

Research Articles

Sex Differences in Mathematics and Reading Across Italy: The Role of Socioeconomic Factors

Lorenzo Esposito¹, Enrico Toffalini², David Giofrè¹

¹ Disfor, University of Genoa, ² DPG, University of Padua

Keywords: sex differences, mathematics, reading, socioeconomic factors, meta-analysis

<https://doi.org/10.65550/001c.155775>

Intelligence & Cognitive Abilities

Vol. 2, Issue 1, 2026

This meta-analysis examined over 26 million observations from Italy's INVALSI surveys, conducted between 2009 and 2023, including students in second, fifth, eighth, tenth, and thirteenth grade across the entire Italian territory.

Preliminary analyses revealed consistent patterns: boys outperformed girls in mathematics, $d = .25$, while girls outperformed boys in reading, $d = -.29$. Regarding mathematics, the geographical gradient was strongly correlated with average performance, $r = -.528$, with a North-South difference corresponding to approximately 2.77 IQ units. Furthermore, the sex difference in mathematics was strongly associated with the geographical gradient, $r = -.817$, indicating larger gaps in Northern provinces, and positively correlated with grade level, $r = .386$, showing that the boys' advantage widens in higher grades. As for reading, the geographical gradient was similarly correlated with average performance $r = -.603$, with a North-South difference of approximately 3.44 IQ units. The sex difference in reading showed a weaker association with the geographical gradient, $r = -.181$, but a strong negative correlation with grade level, $r = -.668$, indicating that the girls' advantage increases substantially in higher grades.

Meta-regression models identified the geographical gradient, employment rate ratio, university enrollment ratio, and mean achievement score as key moderators. Specifically, for mathematics, higher women's university enrollment was associated with smaller sex differences, whereas higher employment equality and higher overall achievement were associated with larger sex differences, favoring boys. For reading, higher women's university enrollment and higher mean achievement were associated with larger girls' advantages. Collectively, these results suggest that while high-performing environments tend to amplify sex differences across domains, economic productivity and specific opportunity structures can act as protective factors, offering new insights into the gender-equality paradox.

1. Introduction

Research on sex differences in mathematics and reading remains a topic of ongoing debate, as men and women perform similarly in several domains, such as education. On the one hand, some studies showed differences in favor of boys in mathematics in certain countries, though not universally (e.g., Halpern, 2012; Maccoby & Jacklin, 1977). Conversely, girls generally outperform boys in reading across almost all countries (Stoet & Geary, 2018). Beyond reflecting curriculum mastery, it has been argued that standardized measures of academic achievement may also serve as proxies for general cognitive ability. Consistent with this view, research showed that performance across different academic domains is strongly correlated, suggesting the influence of a common underlying cognitive factor (Rindermann, 2007). Recent neurodevelopmental evidence sup-

ports this view, identifying common cognitive control networks engaged in both mathematics and reading, alongside domain-specific spatial and language contributions (Ünal et al., 2025). Adopting this perspective, analyzing these patterns provides insight into how macro-structural factors might interact with cognitive abilities.

Several factors were proposed to explain differences in cognitive abilities. Since academic competencies like reading and mathematics do not emerge without formal schooling, they inherently involve a strong environmental component. However, research suggested that biological factors may contribute to these differences. Biological aspects, such as brain structure, prenatal hormone exposure, and specific sex differences in the language system and spatial abilities (e.g., mental rotation), were highlighted as possible evolutionary or physiological precursors that might impact learning trajectories (Eals & Silverman, 1994; Finegan

et al., 1992; van der Linden et al., 2017). Other theories, which argued that innate differences alone cannot fully account for sex differences, emphasized societal and cultural influences as explanations for sex differences (e.g., Halpern, 2012). For example, social and educational contexts, such as Gross Domestic Product (GDP) and economic opportunities, were proposed as contributors to the observed sex differences (Breda & Napp, 2019; Ceci et al., 2009; Eagly & Wood, 1999; Else-Quest et al., 2010). For example, higher GDP is often connected with better-resourced schools, higher-quality instruction, and increased access to educational tools and extracurricular activities (Else-Quest et al., 2010). Similarly, increased economic opportunities for women, such as increased women's participation in research and labor markets, can help to overcome structural and societal barriers. According to Social Structural Theory, these factors affect the division of work and society expectations, impacting both parental and individual conceptions of what talents are desirable and attainable (Eagly & Wood, 1999). Also, according to Expectancy-Value Theory, when girls perceive larger returns on mathematics investment, both in education and possible jobs, they are more inclined to value and engage in mathematical study (Eccles, 1983).

Taken together, the evidence suggests that neither biological nor sociocultural explanations are sufficient on their own, and that sex differences can only be understood by considering the influence of both.

1.1. Relationship between national development and sex differences

As noted above, societal and cultural factors were proposed as explanations for sex differences. For example, some evidence suggested that the level of gender equality, defined as the parity in access to resources, opportunities, and rights across political, economic, and social domains (e.g., Global Gender Gap Index), was linked to the sex differences in mathematics, with more developed countries showing smaller differences (Guiso et al., 2008; Marks, 2008; Reilly, 2012; Román-González et al., 2017). Other studies reported no significant or mixed relationships between gender-equality indices and sex differences in mathematics and achievement in general (Else-Quest et al., 2010; Hamamura, 2012; Ireson, 2017; Munir & Winter-Ebmer, 2018; Penner, 2008). Intriguingly, Stoet and Geary (2018), using data from the Programme for International Student Assessment (PISA), found a negative relationship between sex differences in intraindividual academic strengths and the level of gender equality. This pattern, termed the “educational-gender-equality paradox”, suggested that relative to their own performance in reading, boys outperform girls in mathematics in most countries even those with high levels of gender equality. This phenomenon was recently replicated using PISA and TIMSS data, confirming the robustness of these intraindividual patterns across different assessments (Balducci et al., 2024). This seemingly paradoxical result aligns with expectancy-value theory (Eccles, 1983), which proposes that in more gender-equal countries, where economic pressures

are lower and life satisfaction is higher, individuals have greater freedom to pursue educational paths that match their personal interests, rather than being constrained by economic pressures. Specifically, this framework also posits that academic choices are driven by intraindividual comparisons of strengths (e.g., relative competence in science versus reading) and domain-specific interests. Consequently, in contexts where individuals can prioritize interests over contextual pressures, sex differences in STEM engagement could become more pronounced.

In this framework, a recent study by Stoet and Geary (2022) confirmed that women are more likely to pursue people-oriented careers, while men tend to choose things-oriented occupations. Also, they found that sex differences are more pronounced in countries with higher levels of gender-equality, further supporting the gender-equality paradox previously observed in university course enrollment and professional choices (Lippa et al., 2014). Other research indicates that the magnitude of sex differences in academic performance, particularly in mathematics and reading, can vary across countries and could be influenced by living conditions (Stoet & Geary, 2013). For instance, in countries with higher living standards, the sex difference in mathematics tends to be smaller, whereas the difference in reading comprehension is larger. These findings suggest that living conditions can affect academic disparities (Herlitz et al., 2023). In addition, Balducci et al. (2024), using recent PISA and Trends in International Mathematics and Science Study (TIMSS) data, suggested that in less gender-equal countries, pressures related to life quality might promote greater engagement of women in STEM fields, thus closing the gap between men and women (Balducci, 2023).

Socioeconomic factors, such as Gross Domestic Product (GDP) and economic opportunities, were proposed as contributors to the sex differences (Breda & Napp, 2019; Ceci et al., 2009; Eagly & Wood, 1999; Else-Quest et al., 2010). Notably, sex differences in academic achievement show distinct patterns in mathematics and reading in relation to economic development. For example, in mathematics, the boys' advantage generally decreases in countries with higher living conditions, including higher national wealth or better educational resources (Herlitz et al., 2023). However, cross-national studies, including PISA analyses, suggested that macroeconomic indicators, such as per capita GDP, are not always correlated with sex gaps in mathematics (Marks, 2008). In a recent work, Giofrè and colleagues (2020), using INVALSI data, examined the sex differences in mathematics and reading in Italy, and found a considerable association between sex differences in mathematics and the GDP, with northern regions, which are wealthier than southern regions, showing larger sex gaps.

1.2. The Italian “situation”

Italy provides a valuable opportunity to study sex differences in mathematics and reading achievement, as it is widely recognized that significant social and cultural differences exist among Italy's territories (Felice, 2011). The difference in social opportunities between men and women is particularly pronounced in the southern part of the country

(Checchi & Peragine, 2010). In fact, persistent disparities between Northern and Southern (and Isles) regions were documented across several areas, including added value (Proietti, 2005), unemployment rate (Carmeci & Mauro, 2002), industrial development (Dunford, 2008), socioeconomic resilience (Cellini & Torrisi, 2014), labor (Patacchini, 2008) and agricultural productivity (Esposti, 2011).

For these reasons, it would be misleading to treat Italy as a uniform entity (Felice, 2011). Instead, differences should be considered and examined at more detailed levels, such as regional or provincial. Regions in Italy, in fact, have separate governance and reflect distinct historical and cultural backgrounds, and currently there are variations in dialects, cuisine, artistic traditions, and other cultural dimensions.

1.3. The current study and hypotheses

For the present meta-analysis, an exceptionally large number of data points, over 23 million observations in mathematics and reading, were drawn from the INVALSI survey, while relevant social and economic moderating factors were obtained from the Italian National Institute of Statistics (ISTAT). Each year, INVALSI (Istituto Nazionale per la VALutazione del Sistema di Istruzione [National Institute for the Assessment of the Instruction System]) evaluates the entire Italian child population in the 2nd, 5th, 8th, 10th, and 13th grades. In the Italian school system, children typically attend the 2nd grade at seven years of age, the 5th grade at ten, the 8th grade at thirteen, the 10th grade at fifteen, and the 13th at eighteen. Mathematics items cover different domains, including algebra, geometry, and statistics. Reading literacy items are divided into two main areas: reading comprehension and grammar. The INVALSI data spans multiple years, geographic regions, and grades, all within the same school system that follows a nationally regulated curriculum.

Considering previous evidence on this topic, sex differences favoring boys in mathematics are expected, whereas a girls' advantage is hypothesized for reading (Giofrè et al., 2020). It is also reasonable to expect regional differences across Italy, with northern areas showing an advantage in both mathematics and reading. Additional analyses were conducted to examine the contribution of socio-economic factors to the sex differences and their association with geographical region. In particular, we aimed to explore the so-called "gender-equality paradox", which suggests that areas with higher levels of gender-equality may also exhibit larger sex differences (Stoet & Geary, 2018). Within this framework, a positive relationship between socio-economic development and sex differences in mathematics and reading achievement is expected. To assess whether the observed associations extend to overall academic performance, relationships between socio-economic factors and "general achievement ability" were also tested.

2. Methods

2.1. Materials

Data from the INVALSI surveys on the Italian student population (grades 2, 5, 8, 10, and 13) for mathematics and reading from 2009 to 2023 were analyzed. Tasks differ for each grade and year of survey. Each response was coded as either correct ("1") or incorrect ("0"). For both mathematics and reading literacy, the total scores were created based on IRT (Item Response Theory) via WLE method (Weighted Likelihood Estimation). More information about the test, with examples, is available at the INVALSI website (<https://INVALSI-areaprove.cineca.it/>). Crucially, the WLE scores provided by INVALSI are derived from a Rasch Model calibration (Item Response Theory). INVALSI procedures aim to minimize Differential Item Functioning (DIF) across subgroups through rigorous pre-testing (INVALSI, 2015, 2016, 2017). Therefore, we assume approximate measurement invariance across sex, although we acknowledge that DIF cannot be completely ruled out. The study analyzed secondary, de-identified population data provided by the INVALSI. According to national regulations for secondary data analysis, Institutional Review Board (IRB) review was not required.

A total of 26,645,542 observations for mathematics and 26,553,867 observations for reading were obtained. Data for the 2019-2020 academic year was unavailable due to the COVID-19 pandemic. In addition, in years 2009-2012 data was not available for "Bolzano", "Monza e della Brianza", "Fermo", "Forlì-Cesena", "Pesaro e Urbino", "Barletta-Andria-Trani", and "Sud Sardegna"; in years 2013-2014 for "Monza e della Brianza", "Fermo", and "Sud Sardegna"; in years 2014-2017 for "Sud Sardegna".

For the purpose of this study, it was decided to build a "geographical gradient" variable able to synthesize the geographical coordinates of the provinces in single quantitative value. To this aim, the geographical gradient was built by labelling provinces which were ordered simultaneously by latitude (from North to South) and longitude (from West to East) with a progressive number (from 1 to 107). To do so, an angle of 40° was chosen, then longitude and latitude coordinates were rotated clockwise. Subsequently, each province was sorted by its new latitude coordinate, corresponding to its orthogonal projection on the new axis, from the northernmost to the southernmost. Each province was assigned a numerical value according to this order. For example, "Venezia", which is a northern province was assigned "40", while "Napoli" which is a southern province was assigned "85".

Relevant and well-known differences among northern and southern Italian regions concerning socio-economic factors were analyzed. We therefore collected data on a series of socioeconomic factors related to the geographical gradient to input as moderators in the meta-analyses. This data was retrieved and elaborated from the Italian National Institute of Statistics, (ISTAT) website (<https://esploradati.istat.it/databrowser/>).

The added value per inhabitant, defined as the contribution of the local production system to the national GDP

(a proxy for economic productivity), was measured year-by-year from 2009 to 2022 (most recent available data) for each Italian province. High correlations were observed among the yearly measurements, with a minimum $r = .87$, indicating that the values did not significantly differ from year to year. To simplify the analysis and avoid redundancy without losing meaningful information, the mean value of the measurements across years was computed. The added value is expressed in tens of thousands of euros to facilitate interpretability.

The employment rate was measured year-by-year from 2009 to 2023 (most recent available data) for each Italy's province (for 15–64 years-old people) and represents the ratio between the women's employment rate and men's employment rate. Again, given the high correlation among the years, minimum $r = .88$, the mean value of the measurements across years was computed.

The percentage of women/men in town administrations was measured year-by-year from 2009 to 2023 (most recent available data) for each Italy's province and represents the ratio between the percentage of women and men town administrators. The moderate correlations among the years, minimum $r = .43$, did not suggest an overlap among the yearly measurements. To avoid distortion in the data, only the most updated info was considered, instead of computing the mean value.

University enrollment rate was measured year-by-year from 2014 to 2020 (most recent available data) for each Italy's province and represents the ratio between the percentage of women and men who chose to enroll at university. Given the moderate-to-high correlations among the years, minimum $r = .65$, the mean value of the measurements across years was computed.

The hourly gross wage for employees was measured year-by-year from 2014 to 2021 (most recent available data) for each Italy's province and represents the ratio between women's wage and men's wage. Again, given the high correlation between the years, minimum $r = .90$, the mean value of the measurements across years was computed. It is worth mentioning that for all the above-mentioned ratios, the natural logarithm was computed to enhance interpretability of the statistical analyses.

Two additional moderators were computed. First, to account for potential selection bias in upper secondary education (grades 10 and 13), we collected data on the share of students enrolled in STEM-oriented tracks (e.g., scientific high schools and technical institutes) for each province, serving as a proxy for track composition. Second, to test whether sex differences are moderated by the overall performance level, the mean provincial achievement score, averaged across boys and girls, was computed.

2.2. Data analysis

Preliminary descriptive analyses included Pearson correlations between geographical gradient and achievement gaps, and the conversion of effect sizes into standard IQ units to facilitate comparison with literature on cognitive abilities literature. A meta-analytic approach was used to estimate sex differences in mathematics and reading. Such

an approach was used even though we had virtually the entire population over the mentioned time span, because data can be still considered as a sample of a much wider set of possible observations. Meta-analyses were conducted by using the “metafor” package (Viechtbauer, 2010) within the RStudio environment (R Core Team, 2023; RStudio Team, 2023).

For each year and grade level in every Italian province, the standardized difference in performance between boys and girls (Cohen's d) and its variance were computed following Borenstein et al. (2011) formula. Boys and girls were coded as the first and second group respectively, therefore a positive Cohen's d indicates a mean difference favoring boys, whereas a negative Cohen's d indicates a mean difference favoring girls. In the present work, the statistical unit is represented by the effect size for a given province in a given survey year. Therefore, since multiple survey years are examined in the meta-analyses, multiple effect sizes were coded for each year, and multiple effect sizes were coded for each province.

Firstly, meta-analytic models were fitted for each grade level separately to facilitate interpretation of the sex difference in mathematics and reading. Random-effects estimation was applied to address heterogeneity across survey years (Borenstein et al., 2011). This allowed to obtain initial meta-analytic estimates of the sex difference over the grade levels considered. In a second step, a series of possible moderators was tested. At all stages, both provinces and survey years were included as crossed random effects. Throughout these steps, model evaluation and goodness of fit relied on multiple indices. As for absolute indices, “ τ ” (tau) and R^2 were examined. The first was considered an indication of the heterogeneity; in meta-analyses with multiple levels of random effects (i.e., crossed effects), this index was computed as the sum of the single heterogeneity indices for each random level (Konstantopoulos, 2011). The second index was considered an indication of the variance explained by the model, and it was computed starting from the “ τ^2 ”, as the difference between the heterogeneity of the model with less parameters and the one with more parameters, divided by the heterogeneity of the model with less parameters, e.g., $(\tau_1^2 - \tau_2^2) / \tau_1^2$ (Viechtbauer, 2010). Additionally, Intraclass Correlation Coefficients (ICCs) were computed for both province and year levels to explicitly quantify the proportion of variance attributable to geographical and temporal clustering. Variable selection followed a two-step procedure to ensure stability. First, an initial screening was performed using a stepwise selection based on the Bayesian Information Criterion (BIC) to identify a parsimonious set of predictors (Burnham & Anderson, 2007). Second, to address potential multicollinearity among the correlated socioeconomic and geographical indicators, and to validate the stepwise selection, we employed a Ridge regression. This allowed us to identify any suppressed effects (e.g., economic factors masking geographical ones) that the conservative BIC threshold might have discarded. Also, to address potential multicollinearity, we examined the moderator correlation matrix and Vari-

ance Inflation Factors (VIFs) (see Supplementary Materials).

The comparison between the best model and the null model was performed via likelihood ratio test for nested models, based on the χ^2 (chi-square) distribution, where statistically significant results indicate a difference between two models; this was done to ensure that the final model, taken together, significantly explains any heterogeneity. Parameters were estimated considering their weights, computed as $w_i = 1/(\tau^2 + v_i)$, where v_i is the sampling variance.

To verify the robustness of the selected models against potential spatial misspecification, we performed a sensitivity spatial analysis in which the linear geographical gradient was replaced by a two-dimensional thin-plate regression spline on longitude and latitude. The model retained the random effect for survey years to account for the repeated measures structure. Residual correlation was subsequently assessed using Moran's I test (see Supplementary Materials). To understand the individual contribution of each moderator to the final model, a series of additional meta-analyses was conducted, in which each moderator was entered in isolation. To address the potential limitation of averaging moderators across multi-year windows, we conducted a temporal robustness check (see Supplementary Materials). We re-estimated the key meta-regression models using year-specific, lagged (t-1) values for the moderators to account for temporal precedence. For the employment rate ratio, historical data were available for the entire study period, allowing for a complete lagged analysis. However, for the university enrollment ratio, historical data availability limited the lagged analysis to specific survey waves. Therefore, the models for mathematics, examining grade 13, covered the 2018 and 2020 waves, while the analysis for reading, examining grade 10, was restricted to the 2017 and 2018 waves. Also, to rule out the possibility that the aggregate results were driven by specific influential outliers, a leave-one-out influence analysis was performed by iteratively removing one province at a time and re-estimating the models (see Supplementary Materials).

Finally, to capture the general pattern of sex differences across subjects (e.g., mathematics and reading), we performed a Principal Component Analysis (PCA) on the SMDs in mathematics and reading for each province-year combination. The composite SMD was then used in a multilevel meta-analytic model to serve as a proxy for "general scholastic sex-difference factor", aligning the analysis with the broader literature on general cognitive differences.

3. Results

3.1. Preliminary analysis

A preliminary analysis showed that mathematics and reading scores were strongly and similarly correlated across all grades (see Table 1). Results from the first series of meta-analyses on sex differences showed a boys' advantage in mathematics increasing with grade; conversely, a similar advantage for girls was observed for reading (Figure 1). However, while the sex difference in reading performance

was relatively homogeneous and increased steadily with grade across Italian provinces, the sex difference in mathematics performance followed a sharp geographical gradient, with the most northern Italian provinces presenting a larger sex difference than the most southern provinces (see Figure 2). The geographical gradient showed a strong correlation with overall mean math achievement, $r = -.528$. In other words, the performance gap between the northernmost and southernmost provinces corresponds to approximately 2.770 IQ units, with southernmost provinces showing lower IQ units. As for reading, again the geographical gradient showed a strong correlation with overall mean reading achievement, $r = -.603$. In fact, the performance gap between the northernmost and southernmost provinces corresponds to approximately 3.444 IQ units, with southernmost provinces showing lower IQ units. The geographical gradient was also significantly associated with the magnitude of the sex difference in mathematics, $r = -.817$, suggesting that sex differences tend to widen in northern provinces, whereas it was weakly correlated with sex difference in reading, $r = -.181$. Finally, consistent with the developmental trend observed (see Figure 1), grade level was positively correlated with the sex difference in mathematics, $r = .386$, meaning that the difference tends to widen in higher grades, with boys showing higher performance; conversely, in reading, the sex difference tends to favor girls in higher grades, $r = -.668$. These strong correlations across domains suggest the presence of a common underlying component, aligning with the perspective that these scores can be interpreted as indicators of general scholastic ability.

3.2. Analysis of moderators

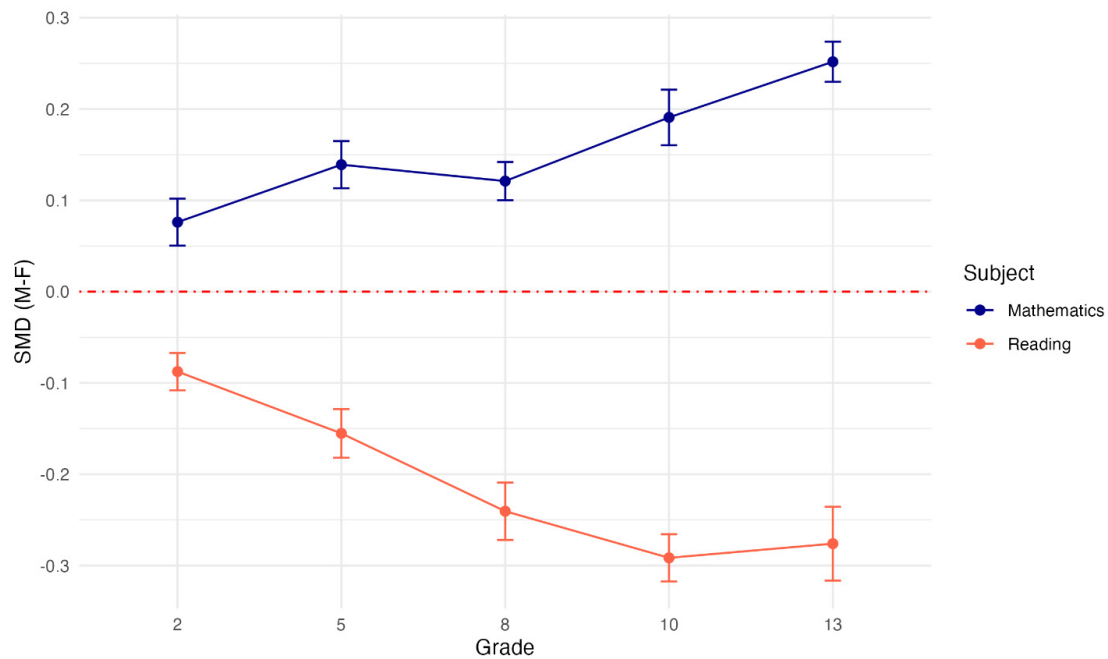
In-depth analyses were conducted on data for the mathematics performance of 13th graders, because this is the grade in which the sex differences observed are the largest, $d = 0.252$, indicating that the average boy scored higher than approximately 60% of girls. Focusing on the grade with the largest achievement gap allowed us to more precisely identify the factors influencing sex differences. Firstly, a null model with only random effects (i.e., provinces and survey years), was performed. Secondly, to address the high heterogeneity, several moderated meta-analyses were performed, $BIC = -906.764$, $\tau = .093$, $\sigma_1 = 0.092$, $\sigma_2 = 0.013$, $ICC_1 = .980$, $ICC_2 = .020$. As for reading, further analyses were conducted on data of 10th graders, $d = -0.292$, indicating that the average girl scored higher than approximately 61.5% of boys. Again, due to the high level of heterogeneity, moderated meta-analyses were conducted to address this issue, $BIC = -1145.219$, $\tau = .007$, $\sigma_1 = 0.048$, $\sigma_2 = 0.037$, $ICC_1 = 0.626$, $ICC_2 = 0.373$.

Preliminarily, we analyzed the correlation matrix of all potential moderators. As expected, the geographical gradient showed strong associations with key socio-economic indicators, particularly with the employment rate ratio, $r = -.750$, and mean added value, $r = -.744$. Strong correlations were also observed among economic variables; notably, the mean added value was highly correlated with the employment rate ratio, $r = .800$, and the university enroll-

Table 1. Correlations between mathematics and reading scores; additionally, SMD listed by grade level.

Grade	Correlations		Mathematics		Reading	
	<i>r</i>	95% CI	SMD	95% CI	SMD	95% CI
2 nd	.873	[.861, .885]	0.076	[0.050, 0.102]	-0.088	[-0.108, -0.067]
5 th	.691	[.661, .717]	0.139	[0.113, 0.165]	-0.155	[-0.182, -0.129]
8 th	.769	[.747, .790]	0.121	[0.100, 0.142]	-0.24	[-0.272, -0.209]
10 th	.919	[.899, .921]	0.191	[0.160, 0.221]	-0.292	[-0.317, -0.266]
13 th	.960	[.958, .971]	0.252	[0.230, 0.274]	-0.276	[-0.316, -0.236]

Note. Correlations between mathematics and reading scores are reported. Standardized mean difference between boys and girls are reported in mathematics and reading. SMD = Standardized Mean Difference.

**Figure 1. Graphical representation of the SMD between boys and girls in mathematics (blue) and reading (red) as a function of grade.**

Note. SMD = Standardized Mean Difference.

ment ratio, $r = -.742$. Furthermore, the university enrollment ratio showed a substantial negative correlation with the employment rate ratio, $r = -.601$, while the wages ratio correlated moderately with added value, $r = -.663$, and the geographical gradient, $r = .565$. Conversely, the ratio of women administrators showed generally weaker associations with the other predictors, ranging from $r = -.360$ to $r = .342$. Crucially, the mean achievement score showed very strong correlations with mean added value, $r = .794$, and the geographical gradient, $r = -.753$, indicating that overall performance is linked to the economic and geographical context. On the other hand, the percentage of students enrolled in STEM tracks showed negligible correlations with the main socio-economic and geographical predictors, with coefficients ranging from $r = -.118$ to $r = .147$, suggesting that track composition operates independently of the broader socio-economic framework. To address the stability of the variable selection in the presence

of correlated predictors, we inspected the VIFs for the full set of candidate moderators (see Supplementary Materials).

Subsequently, we performed a series of meta-regression models to identify the best-fitting model explaining sex differences across provinces, followed by additional analyses aimed at assessing the unique contribution of each moderator (see Table 2). As for mathematics, the employment rate ratio, university enrollment ratio together with the geographical gradient and mean achievement score, resulted the best moderators of the sex difference. To address potential multicollinearity and verify the stability of the stepwise selection, we performed a penalized meta-regression using the Ridge method, including all candidate moderators (see Supplementary Materials). Our final model including the employment rate and university enrollment ratios together with the geographical gradient and averaged total score, explained about 79.815% of the variance, $BIC = -1044.794$, $\tau = 0.042$, $\sigma_1 = 0.042$, $\sigma_2 = 0.001$, $ICC_1 = .999$, $ICC_2 = .000$.

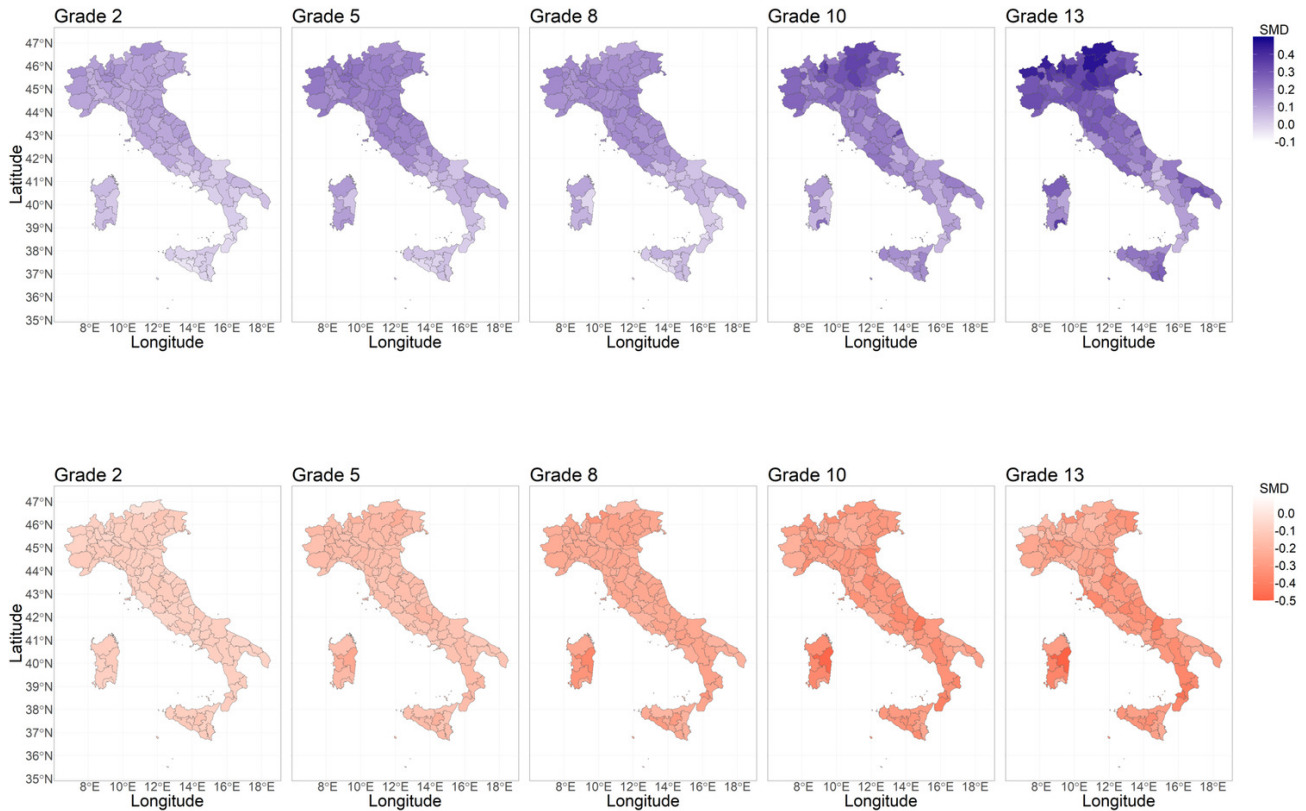


Figure 2. Spatial distribution of sex differences in mathematics (blue) and reading (red) across Italian provinces.

Specifically, the sex difference significantly decreased as the employment rate ratio, $\beta = -0.032$, $SE = 0.008$, 95% CI $[-0.046, -0.017]$, and the university enrollment ratio, $\beta = -0.052$, $SE = 0.006$, 95% CI $[-0.063, -0.041]$, increased. Furthermore, a significant geographical gradient was observed, $\beta = -0.036$, $SE = 0.008$, 95% CI $[-0.051, -0.021]$. Finally, the averaged total score showed a positive association with the sex difference, $\beta = 0.033$, $SE = 0.004$, 95% CI $[0.025, 0.040]$, suggesting that provinces with higher mean achievement tended to exhibit larger boys' advantages (see Supplementary Materials). Subsequently, we tested moderators again, one at a time. Regarding the geographical gradient, the sex difference significantly decreases as the gradient increases, $\beta = -0.064$, $p < .001$. Interestingly, these findings show that the more men and women are equally employed the more the sex difference increases, $\beta = 0.048$, $p < .001$. Also, the sex difference in mathematics decreases as the women-to-men university enrollment ratio increases, $\beta = -0.073$, $p < .001$. Although women's enrollment is typically higher than men's, regions where this advantage is larger tend to exhibit smaller sex gaps in math. Finally, the mean achievement score was positively associated with the gender gap, $\beta = 0.059$, $p < .001$, indicating that provinces with higher overall performance levels tend to show more pronounced boys' advantages (see Supplementary Materials).

As for reading, the stepwise selection found the university enrollment ratio together with the mean score, to be

the best moderators of the sex difference. Our final model including the university enrollment ratio with the mean achievement score, explained about 6.046 % of the variance, $BIC = -1210.385$, $\tau = 0.058$, $\sigma_1 = 0.046$, $\sigma_2 = 0.036$, $ICC_1 = .614$, $ICC_2 = .386$. Specifically, the university enrollment ratio, $\beta = -0.033$, $SE = 0.005$, 95% CI $[-0.042, -0.024]$, and the mean achievement score, $\beta = -0.021$, $SE = 0.003$, 95% CI $[-0.027, -0.015]$, were negatively associated with the effect size, indicating larger girls' advantages (i.e., the effect size becomes more negative) in contexts with higher women-to-men university enrollment ratio and mean performance. Subsequently, we tested moderators again, one at a time (Table 2). The university enrollment ratio reached the threshold for statistical significance, $\beta = -0.025$, $p < .001$, as well as the mean achievement showed a significant negative association, $\beta = -0.019$, $p < .001$, indicating that in higher-performing provinces, the girls' advantage tends to be larger. As mentioned above, spatial sensitivity analysis and lagged temporal models were performed, highlighting robust and consistent results (see Supplementary Materials).

3.3. General scholastic sex-difference factor

A PCA was conducted across mathematics and reading by each province and assessment's year combination. The first principal component accounted for the majority of variance, $\lambda_{Math} = -.707$, $\lambda_{Reading} = -.707$, indicating that this

Table 2. Fit measures and standardized coefficients of the final predictors for mathematics (top) and reading (bottom)

Moderator	BIC	τ	σ_1	σ_2	ICC ₁	ICC ₂	R ²	χ^2	p(χ^2)	β	SE	95% CI	p	ΔR^2
<i>Mathematics</i>														
Null Model	-906.764	0.093	0.092	0.013	.980	.020								
Geographical Gradient	-966.472	0.067	0.066	0.012	.964	.036	47.526	65.767	< .001	-0.064	0.007	[-0.077, -0.051]	< .001	5.081
Employment Rate Ratio	-933.448	0.079	0.078	0.013	.974	.026	27.192	32.743	< .001	0.048	0.008	[0.033, 0.064]	< .001	4.620
University Enrollment Ratio	-999.449	0.057	0.055	0.012	.950	.050	63.193	98.744	< .001	-0.073	0.006	[-0.084, -0.062]	< .001	20.024
Mean Achievement	-979.112	0.063	0.062	0.010	.974	.026	54.002	78.407	< .001	0.059	0.006	[0.049, 0.07]	< .001	4.683
<i>Reading</i>														
Null Model	-1145.219	0.060	0.047	0.037	.627	.373								
University Enrollment Ratio	-1169.528	0.054	0.041	0.037	.551	.449	17.19	31.161	< .001	-0.025	0.004	[-.033, -.017]	< .001	6.046
Mean Achievement	-1175.617	0.067	0.055	0.036	.701	.299	0	37.251	< .001	-0.019	0.003	[-0.025, -0.013]	< .001	0

Note. ^a Values reported as .000 are due to rounding to three decimal places. Employment rate ratio and university enrollment ratio are expressed as the logarithm of the F/M ratio. σ_1 = heterogeneity at province level. σ_2 = heterogeneity at year level. ICC₁ = intraclass correlation at province level. ICC₂ = intraclass correlation at year level.

component captures a general scholastic sex-difference factor. So, the first principal component was considered for the meta-regression, including province-level and year-level random effects. First, a null model was performed showing no statistical difference in the general scholastic sex-difference factor, $d = 0.001$, $p = .993$, $BIC = 47470.225$, $\tau = 1.109$, $\sigma_1 = 1.103$, $\sigma_2 = 0.114$, $ICC_1 = .989$, $ICC_2 = .011$. Subsequently, a moderated meta-analysis was performed including all the focal moderators. Results showed that higher women-to-men employment rate ratio was associated with larger general scholastic sex difference favoring girls, $\beta = -0.554$, $SE = 0.108$, $p < .001$. Similarly, higher women-to-men university enrollment ratios predicted larger general scholastic sex difference favoring girls, $\beta = -0.647$, $SE = 0.079$, $p < .001$. The geographical gradient remained a significant predictor, $\beta = -0.424$, $SE = 0.103$, $p < .001$, indicating that northern provinces continue to show larger sex differences than southern provinces. Finally, the averaged total score showed a positive association, $\beta = 0.401$, $SE = 0.017$, $p < .001$. Residual heterogeneity was approximately $\tau = 0.630$, $\sigma_1 = 0.622$, $\sigma_2 = 0.106$, $ICC_1 = .974$, $ICC_2 = .026$. This model explained approximately 67.720% of heterogeneity compared to the null model, $BIC = 44481.1576$, $\chi^2(4) = 3015.426$, $p < .001$.

4. Discussion

The topic of sex differences in mathematics and reading is particularly relevant today, since these skills are linked to educational achievement, professional qualifications, adult socioeconomic status, and interest in scientific disciplines (Ritchie & Bates, 2013). In this regard, the INVALSI survey, covering millions of students across all Italian provinces over multiple years, allows for a detailed and reliable examination of sex differences in mathematics and reading. At the same time, it offers new insights within a single country, in which the education system is the same in the whole country, despite territorial differences in school achievement and other societal aspects. From this perspective, the Italian context provides a unique opportunity to investigate the gender-equality paradox: despite a uniform legal and institutional framework, regional economic heterogeneity leads to situations in which apparent gender equality, such as similar wage, reflects equal disadvantage due to poverty rather than genuine equality.

Our findings showed that boys outperform girls in mathematics, and girls perform better in reading; interestingly, these sex differences tend to increase with grade level, demonstrating more considerable differences in higher grades. Notably, this study found considerable heterogeneity in sex differences across provinces, particularly in mathematics. Specifically, northern Italian provinces showed a larger sex difference favoring boys, whereas southern provinces showed smaller differences. These results are consistent with previous literature highlighting regional differences (Giofrè et al., 2020) and suggest that the socioeconomic environment could influence the size of sex differences in academic achievement. Specifically, the pattern observed along the Italian country mirrors cross-national findings, where larger sex differences in mathematics are

frequently observed in more developed and gender-equal contexts (Stoet & Geary, 2013).

4.1. Current findings

As for mathematics, the employment rate and the university enrollment ratios were included alongside the geographical gradient and mean achievement score. Interestingly, our results suggest that differences among provinces could still account for a substantial portion of the heterogeneity in the mathematics sex difference, which is not explained by the socioeconomic variables.

Regions with higher women-to-men employment equality (predominantly in the North) nonetheless display larger sex differences in mathematics, while regions with a more unfavorable women's employment ratio (predominantly in the South) show smaller gaps. This result is not entirely in line with the view that greater equality is associated with reduced sex differences in achievement (Guiso et al., 2008; Marks, 2008; Reilly, 2012; Román-González et al., 2017). Conversely, it is consistent with the gender-equality paradox, according to which gender-equality is associated with larger sex gaps in achievement. A potential caveat is that, in Italy, official employment data could underestimate women's actual occupation, especially in southern regions where undeclared work is more common. So, the women-to-men ratio could appear lower than it really is, thereby biasing the association between gender equality and mathematics achievement. Moreover, it is critical to distinguish between the univariate and multivariate relationships observed with the employment rate ratio. While the unadjusted ecological association revealed a positive relationship with the gender gap, consistent with the gender-equality paradox, its effect shifted negative in the full multivariate model. This reversal is most likely due to a suppression or confounding effect, which is principally driven by the significant association between the geographical gradient and achievement. As a result, the partial negative estimate in the full model should be interpreted considering this interdependence.

University enrollment negatively predicted the sex difference in mathematics, specifically, the higher the women-to-men enrollment ratio, the smaller the mean difference in mathematics. Since in most cases the women-to-men university enrollment ratio is higher than one, that is, women are generally more enrolled than men, this finding indicates that the disproportion favoring females in enrollment is associated with a reduced sex gap in mathematics. In other words, this evidence suggests that when women are overrepresented in university enrollment then the sex difference in mathematics tends to close. These results could be explained by multiple factors, for example, territories with more women's enrollment could represent cultural circumstances that favor women's educational success. As a consequence, higher women representation in tertiary education appears to weaken the boys' advantage in mathematics. This last finding could look partly in contrast with the "gender-equality paradox", according to which, a higher women representation in tertiary education should be associated with larger gaps in achievement. How-

ever, in our results, we observed the opposite trend, specifically, the overrepresentation of women in university was associated with smaller sex difference in mathematics. As mentioned above, these results require careful interpretation and should be contextualized in the Italian framework. For example, internal migration from the South to the North could play a role, for instance, students who relocate for university may come from higher socioeconomic backgrounds. Such societal factors could concentrate high-performing girl students in northern regions, potentially contributing to regional variations in the women-to-men enrollment ratio and influencing the observed sex differences in mathematics. Also, while gender equality initiatives are the same throughout the country, the North is wealthier and more developed than the South, which has higher rates of poverty and unemployment. In a recent study, Giofrè et al. (2020) showed that regional GDP per capita is associated with sex differences in mathematics, in fact, the boys' advantage was larger in the wealthier northern regions and smaller in the poorer southern regions. However, GDP explained only part of the observed heterogeneity, suggesting that socioeconomic conditions could influence the sex gap. Thus, the apparent equality in some southern locations could not be acknowledged as evidence of gender equality, but rather of equal disadvantage, highlighting how socioeconomic poverty could distort the interpretation of sex gaps. In other words, this divergence suggests a critical paradox, in fact, in Northern regions, a sex gap could reflect a more equal environment, which is favoring individual choice and interest, whereas in poorer and disadvantaged Southern regions, apparent gender equality could arise from a lack of opportunity where choices are impacted by economic pressures (Eccles, 1983).

Moreover, regarding the association between university enrollment and secondary school performance, it is crucial to clarify the direction of causality. Since tertiary education occurs after secondary school, high women's enrollment rates cannot directly cause prior achievement. Instead, we interpret this variable as a proxy for the local opportunity structure and educational aspirations. It is worth noting that Italian universities do not require a specific number of mathematics courses for enrollment, although admission to many university programs depends on passing entrance tests that frequently include a mathematics component. However, the link between university enrollment and math performance likely reflects the tracking system. Students who intend to pursue university education predominantly enroll in academic high schools, known as *Licei*. These schools typically offer a curriculum designed to provide strong theoretical preparation, including mathematics, to facilitate university admission. Consequently, in provinces with high female university enrollment, a larger proportion of girls likely attended these academic tracks, receiving more extensive mathematics instruction. While internal migration from South to North is a significant factor in Italy, it typically involves students moving after graduation; therefore, it is unlikely to directly influence secondary school differences unless entire families migrate earlier. Consequently, rather than a migration artifact, higher

women representation in tertiary education appears to weaken the boys' advantage in mathematics by reflecting an environment where girls are motivated to invest in STEM-related skills earlier in their trajectory.

Finally, the mean achievement score appeared as a significant predictor, in fact, provinces with higher mean achievement levels had greater sex gaps favoring boys. This finding is consistent with the larger literature on the gender-equality paradox, which suggests that in higher-achieving environments, where educational resources are typically more available, the boys' advantage in mathematics tends to increase. Taken together, these findings lead to an empirical puzzle: while more women's university enrollment might be a protective factor, reducing the difference, higher overall achievement and employment parity are associated with a larger boys' advantage.

It is worth noting that these relationships emerged regardless of the share of students enrolled in STEM-oriented tracks (e.g., scientific high schools) in each province. This suggests that the observed regional disparities are likely not just the result of selection bias into distinct curricular pathways, confirming that broader socioeconomic factors have a direct impact on the sex gap.

As for reading, university enrollment was the only socioeconomic factor that moderated the sex difference in reading, and both the geographical gradient and the employment ratio showed no significant effects on reading sex difference. In other words, higher women's enrollment was associated with the difference in reading beyond the influence of other socioeconomic variables and the geographical gradient, with girls outperforming boys in reading as the number of women in tertiary education increases compared to number of men. One plausible explanation is that women's engagement in tertiary education reflects broader provincial-level cultural and educational conditions that shape gendered patterns of achievement much earlier in development, rather than having a direct influence on secondary school performance. In this sense, tertiary enrollment likely functions as a contextual marker, namely, provinces with higher women's enrollment may already provide stronger educational support for girls, higher academic expectations, and environments that reduce gendered stereotypes, which could enhance girls' skills before the transition to university. This pattern is consistent with cross-national evidence suggesting that reading achievement is a strong predictor of university enrollment (Stoet & Geary, 2020). Consequently, the female advantage in reading likely contributes directly to the higher rates of female participation in tertiary education. It is worth noting that higher GDP and gender equality often covary across regions and countries. It might not be equality per se that drives the association, although it can loosen social and educational expectations, but rather the level of economic resources that allows individuals to pursue their strengths and interests (Stoet & Geary, 2018).

Sex differences in reading were associated with mean achievement, in fact, the higher the mean achievement, the more the girls' advantage increased. However, the amount of variance explained by this moderator was virtually null.

This partly contrasts with what we found in mathematics, where mean achievement was significantly associated with the sex difference and explained a portion of variance. Again, these relationships emerged regardless of the share of students enrolled in STEM-oriented tracks in each province. This suggests that the observed regional disparities are likely not the result of selection bias into distinct curricular pathways.

When considering both domains, a clear pattern emerges regarding how the environment shapes the gap. The mean achievement score acts as an amplifier for the stronger group. In mathematics, provinces with higher overall achievement show a larger advantage for boys. Similarly, in reading, provinces with higher achievement show a larger advantage for girls. This suggests that high-performing contexts tend to magnify the pre-existing strengths of the group that is already ahead. On the other hand, university enrollment consistently shifts the balance toward females. In mathematics, higher female enrollment is associated with a reduced male advantage, likely reflecting that in these contexts, girls intending to pursue university engage in more rigorous preparation. In reading, it corresponds to a larger female advantage, allowing girls to pull further ahead. This suggests that contexts where women invest more in tertiary education consistently characterize environments that support female academic performance across all subjects.

Finally, the results concerning the general scholastic sex-difference factor, defined as a composite factor that integrates both mathematics and reading, provide a strong validation of our theoretical framework. By examining a single underlying dimension, we found that the socioeconomic factors identified in the separate domains are not isolated patterns but could reflect a broader systemic dynamic. The fact that women-to-men employment and university enrollment ratios significantly predicted this general sex difference indicates that regional opportunity structures exert a direct influence on the overall gender gap. This supports the idea that standardized achievement measures capture a shared underlying component of academic differentiation that seems to be impacted by socioeconomic factors. In this sense, the gender equality paradox appears to extend beyond subject-specific outcomes and emerges at the level of systemic academic disparities, shaped by the broader configuration of opportunities available to women and men in each region.

4.2. Limitations

A limitation is the potential effect of item format on sex comparisons. While the INVALSI tests are designed using strong psychometric approaches, including the Rasch Model, to assure internal validity and unidimensionality (INVALSI, 2015, 2016, 2017), the technical documentation does not include specific DIF results by sex. Analytical findings (INVALSI, 2018, 2019, 2021, 2022) show that sex gaps may be altered by test structure, with multiple-choice questions favoring males and open-ended ones favoring females. Furthermore, fluctuations in performance gaps across ability levels and administration types indicate that

functional non-invariance due to item format cannot be completely ruled out.

4.3. Future research

Future research should aim to further clarify the factors underlying sex differences in academic performance across Italy. In particular, it would be valuable to consider additional aspects of the Italian geography and socioeconomic context to refine or replace the current geographical gradient, including variables such as altitude, urban versus rural classification, or the prevalence of mental-health conditions. Moreover, more detailed measures of the educational environment should be considered, such as curriculum classification (STEM versus non-STEM tracks) in tertiary education and career typology (things-oriented versus people-oriented employments). By integrating these finer-grained indicators, future studies could provide a more comprehensive understanding of the mechanisms shaping sex differences in both mathematics and reading, thus providing more targeted educational policies.

4.4. Implications

These results have relevant implications for understanding educational disparities in Italy and similar contexts. The data suggest that addressing the sex difference in mathematics and reading involves more than simply ensuring equal access to education; rather, achievement gaps appear closely linked to specific regional socioeconomic conditions and the structure of educational opportunities. The finding that sex differences in mathematics are larger in northern regions implies that in developed contexts, societal norms or labor market structures may differentially influence educational choices.

Our results also indicate that university enrollment rates, which are higher for women in southern Italy, play a significant role in shaping sex differences in both mathematics and reading. This association suggests that higher levels of women's participation in higher education act as a protective factor against sex differences. Consequently, mechanisms that foster university access for women, particularly in developed areas where it is relatively lower, could represent a key leverage point for reducing disparities in academic achievement.

It is worth noting that in our study, among the socioeconomic moderators, we considered both employment rate and university enrollment. These two indicators are typically combined when calculating the Global Gender Gap Index (World Economic Forum, 2025) for countries worldwide. However, when examining the Italian context, we observed a clear divergence: women are less employed than men, and they show higher enrollment rates in tertiary education. This divergence presents an empirical puzzle, suggesting that the relationship between education and employment is complex and varies across regions. We speculate that these two phenomena may be interrelated, reflecting structural and cultural factors that shape opportunities differently across Italy's territories. Consequently, interpreting gender gaps using the Global Gender Gap In-

dex alone may be misleading, as it does not capture the nuanced interplay between educational achievement and employment within specific regional contexts.

Finally, our results indicate that the boys' advantage in mathematics is strongest in Northern, wealthier provinces characterized by higher overall achievement. Conversely, the gap is weakest in Southern areas, where economic constraints are more prevalent. For reading, the girls' advantage is pervasive but tends to be strongest in provinces with higher overall achievement. Therefore, to transport this design to other national contexts, future research would require several key features: population-level assessments similar to INVALSI; granular, regional-level gender equality indicators rather than national averages; and distinct curricular tracking pathways, which likely interact with these socio-economic sorting mechanisms. Generalizing these results to more homogeneous or non-tracked systems should be done with caution.

5. Conclusions

This study found consistent sex differences in mathematics and reading across Italian provinces: boys outperform girls in mathematics, while girls outperform boys in reading, with differences widening at higher grade levels. A clear geographical gradient was observed, with northern provinces showing larger mathematics differences favoring boys. Socioeconomic factors, such as employment rate and university enrollment, alongside mean achievement levels, played a crucial role in shaping these gaps. Finally, these patterns were confirmed at the level of general scholastic sex-difference factor, highlighting how regional opportunity structures influence overall cognitive outcomes.

.....

CRedit authorship contribution statement

Lorenzo Esposito: Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review and editing. Enrico Toffalini: Methodology, Supervision, Writing – original draft, Writing – review and editing. David Giofrè: Conceptualization, Project administration, Investigation, Validation, Supervision, Writing – review and editing.

Declaration of Interest statement

The authors declare no conflicts of interest.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Ethical statement

This study is based on fully anonymized secondary data from national standardized assessments and thus did not require formal ethical approval under Italian regulations. The research adheres to international ethical standards, including the Committee on Publication Ethics (COPE), the American Psychological Association (APA), and the Declaration of Helsinki. All procedures ensured privacy, confidentiality, and responsible data use in accordance with ethical best practices.

Submitted: October 16, 2025 CDT. Accepted: January 17, 2026 CDT. Published: February 04, 2026 CDT.



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-4.0). View this license's legal deed at <http://creativecommons.org/licenses/by/4.0> and legal code at <http://creativecommons.org/licenses/by/4.0/legalcode> for more information.

References

- Balducci, M. (2023). Linking gender differences with gender equality: A systematic-narrative literature review of basic skills and personality. *Frontiers in Psychology*, 14(February), 1–16. <https://doi.org/10.3389/fpsyg.2023.1105234>
- Balducci, M., Larose, M.-P., Stoet, G., & Geary, D. C. (2024). The gender-equality paradox in intraindividual academic strengths: A cross-temporal analysis. *Psychological Science*, 35(11), 1246–1259. <https://doi.org/10.1177/09567976241271330>
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. (2011). *Introduction to meta-analysis*. John Wiley & Sons. <https://doi.org/10.1002/9780470743386>
- Breda, T., & Napp, C. (2019). Girls' comparative advantage in reading can largely explain the gender gap in math-related fields. *Proceedings of the National Academy of Sciences of the United States of America*, 116(31), 15435–15440. <https://doi.org/10.1073/pnas.1905779116>
- Burnham, K., & Anderson, D. (2007). Monte Carlo insights and extended examples. *Model Selection and Multimodel Inference*, 206–266. https://doi.org/10.1007/978-0-387-22456-5_5
- Carmeci, G., & Mauro, L. (2002). The convergence of the Italian regions and unemployment: Theory and evidence. *Journal of Regional Science*, 42(3), 509–532. <https://doi.org/10.1111/1467-9787.00269>
- Ceci, S. J., Williams, W. M., & Barnett, S. M. (2009). Women's underrepresentation in science: Sociocultural and Biological Considerations. *Psychological Bulletin*, 135(2), 218–261. <https://doi.org/10.1037/a0014412>
- Cellini, R., & Torrì, G. (2014). Regional resilience in Italy: A very long-run analysis. *Regional Studies*, 48(11), 1779–1796. <https://doi.org/10.1080/00343404.2013.861058>
- Checchi, D., & Peragine, V. (2010). Inequality of opportunity in Italy. *Journal of Economic Inequality*, 8(4), 429–450. <https://doi.org/10.1007/s10888-009-9118-3>
- Dunford, M. (2008). After the three Italies the (internally differentiated) North-South divide: Analysing regional and industrial trajectories. In *Annales de Géographie* (Vol. 117, Issue 664). <https://doi.org/10.3917/ag.664.0085>
- Eagly, A. H., & Wood, W. (1999). The origins of sex differences in human behavior: Evolved dispositions versus social roles. *Evolution, Gender, and Rape*, 54(6), 408–423. <https://doi.org/10.1037/0003-066X.54.6.408>
- Eals, M., & Silverman, I. (1994). The Hunter-Gatherer theory of spatial sex differences: Proximate factors mediating the female advantage in recall of object arrays. *Ethology and Sociobiology*, 15(2), 95–105. [https://doi.org/10.1016/0162-3095\(94\)90020-5](https://doi.org/10.1016/0162-3095(94)90020-5)
- Eccles, J. (1983). Expectancies, values and academic behaviors. In *Achievement and achievement motives: Psychological and sociological approaches* (pp. 75–146). Freeman.
- Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, 136(1), 103–127. <https://doi.org/10.1037/a0018053>
- Esposti, R. (2011). Convergence and divergence in regional agricultural productivity growth: Evidence from Italian regions, 1951–2002. *Agricultural Economics*, 42(2), 153–169. <https://doi.org/10.1111/j.1574-0862.2010.00508.x>
- Felice, E. (2011). Regional value added in Italy, 1891–2001, and the foundation of a long-term picture. *Economic History Review*, 64(3), 929–950. <https://doi.org/10.1111/j.1468-0289.2010.00568.x>
- Finegan, J. A. K., Niccols, G. A., & Sitarenios, G. (1992). Relations between prenatal testosterone levels and cognitive abilities at 4 years. *Developmental Psychology*, 28(6), 1075–1089. <https://doi.org/10.1037/0012-1649.28.6.1075>
- Giofrè, D., Cornoldi, C., Martini, A., & Toffalini, E. (2020). A population level analysis of the gender gap in mathematics: Results on over 13 million children using the INVALSI dataset. *Intelligence*, 81(June). <https://doi.org/10.1016/j.intell.2020.101467>
- Guiso, L., Monte, F., & Sapienza, P. (2008). Differences in test scores correlated with indicators of gender Equality. *Science*, 320(May), 1–2.
- Halpern, D. F. (2012). *Sex differences in cognitive abilities* (4th ed.). Psychology Press.
- Hamamura, T. (2012). Power distance predicts gender differences in math performance across societies. *Social Psychological and Personality Science*, 3(5), 545–548. <https://doi.org/10.1177/1948550611429191>
- Herlitz, A., Hönig, I., Hedebrant, K., & Asperholm, M. (2023). A systematic review and new analyses of the gender-equality paradox. *Perspectives on Psychological Science*. <https://doi.org/10.1177/17456916231202685>
- INVALSI. (2015). *Rapporto tecnico prove INVALSI 2015*. INVALSI.
- INVALSI. (2016). *Rapporto tecnico prove INVALSI 2016*. INVALSI.
- INVALSI. (2017). *Rapporto tecnico prove INVALSI 2017*. INVALSI.
- INVALSI. (2018). *Rapporto prove INVALSI 2018*. INVALSI.
- INVALSI. (2019). *Rapporto prove INVALSI 2019*. INVALSI.
- INVALSI. (2021). *Rapporto prove INVALSI 2021*. INVALSI.
- Ireson, G. (2017). Gender achievement and social, political and economic equality: a European perspective. *Educational Studies*, 43(1), 40–50. <https://doi.org/10.1080/03055698.2016.1237868>
- Konstantopoulos, S. (2011). Fixed effects and variance components estimation in three-level meta-analysis. *Research Synthesis Methods*, 2(1), 61–76. <https://doi.org/10.1002/jrsm.35>

- Lippa, R. A., Preston, K., & Penner, J. (2014). Women's representation in 60 occupations from 1972 to 2010: More women in high-status jobs, few women in things-oriented jobs. *PLoS ONE*, 9(5). <https://doi.org/10.1371/journal.pone.0095960>
- Maccoby, E. E., & Jacklin, C. N. (1977). *The Psychology of Sex Differences*.
- Marks, G. N. (2008). Accounting for the gender gaps in student performance in reading and mathematics: Evidence from 31 countries. *Oxford Review of Education*, 34(1), 89–109. <https://doi.org/10.1080/03054980701565279>
- Munir, F., & Winter-Ebmer, R. (2018). Decomposing international gender test score differences. *Journal for Labour Market Research*, 52(1). <https://doi.org/10.1186/s12651-018-0246-8>
- Patacchini, E. (2008). Local analysis of economic disparities in Italy: A spatial statistics approach. *Statistical Methods and Applications*, 17(1), 85–112. <https://doi.org/10.1007/s10260-007-0054-8>
- Penner, A. M. (2008). Gender differences in extreme mathematical achievement: An international perspective on biological and social factors. *American Journal of Sociology*, 114(SUPPL. 1), 138–170. <https://doi.org/10.1086/589252>
- Proietti, T. (2005). Convergence in Italian regional per-capita GDP. *Applied Economics*, 37(5), 497–506. <https://doi.org/10.1080/0003684042000318173>
- R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. <https://www.R-project.org/>
- Reilly, D. (2012). Gender, culture, and sex-typed cognitive abilities. *PLoS ONE*, 7(7), 15–16. <https://doi.org/10.1371/journal.pone.0039904>
- Rindermann, H. (2007). The g-factor of international cognitive ability comparisons: The homogeneity of results in PISA, TIMSS, PIRLS and IQ-tests across nations. *European Journal of Personality*, 21(5), 667–706. <https://doi.org/10.1002/per.634>
- Ritchie, S. J., & Bates, T. C. (2013). Enduring links from childhood mathematics and reading achievement to adult socioeconomic status. *Psychological Science*, 24(7), 1301–1308. <https://doi.org/10.1177/0956797612466268>
- Román-González, M., Pérez-González, J. C., & Jiménez-Fernández, C. (2017). Which cognitive abilities underlie computational thinking? Criterion validity of the Computational Thinking Test. *COMPUTERS IN HUMAN BEHAVIOR*, 72, 678–691. <https://doi.org/10.1016/j.chb.2016.08.047>
- RStudio Team. (2023). *RStudio: Integrated Development Environment for R*. <http://www.posit.co/>
- Stoet, G., & Geary, D. C. (2013). Sex differences in mathematics and reading achievement are inversely related: Within- and across-nation assessment of 10 years of PISA data. *PLoS ONE*, 8(3). <https://doi.org/10.1371/journal.pone.0057988>
- Stoet, G., & Geary, D. C. (2018). The gender-equality paradox in Science, Technology, Engineering, and Mathematics Education. *Psychological Science*, 29(4), 581–593. <https://doi.org/10.1177/0956797617741719>
- Stoet, G., & Geary, D. C. (2020). Gender differences in the pathways to higher education. *Proceedings of the National Academy of Sciences of the United States of America*, 117(25), 14073–14076. <https://doi.org/10.1073/pnas.2002861117>
- Stoet, G., & Geary, D. C. (2022). Sex differences in adolescents' occupational aspirations: Variations across time and place. *PLoS ONE*, 17(1 January), 1–18. <https://doi.org/10.1371/journal.pone.0261438>
- Ünal, Z. E., Park, Y., Simsek, E., Menon, V., & Geary, D. C. (2025). Neurodevelopmental commonalities in the cognitive control networks for mathematics and reading in meta-analysis of 3308 participants. *Nature Communications*, 16, 8398. <https://doi.org/10.1038/s41467-025-63259-8>
- van der Linden, D., Dunkel, C. S., & Madison, G. (2017). Sex differences in brain size and general intelligence (g). *Intelligence*, 63(October 2016), 78–88. <https://doi.org/10.1016/j.intell.2017.04.007>
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor. *Journal of Statistical Software*, 36(3), 1–48. <https://doi.org/10.18637/jss.v036.i03>
- World Economic Forum. (2025). *Global Gender Gap Report 2025*. <https://www.weforum.org/publications/global-gender-gap-report-2025/>

Supplementary Materials

Supplementary Material

Download: <https://icajournal.scholasticahq.com/article/155775-sex-differences-in-mathematics-and-reading-across-italy-the-role-of-socioeconomic-factors/attachment/326592.docx>
