

## **24° Congresso internazionale del CIRIEC**

*Napoli Italia 30 settembre - 2 ottobre 2002*

### **PISA: L'INDICATORE PENDENTE? DIFFICOLTA' DI INTERPRETAZIONE DEL PROGRAMMA OCSE PER LA VALUTAZIONE DELLE CAPACITA' DEGLI STUDENTI DI 15 ANNI**

**Stefania Gabriele**

ISAE – Istituto di Studi e Analisi Economica

**Stefano Zolea**

Regione Campania

(Italia)

*Comunicazione libera*

Il "Programme for International Student Assessment", PISA nell'acronimo, è uno studio molto ampio, finalizzato principalmente a valutare, in prospettiva comparata tra paesi, le capacità di comprensione di un testo e la dimestichezza con concetti matematici e scientifici dei 15enni che frequentano la scuola. L'indagine è rivolta anche alle possibili cause delle differenze riscontrate tra i giovani.

Il programma ha fatto uso di consistenti risorse e si pone grandi ambizioni, in particolare vuole rappresentare un punto di riferimento importante per chi sarà chiamato a intervenire o a disegnare le politiche per la scuola. Il primo capitolo del rapporto "Knowledge and Skills for Life. First results from PISA 2000" afferma che il rigore con cui è stato condotto il lavoro, il coinvolgimento di esperti, i controlli sulla qualità delle traduzioni, sul campionamento e sulla raccolta dei dati assicurerebbero un "elevato livello di validità e affidabilità".

La presente comunicazione è rivolta a evidenziare alcuni problemi sottostanti al tentativo di confrontare le abilità di soggetti appartenenti a comunità di lingua e cultura diversa. I dubbi che emergono suggeriscono che gli indicatori calcolati dal PISA vanno considerati con una certa prudenza, soprattutto qualora si voglia ragionare sulle politiche da intraprendere.

## **LE GRADUATORIE DEI PAESI EMERSE DAL PROGETTO PISA**

Il progetto PISA è stato realizzato per la prima volta nel 2000 e ha coinvolto 32 paesi, di cui 28 appartenenti all'OCSE (in seguito è stato esteso ad altre nazioni, ma i risultati non sono ancora disponibili). Sono stati intervistati più di 250mila studenti (su circa 17 milioni di allievi 15enni). I campioni nazionali coprono tra il 95% e il 97% della popolazione scolastica 15enne<sup>1</sup>, se si eccettuano tre paesi (Brasile, Lussemburgo e Polonia).

Obiettivo dichiarato del programma OCSE per la valutazione internazionale degli studenti è verificare fino a che punto i ragazzi siano in grado di usare il sapere e le abilità acquisite per affrontare le sfide della vita reale, non tanto quanto abbiano assimilato determinati programmi di studio. Il progetto è rivolto a considerare la comprensione dei concetti, la padronanza dei processi e la capacità di far fronte a diverse situazioni. Questa concezione ampia del sapere e delle abilità è sintetizzata nel termine "literacy", difficilmente traducibile in italiano, ma che fa riferimento ad un concetto più vasto di quello che si intende con la parola "alfabetizzazione".

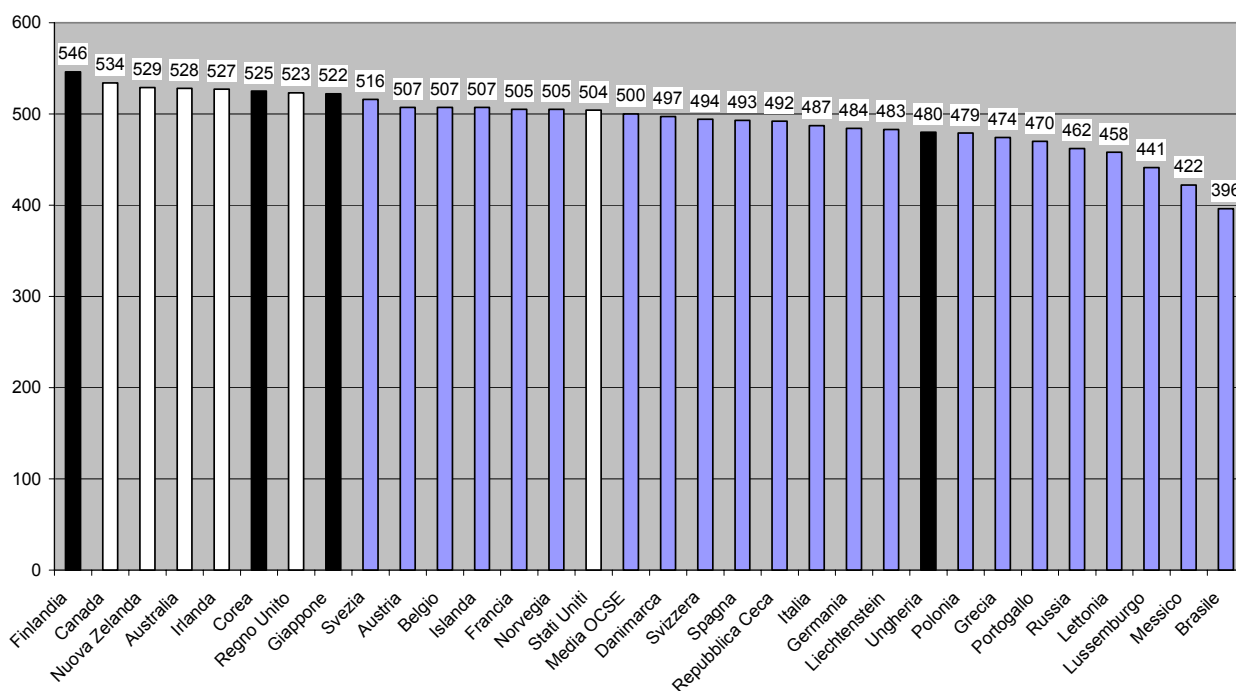
---

<sup>1</sup> Per l'esattezza, sono stati intervistati giovani di età compresa tra 15 anni e tre mesi e 16 anni e due mesi.

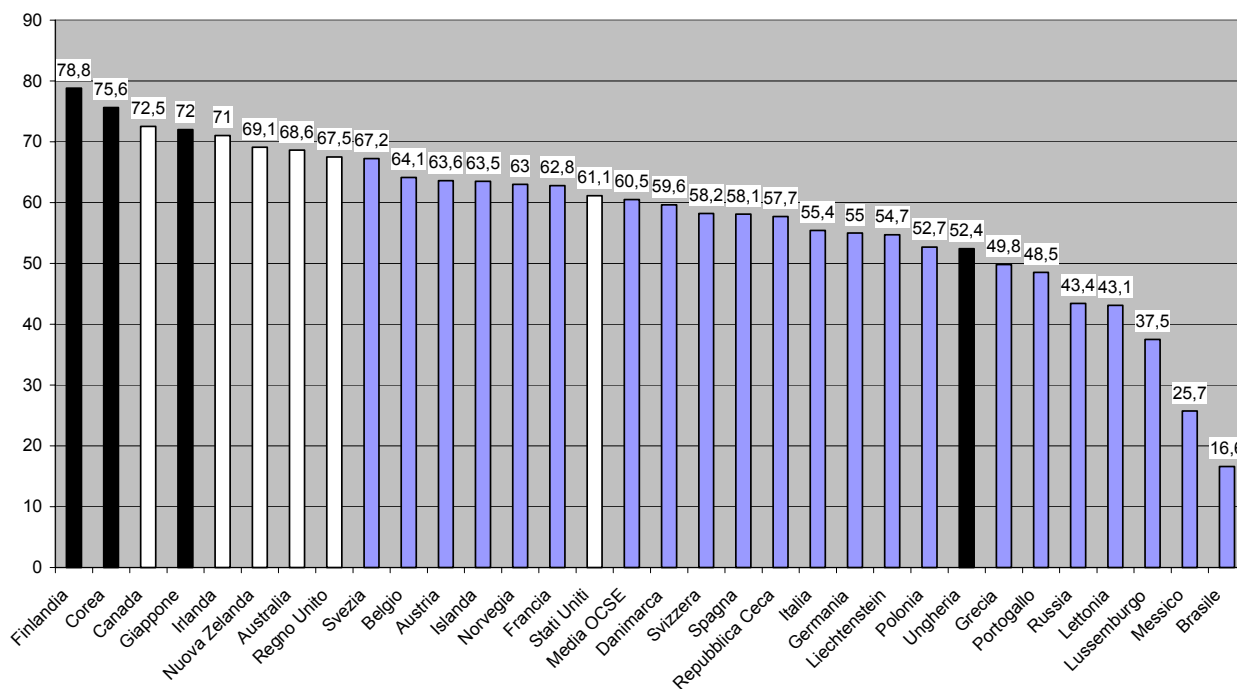
Le valutazioni sono concentrate su tre materie, innanzitutto la lettura, un aspetto cui è stata attribuita un'attenzione particolare in questa prima inchiesta, e poi la matematica e le scienze. Queste ultime due saranno analizzate più attentamente in occasione delle prossime indagini, che verranno realizzate rispettivamente nel 2003 e nel 2006.

Il grafico 1 mostra la graduatoria dei paesi esaminati in base al punteggio medio raggiunto dagli studenti sottoposti alla prova di lettura. La media di tutti i paesi OCSE è indicata in 500 punti. Il grafico 2 presenta una classifica costruita sulla base della percentuale di studenti che raggiungono almeno il terzo livello di competenza, sui cinque individuati nello studio (che corrisponde ad un punteggio compreso tra 481 e 552 punti). I grafici 3 e 4 riportano le graduatorie dei valori medi rispettivamente per matematica e scienze. In tutti questi grafici le barre che rappresentano i paesi sono di colore bianco quando la lingua parlata è l'inglese, grigio-azzurro per le altre lingue indoeuropee, nero per quelle non indoeuropee. Prima di spiegare perché abbiamo evidenziato le caratteristiche linguistiche, richiamiamo brevemente le linee fondamentali di impostazione del progetto PISA.

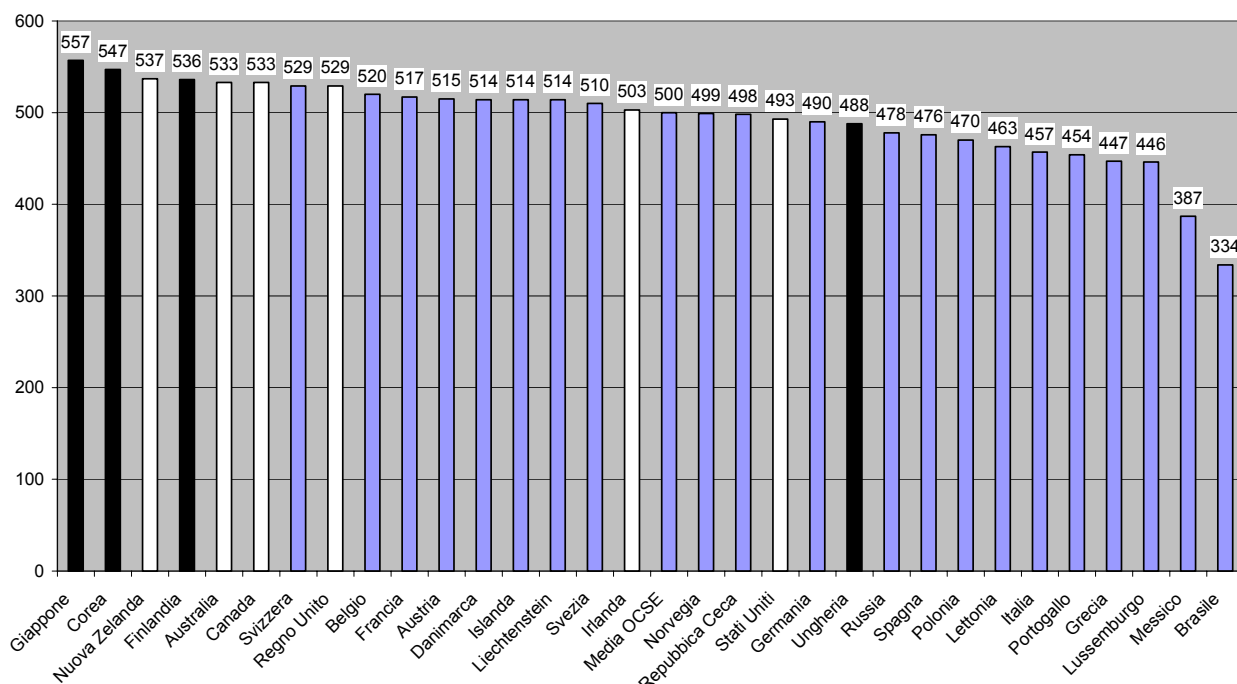
**Graf. 1 - Lettura (graduatoria dei valori medi)**



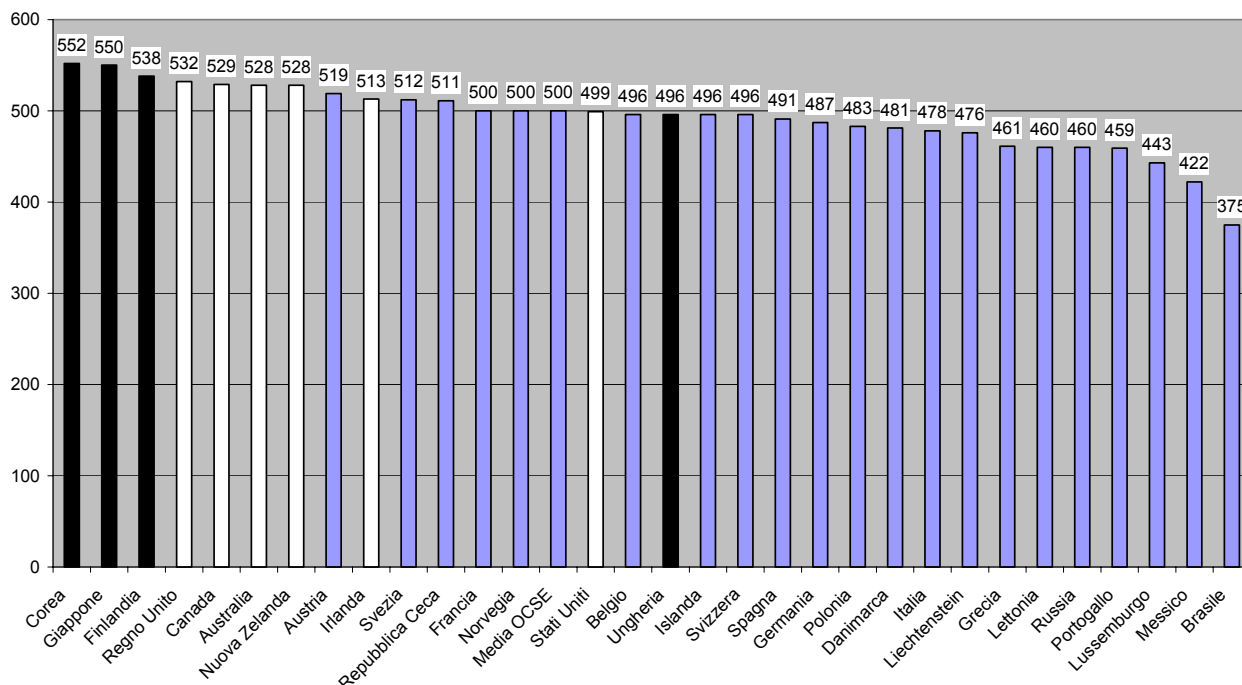
Graf. 2 - Lettura (% di studenti che raggiungono almeno il terzo livello)



Graf. 3 - Matematica (graduatoria dei valori medi)



Graf 4 - Scienze (graduatoria dei valori medi)



## LE CARATTERISTICHE DEI TEST

L'esame delle capacità dei ragazzi è stato realizzato attraverso la somministrazione di alcuni test scritti e sono state concesse due ore di tempo per rispondere alle domande. In effetti sono stati preparate complessivamente 140 prove, per un totale di sette ore di lavoro, ma a diversi allievi sono state proposte diverse combinazioni di test. Gli item sono stati rivisti da specialisti nazionali nelle materie e da esperti di valutazione. I Paesi coinvolti hanno convalidato l'applicabilità culturale, la coerenza con i programmi di studio e l'adeguamento del grado di difficoltà. Le prove sono state effettuate nelle scuole.

I test si presentano in taluni casi in forma continua (e tra questi si distinguono i testi narrativi, descrittivi e argomentativi), ma a volte includono diagrammi, grafici, moduli, elenchi; essi richiedono dunque l'elasticità necessaria a saper trarre informazioni da fonti con caratteristiche diverse. Non si mira a verificare le competenze elementari legate alla capacità di leggere, che dovrebbero essere state assimilate molto prima dei 15 anni, ma piuttosto quanto il giovane sia in grado di reperire le informazioni, di comprendere un testo, di interpretarlo, di riflettere su quanto letto, di valutare e argomentare la propria opinione. Inoltre, si richiede di individuare il contesto a cui il testo si riferisce, distinguendo documenti di tipo e con finalità diverse (una lettera privata o un documento ufficiale, eccetera). Quanto alla

matematica, si intende misurare la capacità di analizzare, ragionare e comunicare le proprie idee sapendo porre, formulare e risolvere problemi matematici in contesti diversi. Ai fini della valutazione, sono state individuate tre grandi dimensioni: riguardo al contenuto della matematica, sono stati privilegiati i concetti delle variazioni e relazioni e degli spazi e le forme; quanto ai processi, ci si riferisce alla riproduzione di semplici calcoli o definizioni di base, all'individuazione delle procedure matematiche per risolvere problemi facili, alla riflessione, ovvero al ragionamento matematico, alla generalizzazione e alla comprensione profonda; infine, per l'ultima dimensione considerata, che riguarda le situazioni in cui la matematica viene utilizzata, sono stati scelti problemi simili a quelli che si incontrano nella vita reale. La valutazione della cultura scientifica parte da una definizione di quest'ultima come capacità di utilizzare le conoscenze scientifiche, di identificare i problemi e trarre conclusioni basate sull'evidenza, per comprendere il mondo naturale e i cambiamenti apportati dall'uomo e adottare decisioni a questo proposito. Tre dimensioni sono considerate anche per le scienze: i concetti scientifici, con riferimento alla fisica, alla chimica, alla biologia, alle scienze naturali e spaziali, e con particolare riguardo ai temi della biodiversità, delle forze e del movimento e del cambiamento fisiologico; il processo scientifico, cioè la capacità di constatare, interpretare e utilizzare i fatti, e in particolare l'identificazione del problema, il reperimento di elementi di prova, il raggiungimento di conclusioni, la comunicazione delle stesse e la manifestazione della comprensione dei concetti scientifici; le situazioni scientifiche, ovvero i campi di applicazione, che sono costituiti dalle esperienze della vita reale, e tra cui sono state scelti esempi tratti dalle scienze della vita e della sanità, della Terra e dell'ambiente, della tecnologia.

Oltre ai test, è stato somministrato un questionario con domande riguardanti le caratteristiche degli allievi e dei loro istituti scolastici. In alcuni paesi sono state proposte anche domande relative all'atteggiamento di fronte allo studio e ai metodi di apprendimento adottati, nonché alla familiarità con l'informatica. Un questionario è stato sottoposto anche ai capi di istituto per approfondire alcuni aspetti relativi al funzionamento delle scuole. Le informazioni raccolte sono state elaborate al fine di analizzare le relazioni tra i risultati dei test e le altre variabili dedotte dai questionari, come ad esempio il background familiare, il sesso, il metodo di studio preferito.

## **GLI OBIETTIVI “POLITICI” DEL PISA**

Come sopra accennato, il programma PISA è stato realizzato con l'ambizione di fornire un importante strumento politico: il rapporto di presentazione dei primi risultati, sopra citato, dichiara che l'obiettivo è consentire ai governi di trarre insegnamenti politici dalle evidenze emerse dall'indagine. E' pur vero che i diversi rapporti scaturiti dal progetto esprimono qua e là alcune cautele, e invitano alla prudenza sotto diversi aspetti. Si ammette ad esempio che il punteggio di un paese

superiore a quello di un altro paese non implica automaticamente che il primo goda di un sistema di insegnamento migliore. Al contrario, viene affermato che l'influenza di tutte le esperienze di apprendimento vissute dentro e fuori la scuola, prima e dopo l'età scolare, produce i risultati diversi che emergono dal PISA. Tuttavia, si ritiene che le analisi scaturite dall'indagine possano "orientare le decisioni politiche e l'allocazione delle risorse", perché permettono di individuare i fattori che contribuiscono allo sviluppo delle competenze. Le informazioni prodotte potrebbero essere utilizzate dagli insegnanti e dai decisori politici per identificare i punti deboli e quelli forti del proprio sistema, anche grazie all'analisi delle situazioni individuali degli studenti (di genere, familiari, eccetera), del loro atteggiamento di fronte allo studio (interesse per le materie, propensione alla lettura, eccetera), delle caratteristiche delle scuole. Si sostiene inoltre che il PISA, essendo mirato alle competenze utili nella vita di tutti i giorni, non offre indicazioni limitate all'efficienza interna dei singoli sistemi scolastici, ma piuttosto risultati comparabili tra i paesi e tra le culture. La scelta di condensare i risultati dei test sottoposti agli studenti in graduatorie dei paesi (riferite alla lettura, alla matematica e alle scienze) appare tuttavia fuorviante, e può - sia pure involontariamente - suggerire un'interpretazione delle classifiche riportate come se fossero applicabili direttamente all'efficacia dei sistemi scolastici, mentre esse costituiscono piuttosto, come si è visto, il risultato di un insieme complesso di fattori sociali, culturali, economici che agiscono sulle capacità di apprendimento.

Va insomma certamente condiviso che l'ampiezza del materiale messo a disposizione, la disponibilità di rapporti, dati e informazioni, l'ampiezza delle analisi proposte rende il PISA uno strumento prezioso per chiunque si voglia occupare dei temi da esso trattati<sup>2</sup>. E' tuttavia l'ambizione "politica" che sollecita una verifica a tutto campo per constatare se i fondamenti del progetto siano assolutamente condivisibili. Alcuni osservatori stanno ragionando sulle metodologie statistiche utilizzate sia dal PISA, sia da precedenti studi di confronto internazionale sul "prodotto" dell'istruzione, e suggeriscono di adottare cautela nell'interpretazione dei risultati (Micklewright J. (2002), *Measuring Functional Literacy and Numeracy in the European Union*, *Politica Economica*, a. XVIII, n.1, aprile).

Ma una difficoltà può essere evidenziata anche da un altro punto di vista, ovvero ci si può chiedere se una valutazione comparata dei risultati dell'apprendimento possa essere rigorosamente effettuata, e se essa possa essere considerata affidabile al punto da pensare di tarare su di essa gli interventi di riforma che i governi volessero intraprendere.

---

<sup>2</sup> Lascia qualche perplessità, al contrario, l'assenza di informazioni sui risultati dell'indagine svolta in Italia: non solo il rapporto-paese non è ancora disponibile, ma non vengono neppure forniti dal Ministero dell'Istruzione, su richiesta, esempi dei test in italiano, malgrado molti esempi siano disponibili in altre lingue.

## SULLA TRADUZIONE

Le teorie linguistiche più recenti sostengono la sostanziale neutralità delle lingue rispetto all'espressione del pensiero. Contrariamente alle impostazioni dominanti fino a qualche decennio fa i più sostengono che " **quasi** ogni cosa che possiamo dire in una data lingua può essere tradotta in qualunque altra lingua, mantenendo immutato il pensiero veicolato nella lingua data. (Jackendoff, R. (1990), Linguaggio e natura umana, Bologna, p.250)".

Insomma, emerge un vasto consenso negli specialisti "nel ritenere che le differenze di pensiero suscettibili di essere condizionate da differenza di struttura grammaticale siano **abbastanza** superficiali" (ibid.).

E' d'altronde noto che talvolta anche un ottimo traduttore di settore (p.e. un traduttore scientifico o un traduttore letterario) incontra in alcune traduzioni di testi dei forti limiti di efficacia nel lavoro traspositivo. Addirittura, tale disagio si può verificare nella stessa persona, come è noto essere avvenuto in Nabokof, brillante scrittore di lingua inglese che "si ostinava a scrivere sempre con l'aiuto di dizionari e di grammatiche; spiegava di se stesso -Penso come un genio, scrivo come un grande scrittore e parlo come un bambino-" (Pinker, S. (1997), L'istinto del linguaggio, Milano, p.283).

Bisogna inoltre tenere conto della contraddittorietà dei risultati scientifici ottenuti finora sulla rilevazione di sottili differenze di efficienza di apprendimento e di ragionamento in conseguenza della diversità strutturale tra gli alfabeti fonetici occidentali e quelli dell'estremo oriente; e come non citare pure la *vexata quaestio* dell'influenza delle differenze delle aree corticali del cervello tra uomini e donne.

Anche nei più tenaci sostenitori chomskiani dell'unicità e dell'universalità della grammatica mentale comunque emergono dei dubbi sull'assoluta sovrapposibilità delle varie lingue (nella letteratura consultata si trovano spesso espressioni del tipo: emergono differenze "**abbastanza**" superficiali tra le lingue, **pressoché** ogni forma di pensiero è trasponibile da una lingua all'altra" ecc.). Chiunque conosca la teoria dell'errore sa che un piccolo "bias" di fondo nelle basi di un percorso scientifico può moltiplicare l'errore stesso a livelli incontrollabili nell'interpretazione dei risultati finali del percorso stesso. Dato che nel caso del PISA si tratta di una quantificazione di risultati ottenuti da risposte a test tradotti in molte lingue, non si può escludere una influenza "linguistica" nelle risposte ai test.

E' innegabile che il metodo scelto dall'OCSE per l'ultima versione del PISA tende ad attenuare gli eventuali bias di traduzione, in quanto è stata spesso attuata una procedura di doppia traduzione di conferma per altre lingue a partire dall'inglese e dal francese, ulteriormente supportata da altri interventi di esperti delle varie lingue per

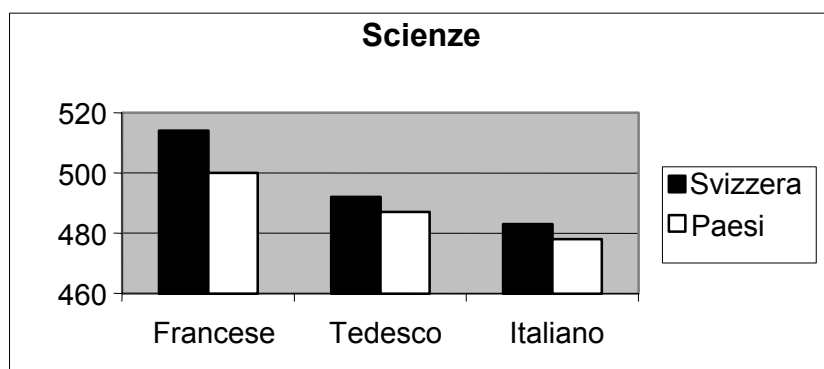
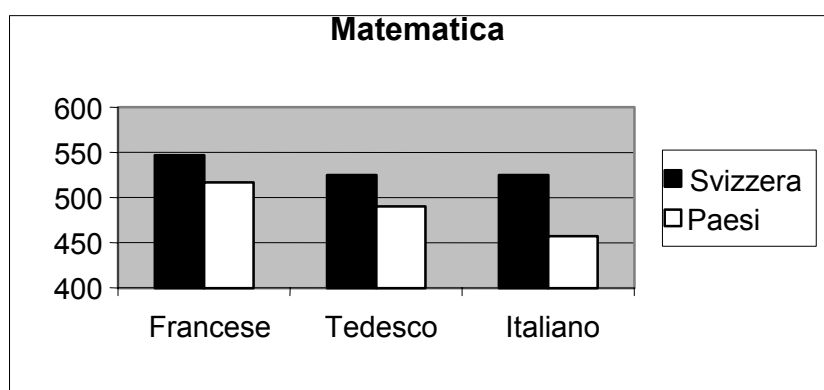
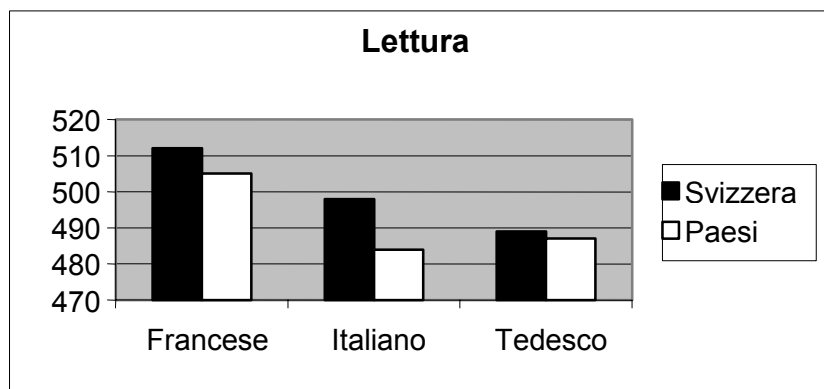


minimizzare gli errori avvenuti in indagini analoghe del passato. Tuttavia, non tutte le procedure appaiono avere avuto lo stesso rigore e rimane l'impressione di test concepiti nella mente di studiosi di lingua madre inglese (o comunque parlanti correntemente l'inglese).

Comunque nel passato severe critiche sono già state condotte da Blum et al. (Blum A., Goldstein H., Guerin-Pace F., An analysis of international comparison of adult literacy, in *Assessment in education: principles, policy and practice*, JULY 2001, vol. 8, n.2, pp.225-246) a proposito della significatività dei test di indagine sulla *literacy* tra vari paesi e tra varie lingue in base all'International Adult Literacy Survey, IALS, pubblicata in un rapporto dell'OCSE (OCDE, (1997), *Literacy skills for the knowledge society*, Paris). Da questa indagine risultava una bassa valutazione della capacità di "leggere" i giornali o di scrivere una lettera da parte dei francesi in comparazione con i britannici, gli americani o gli olandesi. Blum, studiando accuratamente le versioni inglese e francese, rilevò delle fonti di bias nelle traduzioni e una migliore precisione nei test in lingua inglese. Nell'analisi della strutturazione dei test sono emerse evidenze in merito a una possibile fonte di "rumore" nel cosiddetto cultural bias tra francese e inglese. Emergevano inoltre correlazioni statistiche con le lingue tra le risposte di nazioni che avevano comunità di lingua madre diversa (i fiamminghi del Belgio si assimilavano agli olandesi, i francofoni della Svizzera ai francesi, i tedescofoni svizzeri con i tedeschi di Germania): cosa che si è vista sussistere anche per il PISA, come mostra il grafico 5, suggerendo una non indifferenza, qualunque sia la motivazione, del bias linguistico in questo tipo di indagini.

Un ultimo particolare può suscitare interesse: l'ottima posizione in graduatoria nei vari test del PISA per ciò che concerne la Corea e il Giappone, due paesi caratterizzati da lingue profondamente differenti da quelle indoeuropee (anche la buona posizione in graduatoria dei parlanti in finnico, altra lingua non indoeuropea, è interessante), a segnalare in ogni caso la possibilità di un'influenza del bias linguistico.

**Graf. 5 - Graduatoria PISA per le lingue francese, tedesco e italiano parlate nei paesi di origine e in Svizzera**



## PERPLESSITA' SU ALCUNI TEST DI SCIENTIFIC LITERACY

La valutazione della cultura scientifica intrapresa dal PISA, come si è accennato, non intende limitarsi tanto alla stretta capacità di utilizzare le conoscenze scientifiche, quanto all'abilità a sviscerare alcune questioni e ad attualizzarle nella realtà quotidiana.

C'è da osservare che in una batteria di test cui rispondere in tempo limitato sono state pertanto escluse altre caratteristiche ritenute importanti del pensiero e dell'agire scientifico contemporaneo come lo scetticismo, la natura probabilistica di gran parte della conoscenza scientifica e la natura accumulativa talvolta contraddittoria della prassi e della sperimentazione scientifica (M. Ageno, (1980), Lezioni di Biofisica, Bologna, p. 17).

E' noto pure che le impostazioni neodarwiniane più diffuse ed in auge nella comunità scientifica possono creare e creano problemi nell'insegnamento di alcuni argomenti della biologia presso le scuole di alcuni paesi per la forte presenza di istanze creazioniste e lamarckiane. E' noto pertanto che nei test di scientific literacy vige spesso una forte autocensura nella predisposizione di questionari che si riferiscano al neo-darwinismo, creandosi così un "vulnus" sulla verifica di uno dei presidi più importanti del pensiero scientifico moderno.

Si rileva infine nel PISA un esempio di un test che fornisce valutazioni positive ad una risposta di natura politica e non di natura scientifica. E' il caso della risposta al test relativo alla diminuzione dello strato di ozono troposferico. La domanda prevede una risposta sulla base della validità del principio di precauzione, che è notoriamente un principio assolutamente politico (Cameron J. (1995), Addressing uncertainty. Law, policy and the development of the precautionary principle, in Dente, B (ed.), Environmental Policy in Search of New Instruments, Dordrecht, pp.95-142; Godard O. (1997), Le principe de précaution, Paris 1997, p.63). Emergono pertanto degli elementi di perplessità su cosa effettivamente si stia tentando di misurare con il PISA quanto alla *scientific literacy*.

