatto per non più di un anno, non si differenzia significativamente rispetto a quello di chi invece l'ha frequentata per due o più anni.

Le variabili del terzo gruppo si riferiscono all'atteggiamento dello studente verso la scuola (ATTSCHL), al modo in cui gli studenti affrontano lo studio (MEMOR, ELAB, CSTRAT) e alle strategie metacognitive che essi ritengono più efficaci per comprendere e ricordare un testo (UNDREM) e per riassumerne il contenuto (METASUM). L'atteggiamento dello studente nei confronti della scuola non sembra esercitare alcun effetto significativo sulle prestazioni in lettura e altrettanto dicasi per l'uso di strategie di elaborazione delle informazioni per studiare (ad es., tentando di collegare le nuove informazioni con altre apprese precedentemente o con le proprie esperienze). Il far conto sul cercar di imparare a memoria quello che si sta studiando ha un significativo effetto negativo sul risultato nella prova di comprensione, mentre il ricorso a strategie di controllo del proprio apprendimento (ad es. cercar di ricordare i punti più importanti di ciò che si studia o di reperire informazioni aggiuntive quando ci si accorge di non aver capito qualcosa) incide positivamente. Anche la consapevolezza delle strategie più efficaci per comprendere e ricordare un testo e soprattutto per farne il riassunto appare collegata in senso positivo con il rendimento in lettura.

Le variabili relative all'accesso alle risorse tecnologiche e alla disponibilità e all'uso a casa delle stesse (modello 1_4) mostrano – fatta eccezione per l'indicatore ICTHOME – effetti significativi sul punteggio in lettura, anche se con segni diversi. Mentre l'accesso alle tecnologie, la fiducia nella propria capacità di affrontare con successo compiti d'alto livello nell'uso del computer e l'atteggiamento verso di esso sono associati con risultati in lettura più elevati, l'utilizzo del computer per svolgere i compiti a casa e per divertirsi, in misura superiore alla media, si associa a risultati più bassi, in particolare nel primo caso. Le TIC appaiono dunque, rispetto alla competenza nella lettura, come uno strumento a doppio volto, ora capace di esercitare un'influenza positiva, ora decisamente negativa.

Nel modello 1_6 sono prese in considerazione le variabili più direttamente legate ai comportamenti verso la lettura dello studente e non a caso esse hanno tutti effetti significativi sul rendimento nella prova di comprensione. Se uno studente dichiara di dedicare alla lettura per proprio piacere almeno mezz'ora al giorno, ciò si traduce in un aumento di 17 punti rispetto a chi afferma di non leggere affatto e il punteggio cresce di altri 17 punti se lo studente legge per più di mezz'ora al giorno. Anche il piacere che lo studente trae dalla lettura comporta un aumento significativo di ben 19 punti in più, mentre più deboli, ma pur sempre significativi, sono gli effetti connessi al grado di varietà delle letture cui l'alunno si dedica con regolarità e al leggere via Internet. Curiosamente, è invece negativo l'effetto associato al frequentare le biblioteche, esterne o interne alla scuola, per prendere libri in prestito, fare i compiti, ecc. (-15,4 punti). Ciò è probabilmente dovuto a due ordini di ragioni: da un lato, la presenza delle altre variabili che condizionano l'effetto di LIBUSE; se infatti si considera l'indicatore isolatamente, il suo coefficiente lineare è trascurabile e non significativo. D'altro lato, il suo valore

²³ La variabile è stata preliminarmente standardizzata con media eguale a 0 e deviazione standard eguale a 1.

²⁴ Le risposte degli studenti sul titolo di studio che pensavano di poter conseguire in futuro sono state codificate secondo una scala da 0 a 6, in cui lo 0 corrisponde alle risposte non date e i valori da 1 a 6 alla classificazione ISCED dei titoli di studio, da ISCED 2 (diploma di scuola secondaria inferiore) a ISCED 5A (laurea specialistica o dottorato). La variabile è stata quindi riscalata sottraendo il valore mediano (=3).

medio è più alto per gli studenti del gruppo linguistico tedesco o ladino rispetto a quelli del gruppo italiano (-0,08 di contro a -0,16), cosicché esso di fatto funziona come marcatore degli studenti di lingua italiana, il cui punteggio in lettura è più basso rispetto a quello dei coetanei di lingua tedesca o ladina.

L'ultimo gruppo di variabili esaminato riguarda l'influenza dei genitori sulla competenza in lettura dei figli. Le due variabili che mostrano un effetto significativo in senso positivo sui risultati sono il grado di motivazione personale dei genitori a leggere (MOTREAD) e la presenza in famiglia di libri, giornali e riviste a disposizione anche dei figli (READRES), mentre l'aver seguito i figli nelle prime fasi di apprendimento della lettura a scuola (PRESUPP) non ha un'associazione significativa con i risultati nella prova di comprensione. A sorpresa, la variabile CURSUPP, che integra le risposte dei genitori a domande sulla frequenza con cui discutono con i figli di temi d'attualità, di libri, film o programmi televisivi e sulla frequenza con cui fanno con loro una serie di attività quali cenare insieme, parlare, andare in libreria o in biblioteca, assisterli nei compiti, ecc., sembra avere un significativo effetto negativo sul rendimento in lettura. Tale esito è però ingannevole in quanto la variabile, anche in questo caso, tende a identificare gli studenti del gruppo linguistico italiano rispetto a quelli del gruppo tedesco/ladino, che hanno in lettura un punteggio medio più elevato ma presentano sull'indicatore in questione un valore medio nettamente più basso rispetto ai primi (-0,37 versus 0,80). Questa diversità di comportamenti fra i genitori del gruppo tedesco rispetto a quelli del gruppo italiano dipende probabilmente, oltre che da ragioni culturali, dal più basso livello medio dell'indice di status (in particolare, per ciò che riguarda la componente costituita dal livello d'istruzione) dei primi rispetto ai secondi²⁵.

Nel modello 1_7 le variabili dimostrate significative nei modelli precedentemente stimati sono state inserite insieme e ne sono stati ricalcolati gli effetti. Come si può vedere dalla tabella 1, alcune variabili, come ad esempio, lo status socio-economico-culturale – ma anche l'uso di strategie di elaborazione, l'utilizzo del computer per divertimento e il leggere on-line – cessano di essere significative, mentre tutte le altre confermano la significatività del proprio effetto sul rendimento in lettura anche se, generalmente, in misura più ridotta. In particolare, l'effetto del genere si ridimensiona scendendo da 31 punti a poco più di 10 punti.

Per concludere, è da sottolineare che le variabili esplicative prese in considerazione nel modello 1-7 riducono la varianza all'interno delle scuole di una quota pari quasi al 37%. La notevole diminuzione che si determina anche per quanto riguarda la varianza tra le scuole (64% circa) attesta la non uniforme distribuzione fra di esse delle caratteristiche degli alunni e non rappresenta un autentico effetto contestuale (Hox, 2002).

10.4 Gli effetti delle caratteristiche di scuola

La tabella 2 riporta i risultati della stima dei modelli contenenti solamente le variabili esplicative di secondo livello, ossia le variabili di scuola.

Tabella 2: Effetti delle variabili di scuola sui risultati in lettura PISA 2009

	Modello 00	Modello 2_1	Modello 2_2	Modello 2_3	Modello 2_4	Modello 2_5	Modello 2_6	Modello 2_7
Intercetta	489,7	505,7	569,4	504,1	483,8	479,8	493,2	545,4
Livello Scuola								
UBIC		-13,2						-
COMPET		-26,2						-
LINGSC			-40,2***					-40,4***
TECNICO			-51,9***					-20,3*

²⁵ L'Escs medio degli studenti tedeschi e ladini è -0,27 rispetto a un valore di 0,00 fra gli studenti italiani.

	Modello 00	Modello 2_1	Modello 2_2	Modello 2_3	Modello 2_4	Modello 2_5	Modello 2_6	Modello 2_7
PROFESS			- 106,6***					-70,2***
CFP			- 135,6***					-86,6***
SCHESCS				110,5***				40,2***
SCHSIZE				-0,7				-
SCHSIZE ²				0,0				-
PCGIRLS				4,2***				1,6
PCIMMIG				-18,4***				-8,9*
STRATIO					2,1			-
TCSHORT					6,4			-
RATCOMP					-2,6			-
SCMATEDU					28,5***			3,1
ICTSCH					31,1			-
USESCH					-79,3**			-11,9
TCHPARTI						24,0**		0,5
LDRSHP						5,9		-
EXCURACT						14,4**		-6,2*
TEACBEHA						-4,8		-
STUDBEHA						22,0*		4,9
STUDREL						-23,6		-
DISCLIMA						31,3**		-2,6
PRESSGE						25,7**		4,9
PQSCHOOL						25,1*		-3,6
PARINVOL						-26,5		-
LORESET							-21,6**	-11,5*
STIMREAD							-34,5	-
STRSTRAT							35,7	-
RFSINTRP							76,0***	1,9
RFSNCONT							-11,8	-
RFSTRLIT							-13,8	-
RFSFUMAT							-74,7***	-22,8**
Componenti casuali								
Varianza Livello1	4416,1	4416,1	4413,8	4423,0	4415,5	4414,7	4419,0	4423,0
Varianza livello 2	3886,5	3741,5	638,4	1219,2	3233,0	2491,7	2815,1	375,3
Varianza tra (%)	46,8							
Rid. Var. entro(%)		0,0	0,1	-0,2	0,0	0,0	-0,1	-0,2
Rid. Var. tra (%)		3,7	83,6	68,6	16,8	35,9	27,6	90,3

* 0.05 < p-value ≤ 0.10; ** 0.01 < p-value ≤ 0.05; *** p-value ≤ 0.01

Anche nel caso delle variabili di scuola, i dati della tabella 2 si leggono in modo del tutto analogo a quelli riportati nella tabella 1. Più precisamente, ciascun valore esprime l'effetto di una variazione unitaria, sulla media in lettura di scuola, delle variabili inserite nei diversi modelli. Il primo modello vede l'introduzione delle variabili relative al contesto esterno che, come si può notare, non esercita un effetto significativo sul risultato medio di scuola. L'ubicazione della scuola in un piccolo centro o in un medio-grande non si riflette in modo significativo sui risultati in lettura della scuola, né, tantomeno, la presenza sul territorio di almeno un'altra scuola dello stesso tipo produce una variazione statisticamente significativa dei risultati.

Il secondo modello prende invece in esame le variabili relative alla tipologia di scuola, intesa sia rispetto al gruppo linguistico di appartenenza, sia in relazione all'indirizzo di studio. L'istituto tipo, posto alla base del confronto, è un liceo con lingua d'insegnamento tedesca²⁶. Dal modello 2_2 si evince chiaramente che le scuole di lingua italiana conseguono risultati significativamente più bassi di quelle di lingua tedesca, con un divario di oltre 40 punti della scala di lettura e tale differenza permane praticamente invariata anche quando si introducono le variabili degli altri modelli, ossia tutte quelle che esercitano un effetto significativo sul risultato medio di scuola. I risultati significativamente più bassi degli istituti tecnici (circa 52 punti), degli istituti professionali (circa 107 punti) e della formazione professionale (circa 136 punti) permangono anche quando si considerano altre variabili, ma vengono da esse in parte mitigati, specie quando si introduce la variabile che permette di tenere in considerazione la diversa origine socio-economico-culturale degli allievi di ciascun tipo di scuola. Il modello 2_3 indica infatti che l'innalzamento di una deviazione standard dell'indice medio di scuola dello status socio-economico-culturale si associa a un aumento di oltre 110 punti, quando viene considerato di per sé, ma che continua a segnare una differenza di oltre 40 punti sul risultato medio di scuola quando lo si esamina insieme ad altre variabili. Considerazioni in parte diverse valgono quando si considerano altre variabili relative alla composizione del corpo studentesco. Si nota infatti che la dimensione della scuola non esercita alcun effetto significativo, né in termini lineari né quadratici. Ciò significa che i dati non mettono in evidenza una relazione diretta di alcun tipo tra la dimensione della scuola e i risultati medi in lettura. Un incremento del 10% di presenza di studentesse, rispetto alla media provinciale, se singolarmente considerato, si associa a un modesto effetto positivo sul punteggio medio in lettura, ma tale effetto scompare quando si introducono altre variabili esplicative. Diverso è invece l'effetto dell'incremento della presenza di allievi di origine immigrata. Esso si associa a una diminuzione significativa del punteggio medio in lettura, sia se considerato solamente insieme alle altre variabili relative alla composizione del corpo studentesco (modello 2_3), sia insieme a tutte le altre variabili che mostrano una relazione significativa con il punteggio medio in lettura (modello 2_7). Come sarà più diffusamente illustrato nei paragrafi successivi, la competenza relativa alla comprensione della lettura pare avere comportamenti diversi rispetto alle competenze matematiche e scientifiche che sono state l'oggetto principale di indagine nel 2003 e 2006, rispettivamente. Infatti, si nota che le variabili relative alle risorse umane e materiali della scuola, quelle relative ai processi e all'insegnamento della lingua del test tendono a esercitare un effetto statisticamente significativo sul punteggio medio in lettura solo quando sono considerate insieme alle variabili della stessa tipologia, mentre raramente l'associazione con il risultato medio di scuola in lettura rimane significativa quando le predette variabili vengono inserite nell'ultimo modello (modello 2_7) in cui sono considerate anche le variabili relative alla tipologia di scuola e alla composizione del corpo studentesco. Più precisamente, tra le variabili relative alle risorse umane e materiali nessuna esercita un effetto significativo sul punteggio medio in lettura, neanche la qualità delle attrezzature per l'insegnamento e l'uso a scuola delle TIC che paiono associarsi positivamente ai risultati in lettura quando sono considerate insieme solamente a variabili dello stesso tipo (modello 2_4). Anche le variabili riferite ai processi mostrano un comportamento molto simile a quello delle variabili illustrate nell'analisi del modello 2_4. L'unica variabile che esercita un effetto statisticamente significativo e negativo sugli esiti in lettura è quella relativa all'indicatore che identifica l'offerta di attività extracurricolari. Con buona probabilità l'accesso alle predette attività, a parità di altre condizioni, in misura superiore alla media tende a individuare studenti più deboli in termini di abilità e competenza di lettura. Tra le variabili relative all'insegnamento della lingua del test solo quelle relative al numero di ore d'insegnamento settimanale e alla lettura di testi funzionali si associano significativamente e negativamente al livello medio di scuola della competenza in lettura. Il numero delle ore d'insegnamento nella lingua del test pare avere un effetto controintuitivo poiché un'ora di insegnamento in più della media si associa negativamente alla competenza di lettura. Tale effetto sembra spiegarsi col fatto che valori superiori alla media di questa variabili tendono a identificare maggiormente allievi che frequentano tipi di scuola

²⁶ Le scuole del gruppo linguistico ladino sono state accumulate a quelle del gruppo linguistico tedesco.

con livelli di *performance* più bassi e in misura maggiore scuole del gruppo linguistico italiano che, come mostrato, raggiunge livelli di competenza in lettura meno elevati di quello tedesco/ladino. Più facilmente interpretabile è l'effetto della lettura dei testi funzionali. Infatti, essa si associa negativamente ai risultati sia nel modello 2_6 sia nel modello 2_7 poiché essa identifica principalmente allievi degli istituti professionali e della formazione professionale, anche se è importante notare che tale variabile continua a essere legata negativamente e in modo significativo alla competenza in lettura anche quando si tengono sotto controllo le variabili relative al tipo di scuola.

10.5 I modelli con variabili esplicative di primo e secondo livello

La tabella 3 riporta i risultati della stima dei modelli che includono i predittori di primo e secondo livello rivelatisi significativi nelle analisi precedenti. Per valutare meglio gli effetti delle variabili di scuola sulla prestazione in lettura degli studenti, queste, come già detto al paragrafo 2, sono state inserite una alla volta nei modelli successivamente stimati.

Tabella 3: Effetti delle variabili degli studenti e delle scuole sui risultati in lettura PISA 2009

	Modello 00	Modello 3_1	Modello 3_2	Modello 3_3	Modello 3_4	Modello 3_5	Modello 3_6	Modello 3_7
Intercetta	489,7	492,5	533,0	524,7	525,3	527,7	522,9	519,9
Livello Studente								
FEMMINA		10,8**	9,8**	10,0**	10,1**	10,1**	10,3**	10,2**
IMMIG		-41,8***	-42,4***	-42,5***	-41,1***	-41,1***	-41,0***	-41,1***
LINGCASA		-8,3**	-9,0**	-9,1**	-9,2**	-9,2**	-9,2**	-9,4**
RITARDO		-30,7***	-29,8***	-29,7***	-29,6***	-29,6***	-29,4***	-29,4***
VOTO_LT		9,3***	9,5***	9,4***	9,2***	9,4***	9,3***	9,2***
TITOLO		4,7***	3,7***	3,6***	3,7***	3,7***	3,7***	3,6***
UNDREM		5,6***	5,5***	5,5***	5,4***	5,3***	5,3***	5,3***
METASUM		10,9***	10,5***	10,5***	10,6***	10,6***	10,6***	10,6***
MEMOR		-6,2***	-6,0***	-5,9***	-6,0***	-6,0***	-6,1***	-6,0***
ICTRES		5,5***	5,4***	5,4***	5,4***	5,4***	5,4***	5,3***
HOMSCH		-10,1***	-10,0***	-10,1***	-10,1***	-10,1***	-10,1***	-10,2***
HIGHCON		5,0**	4,8**	4,8**	4,9**	4,9**	4,8**	4,8**
ATTCOMP		3,6**	3,6**	3,6**	3,6**	3,6**	3,6**	3,6**
TEMPO_LET1		12,7***	12,8***	12,7***	13,0***	13,0***	12,9***	13,0***
TEMPO_LET2		17,9***	18,3***	18,2***	18,4***	18,4***	18,3***	18,4***
LIBUSE		-9,1***	-9,6***	-9,5***	-9,5***	-9,5***	-9,4***	-9,3***
JOYREAD		11,0***	10,5***	10,4***	10,5***	10,5***	10,5***	10,6***
DIVREAD		6,1***	6,0**	5,9**	5,8**	5,8**	5,8**	5,8**
MOTREAD		3,2**	3,1**	3,1**	3,0*	3,2*	3,1*	3,0*
READRES		3,4**	3,2**	3,1**	3,1**	3,1**	3,2**	3,2**
CURSUPP		-6,4***	-6,4***	-6,4***	-6,4***	-6,4***	-6,4***	-6,4***
Livello Scuola								
LINGSC		-12,9	-22,6**	-25,3***	-15,1*	-16,2*	-8,6	-15,9*
TECNICO			-19,4***	-10,5	-8,7	-11,2	-3,9	-0,1

	Modello 00	Modello 3_1	Modello 3_2	Modello 3_3	Modello 3_4	Modello 3_5	Modello 3_6	Modello 3_7
PROFESS			-64,2***	-51,5***	-50,2***	-53,1***	-46,9***	-40,5***
CFP			-75,6***	-58,4***	-55,4***	-59,1***	-58,3***	-49,5***
SCHESCS				22,7**	22,3**	20,1*	21,3*	24,7**
PCIMMIG					-8,7***	-8,6***	-8,8***	-7,8**
EXCURACT						-2,0	-1,1	-1,2
LORESET							-10,1**	-9,3**
RFSFUMAT								-21,0**
Componenti casuali								
Varianza Livello1	4416,1	2804,7	2801,1	2801,6	2805,4	2805,5	2806,0	2806,6
Varianza livello 2	3886,5	1433,7	416,7	393,3	319,0	322,8	301,1	270,5
Varianza between (%) ICC	46,8							
Riduz. varianza within (%)		36,5	36,6	36,6	36,5	36,5	36,5	36,4
Riduz. varianza between (%)		63,1	89,3	89,9	91,8	91,7	92,3	93,0

* 0.05 < p-value ≤ 0.10; ** 0.01 < p-value ≤ 0.05; *** p-value ≤ 0.01

Per quanto riguarda le variabili a livello studente, come si può constatare confrontando la tabella 3 con la 1, il quadro degli effetti ad esse connessi rimane sostanzialmente immutato, così come la percentuale di varianza spiegata all'interno delle scuole. È dunque sui predittori a livello scuola che si focalizzerà qui l'attenzione. Sono d'altronde questi ultimi che rivestono il maggiore interesse, sia per cercar di comprendere le diversità di prestazione fra una scuola e l'altra, sia per l'impostazione e la valutazione delle politiche educative.

Nel primo dei modelli calcolati (3_1), è stato inserito come variabile esplicativa soltanto il gruppo linguistico di appartenenza dell'istituto. Se la scuola è di lingua italiana, ciò comporta, a parità di caratteristiche individuali degli alunni, un abbassamento del risultato medio in lettura dello studente di circa 13 punti, ma tale differenza non è statisticamente significativa.

Lo scarto diviene invece significativo – crescendo, nel modello 3_2, da 13 fino a quasi 23 punti e nel modello 3_3 fino a 25 punti circa - quando si tiene sotto controllo il tipo d'indirizzo e lo status socio-economico-culturale medio della scuola. Nel primo caso (modello 3_2), l'aumento è dovuto al fatto che il punteggio medio in lettura degli studenti tedeschi e ladini rispetto ai loro coetanei italiani differisce significativamente per quanto riguarda gli istituti tecnici, gli istituti professionali e le scuole di formazione professionale, mentre per quanto riguarda i licei (categoria di riferimento), come si può vedere dalla tabella che segue, la differenza non è significativa.

Tab. 4: Medie in lettura e differenze tra le medie per gruppo linguistico e tipo di scuola

	Gruppo linguistico tedesco/Ladino		Gruppo linguistico italiano		Differenza tra le medie		
	Media	Err. St.	Media	Err. St.	GT/L - GI	Err. St.	t-ratio
Licei	561	2,8	551	4,6	10	5,6	1,76
Ist. Tecnici	517	3,5	485	6,8	32	7,2	4,42***
Ist. Professionali	474	4,8	391	10,0	83	11,1	7,51***
Scuole professionali	438	4,3	381	8,6	57	9,1	6,22***
Tutti gli indirizzi	498	1,7	486	3,6	12	4,1	2,95**

* 0.05 < p-value ≤ 0.10; ** 0.01 < p-value ≤ 0.05; *** p-value ≤ 0.01

Nel secondo caso (modello 3_3), l'ulteriore crescita della differenza tra il punteggio di uno studente del gruppo tedesco/ladino rispetto a uno studente del gruppo italiano dipende dal fatto che, come sopra già osservato, l'indice di status degli alunni del secondo gruppo è mediamente più alto rispetto a quello degli studenti del primo, sia a livello provinciale che nei vari tipi di scuola (tranne negli istituti tecnici), come si può vedere dalla tabella seguente.

Tab. 5: Escs medio per gruppo linguistico e tipo di scuola

	Gruppo linguistico tedesco/ladino		Gruppo linguistico italiano	
	ESCS medio	Err. St.	ESCS medio	Err. St.
Licei	0,16	0,04	0,29	0,06
Istituti tecnici	-0,20	0,04	-0,23	0,09
Istituti professionali	-0,43	0,05	-0,25	0,15
Scuole professionali	-0,61	0,05	-0,37	0,08
Tutti gli indirizzi	-0,27	0,02	0,00	0,04

Un secondo aspetto da sottolineare, anche se ben noto dalle analisi in precedenza condotte sui dati PISA in ambito locale e internazionale, è lo scarto che esiste tra l'effetto dello status personale dell'alunno sul livello raggiunto nelle competenze di base e quello prodotto dalla media dell'indice di status economico-sociale della scuola. Nel nostro caso, l'effetto netto dell'origine familiare sulla prestazione in lettura a livello individuale è praticamente nullo quando la variabile sia considerata, come nel modello 1_7 di tabella 1, insieme alle altre variabili che caratterizzano gli studenti, tanto che è stata eliminata dall'analisi. Tuttavia, anche se si guarda all'unico modello (vedi modello 1_1 di tabella 1) in cui lo status personale dello studente risulta significativo, la grandezza dell'effetto ad esso associato è decisamente minore rispetto a quella dell'effetto connesso allo status medio della scuola²⁷. Questo fenomeno non riguarda solo la provincia di Bolzano, ma esso è comune alla maggior parte dei paesi OCSE. Come il Rapporto Internazionale osserva: «In quasi tutti i paesi, indipendentemente dalla loro propria condizione socio-economica, gli studenti che frequentano scuole in cui la condizione socio-economica media è favorevole, tendono ad avere risultati migliori di quelli conseguiti quando sono iscritti a scuole che reclutano alunni in situazione di svantaggio. Nella maggioranza dei paesi OCSE, la relazione tra lo status socio-economico-culturale medio degli studenti d'una scuola e il loro rendimento è più stretta della relazione tra la condizione socio-economica personale di uno studente e il suo rendimento» (OCSE, 2009, Vol. V, p. 90: TdA.A).

E ancora: «Consideriamo per ipotesi due studenti che vivano in famiglie con uno status socio-economico medio, così come misurato dall'indice ESCS di PISA. Uno dei due studenti frequenta una scuola con uno status medio buono, ad esempio superiore alla media OCSE di un quarto di deviazione standard. La maggior parte dei compagni di questo studente proviene da famiglie benestanti. L'altro studente frequenta una scuola più svantaggiata sotto il profilo socio-economico: la media dell'indice di status è di un quarto di deviazione standard inferiore alla media OCSE, cosicché lo studente viene da una famiglia più agiata di quella di molti dei suoi compagni. I risultati indicano che il primo studente avrà prevedibilmente, in media fra i paesi OCSE, un punteggio in lettura di 32 punti più alto del secondo studente» (ibidem: TdA.A).

Un terzo elemento che val la pena di rilevare, infine, è che, mettendo sotto controllo, nel modello 3_3, l'ESCS medio di scuola, si riduce la differenza di prestazione fra gli studenti di un liceo e quella degli studenti degli altri indirizzi, non solo ma, nel caso degli studenti dell'istituto tecnico, tale differenza non è più significativa (mentre continua ad esserlo per gli studenti dell'istruzione e formazione professionale). Il cambiamento che si osserva nell'effetto prodotto dalla variabile "tipo d'indirizzo", a seconda che si controlli o meno per il livello medio di *background* degli alunni dell'istituto, pone la

²⁷ Si deve anche tener conto che nel modello 3_3 l'effetto dell'ESCS medio di scuola è temperato dalla presenza della variabile relativa all'indirizzo, che sconta le differenze di capacità e preparazione fra gli alunni dei vari tipi di scuola. Senza tale variabile, se cioè si stima un modello con, a livello 2, le sole variabili rappresentate dall'ESCS medio e dal gruppo linguistico di appartenenza della scuola, l'effetto dell'ESCS ammonta a 73,5 punti (nello stesso tempo la differenza, a parità di status, di un alunno tedesco o ladino rispetto a un alunno italiano diviene pari a +31 punti).

domanda su quale sia la relazione tra queste due variabili: si potrebbe pensare che la scelta dell'indirizzo, operata al momento dell'iscrizione alla scuola secondaria superiore, funzioni come una variabile mediatrice dell'influenza dell'origine sociale sui risultati in lettura ma anche che il diverso livello di capacità che caratterizza gli studenti a seconda dell'indirizzo prescelto intervenga a moderare l'effetto del *background*.

Interessante è anche la variazione che si produce quando, nel modello 3_4, si introduce la variabile rappresentata dalla percentuale di studenti d'origine immigrata presenti nel corpo studentesco della scuola: l'aumento di un punto percentuale degli alunni immigrati rispetto alla situazione di una scuola in cui non vi siano alunni stranieri – che è in questo caso la base di riferimento – riduce il divario di prestazione in lettura fra gli alunni del gruppo linguistico italiano da un lato e tedesco/ladino dall'altro di una decina di punti, controbilanciando l'effetto dell'indice di status socio-economico medio dell'istituto che, come abbiamo visto, tende invece ad ampliarlo. Come si può vedere dalla tabella seguente, che riporta, separatamente per ognuno dei due gruppi linguistici, la percentuale di alunni immigrati sul totale degli alunni di ciascun tipo di scuola e sul totale degli studenti di tutti gli indirizzi, non solo gli alunni d'origine straniera – complessivamente pari al 5,2% degli studenti di scuola superiore della provincia campionati – si concentrano prevalentemente nelle scuole di lingua italiana (13,2% versus 3,3%), ma la loro distribuzione nei vari indirizzi è differente.

Tab. 6: Percentuale di alunni d'origine immigrata nei vari tipi di scuola dei due gruppi linguistici (dati pesati)

	Gruppo linguistico tedesco/ladino		Gruppo linguistico italiano	
	Percentuale media studenti immigrati sugli iscritti	Err. St.	Percentuale media studenti immigrati sugli iscritti	Err. St.
Licei	3,5	0,78	5,3	1,47
Istituti tecnici	2,8	0,92	19,6	4,85
Istituti professionali	3,2	0,71	21,3	5,40
Scuole professionali	3,6	1,04	21,5	4,13
Tutti gli indirizzi	3,3	0,47	13,2	1,61

Nelle scuole del gruppo tedesco/ladino, gli stranieri sono distribuiti in proporzione relativamente uniforme nei vari indirizzi (ad es., la percentuale di immigrati nei licei è praticamente uguale a quella delle scuole professionali), mentre nelle scuole del gruppo italiano essi tendono a concentrarsi soprattutto nell'istruzione tecnica e nell'istruzione e formazione professionale, il che fa pensare che le caratteristiche degli alunni in questione siano probabilmente diverse in un caso e nell'altro. Se si fa l'esercizio di eliminare dal campione gli studenti di cittadinanza non italiana, il punteggio medio in lettura degli studenti del gruppo tedesco/ladino cresce di soli 2 punti (da 498 a 500) mentre nel gruppo italiano aumenta di 12 punti (da 486 a 498) e la differenza tra gli uni e gli altri non è più significativa²⁸.

Negli ultimi tre modelli (da 3_5 a 3_7) sono introdotte una dopo l'altra le poche variabili di scuola sopravvissute che non hanno a che fare con il contesto esterno o interno dell'istituto. Di esse, l'offerta di attività extra-curricolari cessa di esser significativa quando viene considerata insieme alle variabili di livello 1, mentre restano significative le variabili relative al numero medio di ore settimanali di lezione della lingua del test e alla lettura per la scuola di testi di tipo funzionale, sebbene il loro effetto appaia notevolmente ridotto rispetto al modello 2_7, in particolare per quanto riguarda l'ultima delle tre (cfr. tabella 2). Paradossalmente, l'aumento di un'ora dell'orario settimanale di lezione rispetto alla media provinciale determina un abbassamento del punteggio in lettura tra i 10 e i 9 punti, a seconda del modello, come si può vedere dalla tabella 3. Inoltre, l'introduzione di questa variabile nel modello 3_6 rende non più significativa la differenza di prestazione tra gli alunni del gruppo tedesco/ladino rispetto a quelli del gruppo italiano. Quest'ultimo risultato è da ricondurre al fatto che il numero di ore

²⁸ L'errore standard della differenza del punteggio in lettura dei due gruppi, pari a 2 punti, è 4,6; il p-value associato al rapporto tra la statistica e il suo errore (0,44) è dunque ben inferiore alla soglia del 5%.

settimanali dedicate all'insegnamento della lingua del test è nelle scuole del gruppo italiano più alto che nelle scuole del gruppo tedesco/ladino, come si può vedere dalla tabella che segue.

Tab. 7: Numero medio di ore settimanali di lezione della lingua del test per gruppo linguistico e tipo di scuola

	Media ore	Err.st.	Media ore	Err.st.
Licei	3,4	0,02	3,8	0,05
Istituti tecnici	4,0	0,03	4,0	0,04
Istituti professionali	3,3	0,06	5,7	0,11
Scuole professionali	3,2	0,05	3,6	0,08
Tutti gli indirizzi	3,5	0,02	4,2	0,04

Nel modello finale (3_7) il gruppo linguistico torna a esser significativo, anche se alla soglia meno elevata di probabilità (90%). L'effetto negativo della variabile costituita dalla lettura di testi non continui riflette il fatto che questo indicatore ha in genere un valore medio più basso nei licei e tende a crescere man mano che si passa agli istituti tecnici e da questi all'istruzione e formazione professionale²⁹. Non a caso esso, come si può constatare dalla tabella 3, riduce anche di alcuni punti la differenza di prestazione in lettura rispetto ad uno studente di liceo, assorbendo in parte l'effetto dovuto all'indirizzo dell'istituto.

10.6 Considerazioni conclusive

Per sintetizzare gli spunti principali di riflessione suggeriti dall'analisi effettuata, rileviamo, innanzitutto, che il modello finale (3_7) spiega³⁰ il 36% della varianza tra gli alunni all'interno delle scuole e ben il 93% della varianza tra le scuole. Come quasi sempre accade, la quota di varianza tra alunni spiegata è minore di quella tra istituti. Rimane tuttavia anche in questo secondo caso un 7% di varianza residua, non spiegata dalle variabili considerate nei modelli stimati e attribuibile a caratteristiche peculiari delle scuole del campione esaminato.

Le variabili di contesto della scuola (indirizzo dell'istituto e composizione del corpo studentesco) sono, come emerso in numerose ricerche, quelle che hanno il peso maggiore nel determinare il rendimento degli alunni, al netto delle loro caratteristiche personali. Esse rendono conto di circa il 92% della varianza tra scuole (modello_3_4).

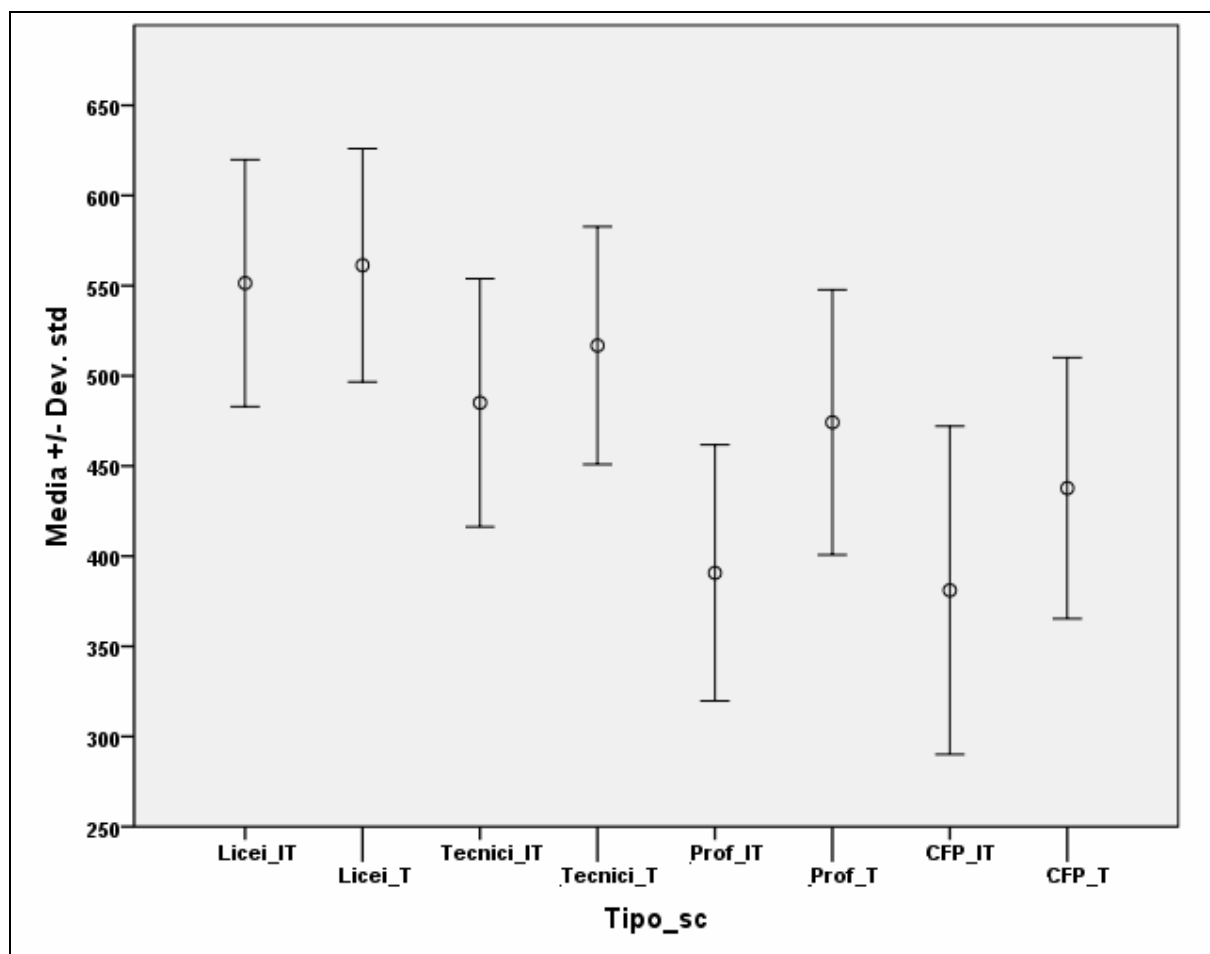
Qualche riflessione merita il ruolo svolto dal gruppo linguistico di appartenenza della scuola: da solo, come s'è visto, esso non differenzia in modo significativo i risultati in lettura degli alunni dei due gruppi, mentre discrimina tra la prestazione degli studenti italiani – abbassandola di più di 20 punti – e quella dei coetanei tedeschi o ladini quando si tiene sotto controllo il tipo d'indirizzo dell'istituto. Val la pena di soffermarsi su questo punto per cercar di spiegare come ciò accada. Se si va a vedere la distribuzione percentuale (pesata) degli alunni nei diversi tipi d'indirizzo separatamente nei due gruppi linguistici, si può constatare che, mentre nei licei di lingua tedesca/ladina si trova il 28% degli alunni, nel gruppo linguistico italiano tale quota sale al 49%. La restante popolazione di studenti del gruppo linguistico tedesco/ladino si ripartisce per il 27% negli istituti tecnici, per l'11% negli istituti professionali e per il 34% nelle scuole professionali. Le corrispondenti quote nel gruppo linguistico italiano ammontano rispettivamente al 19%, al 17% e al 15%. Posto che la competenza in lettura sia distribuita normalmente fra gli studenti, se si tien conto che le scelte dell'indirizzo riflettono in una qualche misura il grado di capacità e preparazione degli alunni in base a una gerarchia di prestigio culturale che vede in testa i licei, seguiti dagli istituti tecnici, e quindi, nell'ordine, dagli istituti professionali e dalle scuole di formazione professionale, è facilmente intuibile, come mostrato in figura

²⁹ I valori medi che la variabile assume nei vari tipi di scuola sono i seguenti: licei=-0,25; istituti tecnici=-0,05; istituti professionali=0,01; scuole professionali=0,09.

³⁰ "Spiegare" è qui assunto come equivalente di "ridurre la varianza" e non deve essergli direttamente attribuito un significato causale.

4, quale sia la logica conseguenza della diversa proporzione con cui gli studenti dei due gruppi si distribuiscono nei vari indirizzi.

Figura 4. Punteggi medi in lettura nei diversi tipi di scuola per gruppo linguistico³¹



Il crearsi, attraverso un processo “a cascata”, di un differenziale nel livello medio di competenza degli studenti a favore delle scuole di lingua tedesca/ladina diviene infatti gioco forza inevitabile, pur nel rispetto dell’ordine gerarchico fra i diversi indirizzi all’interno di ciascuna delle due comunità. In altre parole, le scuole dei due gruppi linguistici, in particolare per quanto riguarda l’istruzione tecnica e professionale, intercettano porzioni diverse della curva di distribuzione delle abilità. Ad esempio, gli studenti degli istituti tecnici tedeschi o ladini si troveranno presumibilmente, per la maggior parte, nella metà superiore di tale curva, mentre i loro coetanei italiani si troveranno per la maggior parte nella parte inferiore.

Rimane il fatto che, in ogni caso, tra gli studenti del gruppo tedesco/ladino complessivamente considerati e gli studenti italiani nel loro insieme esiste una differenza significativa di 12 punti (vedi tabella 4) ma tale differenza, come abbiamo visto nel paragrafo precedente, è in gran parte dovuta alla maggior presenza di studenti d’origine immigrata nelle scuole di lingua italiana e al diverso livello di abilità da essi posseduto rispetto a quelli che frequentano le scuole di lingua tedesca. Come si ricorderà, se dal campione si tolgono gli studenti stranieri, la differenza tra i due gruppi si riduce a soli due punti e cessa di essere significativa. Ovviamente, l’interpretazione qui avanzata non esclude affatto l’intervento dell’azione di altri fattori connessi al funzionamento delle scuole, che dovrebbero però a questo punto essere oggetto di ricerche mirate non essendo sufficienti, o non pertinenti nel caso in questione, gli elementi d’informazione forniti dal questionario-scuola di PISA.

³¹ In ogni barra verticale il pallino rappresenta il punteggio medio in lettura, mentre l’estremo inferiore di ciascun segmento individua il punteggio medio meno la deviazione standard e l’estremo superiore il punteggio medio più la deviazione standard.

Il secondo aspetto che in sede conclusiva intendiamo rimarcare è l'importanza, principalmente a livello di scuola, del livello socio-economico-culturale della famiglia di provenienza degli studenti. Mentre sul piano individuale l'effetto dell'ESCS cessa di essere significativo quando si introducono altre variabili oltre a quelle socio-demografiche, a livello di scuola il ruolo giocato da questo indicatore rimane significativo in tutti i modelli specificati. Come è stato già sottolineato nel paragrafo precedente, la composizione del corpo studentesco genera un effetto "contestuale" che va al di là dell'effetto del *background* personale del singolo alunno. Tuttavia, è opportuno rilevare che il disegno di ricerca di PISA enfatizza il ruolo svolto dallo status socio-economico medio degli studenti d'un istituto perché manca una misura del livello di competenza/capacità degli studenti all'entrata della scuola secondaria superiore. Non a caso, controllando per la variabile costituita dall'indirizzo dell'istituto, che si può considerare, entro certi limiti, come una *proxy* del livello di capacità degli studenti, la dimensione dell'effetto muta (vedi paragrafo 5 e nota 26). Ancora una volta, indagini più specifiche svolte a livello locale, che, ad esempio, valutassero le competenze degli studenti all'uscita dalla scuola media, prima di iniziare la scuola superiore, potrebbero aiutare a chiarire meglio il quadro e a far emergere il contributo di ciascuno di questi due fattori e la loro relazione reciproca.

Un terzo e ultimo punto da sottolineare è l'insuccesso nell'individuazione di variabili di scuola, diverse dalle caratteristiche degli studenti reclutati da un istituto, che abbiano una relazione positiva con la competenza in lettura. Le uniche due variabili di questo genere con un effetto significativo hanno con il rendimento nella prova di comprensione un rapporto inverso. Del tutto controintuitiva appare, in particolare, la relazione negativa che si osserva tra il numero di ore settimanali di lezione della lingua del test e il punteggio in lettura. Per tentare di avanzare qualche ipotesi di spiegazione, oltre a sottolineare il fatto che il numero di ore dedicato dalle scuole all'insegnamento della lingua del test non è un dato obiettivo ma è ottenuto dall'aggregazione delle risposte al questionario-studente, bisogna anche considerare che la comprensione della lettura, a differenza della competenza matematica o scientifica, non corrisponde a un contenuto disciplinare specifico, ma è un'abilità in qualche modo trasversale, il cui sviluppo è influenzato da fattori soggettivi ed extra-scolastici in misura maggiore di altre abilità o competenze. Alla luce dei risultati della nostra analisi, si potrebbe affermare che lo sviluppo di questa capacità è governato assai più dal "principio del piacere" che non da altre motivazioni su cui la scuola è in grado di far leva e che quello che essa può fare – anche se ci rendiamo conto, nell'atto in cui lo formuliamo, da una parte dell'ovvietà del suggerimento e dall'altra del fatto che la sua messa in pratica sfugge a dettami e regole precostituite – è di cercare di creare un ambiente e un contesto in cui il gusto della lettura abbia la possibilità di nascere e crescere.

APPENDICE 1

Descrizione variabile	Nome variabile	Media	Dev.St	Min	Max	%Val. imp.
LIVELLO 1						
Indice di status socio-economico-culturale	ESCS	-0,18	0,87	-2,82	3,02	0
Indice di status s.e.c. al quadrato	ESCS ²	0,79	1,05	0,00	1,00	0
Lo studente è femmina	FEMMINA	0,50	0,50	0,00	1,00	0
Lo studente è d'origine immigrata	IMMIG	0,06	0,23	0,00	1,00	1
Lo studente parla a casa l'italiano	LINGCASA	0,22	0,42	0,00	1,00	14
La famiglia dello studente ha più di tre figli	NUM_FIGLI	0,10	0,30	0,00	1,00	13
Lo studente non ha frequentato la scuola dell'infanzia	FREQ_SI	0,13	0,34	0,00	1,00	1
Lo studente è in ritardo	RITARDO	0,28	0,45	0,00	1,00	0
Voto nella lingua del test (standardizzato)	VOTO_LT	0,00	1,00	-6,79	2,94	10
Titolo massimo che lo studente si aspetta di conseguire	TITOLO	0,76	1,56	-3,00	3,00	0
Atteggiamento verso la scuola	ATSCHL	-0,06	0,93	-2,99	2,01	5
Uso di strategie basate su comprendere e ricordare	UNDREM	0,26	0,95	-1,88	1,42	2
Uso di strategie basate sul riassumere	METASUM	0,29	0,89	-2,01	1,34	2
Per studiare lo studente impara a memoria	MEMOR	0,02	0,90	-3,02	2,69	0
Per studiare lo studente elabora le informazioni	ELAB	0,04	0,99	-2,41	2,76	0
Per studiare lo studente usa strategie di controllo	CSTRAT	0,15	0,95	-3,45	2,50	0
Accesso alle tecnologie dell'informazione e com. (TIC)	ICTRES	-0,01	0,95	-3,03	1,54	0
Disponibilità delle TIC a casa	ICTHOME	0,22	0,83	-4,19	1,41	1
Uso del computer per fare i compiti a casa	HOMSCH	-0,29	0,93	-1,92	3,04	2
Fiducia nel saper usare il computer in compiti d'alto liv.	HIGHCONF	0,05	0,90	-3,59	1,62	2
Uso del computer per divertirsi	ENTUSE	-0,32	1,10	-3,10	2,99	1
Atteggiamento verso il computer	ATTCOMP	0,25	0,82	-2,44	0,86	2
Lo studente legge per piacere almeno 30' al giorno	TEMPO_LET1	0,25	0,43	0,00	1,00	0
Lo studente legge per piacere più di 30' al giorno	TEMPO_LET2	0,29	0,45	0,00	1,00	0
Frequenzazione biblioteche interne alla scuola o esterne	LIBUSE	-0,10	0,89	-1,85	3,99	0
Piacere tratto dalla lettura	JOYREAD	-0,04	1,15	-3,23	3,50	1
Varietà delle letture fatte regolarmente	DIVREAD	-0,06	0,91	-3,95	4,65	0
Lettura on-line	ONLNREAD	-0,24	1,00	-5,35	3,51	0
Sostegno dei genitori ai figli nell'imparare a leggere	PRESUPP	-0,30	0,81	-3,58	2,75	13
Motivazione personale dei genitori alla lettura	MOTREAD	0,30	1,04	-2,96	1,60	13
Disponibilità risorse per la lettura in casa	READRES	0,35	0,91	-3,00	1,43	13
Sostegno attuale ai figli per l'impegno nella lettura	CURSUPP	-0,26	0,86	-3,95	3,80	13
LIVELLO 2						
La scuola si trova in un centro con più di 15.000 ab.	UBIC	0,41	0,49	0,00	1,00	0
La scuola deve competere con almeno un'altra scuola	COMPET	0,40	0,49	0,00	1,00	0
La scuola è di lingua italiana	LINGSC	0,31	0,46	0,00	1,00	0

Descrizione variabile	Nome variabile	Media	Dev.St	Min	Max	%Val. imp.
La scuola è un Istituto Tecnico	<i>TECNICO</i>	0,26	0,44	0,00	1,00	0
La scuola è un Istituto Professionale	<i>PROFESS</i>	0,14	0,34	0,00	1,00	0
La scuola è un CFP	<i>CFP</i>	0,27	0,45	0,00	1,00	0
Indice medio di status s.e.c. della scuola	<u>SCHESCS</u>	-0,20	0,43	-1,22	0,95	0
Dimensione della scuola (numero studenti/100)	SCHSIZE	3,36	2,50	0,22	14,59	0
Dimensione della scuola al quadrato	SCHSIZE ²	17,45	28,03	0,05	212,9	0
Percentuale di ragazze (%/10)	PCGIRLS	5,02	2,73	0,06	10,00	0
Percentuale di studenti d'origine immigrata	PCIMMIG	0,81	1,42	0,00	9,09	0
Numero studenti per insegnante	STRATIO	7,04	2,48	2,23	11,90	2
Gli insegnanti sono carenti quali-quantitativamente	TCSHORT	0,27	0,86	-1,02	1,90	0
Proporzione computer a uso didattico	IRATCOMP	1,1	0,75	0,10	4,23	2
Qualità delle attrezzature per l'insegnamento	SCMATEDU	0,67	0,94	-1,35	1,93	0
Disponibilità delle TIC a scuola	<u>ICTSCH</u>	0,25	0,37	-1,06	1,03	0
Uso delle TIC a scuola	<u>USESCH</u>	0,32	0,37	-0,70	1,23	0
Partecipazione degli insegnanti	TCHPARTI	0,36	0,59	-2,07	1,57	0
Leadership del capo d'istituto	LDRSHP	-0,29	0,78	-2,08	2,29	1
Offerta di attività extracurricolari	EXCURACT	0,04	0,85	-1,55	2,35	0
Comportamenti degli insegnanti	TEACBEHA	-0,27	0,74	-1,83	1,40	0
Comportamenti degli studenti	STUDBEHA	-0,19	0,81	-1,67	2,36	0
Relazioni fra studenti e insegnanti	<u>STUDREL</u>	0,01	0,37	-1,22	1,14	0
Clima disciplinare	<u>DISCLIMA</u>	0,11	0,55	-1,81	1,53	0
Pressione da alcuni o molti genitori per alti risultati	<i>PRESSGE</i>	0,54	0,50	0,00	1,00	0
Percezione da parte dei genitori della qualità della scuola	<u>PQSCHOOL</u>	0,31	0,44	-0,97	1,53	0
Coinvolgimento dei genitori nella vita della scuola	<u>PARINVOL</u>	-0,17	0,41	-0,90	1,04	0
Media ore settimanali insegnamento della lingua del test	<u>LORESET</u>	3,67	0,70	2,50	7,25	0
Sostegno degli insegnanti all'impegno nella lettura	<u>STIMREAD</u>	-0,21	0,31	-1,08	0,52	0
Qualità del lavoro in classe sulla lettura di testi	<u>STRSTRAT</u>	-0,09	0,25	-0,67	0,49	0
Lettura per la scuola di testi narrativi	<u>RFSINTRP</u>	-0,07	0,35	-0,94	0,96	0
Lettura per la scuola di testi non continui	<u>RFSCONT</u>	-0,34	0,33	-1,20	0,89	0
Lettura per la scuola di testi letterari e di storia d. letter.	<u>RFSTRLIT</u>	-0,20	0,43	-1,14	1,05	0
Lettura per la scuola di testi funzionali	<u>RFSFUMAT</u>	-0,08	0,39	-1,23	1,42	0

Nota: le variabili in corsivo sono variabili categoriali dicotomiche (*dummy*). La condizione cui si fa riferimento nella descrizione di ciascuna di esse è codificata con "1" (con "0" la condizione opposta), per cui la media della variabile corrisponde alla proporzione di soggetti che rientrano nella categoria 1. Le variabili sottolineate sono ottenute per aggregazione dal questionario-studente o dal questionario-genitori.

APPENDICE 2

1) Equazioni generali dei modelli con sole variabili di primo livello:

Livello 1:

$$Y_{ij} = \beta_{0j} + \beta_{1j}(X_{1ij} - \bar{X}_{..}) + \dots + \beta_{kj}(X_{kij} - \bar{X}_{..}) + e_{ij}$$

Livello 2:

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

2) Equazioni generali dei modelli con sole variabili di secondo livello:

Livello 1:

$$Y_{ij} = \beta_{0j} + e_{ij}$$

Livello 2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Z_{1j} + \dots + \gamma_{0h}Z_{hj} + u_{0j}$$

3) Equazioni generali dei modelli con variabili di primo e secondo livello:

Livello 1:

$$Y_{ij} = \beta_{0j} + \beta_{1j}(X_{1ij} - \bar{X}_{..}) + \dots + \beta_{kj}(X_{kij} - \bar{X}_{..}) + e_{ij}$$

Livello 2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Z_{1j} + \dots + \gamma_{0h}Z_{hj} + u_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

Riferimenti bibliografici/Literaturhinweise

- Boletta, R; Pozio, S. (2005): La competenza matematica dei quindicenni. In: Il livello di competenza dei quindicenni in matematica, lettura, scienze e problem solving. PISA 2003 – Risultati dell'Alto Adige, A cura di M.T. Siniscalco. Bolzano, Istituto pedagogico per il gruppo linguistico tedesco 2005.
- Bratti, M., Checchi, D. e Filippin, A. (2008): Da dove vengono le competenze degli studenti? I divari territoriali nell'indagine OCSE PISA 2003, Bologna: Il Mulino.
- Bryk, A. S., Raudenbush, S. W. (2002), *Hierarchical Linear Models: application and data analysis*, (2nd Edition) Thousand Oaks, CA, Sage.
- Casteel, C.P., & Isom, B.A. (1994). Reciprocal processes in science and literacy learning. *Reading Teacher*, 47(7), 538-545.
- Does participation in pre-primary education translate into better learning outcomes at school? – OECD: PISA in focus 2011/I
- Ertl H. (2006): "Educational standards and the changing discourse on education: the reception and consequences of the PISA Study in Germany". *Oxford Review of Education*, vol. 32 n. 5 pp. 619-634.
- Fandiño Pinilla M. I. (2006): Educare alla competenza matematica. In: *Rassegna. Periodico dell'Istituto Pedagogico (Bolzano)*, n. 29, p. 21-28
- Guthrie J. T. e Wigfield A. (2000), "Engagement and Motivation in Reading", in Kamil M. L. e Mosenthal P. B. (a cura di), *Handbook of reading research*, Mahwah, NJ, Erlbaum, 2000.
- Hox, J. J. (2002), *Multilevel Analysis: Techniques and Applications*, Lawrence Erlbaum Associates, Hillsdale, NJ.
- INVALSI (2011): Le competenze in lettura, matematica e scienze degli studenti quindicenni italiani. Rapporto nazionale. PISA 2009. Frascati: INVALSI (http://www.invalsi.it/invalsi/ri/Pisa2009/documenti/RAPPORTO_PISA_2009.pdf - versione del 01.04.2011)
- Kirsch I., de Jong J., Lafontaine D., Mc Queen J., Mendelovits J., Monseur C. (2002), *Reading for change. Performance and engagement across countries. Results from PISA 2000*. Paris, OECD.
- Klieme, E.; Artelt, C.; Hartig, J.; Jude, N.; Köller, O.; Prenzel, M.; Schneider, W.; Stanat, P. (2010) (Hrsg.). *PISA 2009. Bilanz nach einem Jahrzehnt*. Münster/New York/München/ Berlin: Waxmann
- Magno, F. (2009): Der Ansatz zur Beurteilung der naturwissenschaftlichen Kompetenz in PISA 2006. In: *Das Kompetenzniveau der Fünfzehnjährigen im Bereich der Naturwissenschaften, des Lesens und der Mathematik. PISA 2006 – Ergebnisse Südtirols*. Hrsg. von M.T. Siniscalco u. R. Meraner. Bozen, Pädagogisches Institut., S. 25-48 (<http://www.schule.suedtirol.it/pi/themen/documents/pisa2006/PISA2006deutsch.pdf>)
- OECD (1999), *Measuring Student Knowledge and Skills – A New Framework for Assessment*. Paris: OECD
- OECD (2001), *Knowledge and Skills for Life. First results from PISA 2000*, OECD, Paris.
- OECD (2002), *PISA 2000 Technical Report*, R. Adams e M. Wu (eds.), Paris, OECD.
- OECD (2004), *Learning for tomorrow's world. First results from PISA 2003*, Paris, OECD.
- OECD (2006), *Assessing scientific, reading and mathematical literacy. A framework for PISA 2006* OECD. Paris (trad. it.: *PISA 2006. Valutare le competenze in scienze, lettura e matematica. Quadro di riferimento di PISA 2006*, Roma, Armando, 2007).
- OECD (2007a): *PISA 2006. Sciences competencies for tomorrow's world. Volume 1 – Analysis*. Paris: OECD. (http://www.pisa.oecd.org/document/2/0,3343,en_32252351_32236191_39718850_1_1_1_1,00.html).

- OECD (2007b), PISA 2006. Volume 2 – Data / Données. Paris: OECD
(http://www.pisa.oecd.org/document/2/0,3343,en_32252351_32236191_39718850_1_1_1_1,00.html).
- OECD (2007c): Pisa 2006: Schülerleistungen im internationalen Vergleich. Naturwissenschaftliche Kompetenzen für die Welt von morgen. Paris: OECD
- OECD (2008): PISA Technical Report: Second Edition. Paris: OECD
- OECD (2009), PISA 2009 assessment framework. Key competencies in reading, mathematics and science. Paris, OECD.
(<http://www.oecd.org/dataoecd/11/40/44455820.pdf>)
- OECD (2010): Pathways to success: how knowledge and skills at age 15 shape future lives in Canada. Paris: OECD Publishing.
- OECD (2010a): PISA 2009 results: what students know and can do – student performance in reading, mathematics and science. Volume I. Paris: OECD (<http://dx.doi.org/10.1787/9789264091450-en>)
- OECD (2010a-DE): PISA 2009 Ergebnisse: Was Schülerinnen und Schüler wissen und können. Schülerleistungen in Lesekompetenz, Mathematik und Naturwissenschaften. Bd. 1. Paris: OECD
- OECD (2010b), PISA 2009 Results: Overcoming Social Background: Equity in Learning Opportunities and Outcomes. Volume II, Paris: OECD Publishing. (<http://dx.doi.org/10.1787/9789264091504-en>)
- OECD (2010c), PISA 2009 Results: Learning to Learn: Student Engagement, Strategies and Practices. Volume III. Paris: OECD Publishing. (<http://dx.doi.org/10.1787/9789264083943-en>)
- OECD (2010d), PISA 2009 Results: What Makes a School Successful? Resources, Policies and Practices. Volume IV, Paris: OECD Publishing. (<http://dx.doi.org/10.1787/9789264091559-en>)
- OECD (2010e), PISA 2009 Results: Learning Trends: Changes in Student Performance Since 2000. Volume V. Paris: OECD Publishing. (<http://dx.doi.org/10.1787/9789264091580-en>)
- OECD e Statistics Canada (2000), Literacy in the Information Age: Final Report of the International Adult Literacy Survey, Paris, OECD.
- Olson, D. R. (1994): The world on paper: the conceptual and cognitive implications of writing and reading. Cambridge (UK), New York (USA), Melbourne (Aus): Cambridge University Press
- Paccagnella O. (2006), Centering or not centering in multilevel models? The role of the group mean and the assessment of group effects, Evaluation Review, Bd. 30 Nr.1 66-85
- Schwantner, U; Schreiner, C. (2010)a (Hrsg.): PISA 2009 Internationaler Vergleich von Schülerleistungen. Die Studie im Überblick. Graz: Leykam 2010
Download unter: <https://www.bifie.at/buch/1279>
- Schwantner, U; Schreiner, C. (2010b) (Hrsg.): PISA 2009 Internationaler Vergleich von Schülerleistungen. Erste Ergebnisse Lesen, Mathematik, Naturwissenschaft. Graz: Leykam 2010
(Download unter <https://www.bifie.at/buch/1249>)
- Servizio Nazionale di Valutazione (2010) Rilevazione degli apprendimenti – SNV. Prime Analisi. A.S.2009/2010. INVALSI.
- Siniscalco M. T. (a cura di) (2005a), Il livello di competenza dei quindicenni in matematica, lettura, scienze e problem solving. PISA 2003 – Risultati dell’Alto Adige, Bolzano, Istituto pedagogico per il gruppo linguistico tedesco. (versione italiana)
- Siniscalco M. T. (a cura di) (2005b), Das Kompetenzniveau der Fünfzehnjährigen im Bereich der Mathematik, des Lesens, der Naturwissenschaften und des Problemlösens. PISA 2003 – Ergebnisse Südtirol, Bozen, Pädagogisches Institut. (versione tedesca)
- Siniscalco M. T. (2008) “La lettura nelle indagini internazionali e l’approccio OCSE-PISA”, in Siniscalco M. T. Bolletta R., Mayer M. Pozio S. Le valutazioni internazionali e la scuola italiana, Bologna, Zanichelli, 41-177.
- Siniscalco M. T./Meraner, Rudolf (a cura di) (2009a): Il livello di competenza dei quindicenni in scienze, lettura e matematica. PISA 2006 – Risultati dell’Alto Adige. Bolzano, Istituto pedagogico per il gruppo linguistico tedesco. (versione italiana).

<http://www.schule.suedtirol.it/pi/themen/documents/pisa2006/PISA2006italiano.pdf>

Siniscalco M. T./Meraner, Rudolf (a cura di) (2009b): Das Kompetenzniveau der Fünfzehnjährigen im Bereich der Naturwissenschaften, des Lesens und der Mathematik. PISA 2006 – Ergebnisse Südtirols. Bozen, Pädagogisches Institut. (versione tedesca)

<http://www.schule.suedtirol.it/pi/themen/documents/pisa2006/PISA2006deutsch.pdf>

Wellington, J. J.; Osborne, J. (2001): Language and literacy in science education. Bucks (UK):Open University Press

Willms J. D. (2002) Vulnerable children: findings from Canada's National Longitudinal Survey of Children and Youth, University of Alberta Press, Edmonton.