

# Maschi e femmine in Matematica: dove sta la differenza?

**Giorgio Bolondi**

Libera Università di Bolzano

E-mail: giorgio.bolondi@unibz.it

**Chiara Giberti**

Università di Bergamo

**Clelia Cascella**

INVALSI

**Abstract.** Le prove standardizzate su larga scala mettono in luce come esistano in Italia differenze importanti di risultati di apprendimento tra maschi e femmine, a livello di sistema. Le nostre ricerche condotte in questi anni hanno permesso di evidenziare come questo gap non sia uniforme, ma concentrato in alcuni specifici item, e che questo può essere collegato a fattori di tipo didattico. Questi risultati sono rilevanti per la progettazione e l’implementazione dei percorsi di formazione iniziale e in servizio degli insegnanti.

Large scale assessments and surveys highlight that in Italy there exist a relevant gender gap in mathematical achievement between boys and girls, at a systemic level. Our research shows that this gap is centered on specific items, and that this can be related to didactic aspects. These results are relevant to the planning and implementation of pre- and in-service teachers training programs.

**Keywords.** Differenze di genere, Misconcezioni, Contratto didattico, Valutazioni su Larga scala - Gender gap, Misconceptions, Didactic contract, Large-scale assessment.

## 1. Introduzione

La ricerca sugli apprendimenti in Matematica, in funzione della variabile “genere”, ha avuto un grosso sviluppo negli ultimi decenni, dovuto al fatto che in molti paesi del mondo continua a essere rilevata dalle indagini su larga scala una importante differenza di risultati tra maschi e femmine (si vedano ad esempio Buckley, 2016, per una discussione della situazione in Australia, o Di Tommaso et al., 2016, per l’Italia). Le direzioni in cui si è sviluppata la ricerca sono molteplici, e hanno utilizzato quadri teorici di riferimento molto differenti, con differenti *background* sociologici, epistemologici e culturali. Dal punto di vista della didattica della matematica, l’interesse è soprattutto per i fenomeni di apprendimento risultanti, e in particolare si concentra sugli aspetti più tipicamente matematici dell’apprendimento (ambiti, contenuti, processi coinvolti, formulazione, relazione con le pratiche didattiche), e sulla relazione tra questi fenomeni e le altre variabili.

Una domanda molto naturale è la seguente: in una prova, composta da molte domande e consegne, in cui si rileva una differenza tra maschi e femmine, questa differenza è “spalmata” su tutte le domande, o ci sono domande in cui la differenza è più forte/meno forte/o addirittura a vantaggio delle femmine? Se è così, cosa ci dicono queste domande, attraverso le loro caratteristiche?

Cercheremo di dare alcune prime risposte a questa domanda attraverso l’analisi dei risultati delle prove INVALSI, la valutazione su larga scala degli apprendimenti implementata in Italia negli ultimi dieci anni, utilizzando una strategia in due fasi, basate su diverse metodologie.

Una prima analisi sarà basata direttamente sulle percentuali di risposta, attraverso la definizione di uno specifico *indice* (Bolondi et al., 2017) attraverso il quale individuare in quali, tra le oltre 1400 domande esamina-

te, la differenza di genere ha un impatto più forte. Passando poi ad analisi più approfondite dal punto di vista psicometrico, abbiamo utilizzato uno specifico strumento statistico (*DIF analysis*; Cascella et al., 2020) per individuare, in alcune domande, ulteriori proprietà di questa differenza di risultato. In entrambi i casi, abbiamo combinato questi strumenti quantitativi con l’analisi qualitativa, fondata sui risultati della ricerca in didattica della matematica.

Questo approccio di ricerca, che utilizza *mixed method* sia di tipo *exploratory* che *explanatory* (Creswell & Plano-Clark, 2017; Bolondi & Cascella, 2020) si basa sull’idea che le valutazioni su larga scala possano dare informazioni importanti non solo a livello di sistema, ma anche per l’analisi dei processi di insegnamento e apprendimento, facendo emergere *macrofenomeni* che possono essere collegati a quanto la ricerca didattica ha studiato, e indicare nuove piste di lavoro (Bolondi, 2021; Bolondi & Ferretti, 2021). Al tempo stesso, i risultati evidenziati forniscono informazioni importanti sul ruolo che gli stereotipi nella classe possono giocare nel creare le differenze di risultati tra maschi e femmine, mettendo in luce l’importanza di questo tema nella formazione degli insegnanti negli specifici contesti nazionali (Giberti, 2019a).

## 2. Le prove INVALSI

Le prove INVALSI hanno confermato per l’Italia quanto era già stato messo in evidenza dalle ultime indagini OCSE-Pisa e IEA-TIMSS: tra i paesi industrializzati, l’Italia è uno di quelli per cui il gender gap in matematica è più rilevante. Per individuare le cause di questo, la maggior parte delle ricerche oggi si concentra maggiormente su quelli che sono chiamati *external factors*, i fattori di tipo culturale e sociale, dando meno importanza agli *internal factors*, i fattori di tipo biologico (Giberti, 2019b; Di Tommaso et al., 2016). In particolare, per la matematica entrano in gioco pesantemente fattori come le convinzioni di allievi, genitori e docenti, gli stereotipi sociali che permeano il lavoro di classe, l’autopercezione degli allievi e delle allieve, l’ansia (Jacobs & Bleeker, 2004; Riegle-Crumb, 2005; Freyer & Levitt, 2010). Anche le pratiche didattiche degli insegnanti possono avere un impatto differente su maschi e femmine (Leder & Forgasz, 2008).

L’INVALSI, per analizzare i risultati delle prove, utilizza il modello di Rasch: questo permette di posizionare su un’unica scala le difficoltà  $d$  degli item e le abilità  $a$  degli studenti; questi due indici permettono di definire la probabilità che uno studente di abilità  $a$  sia in grado di rispondere correttamente a una domanda di difficoltà  $d$ . La relazione tra questi fattori viene rappresentata sinteticamente da una curva, l’*Item Characteristic Curve* (ICC), che presenta graficamente la relazione tra l’abilità dello studente e la probabilità di rispondere correttamente a una domanda di data difficoltà. Da un punto di vista strettamente psicometrico-teoretico, non dovrebbero esserci differenze nelle curve una volta che si analizzano sottogruppi di studenti. La nostra seconda tecnica consiste proprio nel mettere a fuoco quegli item per i quali il funzionamento (espresso dalla ICC) è diverso per maschi e femmine. In particolare, è possibile individuare per quali item l’abilità dello studente (misurata dal complesso della prova) influenza in maniera differente la probabilità di rispondere correttamente all’item.

## 3. L’indice di Gender Gap

In prima battuta, la “sensibilità” di un item al genere è leggibile nella differenza tra le percentuali di risposte corrette date dai maschi e dalle femmine all’item. Ovviamente, una differenza di performance del 10% tra maschi e femmine ha un significato diverso, se l’item è molto difficile (ad esempio, ha il 15% complessivo di risposte corrette), o se l’item è molto facile (ad esempio, ha il 90% complessivo di risposte corrette). L’indice di Gender Gap è quindi definito come

$$I_{gg} = (M_k - F_k)/P_k$$

dove

- $M_k$  è la percentuale di risposte corrette dei maschi all’item  $k$
- $F_k$  è la percentuale di risposte corrette delle femmine all’item  $k$
- $P_k$  è la percentuale di risposte corrette dell’intera popolazione all’item  $k$ .

L'indice è positivo se i maschi performano meglio delle femmine su quell'item, negativo se sono le femmine ad avere un migliore risultato.

L'analisi delle prove INVALSI mostra come questo indice possa assumere valori molto vari, positivi e negativi. A titolo di esempio, si riportano in tabella i valori relativi alla prova somministrata nel 2009 nelle classi seconde primarie. Il valore Delta indica la difficoltà della domanda, calibrata dal modello di Rasch, su una scala da -4 a +4; l'ambito è l'ambito di contenuti dell'item, così come etichettato dall'INVALSI.

**Tabella 1.** Selezione di item della prova INVALSI di grado 2 del 2009 con il relativo indice *Igg*.

<b>Item</b>	<b>Igg</b>	<b>Ambito di contenuto</b>	<b>Delta</b>	<b>Percentuale di risposte corrette</b>
M10	30	Numeri	0.79	33%
M20	18	Numeri	-0.55	62%
M14	13	Numeri	-0.35	57%
M15	10	Numeri	-0.26	55%
M2	9	Numeri	-0.19	54%
M3	7	Numeri	-0.33	57%
M22	7	Numeri	-0.69	64%
M16	7	Numeri	1.00	30%
M18	6	Spazio e figure	0.82	33%
M23	6	Numeri	-1.09	72%
M13	5	Spazio e figure	-1.38	78%
M17	4	Numeri	-1.46	78%
M21	2	Spazio e figure	0.58	38%
M5	1	Spazio e figure	-0.29	56%
M9	1	Numeri	0.97	30%
M6b	0	Dati e previsioni	-1.65	81%
M1	-1	Numeri	-1.38	77%
M12	-1	Dati e previsioni	-1.96	85%
M6a	-1	Dati e previsioni	-1.93	84%
M11	-2	Dati e previsioni	0.15	47%
M4	-2	Numeri	0.52	39%
M8	-4	Numeri	0.27	44%
M19	-9	Spazio e figure	1.48	22%
M7	-9	Numeri	-0.20	54%

Come si può osservare, già nella prova di seconda primaria sono presenti item per i quali la differenza di risposte corrette tra maschi e femmine è molto marcata, e item in cui le femmine rispondono meglio dei maschi. I dati sono rilevati su un campione molto ampio (in genere, composto da circa 30.000 studenti), per cui sono statisticamente molto robusti.

Si può osservare come vi siano domande con alto indice sia facili sia difficili; in questa prova, le domande con alto indice appartengono tutte all'ambito *Numeri*.

Possiamo affiancare a questi risultati i risultati della stessa coorte di studenti alle prese con la prova nazionale al termine del primo ciclo di istruzione (classe terza secondaria di primo grado), nel 2015.

**Tabella 2.** Selezione di item della prova INVALSI di grado 8 del 2015 con il relativo indice *Igg*.

<b>Item</b>	<b>Igg</b>	<b>Ambito di contenuto</b>	<b>Delta</b>	<b>Percentuale di risposte corrette</b>
D17	26	Spazio e figure	2.81	8
D16_b	18	Numeri	0.36	42
D18	14	Numeri	0.60	38
D3	12	Spazio e figure	-0.75	65
D15_b	12	Relazioni e funzioni	-0.93	69
D10	11	Relazioni e funzioni	-1.04	71
D24	11	Spazio e figure	-0.35	57
D6	10	Dati e previsioni	-0.22	54
D22	9	Numeri	-0.42	59
D14	8	Dati e previsioni	-0.93	69
D21_b	8	Relazioni e funzioni	0.04	49
D1_c	7	Relazioni e funzioni	-1.26	74
D19	6	Numeri	0.30	44
D2_a	6	Dati e previsioni	-1.53	79
D2_b	5	Dati e previsioni	0.07	48
D13	5	Spazio e figure	1.66	20
D27	4	Dati e previsioni	-0.22	54
D16_a	4	Numeri	-0.19	54
D12	4	Dati e previsioni	-0.85	67
D1_a	4	Relazioni e funzioni	-2.53	90
D21_a	3	Relazioni e funzioni	-0.90	68
D20	2	Relazioni e funzioni	-0.16	53
D4	2	Numeri	0.43	41
D7	2	Relazioni e funzioni	-1.05	71
D15_a	2	Dati e previsioni	-2.43	89
D8_b	1	Spazio e figure	-1.23	74
D5	0	Spazio e figure	0.36	42
D1_b	0	Relazioni e funzioni	-2.02	85
D11_a	0	Spazio e figure	-0.52	61
D25_a	0	Relazioni e funzioni	-0.92	68
D9	-1	Numeri	-0.63	63
D28	-2	Relazioni e funzioni	-0.41	58
D26	-2	Spazio e figure	-0.63	63
D8_a	-3	Spazio e figure	-1.64	80

<b>D25_b</b>	-5	Relazioni e funzioni	-0.33	57
<b>D23</b>	-5	Numeri	-0.64	63
<b>D11_b</b>	-12	Spazio e figure	0.59	38

Qui abbiamo domande con gender gap a favore dei maschi in tutti gli ambiti e di livello di difficoltà molto varia, e abbiamo anche domande con gender gap a favore delle femmine in 3 ambiti su 4, e di diversa difficoltà. La stessa analisi compiuta sull'insieme di tutte le prove INVALSI conferma questa prima evidenza: il gender gap non è legato a un particolare ambito di contenuti, e non si forma su uno specifico livello di difficoltà degli item.

Per poter individuare delle caratteristiche comuni agli item che producono l'effetto complessivo dei gender gap è quindi necessario analizzare nel dettaglio le domande, con un approccio di natura qualitativa.

La nostra ricerca (ha scelto quindi alcuni specifici temi come caso da indagare. Evidenze particolarmente interessanti sono emerse analizzando le domande che coinvolgono i problemi presenti nell'allargamento del sistema dei numeri (dai naturali ai negativi e ai razionali). Questo è un tema classico della ricerca in didattica, e sono ben note le misconcezioni che si possono creare (Steinle & Stacey, 2003; Steinle & Stacey, 2004; Roche & Clarke, 2004; Sbaragli, 2012).

In particolare, è stato possibile individuare cluster di item legati a questo nucleo che, anche in prove di diversi livelli scolastici, mostrano un marcato gap a favore dei maschi. Questo in particolare sembra mostrare che alcune difficoltà riscontrate maggiormente dalle femmine nei primi anni di scuola, non siano poi colmate durante il percorso scolastico. Dalle prime analisi si è visto come, ad esempio, le studentesse siano maggiormente influenzate da particolari misconcezioni legate al passaggio tra l'insieme dei numeri naturali e l'insieme dei numeri razionali (Giberti, 2018). Queste misconcezioni portate avanti negli anni diventano veri e propri ostacoli all'apprendimento se non vi è una consapevolezza e un lavoro mirato da parte dell'insegnante per evidenziarle e superarle. Una domanda rappresentativa di questa situazione è la D21a/b della prova di grado 5 del 2012.

Un altro nucleo su cui sono emerse ricorrenti situazioni di forte gender gap tra maschi e femmine è quello della *stima* di misure. Anche in questo caso, le difficoltà delle femmine sono più marcate in diversi livelli scolari. Esempi molto significativi di questa situazione sono la domanda D7b del grado 6 della prova 2013 e la domanda D3 del grado 8 della prova 2015.

In generale, sembra che la differenza di risultati tra maschi e femmine sia più rilevante negli item in cui più forti sono gli elementi che possono far scattare comportamenti di *contratto didattico* (Bolondi et al., 2018), riportando quindi al tema generale degli stereotipi di genere nella classe di matematica.

#### **4. La DIF analysis**

Come si è detto in un paragrafo precedente, da un punto di vista strettamente psicometrico-teoretico, non dovrebbero esserci differenze nelle curve caratteristiche delle domande una volta che si analizzano sottogruppi di studenti. La nostra seconda tecnica consiste proprio nel mettere a fuoco quegli item per i quali il funzionamento (espresso dalla ICC) è diverso per maschi e femmine, pur restando nei range di accettabilità. In particolare, è possibile individuare per quali item l'abilità dello studente (misurata dal complesso della prova) influenza in maniera differente la probabilità di rispondere correttamente all'item.

Anche in questo secondo approccio, siamo partiti dall'analisi degli oltre 1400 item presenti nella banca dell'INVALSI, e dalla selezione di quelli per i quali la differenza di risultato tra maschi e femmine è quantitativamente e statisticamente rilevante. Per gli stessi cluster di item studiati nella fase precedente abbiamo confrontato, tra di loro e con le curve predette dal modello di Rasch, le spezzate empiriche dei risultati dei maschi e delle femmine.

Per l'item D21a citato in precedenza (un item difficile), si osserva che la probabilità di rispondere correttamente è circa la stessa per maschi e femmine di bassa abilità, mentre la differenza di risultato cresce decisamente al crescere dell'abilità del rispondente. In altre parole, se effettivamente la principale difficoltà consiste nella presenza di una misconcezione, allora i risultati sembrano indicare che questa impatta in maniera più marcata sulle ragazze di alta abilità, rispetto ai maschi. Nell'item D21b, in cui la procedura più semplice

per la risposta corretta non era influenzata dalla misconcezione, la differenza tra maschi e femmine è “spalmata” su tutto il range di abilità.

Per quanto riguarda il secondo cluster di domande, quelle relative alla stima di misure, va osservato preliminarmente che i dati generali sembrano mostrare che questa competenza è correlata solo debolmente con il *latent trait*, il carattere latente misurato dalla prova: in altre, parole, sembra essere abbastanza indipendente dalla competenza matematica complessivamente misurata, anche se si tratta di un elemento sottolineato fortemente nelle Indicazioni Nazionali.

Per queste, che presentano un forte gender gap a livello quantitativo, la differenza è quasi su tutti i livelli di abilità, ed è dovuta alla presenza di distrattori che attraggono in maniera molto forte le femmine.

Infine (Cascella et al., 2021) una misconcezione classica che si è evidenziata essere presente in misura maggiore nelle femmine è quella relativa alla *moltiplicazione che aumenta*. Questa misconcezione è legata al modello di moltiplicazione come addizione ripetuta, e quindi al diverso significato che viene attribuito ai due fattori. In italiano, abitualmente, il primo fattore identifica “la cosa che viene moltiplicata” e il secondo “il numero di volte”; la misconcezione quindi scatta relativamente al primo fattore, ad esempio quando il secondo è un numero decimale minore di 1. Il nostro impianto sperimentale ha permesso di verificare, con una tecnica di ancoraggio, che questo succede in maniera significativamente maggiore nel caso delle ragazze.

## 5. Conclusioni

Le nostre ricerche hanno permesso di verificare che il gender gap in matematica non è distribuito uniformemente su tutte le domande o su tutti i livelli di abilità: esso è strettamente legato alle caratteristiche dei singoli item. È possibile, nei casi studiati, collegare le differenze quantitative osservate alle caratteristiche qualitative delle domande e della loro formulazione. In particolare, abbiamo osservato come alcune tipiche misconcezioni producano effetti più marcati nelle femmine rispetto ai maschi, abbassando il livello di risposte corrette delle femmine di alta abilità.

Questo suggerisce interessanti piste per ulteriori indagini di tipo qualitativo.

## References

- Bolondi, G. (2021). What can we learn from large-scale surveys about our students learning of maths? *AAPP- Atti della Accademia Peloritana dei Pericolanti, Classe di Scienze Fisiche, Matematiche e Naturali*, 99, 2021.
- Bolondi, G., & Cascella (2020). A mixed approach to interpret large-scale assessment psychometric results of the learning of mathematics. *La Matematica e la sua didattica*, 28(2), 255-276.
- Bolondi, G., Cascella, C., & Giberti, C. (2017). Highlights on gender gap from Italian standardized assessment in Mathematics. In J. Novotná & H. Moraová (Eds.), 14th International Symposium Elementary Maths Teaching: Equity and diversity in elementary mathematics education. Prague: SEMT, 393-403.
- Bolondi, G., & Ferretti, F. (2021). Quantifying Solid Findings in Mathematics Education: Loss of Meaning for Algebraic Symbols. *International Journal of Innovation in Science and Mathematics Education*, 29(1), 1-15.
- Bolondi, G., Ferretti, F., & Giberti, C. (2018). Didactic contract as a key to interpreting gender differences in maths. *Journal of Educational, Cultural and Psychological Studies*, 18, 415-435.
- Buckley, S. (2016). *Gender and sex differences in student participation, achievement and engagement in mathematics*. [http://research.acer.edu.au/learning\\_processes/18](http://research.acer.edu.au/learning_processes/18).
- Cascella, C., Giberti, C., & Bolondi, G. (2020). An analysis of Differential Item Functioning on INVALSI tests, designed to explore gender gap in mathematical tasks. *Studies in Educational Evaluation*, Vol. 64, March 2020.

- Cascella, C., Giberti, C., & Bolondi, G. (2021). Changing the order of factors does not change the product but does affect students' answers, especially girls' answers. *Educational Science*, 11(5).
- Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and conducting mixed methods research* (3rd ed.). London, UK: Sage.
- Di Tommaso, M. L., Mendolia, S., & Contini, D. (2016). The Gender Gap in Mathematics Achievement: Evidence from Italian Data. *IZA Discussion paper*, n.10053, Bonn.
- Fryer, R. G., & Levitt, S. D. (2010). An empirical analysis of the gender gap in mathematics. *American Economic Journal: Applied Economics*, 2(2), 210-240.
- Giberti, C. (2018). Differenze di genere e misconcezioni nell'operare con le percentuali: evidenze dalle prove INVALSI. *CADMO*, 2, 97-114.
- Giberti, C. (2019a). Bravi e brave in matematica, perché? Convinzioni degli insegnanti e differenze di genere. In B. D'Amore & S. Sbaragli (eds.), *Atti del Convegno Nazionale La didattica della matematica: riflessioni teoriche e proposte concrete*, Castel San Pietro Terme (BO): Incontri con la Matematica XXXIII.
- Giberti, C. (2019b). Differenze di genere in matematica: dagli studi internazionali alla situazione italiana. *Didattica Della Matematica. Dalla Ricerca Alle Pratiche d'aula*, 5, 44 - 69. <https://doi.org/10.33683/ddm.18.5.2>
- Jacobs, J. E., & Bleeker, M. M. (2004). Girls' and boys' developing interests in math and science: Do parents matter? *New directions for child and adolescent development*, 2004(106), 5-21.
- Leder, G., & Forgasz, H. (2008). Mathematics education: new perspectives on gender. *ZDM – The International Journal on Mathematics Education*, 40(4), 513-518.
- Riegle-Crumb, C. (2005). The cross-national context of the gender gap in math and science. *The social organization of schooling*, 227-243.
- Roche, A., & Clarke, D. M. (2004). When does successful comparison of decimals reflect conceptual understanding. *Mathematics education for the third millennium: Towards 2010*, 486-493.
- Sbaragli, S. (2012). Il ruolo delle misconcezioni nella didattica della matematica (Sbaragli Silvia), *I quaderni della didattica. Metodi e strumenti per l'insegnamento e l'apprendimento della matematica*, Napoli: Edises, 2012.
- Steinle, V., & Stacey, K. (2003). Grade-Related Trends in the Prevalence and Persistence of Decimal Misconceptions. *International Group for the Psychology of Mathematics Education*, 4, 259-266.
- Steinle, V., & Stacey, K. (2004). A longitudinal study of students' understanding of decimal notation: An overview and refined results. In *Proceedings of the 27th annual conference of the Mathematics Education Research Group of Australasia* (Vol. 2, pp. 541-548).