

L'uso delle domande aperte per la verifica della competenza matematica. Suggestioni da PISA 2012

The use of open-ended questions to verify the mathematical literacy. Suggestions from PISA 2012

Giorgio Asquini^a

^a *Sapienza, Università di Roma*, giorgio.asquini@uniroma1.it

Abstract

Il saggio presenta i risultati degli studenti italiani nell'indagine OECD-PISA 2012. Vengono considerati i risultati per le domande a risposta aperta relative alla literacy matematica. È stato utilizzato il database relativo al compendio degli item cognitivi, raggruppando gli item secondo le categorie matematiche del quadro di riferimento del PISA (contenuto, contesto, processo), a cui si è aggiunta la classificazione per livello di performance. La complessiva debolezza degli studenti italiani nelle domande aperte è determinata in particolare dall'alto tasso di omissioni e dai risultati molto scarsi in alcune categorie ("cambiamento e relazioni", "ambito scientifico", "formulazione matematica"). Gli studenti italiani ottengono risultati migliori in poche categorie ("contesto sociale", "interpretazione e valutazione", "spazio e forma"), che possono rappresentare un punto di partenza per il miglioramento. Viene discusso anche il possibile uso dei dati sulle domande aperte nella formazione degli insegnanti.

Parole chiave: domande aperte; competenza matematica; OECD-PISA; formazione insegnanti.

Abstract

The paper presents the results of the Italian students in OECD-PISA 2012. The results for the open-ended questions related to mathematical literacy are examined. The OECD database compendium for the cognitive item responses has been used, the items were grouped into the categories of PISA mathematical framework (content, context, process), and the classification by performance level has also been considered. The overall weakness of the Italian students in the open-ended questions is explained by the high rate of missing and very poor performance in some categories ("change and relationships", "scientific context", "formulate"). Italian students perform better in a few categories ("social context", "interpreting and evaluating", "space and shape"), but these new findings may represent a starting point for improvement. The possible use of data on open-ended questions in teacher training is also discussed.

Keywords: open-ended questions; mathematical literacy; OECD-PISA; teacher training.

1. Introduzione

La bozza del nuovo regolamento sulle classi di concorso per l'insegnamento prevede per gli insegnanti di tutti i livelli scolastici e di tutti gli ambiti concorsuali (Avvertenze generali dell'Allegato A, il possesso di una serie di requisiti culturali e professionali, tra cui al punto 9 "Padronanza delle tematiche legate alla valutazione (sia interna sia esterna), anche con riferimento alle principali ricerche comparative internazionali e alle indagini nazionali"¹.

Nella versione definitiva varata dal MIUR le Avvertenze sono richiamate in tutte le schede degli ambiti e delle classi di concorso, talvolta con ulteriori specifiche (per l'ambito 4, discipline letterarie si cita esplicitamente l'indagine OECD-PISA). Quindi tutti i nuovi aspiranti insegnanti dovranno studiare e dimostrare in sede di esame di conoscere le principali evidenze delle indagini dell'International Association for the Evaluation of Educational Achievement (IEA) e dell'Organization for Economic Co-operation and Development (OECD) e collegarle alle modalità di valutazione utilizzate nella didattica quotidiana, visto che anche il Servizio Nazionale di Valutazione dichiara esplicitamente di rifarsi ai modelli internazionali (Invalsi, 2015). Diventa quindi sempre più urgente riflettere sulle indagini internazionali, sia per i modelli utilizzati, sia per i risultati, ricordando sempre la loro principale funzione: la comparazione dei diversi sistemi di istruzione.

Negli ultimi anni sono decisamente cresciute le informazioni disponibili sulle indagini, utilizzate anche nelle iniziative di formazione iniziale e in servizio degli insegnanti², ma la sottovalutazione di tali informazioni ha finito per porre maggiormente l'accento sui risultati negativi dei diversi cicli del Programme for International Student Assessment (PISA), attraverso i confronti dei punteggi medi, sempre penalizzanti rispetto agli altri Paesi. Ci si è spinti talvolta a stigmatizzare le eterne differenze macro-territoriali tra nord e sud esistenti in Italia, ma poco si è detto invece sui risultati positivi che ancora ottengono gli studenti del livello primario nelle rilevazioni IEA (Timms e Pirls) e addirittura quasi nulla quando, per la prima ed unica volta, anche i nostri studenti quindicenni hanno mostrato un'ottima performance, come è accaduto in PISA 2012 per il problem solving (Asquini, 2014a).

In questo saggio si propone una riflessione sui risultati italiani in matematica di PISA 2012, considerando in particolare le risposte degli studenti italiani alle domande aperte. L'ipotesi di fondo è che tale tipologia di domande costituisca un problema specifico per i nostri studenti, allo stesso modo di quanto già verificato per l'ambito della lettura (Asquini & Corsini, 2012), con l'aggravante che per la matematica la performance dei nostri studenti è sempre stata tra le peggiori rispetto a quella degli altri Paesi. Saranno approfondite le risposte degli studenti in relazione agli elementi principali che caratterizzano la literacy matematica (OECD, 2013a), in modo da poter verificare i principali punti di debolezza mostrati dagli studenti italiani, ma anche eventuali risultati meno negativi. Sarà così possibile fornire interessanti spunti di riflessione per gli insegnanti, legati all'uso di uno strumento di valutazione ancora poco utilizzato nel panorama nazionale, le domande a risposta aperta, cercando di stabilire uno stretto collegamento fra l'esperienza di ricerca e la pratica didattica (Calvani & Menichetti, 2013).

1

http://www.repubblica.it/scuola/2016/01/18/news/titoli_e_prove_ecco_la_bozza_del_bando_per_il_concorso_dei_docenti_2016-131521518/ (ver. 30.03.2016).

² Ricordiamo le iniziative Invalsi "Piano Nazionale di informazione e sensibilizzazione sull'indagine OCSE-PISA e altre ricerche internazionali" e Indire "Progetto Qualità e Merito".

2. Valutare con le domande a risposta aperta

I vantaggi delle domande a risposta aperta nei test sono noti da tempo (Traub & Fisher, 1977): esse sollecitano nel rispondente l'abilità di auto-spiegazione, rispetto alle domande a risposta chiusa che vengono affrontate puntando maggiormente sulla memoria e sulle conoscenze pregresse (Ozuru, Briner, Kurby & McNamara, 2013). Di fatto si tratta di strumenti diversi e complementari, non bisogna confonderli e considerarli simili, quanto invece cercare di sfruttarne le potenzialità specifiche (Rodriguez, 2002). Nel contesto anglosassone l'uso delle *open-ended questions* o *constructed-response questions*, e la relativa riflessione sono ormai consolidati da circa un trentennio (Ackerman & Smith, 1988; Bennett & Ward, 1993; Ward, Dupree & Carlson, 1987), sempre in affiancamento ai quesiti a risposta chiusa. Per questo quando nel 1999 è stato pubblicato il primo quadro di riferimento di PISA (OECD, 1999) non è stata una sorpresa rilevare che una larga parte delle domande cognitive poste agli studenti erano a risposta aperta. Nel primo ciclo la matematica non era l'ambito principale, ma già ben due terzi dei quesiti proposti richiedeva una risposta scritta da parte dello studente (OECD, 2002) e quando nel 2003 il focus dell'indagine è diventato proprio la literacy matematica, la proporzione delle domande aperte fu mantenuta (OECD, 2005) ed è giunta fino al ciclo 2012 (Stacey & Turner, 2015). È stata proprio l'esperienza OECD che ha dato un forte impulso all'uso di domande aperte in contesti di valutazione più ampi, anche di livello nazionale (Anderson & Morgan, 2008; Morris, 2011).

In Italia il Servizio Nazionale di Valutazione (SNV) ha cominciato dal 2010 ad utilizzarle, con maggiore incidenza proprio per l'ambito matematico: già 12 quesiti nell'edizione 2010, di cui quattro a risposta articolata (Invalsi, 2010). Nell'ultima edizione del 2015, le domande aperte andavano dalle 14 per la classe terza della scuola secondaria di primo grado alle 23 della classe quinta della scuola primaria (Invalsi, 2015), anche se bisogna dire che quasi tutte le domande sono state a risposta aperta univoca, nettamente più semplici da codificare da parte degli insegnanti. Quindi anche nel nostro Paese, sulla spinta del modello di valutazione del SNV, si può ipotizzare una crescente diffusione di tale strumento, che comporta problemi di gestione da parte degli insegnanti molto diversi rispetto agli ormai tradizionali test a risposta chiusa.

Quali sono i problemi che pone l'utilizzo di questo tipo di domande? Sostanzialmente due: la stesura della domanda e il trattamento delle risposte. In entrambi i casi PISA rappresenta un illuminante punto di riferimento, considerando che la comparazione internazionale ha elevato al massimo grado le criticità relative ai due problemi. Nella costruzione della domanda il nodo cruciale è la validità, cioè il riferimento esplicito di ogni quesito all'oggetto della valutazione. PISA lo affronta con una definizione meticolosa del quadro di riferimento, da cui derivare tutti gli strumenti utilizzati. In tal senso la sua pubblicazione, che abbiamo già ricordato ha sempre accompagnato lo svolgimento dell'indagine, rappresenta una vera e propria dichiarazione di intenti: cosa vado a valutare e come lo farò, il tutto illustrato anche con una esauriente serie di esempi (OECD, 2013a)³.

Più complesso il problema del trattamento delle risposte, in cui bisogna cercare di garantire un sufficiente grado di oggettività nell'assegnazione di un codice di correttezza. Si tratta di un problema tipico delle prove semi-strutturate (Benvenuto, 2003; Domenici, 1996), che l'OECD affronta con una procedura molto rigorosa (OECD, 2014) seguita da una accurata

³ Ricordiamo che il quadro di riferimento di PISA è disponibile anche in italiano (OECD-Invalsi, 2013) e a questa edizione, disponibile online, faremo riferimento nelle successive analisi.

verifica della *reliability* dei gruppi nazionali di codifica, per garantire un corretto confronto internazionale. Certo, si tratta di procedure costose, che possono indurre qualche perplessità sulla convenienza economica di questo tipo di valutazione, ma il guadagno di informazioni sui risultati cognitivi degli studenti è indubbio (Toch, 2006). Sulla scorta del modello PISA è comunque possibile realizzare delle procedure semplificate, ma non meno rigorose, spendibili in contesti di ricerca più semplici e negli stessi contesti scolastici, facendo perno sulla collaborazione degli insegnanti e sul principio di una adeguata intersoggettività, che può benissimo garantire la qualità della valutazione fondata su domande aperte (Asquini, 2014b). Determinante resta la costruzione di una guida per la codifica, che raccolga tutte le informazioni necessarie al codificatore per leggere le risposte e una serie di esempi per risolvere i casi dubbi. Il vantaggio più importante dell'uso di domande aperte per la valutazione scolastica è la possibilità di risalire facilmente alla fonte, cioè allo studente, per approfondire eventuali risposte ambigue. Anzi, il controllo condiviso in classe delle risposte si può trasformare in un'ottima attività di rinforzo dei contenuti e risultare molto utile per capire le possibili strategie di risposta messe in atto dagli studenti (Pozio, 2010a), poiché il feedback su questo tipo di domande risulta molto più efficace rispetto a quelle a risposta chiusa (Attali, 2015).

3. Il quadro di riferimento in matematica di PISA 2012

Per l'analisi dei risultati italiani nelle domande aperte di matematica in PISA 2012 bisogna considerare la struttura del quadro di riferimento (OECD-Invalsi, 2013). La Figura 1 rappresenta una efficace sintesi di tutti gli elementi che definiscono la literacy matematica.

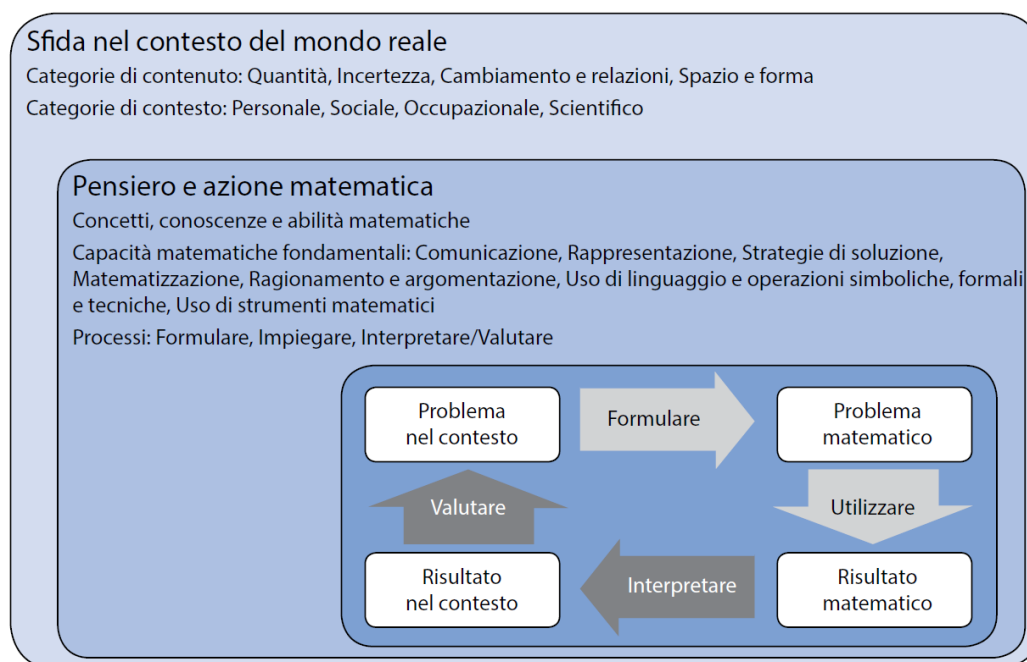


Figura 1. Il modello di literacy matematica in PISA 2012 (ripresa da OECD-Invalsi, 2013, p. 26).

La disposizione dei due livelli chiarisce subito l'approccio di PISA: partire dai problemi reali in cui è necessaria la competenza matematica e poi analizzarne i costrutti specifici, per giungere a un modello ciclico dei processi matematici.

L'analisi della "sfida" considera due categorie, il contenuto e il contesto. Per il contenuto di fatto si tratta di una rivisitazione, molto ampliata, delle denominazioni disciplinari classiche (OECD-Invalsi, 2013). I contesti rappresentano le situazioni reali in cui è possibile esercitare la competenza matematica, da quello squisitamente personale, a quello intellettuale, passando per le situazioni sociali in cui si è coinvolti, fino allo specifico ambiente lavorativo (ibidem).

Per i costrutti interni invece la costruzione dei quesiti di PISA ha considerato in particolare i tre processi matematici in cui pensiero e azione matematica confluiscono (ibidem): formulazione di situazioni in forma matematica; utilizzo di concetti, fatti, procedimenti e ragionamenti matematici; interpretazione, applicazione e valutazione dei risultati matematici.

A queste tre categorizzazioni fondamentali si può aggiungere quella relativa ai livelli di difficoltà delle domande. Turner (2012) analizzando proprio i quesiti di PISA verifica come la difficoltà degli item sia collegata al tipo di capacità fondamentali sollecitate, per cui un quesito facile sarà affrontato in modo diretto con poche conoscenze fondamentali, mentre un quesito molto difficile richiederà il concorso di diverse capacità (dalla matematizzazione alla rappresentazione, dall'uso di strumenti alla comunicazione) e un approccio di tipo problematico. Da notare subito che le domande chiuse non riescono a coprire tutti i livelli previsti da PISA: per i livelli più alti gli unici strumenti utilizzati sono proprio le domande aperte.

Il pacchetto completo di quesiti di PISA è stato costruito cercando di rappresentare in modo adeguato tutte le componenti delle quattro categorie citate, per cui è possibile fare dei confronti puntuali. Per le categorie di contenuto e di processo sono previste delle scale specifiche, accanto a quella di literacy matematica (OECD, 2013b), in cui il punteggio medio italiano è quasi sempre significativamente sotto la media OECD. Ma in questa sede si è scelto di approfondire il risultato italiano verificando l'andamento delle risposte nei singoli item. L'OECD (2013c) mette a disposizione il "Compendium for the cognitive item responses", che raccoglie tutte i dati di risposta dei Paesi partecipanti, compresi gli errori standard, essenziali per la verifica di significatività delle differenze. Inoltre nel rapporto tecnico (OECD, 2014) tutti gli item di matematica sono classificati secondo le quattro categorizzazioni espresse. Quindi si è proceduto a raggruppare i dati relativi agli item nelle diverse categorie, con l'obiettivo di identificare gli elementi di criticità maggiori, ma anche di verificare la presenza di alcune anomalie positive dei risultati degli studenti italiani. In tal modo sarà possibile fornire fondati spunti di riflessione al tema del miglioramento della didattica in matematica.

Una precisazione doverosa è che l'analisi è stata svolta solo sulle prove tradizionali di tipo cartaceo, non sulle domande aperte somministrate per via computerizzata, per i seguenti motivi: il numero delle domande aperte computerizzate è ridotto⁴, e non permette di approfondire l'analisi a tutte le diverse categorie del quadro di riferimento; non è stato ancora rilasciato un numero adeguato di prove esemplificative di tipo computerizzato, per cui non è possibile concretizzare con esempi le diverse categorie del quadro di riferimento; la somministrazione computerizzata comporta specifici problemi di analisi legati all'uso dello strumento, che possono interferire con gli obiettivi di questo saggio. Il tema delle prove computerizzate è comunque richiamato nelle conclusioni.

⁴ Per la prova computerizzata sono previste 26 domande, articolate in 33 item, contro le 43 domande (59 item) della versione cartacea.

Il primo confronto ha riguardato le due tipologie principali di domanda, a risposta chiusa e a risposta aperta. Nella Figura 2 sono confrontate le percentuali di risposta degli studenti italiani rispetto al dato OECD di riferimento, distinguendo domande chiuse e aperte. I tre valori rappresentano le omissioni di risposta, le risposte errate (codice 0 nella definizione PISA) e le risposte corrette (codice 1).

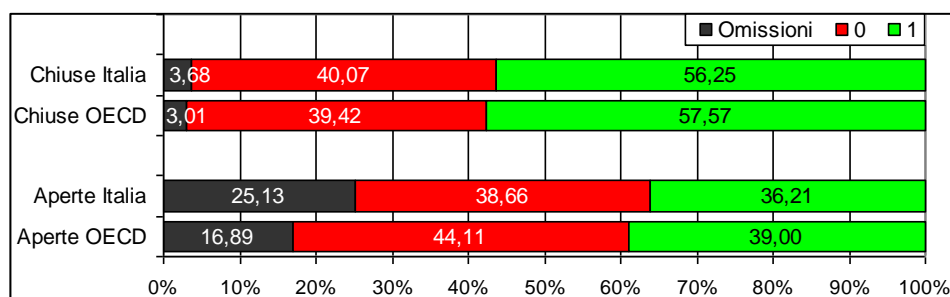


Figura 2. Distribuzione delle risposte per le domande chiuse e aperte di matematica in PISA 2012 (elaborazione da OECD, 2013c).

Le differenze fra le due modalità di risposta sono evidenti. Per le domande chiuse i nostri valori percentuali sono abbastanza vicini a quelli OECD, con una frequenza di omissioni appena superiore e poco più di un punto di differenza nelle risposte corrette. Per le domande aperte invece lo scarto delle risposte esatte sale a quasi tre punti, ma soprattutto le omissioni di risposta rappresentano un quarto del totale, nettamente superiori al dato medio OECD. Il confronto può essere approfondito verificando l'andamento sulle singole domande, in particolare sul numero degli item in cui gli studenti italiani hanno un risultato significativamente diverso dal dato OECD. Nella Figura 3 sono riportati nella parte sinistra i valori assoluti e percentuali degli item che presentano un vantaggio significativo a favore dell'OECD o dell'Italia per quanto riguarda le risposte corrette, nonché il numero dei quesiti in cui non risulta una differenza significativa. Nella parte destra invece le differenze significative riguardanti le omissioni di risposta.

	Risposte corrette Item chiusi		Risposte corrette Item aperti		Omissioni domande chiuse		Omissioni domande aperte	
	VA	%	VA	%	VA	%	VA	%
A favore dell'OECD	16	48	34	58	11	33	42	98
A favore dell'Italia	9	28	9	15	1	3	0	0
Nessuna differenza	8	24	16	27	21	64	1	2
Totale	33	100	59	100	33	100	43	100

Figura 3. Distribuzione degli item per differenze significative, confronto domande chiuse e aperte di matematica in PISA 2012 (elaborazione da OECD, 2013c).

Bisogna chiarire che per le omissioni delle domande aperte si fa riferimento al numero effettivo delle domande (43), di cui però ben 16 prevedono due livelli di risposta corretta, per cui, secondo quanto previsto dall'OECD, ognuna di queste domande conduce a due item diversi, ognuno con uno specifico livello di difficoltà. Questo spiega perché gli item aperti sono 59 (per le domande chiuse il doppio livello di correttezza non è mai previsto, quindi domande e item coincidono).

In 9 item su 33 chiusi gli studenti italiani vanno meglio in modo significativo, mentre in 16 sono nettamente peggiori. Ma la situazione si aggrava nettamente sugli item aperti: in oltre

la metà il dato OECD è significativamente migliore. Tuttavia quei 9 item in cui gli studenti italiani migliorano il dato OECD meritano un approfondimento: un qualche punto di forza da qualche parte sembra esserci.

Impietoso invece il confronto sulle omissioni: solo per una domanda i nostri studenti omettono la risposta in modo simile alla media, in tutte le altre il saltare la domanda è una vera e propria strategia di (non) risposta. Il fenomeno era già emerso nelle precedenti edizioni di PISA (Pozio, 2010b) e anche nel ciclo 2012 di indagine gli studenti italiani si confermano fra i maggiori non rispondenti dell'OECD⁵.

Queste differenze orientano le successive analisi per le diverse categorie: si considereranno solo gli item relativi alle domande aperte; non saranno più analizzate nel dettaglio le omissioni di risposta, che evidentemente sono significativamente più alte in tutte le categorie.

4. Contenuti e contesti matematici

Per la lettura dei risultati disaggregati per categorie è importante il costante rimando sia al quadro di riferimento (OECD-Invalsi, 2013), sia al volume degli esempi di prove PISA che l'Invalsi ha messo a disposizione sul suo sito (Invalsi, 2008), in cui per ogni item sono illustrate categoria e livello di difficoltà. Questa triangolazione è fondamentale per trasformare la lettura dei dati in una occasione di formazione per gli insegnanti dell'area matematica dell'intera fascia dell'obbligo, poiché le difficoltà delle domande sono progressive rispetto a contenuti, contesti e processi, permettendo di ricavare informazioni utili a tutti i livelli.

La Figura 4 presenta l'articolazione delle risposte relativa al contenuto matematico.

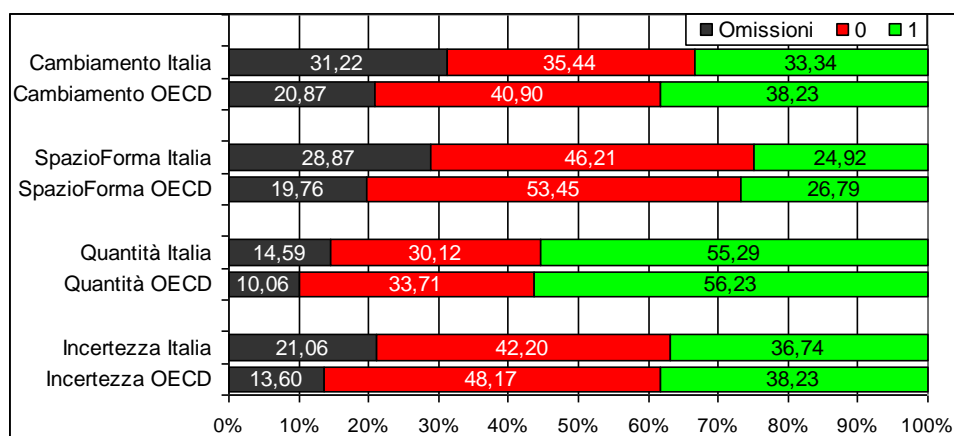


Figura 4. Distribuzione delle risposte per le categorie di contenuto matematico (domande aperte) in PISA 2012 (elaborazione da OECD, 2013c).

⁵ Per completare il discorso sulle omissioni bisogna aggiungere che in media il 2,22% degli studenti italiani non riesce a rispondere alle domande aperte poste alla fine del fascicolo, contro l'1,75% del dato medio OECD. Questa categoria particolare di non rispondenti viene esclusa dall'analisi, ma suggerisce un problema ulteriore degli studenti italiani, circa l'uso del tempo per affrontare un test.

Possiamo notare subito che lo scarto maggiore per le risposte corrette (quasi cinque punti) riguarda la categoria “cambiamento e relazioni”, che si fonda sugli elementi di tipo algebrico, spesso rappresentati nei quesiti anche in forma grafica. Il quadro di riferimento rimanda per questa categoria alla prova “*andatura*”, ma per capire meglio come questo tipo di quesiti metta in difficoltà gli studenti italiani è interessante anche la prova “*l'automobile migliore*”, che presenta un item facile e uno difficile in cui i nostri studenti sono andati decisamente male, pur trattandosi di un problema comparativo stimolante. Le differenze per le altre tre categorie risultano nettamente più contenute (meno di due punti), ma è opportuna la verifica sui singoli item, come illustrato nella Figura 5.

	Cambiamento e relazioni		Spazio e forma		Quantità		Incertezza e dati	
	VA	%	VA	%	VA	%	VA	%
A favore dell'OECD	13	69	8	50	5	42	8	67
A favore dell'Italia	1	5	5	31	2	16	1	8
Nessuna differenza	5	26	3	19	5	42	3	25
Totale	19	100	16	100	12	100	12	100

Figura 5. Distribuzione degli item per differenze significative, confronto categorie di contenuto matematico in PISA 2012 (elaborazione da OECD, 2013c).

Per “cambiamento e relazioni” si conferma la netta distanza dall'OECD, ma in situazione molto simile risulta anche la categoria “incertezza e dati”, con ben due terzi dei quesiti in netto svantaggio. Possiamo dire che si tratta della seconda criticità sui contenuti: probabilità e statistica sono temi su cui i nostri studenti faticano nel confronto internazionale. In questo caso la prova rilasciata “*punteggi di una verifica*”, rappresenta bene le difficoltà che si incontrano per prevedere fenomeni sulla base dei dati.

Nelle altre due categorie invece si conferma la relativa distanza dal dato medio: gli item in cui gli studenti italiani vanno peggio sono la metà o meno, ma controbilanciati da alcuni vantaggi significativi. Da segnalare che dei nove item in cui superiamo l'OECD, cinque sono della categoria “spazio e forma”, l'unica infatti in cui il punteggio medio italiano (aiutato evidentemente dai buoni risultati nei quesiti chiusi) è in linea con quello OECD. Non possiamo quindi parlare di punto di forza, ma almeno nei quesiti di tipo geometrico, di rappresentazione degli spazi, i nostri studenti reggono meglio il confronto. Si può ipotizzare in questo caso che un rinforzo di motivazione a rispondere (le omissioni restano comunque alte) possa portare in tempi brevi a un netto miglioramento dei risultati.

Passiamo alle categorie previste per i contesti matematici. Nella Figura 6 una prima sorpresa: nelle domande relative al contesto sociale, inteso come le diverse comunità di appartenenza dello studente, fino alla dimensione nazionale o globale, il dato italiano è in linea con quello OECD.

Il confronto è decisamente penalizzante per gli altri tre contesti, in particolare quello scientifico, che raccoglie item di diversi ambiti (meteorologia, fisiologia, ecologia) con situazioni problematiche che devono essere affrontate da un punto di vista matematico. Ottimi esempi sono le domande delle prove “*la crescita*” e “*rifiuti*”, che rappresentano due contesti scientifici in cui i nostri studenti non riescono a sfruttare le loro capacità matematiche in modo adeguato. I dati relativi al confronto sui singoli item confermano l'analisi (Figura 7), in particolare per la debolezza in tre quarti dei molti item di contesto scientifico.

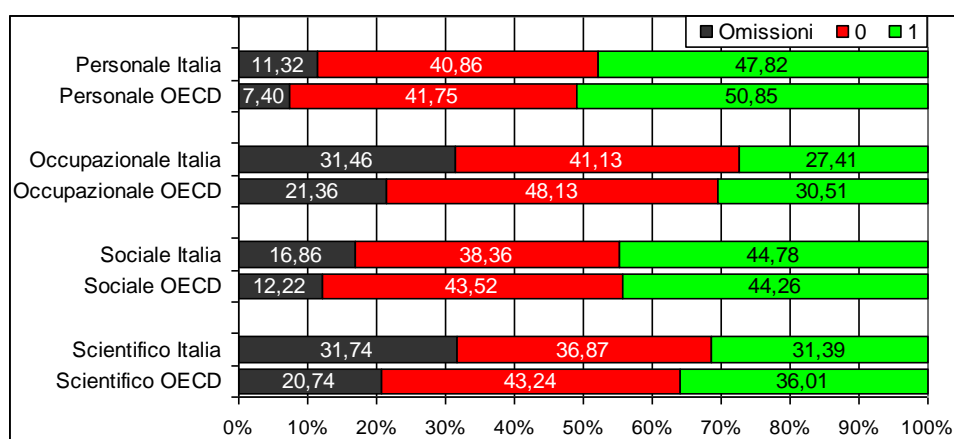


Figura 6. Distribuzione delle risposte per le categorie di contesto matematico (domande aperte) in PISA 2012 (elaborazione da OECD, 2013c).

	Personale		Occupazionale		Sociale		Scientifico	
	VA	%	VA	%	VA	%	VA	%
A favore dell'OECD	5	62	7	47	4	27	18	75
A favore dell'Italia	0	0	2	13	4	27	3	12,5
Nessuna differenza	3	38	3	20	7	46	3	12,5
Totale	8	100	12	100	15	100	24	100

Figura 7. Distribuzione degli item per differenze significative, confronto categorie di contesto matematico in PISA 2012 (elaborazione da OECD, 2013c).

Nella categoria “sociale” c’è un perfetto equilibrio Italia-OECD, si amplifica la debolezza della categoria “personale”, che pure comprende la sfera di vita più vicina allo studente (famiglia, amici): da una parte il maggior interesse è dimostrato dalle percentuali di omissioni, che sono relativamente contenute, ma dall’altra i risultati restano scadenti e in nessun item si migliora il dato OECD. Un chiaro esempio delle difficoltà incontrate dagli studenti italiani è rappresentato dalla prova “*skateboard*”, che tocca una componente matematica molto personale (il risparmio di spesa per un acquisto), ma che ha messo in forte difficoltà i nostri studenti.

5. Processi matematici e livelli di performance

Seguendo la logica del quadro di riferimento ci spostiamo al confronto per le diverse categorie di processo matematico. In questo caso spicca immediatamente il buon risultato nelle domande aperte che richiedono “interpretazione, applicazione e valutazione dei risultati matematici” (Figura 8).

La sostanziale equivalenza rispetto al dato OECD per le risposte corrette è accompagnata da uno scarto più contenuto anche per le omissioni, mentre per le altre due categorie le non risposte abbondano. Se consideriamo lo schema di Figura 1, i processi di interpretazione e valutazione risultano essere su un piano più complesso, poiché prevedono, per esempio, di considerare una soluzione matematica rispetto a una situazione reale, fino a spiegare la possibile non efficacia o i limiti di una soluzione matematica. Il discreto risultato è confermato dal confronto dei punteggi medi già presente nel rapporto OECD (2013b): solo

per questa categoria di processo non risultano differenze significative. In questo caso quindi esemplifichiamo un punto di non debolezza dei nostri studenti con la prova “*popolarità del presidente*”, che chiede di considerare l’attendibilità di diversi sondaggi rispetto a un contesto reale, con brillanti risultati per i nostri studenti.

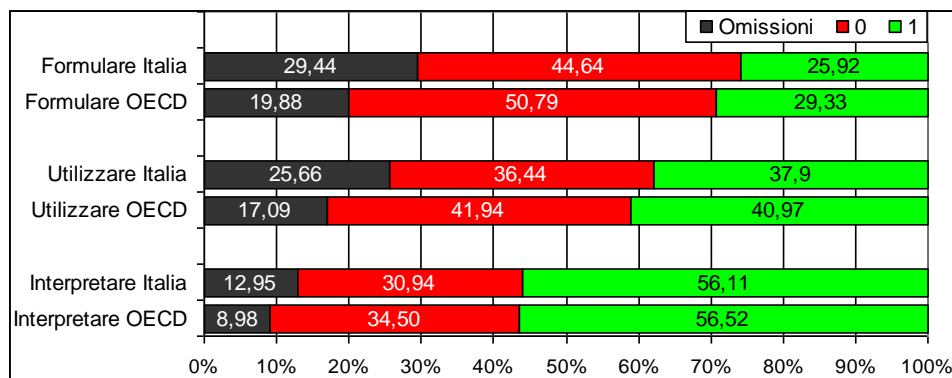


Figura 8. Distribuzione delle risposte per le categorie di processo matematico (domande aperte) in PISA 2012 (elaborazione da OECD, 2013c).

Per l’analisi delle altre due categorie è opportuno scendere nel dettaglio dei singoli item (Figura 9), dove si può subito notare che la categoria in cui andiamo meglio è purtroppo anche la meno rappresentata nel complesso delle domande aperte, con solo 9 item su 59.

	Formulazione di situazioni in forma matematica		Utilizzo di concetti, fatti, procedimenti e ragionamenti matematici		Interpretazione, applicazione e valutazione dei risultati matematici	
	VA	%	VA	%	VA	%
A favore dell’OECD	16	73	16	57	2	22
A favore dell’Italia	3	13,5	6	21,5	0	0
Nessuna differenza	3	13,5	6	21,5	7	78
Totale	22	100	28	100	9	100

Figura 9. Distribuzione degli item per differenze significative, confronto categorie di processo matematico in PISA 2012 (elaborazione da OECD 2013c).

Gli studenti italiani risultano molto più deboli nella “formulazione di situazioni in forma matematica”, con uno svantaggio significativo in tre quarti degli item. È evidente il collegamento con la criticità già rilevata nel contesto scientifico, che richiede spesso di leggere una situazione da un punto di vista matematico, ma la debolezza si estende anche a contesti più vicini agli studenti. Negativo anche il confronto per la categoria “utilizzo di concetti, fatti, procedimenti e ragionamenti matematici”, ma i sei quesiti in cui si supera il dato OECD sembrano un segnale di miglioramento, anche se in quesiti semplici e molto vicini all’esperienza dello studente, come nella prova “*scelte*”, in cui si richiede di risalire alle possibili combinazioni di ingredienti di una pizza, i nostri studenti non brillano.

Concludiamo il confronto con i livelli di performance. PISA definisce sette livelli⁶, ma per molte analisi raggruppa gli studenti in quattro gruppi di performances: “top” per i livelli

⁶ Oltre ai sei livelli standard va considerato quello degli studenti che non raggiungono neanche la soglia inferiore del livello 1, per cui si ha un settimo livello definito Inferiore a 1 (OECD, 2014).

cinque e sei, “strong” per il livello quattro, “moderate” per i livelli due e tre, “lowest” per i livelli uno e inferiore a uno (OECD, 2013b). Poiché la scala di riferimento è la stessa per item e studenti⁷, è possibile raggruppare gli item allo stesso modo, con quattro categorie distinte, rappresentate nella Figura 10.

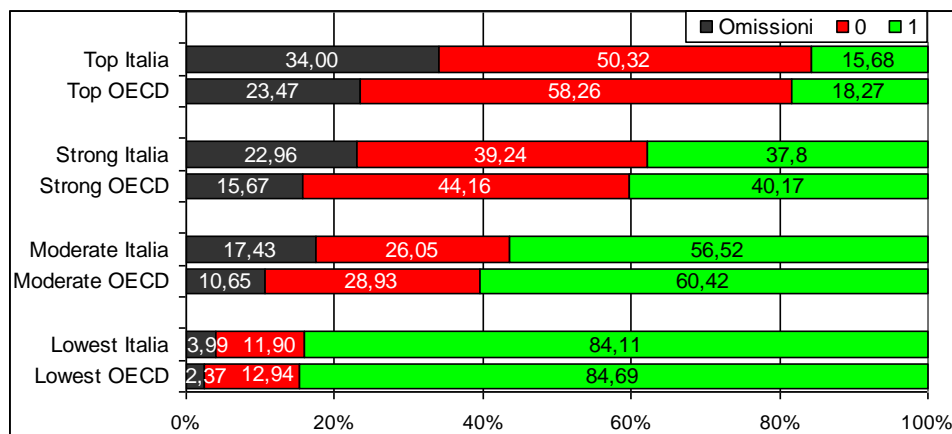


Figura 10. Distribuzione delle risposte per le categorie di performance matematica (domande aperte) in PISA 2012 (elaborazione da OECD, 2013c).

L'unica categoria in cui gli studenti italiani sono vicini al dato OECD è quella delle domande più facili, che però è anche la meno rappresentata, poiché la maggior parte delle domande facili sono a risposta chiusa. Al contrario la categoria più numerosa (27 item su 59) è quella delle domande molto difficili (“top”), che però non comporta il nostro peggior risultato, raggiunto nelle domande abbastanza facili (“moderate”). Da notare anche la parallela crescita delle omissioni in relazione alla difficoltà delle domande. L'analisi dei singoli item nella Figura 11 conferma le differenze.

	Top item		Strong item		Moderate item		Lowest item	
	VA	%	VA	%	VA	%	VA	%
A favore dell'OECD	16	59	6	55	11	65	1	25
A favore dell'Italia	6	22	1	9	2	12	0	0
Nessuna differenza	5	19	4	36	4	24	3	75
Totale	27	100	11	100	17	100	4	100

Figura 11. Distribuzione degli item per differenze significative, confronto categorie di performance matematica in PISA 2012 (elaborazione da OECD, 2013c).

Le difficoltà maggiori sono nelle domande “moderate”, come per esempio quelle della prima domanda della prova “*esportazioni*”, che richiede la semplice lettura di un dato da un grafico. Purtroppo anche in oltre metà degli item più difficili si rileva un significativo svantaggio dei nostri studenti, esemplificabile con la prova “*area di un continente*”, che prevede una sola domanda, ma con due livelli di risposta (due item distinti quindi) entrambi “top”: sono pochissimi i nostri studenti che riescono a dimostrare una efficace capacità di stima delle misure.

⁷ I punteggi di PISA sono calcolati con il modello Item Response Theory (IRT) di Rasch, che rappresenta su una sola scala la facilità dell'item e l'abilità dello studente (Asquini, 2006).

6. Conclusioni

È possibile riepilogare le principali evidenze emerse dall'analisi dei risultati degli studenti italiani nelle domande aperte di matematica in PISA 2012:

- lo scarto rispetto al dato medio OECD è nettamente più marcato per le domande aperte rispetto a quelle chiuse, evidenziando la minor dimestichezza degli studenti italiani con questo tipo di domande;
- particolarmente alta risulta la percentuale delle omissioni nelle domande aperte, problema ormai storico per gli studenti italiani in PISA, suggerendo che l'ultimo decennio di sempre maggior diffusione di test a risposta chiusa nella scuola italiana abbia rinforzato la tendenza a non rispondere appena si deve andare oltre la crocetta;
- la debolezza degli studenti italiani è particolarmente accentuata in alcune categorie: "cambiamento e relazioni" per il contenuto, domande di ambito scientifico, quesiti che richiedono di formulare matematicamente situazioni reali;
- le difficoltà non risultano particolarmente legate al livello di difficoltà delle domande, poiché anche in quelle facili gli studenti italiani scartano significativamente dal dato OECD;
- esistono alcuni elementi di non-debolezza, che possono rappresentare un possibile punto di partenza per migliorare la capacità di affrontare un problema di tipo matematico in modo più articolato: domande legate al contesto sociale, quesiti che richiedono di interpretare e valutare i risultati matematici, contenuti di ambito geometrico.

È possibile anche indicare alcuni possibili sviluppi di indagine, per poter fornire informazioni sempre più dettagliate e utili a tutti i portatori di interesse nella scuola, che è poi uno degli obiettivi delle indagini PISA (OECD, 1999):

- estendere l'analisi alle domande aperte della prova computerizzata in matematica di PISA 2012, per la quale già sappiamo che il risultato medio italiano è stato non significativamente diverso da quello OECD (Invalsi, 2013). Un primo riscontro sul complesso delle domande aperte, svolto sempre sui dati del "Compendium for the cognitive item responses" (OECD, 2013c), ha permesso di verificare che il numero delle omissioni si allinea al dato OECD e il dato sulle risposte corrette presenta uno scarto ridotto (25,24 contro 26,15), suggerendo che le prove computerizzate hanno un effetto positivo sugli studenti italiani, sia nella motivazione a rispondere sia nella performance;
- verificare le differenze esistenti fra maschi e femmine nelle risposte aperte, partendo dalla significativa differenza di genere riscontrata per il nostro Paese (Invalsi, 2013) e ipotizzando che la motivazione a rispondere a questo tipo di domande costituisca uno dei motivi principali della differenza di risultato;
- ricostruire la serie storica dei risultati relativi alle domande aperte di matematica in PISA 2012, con particolare attenzione all'andamento nei link item, cioè le domande utilizzate in tutte le edizioni, che permettono un confronto preciso nel tempo. Ovviamente si tratta di un gruppo ristretto di item, che non consente di dettagliare le diverse categorie, ma permette di verificare la tendenza;
- approfondire, con opportuni studi di caso, l'approccio degli studenti alle domande aperte sulla scorta del lavoro svolto da Pozio (2010b), per fornire indicazioni didattiche più concrete sull'uso di tale strumento nella pratica didattica, secondo una logica di intervento informato, mirato ed efficace (Zan, 2007).

Bibliografia

- Ackerman, T.A., & Smith, P.L. (1988). A comparison of the information provided by essay, multiple-choice, and free response writing tests. *Applied Psychological Measurement, 12*(2), 117–128.
- Anderson, P., & Morgan, G. (2008). *Developing tests and questionnaires for a national assessment of educational achievement*. Washington, DC: The World Bank. http://siteresources.worldbank.org/EDUCATION/Resources/278200-1099079877269/547664-1222888444288/National_assessment_Vol2.pdf (ver. 30.03.2016).
- Asquini, G. (2006). Che cosa ci dicono le risposte delle domande: item analisi classiche e item response theory. *Notiziario U.M.I. Novembre 2006 (parte b)*, 5(13), 13–22.
- Asquini, G. (2014a). Lo strano caso dei risultati italiani di PISA 2012. *Italian Journal of Educational Research, VII*(13), 13–28.
- Asquini, G. (2014b). Definizione di una procedura di codifica delle domande aperte basata sui modelli delle indagini internazionali. *ECPS Journal - Educational Cultural and Psychological Studies, 10*, 463–483.
- Asquini, G., & Corsini, C. (2012). Come evolve la competenza in lettura dei quindicenni. *Atti del convegno XLIV SLI*, 311–327. Roma: Bulzoni.
- Attali, Y. (2015). Effects of multiple-try feedback and question type during mathematics problem solving on performance in similar problems. *Computer and Education, 86*, 260–267.
- Bennett, R., & Ward, W. (eds.). (1993). *Constructive versus choice in cognitive measurement: issues in constructed response, performance testing, and portfolio assessment*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Benvenuto, G. (2003). *Mettere i voti a scuola. Introduzione alla docimologia*. Roma: Carocci.
- Calvani, A., & Menichetti, L. (2013). Evidence Based Education: superare il gap tra ricerca e pratica. *Form@re - Open Journal per la Formazione in Rete, 13*(2), 1–5. <http://www.fupress.net/index.php/formare/article/view/13252> (ver. 30.03.2016).
- Domenici, G. (1996). *Gli strumenti della valutazione*. Napoli: Tecnodid.
- Invalsi. Istituto Nazionale per la Valutazione del Sistema educativo di Istruzione e di formazione (2008). *Compendio prove PISA. Insegnanti*. http://www.invalsi.it/invalsi/rn/odis/doc/Compendio_prove.pdf (ver. 30.03.2016).
- Invalsi. Istituto Nazionale per la Valutazione del Sistema educativo di Istruzione e di formazione (2010). *Rilevazione degli apprendimenti – SNV Prime analisi*. http://www.invalsi.it/download/rapporti/snv2010/Rapporto_SNV_09_10.pdf (ver. 30.03.2016).
- Invalsi. Istituto Nazionale per la Valutazione del Sistema educativo di Istruzione e di formazione (2013). *OCSE PISA 2012. Rapporto nazionale*. http://www.invalsi.it/invalsi/ri/pisa2012/rappnaz/Rapporto_NAZIONALE_OCSE_PISA2012.pdf (ver. 30.03.2016).
- Invalsi. Istituto Nazionale per la Valutazione del Sistema educativo di Istruzione e di formazione (2015). *Rilevazioni nazionali degli apprendimenti 2014-15. Rapporto*

risultati.

http://www.invalsi.it/invalsi/doc_evidenza/2015/034_Rapporto_Prove_INVALSI_2015.pdf (ver. 30.03.2016).

- La Repubblica (18 gennaio 2016). *Titoli e prove, ecco la bozza del bando per il concorso docenti.*
http://www.repubblica.it/scuola/2016/01/18/news/titoli_e_prove_ecco_la_bozza_del_bando_per_il_concorso_dei_docenti_2016-131521518/ (ver. 30.03.2016).
- Morris, A. (2011). Student standardised testing: current practices in OECD, countries and literature review. *OECD Education Working Papers, 65*. Paris: OECD Publishing.
- OECD. Organisation for Economic Co-operation and Development (1999). *Measuring student knowledge and skills. A new framework for assessment*. Paris: OECD Publishing.
- OECD. Organisation for Economic Co-operation and Development (2002). *PISA 2000. Technical report*. Paris: OECD Publishing.
- OECD. Organisation for Economic Co-operation and Development (2005). *PISA 2003. Technical report*. Paris: OECD Publishing.
- OECD. Organisation for Economic Co-operation and Development (2013a). *PISA 2012. Assessment and analytical framework*. Paris: OECD Publishing.
- OECD. Organisation for Economic Co-operation and Development (2013b). *PISA 2012. Results: what students know and can do – student performance in mathematics, reading and science*. (Vol. 1). Paris: OECD Publishing.
- OECD. Organisation for Economic Co-operation and Development (2013c). *Compendium for the cognitive item responses. Database – PISA 2012*.
- OECD. Organisation for Economic Co-operation and Development (2014). *PISA 2012. Technical report*. Paris: OECD Publishing.
- OECD-Invalsi (2013). *PISA 2012. Quadro di riferimento analitico per la matematica, la lettura, le scienze, il problem solving e la financial literacy*. Roma: Armando Editore. <http://www.invalsi.it/invalsi/ri/pisa2012/documenti/Matematica.pdf> (ver. 30.03.2016).
- Ozuru, Y., Briner, S., Kurby, C.A., & McNamara, D.S. (2013). Comparing comprehension measured by multiple-choice and open-ended questions. *Canadian Journal of Experimental Psychology, 67*(3), 215–277.
- Pozio, S. (2010a). *La risoluzione di prove di competenza matematica*. Roma: Edizioni Nuova Cultura.
- Pozio, S. (2010b). Il problema delle omissioni in PISA. In Invalsi, *PISA 2006. Approfondimenti tematici e metodologici* (pp. 203-218). Roma: Armando Editore.
- Rodriguez, M.C. (2002). Choosing an item format. In G. Tindal & T.M. Haladyna (eds.), *Large-scale assessment programs for all students. Validity, technical adequacy, and implementation* (pp. 213-231). Mahwah, NJ: Lawrence Erlbaum Associates.
- Stacey, K., & Turner, S. (eds.). (2015). *Assessing mathematical literacy: the PISA experience*. Cham: Springer.

- Toch, T. (2006). *Margins of error; the educational testing industry in No Child Left Behind era*. Washington, DC: Education Sector Reports. http://www.educationsector.org/sites/default/files/publications/Margins_of_Error.pdf (ver. 30.03.2016).
- Traub, R., & Fisher, C. (1977). On the equivalence of constructed-response and multiple choice tests. *Applied Psychological Measurement*, 1(3), 355–369.
- Turner, R. (2012). *Some drivers of test item difficulty in mathematics*. Paper presented at the Annual Meeting of the American Educational Research Association (AERA), Vancouver. <http://research.acer.edu.au/cgi/viewcontent.cgi?article=1003&context=pisa> (ver. 30.03.2016).
- Ward, W.C., Dupree, D., & Carlson, S.B. (1987). A comparison of free-response and multiple-choice questions in the assessment of reading comprehension. *ETS Research Report Series*, 1987(1), 1–26.
- Zan, R. (2007). *Difficoltà in matematica. Osservare, interpretare, intervenire*. Milano: Springer.