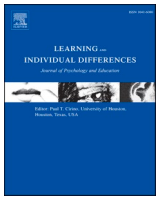
Learning and Individual Differences 122 (2025) 102726

Contents lists available at ScienceDirect

Learning and Individual Differences

journal homepage: www.elsevier.com/locate/lindif



Exploring the mediating role of academic anxiety in mathematics and reading performance among boys and girls: A comprehensive study of Italian fifth graders

Sara Caviola [a,*,1], David Giofrè [b,**,1], Tommaso Feraco [c], Enrico Toffalini [c], Katie Allen [d], David C. Geary [e]

[a] Department of Developmental and Social Psychology, University of Padua, Italy
[b] Disfor, University of Genoa, Italy
[c] Department of General Psychology, University of Padua, Italy
[d] Faculty of Science and Technology (FST), Lancaster University, UK
[e] Psychological Sciences, University of Missouri, USA

ARTICLE INFO

Keywords:
Gender
Mathematics performance
Reading performance
Test anxiety
National assessment

ABSTRACT

There is a long-standing debate regarding the magnitude of gender differences in academic achievement. Using data from a robust and nationally representative sample of 146,227 Italian-fifth-graders, we investigated whether academic anxiety is related to gender differences in mathematics and reading achievement. Across six independent samples, boys had higher performance in mathematics ($ds = -0.13$ to -0.21) and girls had higher performance in reading ($ds = 0.07$ to 0.21) and higher test anxiety ($ds = 0.30$ to 0.37). Meta-analytic procedures indicated these patterns were stable across samples. Path analyses within and across samples suggested about one-third of the academic gender gaps can be accounted for by test anxiety. In particular, with control of test anxiety girls' advantage in reading achievement increased, while boys' advantage in mathematics decreased.
Educational statement: The current study provides an extensive exploration of how test anxiety potentially influences gender differences in mathematics and reading achievement. Test anxiety appears to lower girls' performance on achievement tests and thus underestimating their advantages in reading and overestimating boys' advantages in mathematics. One implication is that efforts to reduce test anxiety will enhance performance on achievement tests, especially for test-anxious girls, and through this provide more accurate estimates of academic competencies.

1. Introduction

Academic learning presents significant challenges for many students, including a spectrum of emotions associated with successes, failures, and test taking (Pekrun & Linnenbrink-Garcia, 2014). These affective components appear to influence students' achievement-related outcomes (Camacho-Morles et al., 2021; Caviola et al., 2022; Cheng & McCarthy, 2018 and Loderer et al., 2020). Generally, positive emotions like enjoyment of learning are associated with achievement gains and negative emotions such as boredom with relatively poor achievement (Goetz & Hall, 2013; Pekrun & Linnenbrink-Garcia, 2014). For many students, achievement testing can trigger considerable anxiety (Robson et al., 2023; Seipp, 1991; von der Embse et al., 2018), especially in difficult areas such as mathematics (Barroso et al., 2021; Caviola et al., 2022). Much of this research has focused on high school and college students (Efklides & Volet, 2005; Linnenbrink, 2006; Linnenbrink-Garcia & Pekrun, 2011; Schutz & Pekrun, 2007), although recent studies have focused on younger students' achievement-related emotions. Even so, for these students less is known about how achievement-related emotions influence academic outcomes than for older students (Hill et al., 2016; Ganley & McGraw, 2016; Sorvo et al., 2017; Szczygieł & Pieronkiewicz, 2022; Wu et al., 2012).

To address this gap, the relation between test anxiety and mathematics and reading achievement was assessed across six independent

* Corresponding author at: Department of Developmental and Social Psychology, University of Padua, via Venezia 8, 35010 Padova, Italy.
** Correspondence to: D. Giofrè, Disfor, University of Genoa, Corso Andrea Podestà, 2, 16128 Genoa, Italy.
E-mail address: david.giofre@unige.it (D. Giofrè).
[1] These authors contributed equally to the work, and their names are listed in alphabetical order.

https://doi.org/10.1016/j.lindif.2025.102726
Received 17 October 2024; Received in revised form 8 May 2025; Accepted 12 May 2025
Available online 19 June 2025

cross-sectional waves of nationally representative samples of fifth-grade Italian children (n = 146,227). The primary goal was to attempt to replicate gender differences in test anxiety and, secondly, to determine whether any such differences might be related to the magnitude of boys' advantage in mathematics and girls in reading attainment. The availability of six large, independent samples offered a unique opportunity to examine the replicability of these gender differences and use meta-analytic techniques to provide a more reliable estimate of these effects.

1.1. Test anxiety and academic achievement

Test anxiety manifests as distress and apprehension associated with taking tests and exams, including concerns about failing and the potential negative consequences that may follow (Zeidner, 1998). It is often conceptualized as a situation-specific personality trait (trait test anxiety; Spielberger et al., 1976), though it can also manifest as a temporary emotional reaction before or during a specific exam (state test anxiety). Test anxiety often increases across grades as academic demands grow (Huberty, 2012; Putwain, Gallard, et al., 2021), but does not always compromise academic performance. At lower levels, anxiety and accompanying physiological changes can enhance performance by sharpening focus and facilitating memory retrieval, as highlighted by the Yerkes-Dodson Law (Jeanneteau & Coutellier, 2022; Yerkes & Dodson, 1908). However, excessive anxiety can undermine memory (Lupien et al., 2007) and can divert attentional focus from the academic task to irrelevant thoughts (e.g., concerns about failure) and undermine motivation to learn (Derakshan & Eysenck, 2009; Eysenck et al., 2007; Eysenck & Calvo, 1992). While test anxiety is distinct from other academic anxieties, such as mathematical anxiety (Caviola et al., 2022), it shares a multidimensional structure composed of cognitive, emotional, and physiological components (Lowe, 2014; Putwain, Gallard, et al., 2021, Putwain, Stockinger, et al., 2021, Putwain, von der Embse, et al., 2021). A widely adopted framework distinguishes between "worry" and "emotionality" (Liebert & Morris, 1967). The former refers to performance-related concerns and fear of failure, often involving negative self-evaluation, whereas the latter encompasses physiological arousal and emotional tension in evaluative contexts.

Pekrun's (2006) control-value theory (CVT) provides a taxonomy of academic emotions and describes how control and value appraisals can act as antecedents of these emotions (Bosshard & Gomez, 2024). Control appraisals include students' achievement-related competence beliefs, expectancies, and attributions, and value appraisals to the perceived importance of a school activity or outcome. Within this framework, achievement emotions are categorized across valence (positive vs. negative), level of physiological arousal (activating vs. deactivating), and focus (e.g., learning tasks) (Pekrun, 2006). Thus, anxiety is an unpleasant, outcome-oriented, and activating emotion and often manifests as test anxiety (e.g., Putwain, Gallard, et al., 2021; Putwain, Stockinger, et al., 2021; Putwain, von der Embse, et al., 2021). Generally speaking, anxious students have lower self-perceptions of their academic abilities compared to their less-anxious peers (Ahmed et al., 2012; Donolato et al., 2020; Nie et al., 2011; Putwain et al., 2010) and experience more difficulty in complex subjects, such as mathematics (Foley et al., 2017; Galla & Wood, 2012; Putwain, 2008; von der Embse & Witmer, 2014). This frequently undermines academic motivation (Cizek & Burg, 2006; Malanchini et al., 2017) in ways that can increase risk of school dropout (Cassady et al., 2022) with repercussions in their later working life (Choe et al., 2019; Hasty et al., 2021; Ramirez et al., 2018).

1.2. Meta-analytic findings for anxiety, educational outcomes, and gender differences

Several studies have identified a link between anxiety and compromised educational outcomes, alongside less positive academic attitudes and an increased risk for depression. Mandler and Sarason (1952) provided the earliest evaluation of test anxiety, but it did not gain prominence until Hembree's (1988) pivotal meta-analysis. They concluded that test anxiety, along with other forms of anxiety, consistently correlates with worse educational outcomes, reduced self-esteem, self-confidence, and generally lower academic motivation. Moreover, their results indicated that girls tend to experience higher anxiety than boys, with this disparity growing across grades. While this meta-analysis provided valuable insights into the association between test anxiety and gender differences in performance, it is important to exercise caution due to the primarily correlational nature of these studies (Geary et al., 2019). Although fewer in number, longitudinal studies suggest a reciprocal relation between test outcomes and anxiety can develop over time (Pekrun et al., 2017a, 2017b; Pekrun et al., 2023a, 2023b; Putwain & Wood, 2023), whereby academic struggles increase anxiety and the anxiety in turn reduces engagement with schooling that leads to further academic issues.

In any case, subsequent research has largely corroborated Hembree's (1988) original conclusions (see also Seipp, 1991). For example, von der Embse et al.' (2018) recent review and meta-analysis confirmed a negative relationship between test anxiety and many educational outcomes, including performance on standardized tests and university entrance exams. The correlations ranged from moderate to large (rs = −0.16 to −0.40). The meta-analysis also confirmed that girls consistently report higher test anxiety than boys across all grade levels (r = 0.21). More recently, Robson et al.' (2023) comprehensive analysis of primary school children again showed a negative correlation between test anxiety and academic achievement for both mathematics (r = −0.21) and literacy (r = −0.20). The study also confirmed that girls exhibit higher test anxiety than boys (Cohen's d = 0.21 but d = 0.08 with adjustment for publication bias), with related studies showing this pattern to be robust across educational contexts and developmental stages (Talbot, 2016; Wigfield et al., 2015). These consistent gender differences in test-related emotional experiences may help explain broader patterns of academic performance.

1.3. Gender differences in mathematics and reading literacy

Gender differences in academic achievement have been well documented across a variety of domains, particularly in mathematics and reading. Girls consistently outperform boys in reading, whereas boys often have an advantage in mathematics. These performance differences are observed internationally and may be related, in part, to girls' higher test anxiety, particularly in mathematics (Cassidy & Johnson, 2002; Deieso & Fraser, 2018). As an example, stable and substantial differences favoring girls in the USA have been consistently observed in reading, and particularly in writing (Reilly et al., 2019; see also Halpern & LaMay, 2000; Hedges & Nowell, 1995; Lynn & Mikk, 2009; Reilly 2012; Reilly et al., 2016; Stanley et al., 1992). Similarly, analyses of USA NAEP data by Reilly et al. (2015) indicate that while boys slightly outnumber girls at the advanced level in mathematics in fourth and eighth grades, this disparity is larger, especially for reading and mathematics as academic strengths, in older students (Stoet & Geary, 2018, 2020).

Despite being a highly developed nation, Italy has significant gender disparities, particularly in women's and men's economic participation and opportunities. According to the World Economic Forum (2023), Italy ranks 79th out of 146 countries in the Global Gender Gap Index, assessing gender equality across economic, political, educational domains. In other words, Italy is highly developed economically but lags other highly developed nations in gender equality and generally shows larger gender differences in achievement. Giofrè et al. (2020) analyzed a large cohort of over 13 million primary students, assessed by the INVALSI (Istituto Nazionale per la VALutazione del Sistema di Istruzione [National Institute for the Assessment of the Instruction System]), in mathematics and reading. They found boys consistently outperformed girls in mathematics, with the gap widening from second to eighth grade. Conversely, girls consistently outperformed boys in reading

across all grade levels, and their advantage also increased across grades.

1.4. Aims and hypotheses

In the current study, we examined gender differences in test anxiety in the context of high-stakes assessments and examined whether it is related to the gender gaps in mathematics and reading. Given the correlational nature of our data, the aim was not to infer causality but rather to provide a more accurate estimation of the associations between test anxiety levels and performance in both mathematics and reading on the INVALSI assessments. We hypothesized that girls would perform better in reading and have higher test anxiety, while boys would perform better in mathematics. The novel contribution here is the assessment of whether academic gender differences are reduced (in mathematics) or amplified (in reading) after controlling for test anxiety. To this end, we analyzed data from six large, representative samples of Italian fifth graders. We used multilevel path analyses, which allowed us to control for the stratification of the sample, while at the same time obtaining reliable estimations of the relation between gender differences in test anxiety and gender differences in reading and mathematics. Finally, meta-analyses across results from the independent samples were used to generate more reliable estimates. To the best of our knowledge, this statistical approach, which allows us to control several important confounding variables, has not been used previously.

2. Methods

2.1. Participants

A sample of 147,482 5th graders who were assessed by the INVALSI between the academic years 2011/2012 and 2016/2017, inclusive. These data only include students in classes that were monitored by members of the INVALSI committee, which ensures higher data reliability and reduces cheating. Given the study's focus on gender differences, we included only students who reported their sex as either male or female. A total of 72 students (0.05 %) were excluded for not meeting this condition. Additionally, we excluded 1,183 students who did not have complete scores for math, reading, or anxiety. Specifically, 2 % of students from the 2011/2012 academic year were excluded due to missing data, while for all other academic years, less than 1 % of participants were excluded for this reason. As shown in Table 1 the final sample included 146,227 students (73,670 males; 72,557 females), with a minimum of 19,383 students in each assessment year. Students were sampled from between 971 and 1642 schools located in 85 to 103 Italian provinces. The median number of students per school varied between 18 and 19 for each year of data collection.

2.2. Materials

The INVALSI are standardized national tests administered at several key stages of the Italian education system to evaluate student competencies in reading, mathematics, and English. It takes place annually in April or May and is administered under controlled conditions in the classroom. Assessments are tailored to each grade level, aligning with national curricular requirements. In Italy, the majority of schools are public, and the Ministry of Education sets national guidelines, educational goals, and standards for these institutions. Fifth-grade students complete their final year of elementary school before transitioning to middle school (primary education level). Their curriculum covers core subjects like Italian, mathematics, science, history, geography, English, and physical education, with an emphasis on digital literacy. Classes typically range from 15 to 25 students, and assessments include written tests, oral exams, group projects, and teacher evaluations.

For fifth graders, the exam is paper-based, and completed during school hours, with each subject examined separately. The mathematics and reading tests each last 75 min. Teachers supervise the process but do not help, and external inspectors may be present to ensure procedural compliance. While individual student performance does not directly impact academic progression, schools are evaluated based on their overall results, often leading to considerable concern among teachers. Consequently, although the test is not officially classified as high-stakes, it assumes high-stakes characteristics. Each response is coded as either correct (1) or incorrect (0). For both mathematics and reading, the items are created based on IRTs. Information and test examples are provided below (more information and sample tests are available on the INVALSI Area Prove).

2.2.1. Mathematics

The INVALSI mathematics test, modeled after the PISA assessment, consists of a single booklet with 30 to 40 multiple-choice and open-ended questions. These assess arithmetic and problem-solving, geometry, measurement, and data interpretation, with items presented in random order.

Arithmetic and *Problem-solving*. Tasks involve basic operations, fractions, and word problems requiring logical reasoning. For example, "Giulia has 450 stickers. She gives 175 to her friend and receives 230 more from another friend. How many stickers does Giulia have now?". Other items assess pattern recognition and proportional reasoning, e.g., "Look at the sequence: 2, 4, 8, 16… What is the next number in this pattern?". *Geometry*. Items test knowledge of perimeters, symmetries, and shape properties. For instance, "A rectangular garden measures 8 meters in length and 6 meters in width. What is the perimeter of the garden?". *Measurement* and *data interpretation*. Tasks involve unit conversions, interpreting graphs using tools like goniometers, and understanding fundamental statistical concepts. For example, "A classroom has 20 students. If 40% of the students bring their lunch from home, how many students buy their lunch at school?"

2.2.2. Reading

The INVALSI evaluates the essential literacy skills of Reading Comprehension and Grammar. The test consists of 30 to 40 questions, including a mix of multiple-choice, true/false, and open-ended formats. The overall goal is to evaluate students' proficiency in Italian, ensuring they are well-prepared for future academic challenges. Unlike the mathematics assessment, which is more integrated, reading comprehension and grammar are assessed in separate sections and cover distinct content areas.

2.2.2.1. Reading comprehension. This assesses students' ability to interpret different kinds of texts, such as narrative or informative, by understanding the main ideas, inferring meanings from context, and recognizing the structure and organization of the text. Questions might involve identifying the theme of a passage or the implications of specific words and phrases within the context.

2.2.2.2. Grammar. This assesses students' understanding of syntactic

Table 1
Boys, girls, and total participants at each year of data collection.

Academic year	11/12	12/13	13/14	14/15	15/16	16/17	Total
Boys	14,618	11,949	12,425	9,854	12,558	12,266	73,670
Girls	14,714	11,931	12,240	9,529	11,967	12,176	72,557
Total	29,332	23,880	24,665	19,383	24,525	24,442	146,227

structures and the correct application of language mechanics, with example questions often requiring the selection of grammatically correct sentences or correcting errors in a given text. It is worth noting that in Italy grammar is generally taught in every school, and is a national requirement.

2.2.3. Test anxiety

Test anxiety is often assessed through trait-like self-report questionnaires. Conversely, when individuals report test-related emotions tied to specific moments before, during, or after an exam, these emotions are considered state test anxiety. The questionnaire is composed of four items focused on state anxiety related to academic testing. Responses were on a 4-point Likert scale, ranging from 1 ("totally disagree") to 4 ("totally agree"): "Before taking the test, I felt worried"; "I was so nervous that I couldn't find the answers"; "During the test, I had the impression I was having difficulty"; and, "I felt calm while taking the test" (reversed coded). Higher score indicates greater anxiety. Reliability was measured using Cronbach's alpha, and ranged from 0.71 to 0.78.

2.3. Statistical analysis

All analyses were conducted using R, version 4.3.1 (R Core Team, 2023). Performance on mathematics, reading, and test anxiety scores were standardized ($M = 0$, $SD = 1$). Pearson's correlations were calculated for each year of data collection.

To summarize the findings across the six years of data collection, we adopted a meta-analytical approach. Initially, we calculated the standardized difference in anxiety between boys and girls (Cohen's *d*), as well as the correlations between anxiety and mathematics and reading performance for each year. We then conducted three different random-effects meta-analytic models: one for sex differences in anxiety to summarize the six Cohen's *d*s, one to summarize the six correlations between anxiety and mathematics, and one to summarize the six correlations between anxiety and reading. Each year was thus treated as an independent study within a classic meta-analysis framework, allowing us to summarize the results of each data collection year into the three main effects of interest. For the meta-analysis of the correlation coefficients, all correlations (*r*) were converted into Fisher *z* values, and we provide the average Fisher *r*-to-*z* transformed correlation values. Mu ($\widehat{\mu}$) values, the meta-analytic standardized difference between boys and girls, is provided for gender differences in test anxiety. Tau (τ) values, which are the estimated standard deviation of true effects across studies, were also reported for each meta-analysis.

We then followed a series of steps for the path analysis that estimated whether test anxiety mediated gender differences in mathematics and reading performance. For each year, we conducted two multilevel path analytic models, treating schools as random factors. The first model only accounted for the effect of gender on the two outcome variables—mathematics and reading—while the effect of test anxiety was set to zero. In the second model, gender predicted both achievement measures and anxiety levels, and test anxiety was allowed to freely mediate the effect of gender on performance.

We were also interested in estimating how gender differences in mathematics and reading performance change when test anxiety is (model 2) or is not considered (model 1). However, the raw differences between the betas from models 1 and 2 can only be manually calculated, as this difference does not have an associated error term. We implemented a bootstrap procedure to estimate standard errors. Specifically, each model was bootstrapped 1,000 times for each year. During each replication, the difference between the first and the second model was calculated. For each year, the mean difference of the bootstrap distribution was then used as the point estimate, and its standard deviation was taken as the associated standard error. Finally, the effects calculated for each cohort were meta-analyzed using random effects models to provide a generalized estimate of each β and their associated standard errors derived from the path analysis.

3. Results

Table 2 shows the minimum and maximum correlations (across years) for the key variables. Girls reported higher levels of test anxiety and had an advantage in reading, whereas boys had an advantage in mathematics. Test anxiety was negatively correlated with mathematics and reading performance.

3.1. Gender differences in test anxiety

Girls consistently reported higher test anxiety, with effect sizes (Cohens *d*) ranging from 0.30 to 0.37 (Fig. 1). The meta-analytic standardized difference derived from the random-effects model was $\widehat{\mu} = 0.33$, 95%CI [0.31; 0.35], $\tau = 0.02$.

3.2. Correlations between test anxiety and mathematics and reading performance

The correlations between test anxiety and mathematics ranged from −0.27 to −0.15 (Fig. 2, panel A). The average Fisher *r*-to-*z* transformed correlation derived from the random-effects model was −0.21, 95%CI [−0.24; −0.18], $\tau = 0.04$. The correlations between anxiety and reading ranged from −0.21 to −0.12 (Fig. 2, panel B). The average Fisher r-to-z transformed correlation from the random-effects model was −0.18, 95% CI [−0.21; −0.15], $\tau = 0.04$.

3.3. Multilevel path analysis

The results from the multilevel path analyses are entirely reported online in Table S01. These results are also meta-analytically summarized and illustrated in Fig. 3 and Fig. 4. All parameters are unstandardized beta coefficients (β). Given that standardized *z*-scores were used for the quantitative variables, these can be interpreted as standardized coefficients. Coefficients for gender differences can be interpreted as standardized mean differences (Cohen's *d*s).

The first set of models (i.e., without the mediation of test anxiety) indicated that girls had consistently lower performance in mathematics, with beta coefficients ranging from −0.21 to −0.13, and higher performance in reading, with coefficients from 0.06 to 0.19 (Table S01). However, the second set of models (i.e., with the mediation of test anxiety) revealed that girls' disadvantage in mathematics was less pronounced, with coefficients between −0.14 and −0.08, and their reading advantage more pronounced, ranging from 0.13 to 0.25 (Table S01).

Both mathematics and reading were negatively related to test anxiety levels, especially in girls, with coefficients from −0.30 to −0.36 (Table S01). Notably, the indirect effects calculated in the second set of models suggested that girls' performance in both subjects was partially and negatively affected by test anxiety compared to that of boys, with betas ranging from −0.05 to −0.08 (Table S01). To sum up, all models consistently revealed similar patterns: girls exhibited higher levels of test anxiety, which in turn was negatively related to their performance in both mathematics and reading performance. There were also notable differences in performance between boys and girls, with boys achieving

Table 2
Range of cross-year correlations (min and max).

	Test Anxiety	Sex	Mathematics
Test Anxiety	–		
Sex	[0.15; 0.18]	–	
Mathematics	[−0.15; −0.26]	[−0.06; −0.10]	–
Reading	[−0.12; −0.21]	[0.03; 0.10]	[0.45; 0.68]

Note. Girls were coded 1 and boys 0. Thus, positive correlations mean higher scores for girls. Point-biserial correlations were computed for gender.

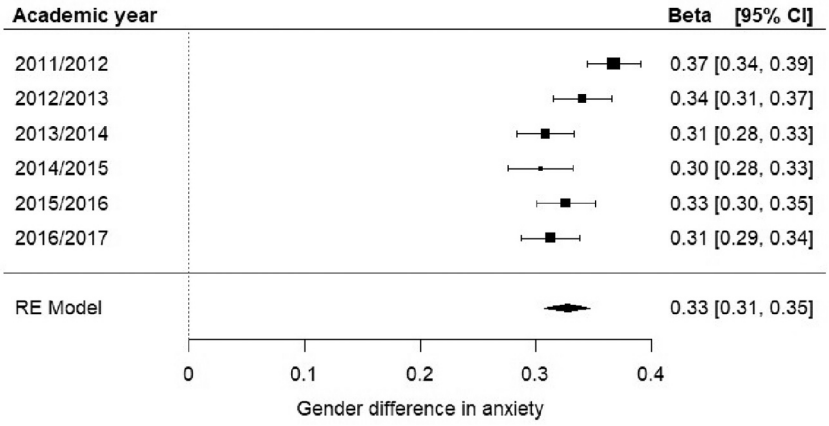


Fig. 1. Standardized gender differences in test anxiety and associated meta-analytic estimates.
Note. Positive values indicate higher test anxiety in girls. The size of the square represents the study weight: studies (yearly samples in our case) with smaller standard errors and somewhat larger sample sizes are given more weight in the calculation of the pooled effect size in the meta-analytic random effect model.

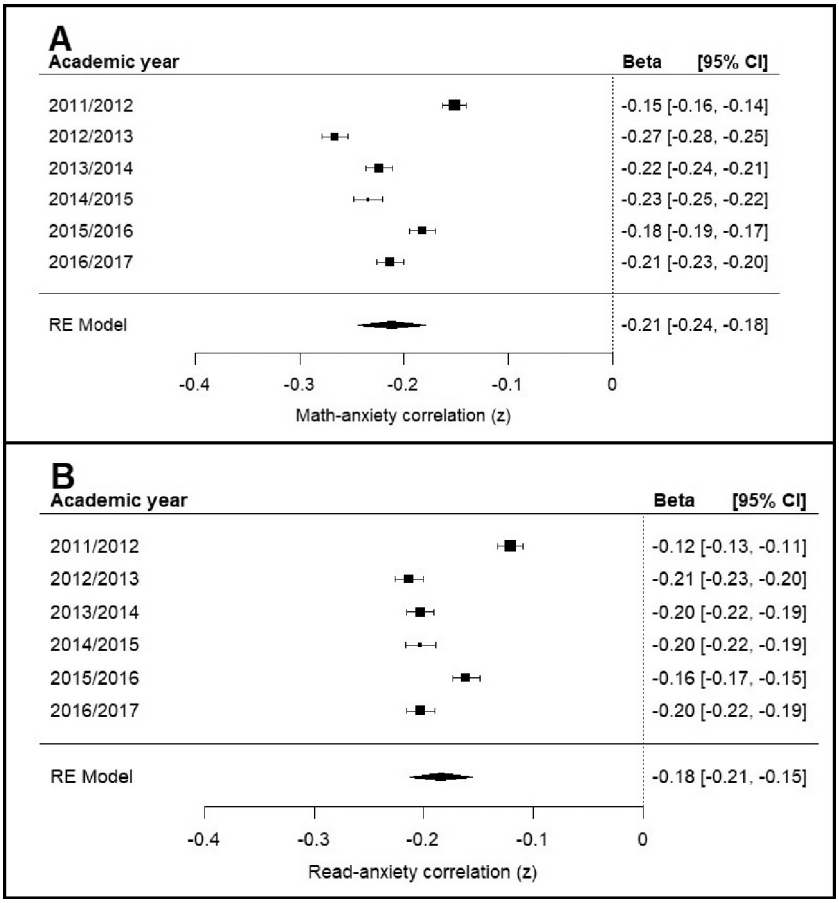


Fig. 2. Meta-analytic correlation between test anxiety and mathematics (panel A) and reading (panel B).
Note. The negative betas indicate higher test anxiety is associated with lower achievement scores. The size of the square represents the study weight: studies (yearly samples in our case) with smaller standard errors and somewhat larger sample sizes are given more weight in the calculation of the pooled effect size in the meta-analytic random effect model.

higher scores in mathematics and lower scores in reading. Importantly, these gender disparities in mathematics decreased, while those in reading increased, when the effect of test anxiety was considered.

3.4. Meta-analytical synthesis

After estimating the standardized association between gender and the other variables for each academic year, we summarized the results comparing models with and without control of test anxiety using a meta-analytic procedure. We also assessed changes in the magnitude of the gender differences in mathematics and reading performance, with control of test anxiety (Table S02).

The initial set of meta-analytic models confirmed that boys outperformed girls on the mathematics test, $\beta = -0.17$, 95%CI [−0.20;

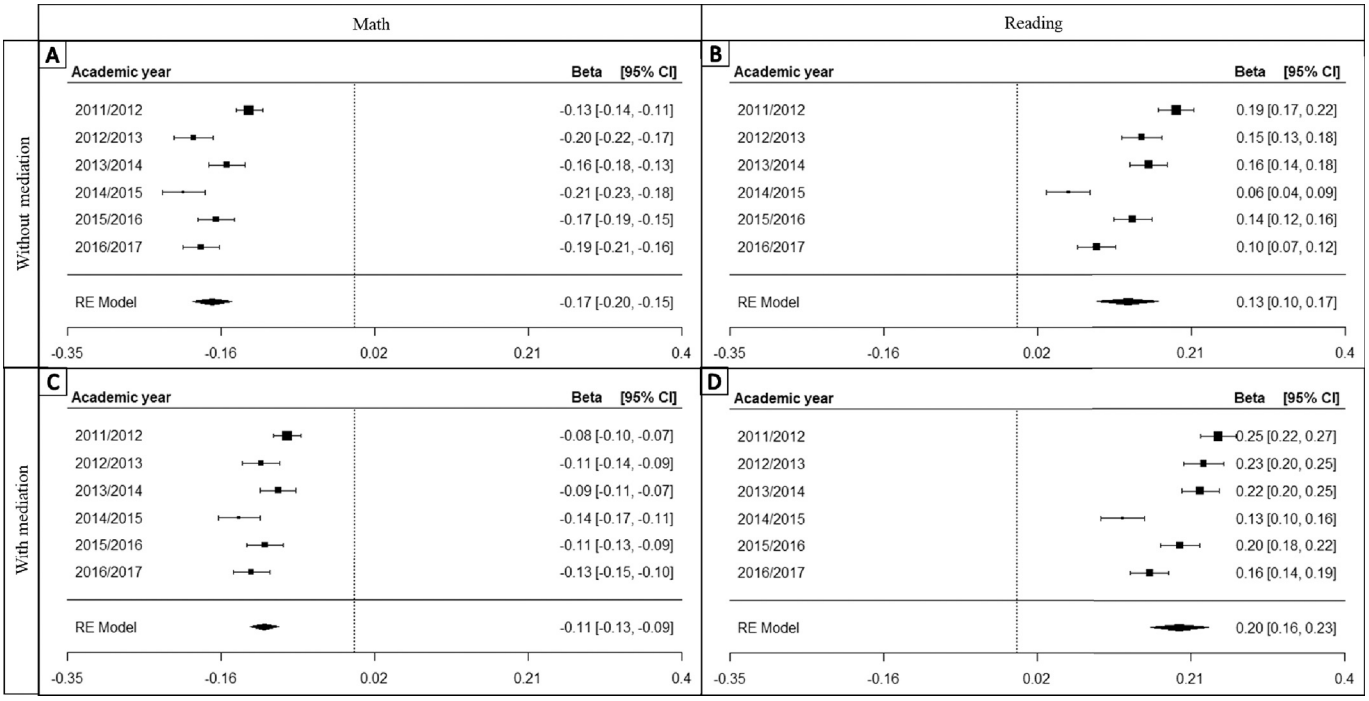


Fig. 3. Meta-analytic gender effects on math and reading performance without (above) and with (below) the mediating effect of anxiety.
Note. The size of the square represents the study weight: studies (yearly samples in our case) with smaller standard errors and somewhat larger sample sizes are given more weight in the calculation of the pooled effect size in the meta-analytic random effect model.

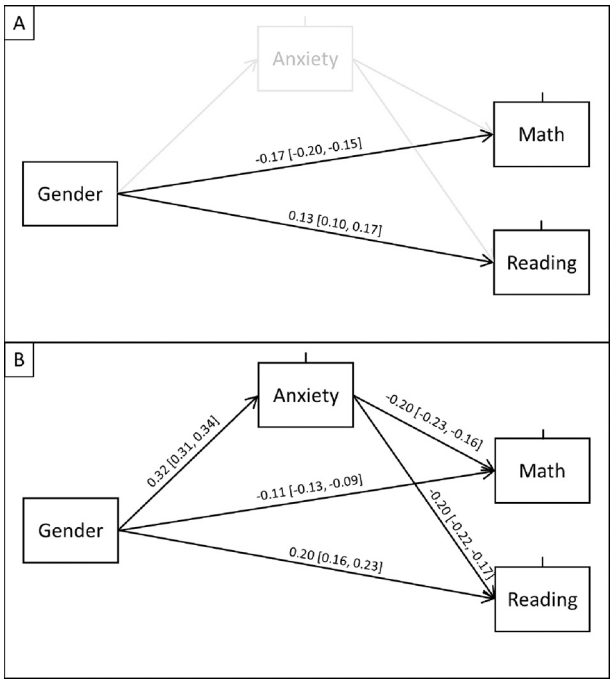


Fig. 4. Relations between gender and performance without (Panel A) and with (Panel B) control of test anxiety. The figures report the path models with meta-analytic estimates and confidence intervals.
Note. Positive paths indicate higher scores for girls. Results of the individual path models are reported online (Table S01).

−0.15], $\tau = 0.03$ (Fig. 3, panel A), and girls outperformed boys on the reading test, $\beta = 0.13$, 95%CI [0.10; 0.17], $\tau = 0.05$ (Fig. 3, panel B). After controlling for gender differences in test anxiety (second set of models), boys' advantage in mathematics decreased, $\beta = -0.11$, 95%CI [−0.13; −0.09], $\tau = 0.02$ (Fig. 3, panel C), and girls' advantage in reading increased, $\beta = 0.20$, 95%CI [0.16; 0.23], $\tau = 0.04$ (Fig. 3, panel D). The meta-analytic results confirmed girls' higher test anxiety levels, $\beta = 0.32$, 95%CI [0.31; 0.34], and revealed it is associated with lower performance in mathematics, $\beta = -0.20$, 95%CI [−0.23; −0.16], and reading, $\beta = -0.20$, 95%CI [−0.22; −0.17]. The path models with the meta-analytic estimates and their confidence intervals are represented in Fig. 4.

Comparisons of the path analytic models without (Fig. 4, panel A) and with (Fig. 4, panel B) control of test anxiety show changes in the magnitude of the achievement gaps. Incorporating test anxiety decreased boys' advantage in mathematics by 0.07 points, $\Delta\beta = 0.07$ [0.06; 0.08], and increased girls' advantage in reading by 0.06 points, $\Delta\beta = -0.06$ [−0.07; −0.06]. Disparities between gender differences in the two models are summarized in an online table (Table S02). Overall, gender differences in mathematics and reading performance remained significant, indicating that test anxiety was a partial mediator.

3.5. Additional analyses

In the INVALSI, the mathematics assessment is organized into different content categories: Arithmetic and Problem-solving, Geometry, and Measurement and Data Interpretation. These categories are assessed collectively, providing an overall measure of mathematical literacy without isolating individual content areas. In many cases, there is substantial overlap between content areas. In contrast, the reading assessment evaluates distinct dimensions of reading literacy. A clear distinction is made between reading comprehension and grammar which are assessed through separate test sections. Given this distinction, we further examined these dimensions. The results, presented in the supplementary online materials (Figs. S04-S05), indicate that girls outperform boys in both reading comprehension ($\widehat{\mu} = 0.12$) and grammar ($\widehat{\mu} = 0.14$), with a statistically significant, albeit modest, advantage in grammar.

Although this study was not explicitly designed to examine differences in variance, we decided to explore this aspect as well. The results

show that boys tend to exhibit greater variability in their scores, meaning they have a higher variance ratio compared to girls. Specifically, in mathematics, the variance ratio is 1.06, indicating that boys have approximately 6 % more variance than girls. Similar findings, though smaller in magnitude, were observed for reading (both reading comprehension and grammar), with an average variance ratio of 1.03 (see Figs. S06-S07 in the Supplementary Online Materials).

4. Discussion

Our findings, replicated across six large samples of fifth graders, confirmed gender differences in test anxiety, and in mathematics and reading performance (e.g., Robson et al., 2023; von der Embse et al., 2018). The magnitude of boys' advantage in mathematics ($d = 0.17$) was somewhat smaller than the recent PISA results (15- to 16-year-olds) for Italy ($d = 0.21$), but this could be due to the age difference for students in the current study and the PISA assessments. The reading gap ($d = 0.13$) was also smaller than the PISA results ($d = 0.19$), but again potentially due to age differences. Indeed, Giofrè et al. (2022) found that gender differences in both mathematics and reading were larger in eighth graders (close to the age of PISA assessments) than fifth graders. Critically, our path models suggested that about a third of the gender differences in mathematics and reading test performance were statistically explained by test anxiety, although a causal relation cannot be inferred from these results.

4.1. Comparison with previous meta-analyses of sex differences in anxiety and achievement

In line with previous meta-analytic studies, our research confirms a negative correlation between test anxiety and mathematics and reading achievement (e.g., Caviola et al., 2022; Robson et al., 2023; von der Embse et al., 2018). Our findings are in close agreement with those of Robson et al.'s (2023) recent meta-analysis, that is, a negative correlation between test anxiety and outcomes in mathematics ($r = -0.21$ here and in Robson et al.) and reading ($r = -0.18$ here and -0.20 in Robson et al.). However, their estimates had much broader confidence intervals than found in the current study which was likely due to the inclusion of many studies with small samples. Our study overcomes this limitation with considerably larger sample sizes—overall almost 3 times larger than the total across the 85 samples included in Robson et al. In other words, our results are likely more robust than those of Robson et al. (2023) but restricted to older elementary students. Indeed, by integrating our data with their findings, we obtained even more robust meta-analytic effects, characterized by narrower confidence intervals (see Supplementary Materials).

Our meta-analytic path models show gender differences in academic performance might be related, in part, to the gender differences in test anxiety. In particular, our findings suggest that girls' higher test anxiety might overestimate boys' advantage in mathematics and underestimate girls' advantage in reading by disrupting performance during test taking. However, it is not clear if test anxiety is a cause or a consequence of variation in academic achievement. Prior studies suggest that struggles with math might have stronger effects on girls' than boys' mathematics anxiety (Geary et al., 2023), but this does not explain the effect we found for reading. One possibility is that girls tend to exhibit higher levels of self-criticism and impose greater academic pressures on themselves (Duchesne & Larose, 2018), a tendency often linked to a fear of evaluation (Ang & Huan, 2006). However, given that the pattern of gender differences in mathematics, reading, and anxiety for Italian students on the PISA are larger than those found for students in most other OECD countries, it is not clear if our findings will generalize to other contexts.

In any case, at least for Italian students it is important to recognize that test anxiety accounted for about one-third of the differences between boys and girls in mathematics and reading test performance. This means that nearly two-thirds of the academic gender differences are unrelated to anxiety, indicating that other factors are at play. Notably, resilience (Donolato et al., 2020) and motivation (Pekrun & Marsh, 2022) may also exert considerable influences on academic outcomes. Recent research also suggests that resilience and self-concept significantly impact both mathematics and reading performance, even when controlling for intelligence (Donolato et al., 2019, 2020). It is thus important to investigate the interaction between pre-existing motivational aspects, achievement, and subsequent anxiety—that conversely—may not be uniformly influenced by gender. Other factors need to be considered as well, including spatial abilities and boys' advantage in mathematics (Geary et al., 2023) and language abilities and girls' advantage in reading and grammar (Reilly et al., 2019).

4.2. Differences in reading and mathematics

Italian girls' advantage in reading is consistent with a substantial body of literature demonstrating that gender differences in reading and related competencies are robust and significant (Reilly et al., 2019; Voyer et al., 2014). While our study did not include writing skills, we showed significant advantages for girls in both reading comprehension and grammar, in keeping with prior studies (see also Kovačević et al., 2007), with the differences being slightly larger in grammar. This finding is particularly noteworthy, as it pertains to fifth-grade students, and the existing literature suggests that such differences may become more pronounced as students' progress through their academic careers (Giofrè et al., 2020; Reilly et al., 2019).

Our findings for mathematics were also consistent with previous studies. Although gender differences in mathematics achievement are generally small on average (Else-Quest et al., 2010; Hutchison et al., 2019; Hyde et al., 2008), our results confirmed the gap in mathematics for Italian students (Giofrè et al., 2020; Stoet & Geary, 2018). Extensive cross-national assessments such as PISA have shown that boys tend to slightly outperform girls in mathematics, with this gap being particularly pronounced among high-achieving students and in some specific countries (Stoet & Geary, 2013). Moreover, recent analyses of intra-individual academic strengths (i.e., best academic subject) suggest that boys are more likely to exhibit relative strengths in mathematics and science, whereas girls tend to excel in reading (Balducci et al., 2024). This pattern, surprisingly, is more pronounced in countries with greater gender equality (Balducci et al., 2024). In any case, the point is these gaps at the mean, albeit small, could have broader effects (e.g., on later occupational choices) through larger gaps in academic strengths (Dekhtyar et al., 2018; Humphreys et al., 1993; Stoet & Geary, 2018).

Broader effects could also emerge through the variance ratio, whereby boys' achievement was more variable, as a group, in both mathematics and reading, but more so in mathematics than reading. This overall result is in line with other studies (Reilly et al., 2015; Reilly et al., 2019) but might underestimate gender differences in variability due to our relatively young sample. Hedges and Nowell's (1995) analysis of mostly older samples revealed larger variance ratios for both reading comprehension ($VR = 1.03$ to 1.16) and mathematics ($VR = 1.05$ to 1.25) than we found.

4.3. The complex interplay between achievement and emotions

Our results can be integrated within the control-value theory of achievement emotions proposed by Pekrun's seminal works (Pekrun, 2000; Pekrun, 2006). This theory posits that achievement is guided, in part, by subjective control over activities (e.g., persistent studying can lead to success), by expectancy beliefs (e.g., how competent students think they are), and by subjective value beliefs (e.g., what students think is important or interesting). Therefore, academic performance, and/or a specific academic competence, may develop over the school years depending on how students engage in activities that: i) they can control, ii) they find interesting, and iii) they feel competent in. Empirical support for the relevance of control-value theory in explaining gender

differences in academic performance is well-documented, both in mathematics (e.g., Guo et al., 2015) and reading (Durik et al., 2006; McGeown et al., 2012). Different degrees of test anxiety might stem from lower levels of perceived control and value beliefs in academic achievement, particularly in mathematic where girls often report lower levels of both perceived control and value (e.g., Pekrun & Stephens, 2010). Overall, the literature seems to indicate that students' emotional patterns can be explained by different degrees of competence beliefs and subjective control-values towards achievement goals, and this is true in both boys and girls.

In a recent study, Geary et al. (2023) uncovered distinct gender patterns in academic development with the transition from middle to secondary school. They observed that although mathematical achievement increased for girls and boys, attitudes towards math declined and anxiety levels rose. Notably, boys with above-average performance in mathematics exhibited resilience against these negative shifts in attitudes and anxiety, and showed gains in utility beliefs. Conversely, high achieving girls still reported higher anxiety and less positive attitudes towards mathematics compared to their male counterparts, even when they excelled in the subject. As noted earlier, across countries girls show higher mathematics anxiety than boys, even with control of mathematics achievement (Stoet et al., 2016). Our findings suggest that the relation between gender differences in test anxiety and achievement is not restricted to mathematics and might lower girls' performance on high-stakes tests across academic domains, although follow-up studies are needed to confirm this and to determine if these patterns hold across grade levels.

4.4. Caveat, limits and future directions

The anxiety measure implemented in this study was a short four-item questionnaire designed to assess state-test anxiety levels, aligning with Pekrun et al.'s (Pekrun et al., 2023a, 2023b) control-value theory. Nonetheless, the reliance on secondary data limited our ability to conduct a more comprehensive assessment of achievement-related emotions that might include positive aspects such as self-esteem, self-concept, resilience, or motivation. Research highlights the existence of distinct forms of anxiety, with specific forms having larger impact on specific subjects (i.e., Caviola et al., 2022), but our results are limited to general test anxiety.

Similarly, our cross-sectional data focused solely on test anxiety in fifth graders which might underestimate gender differences given that anxiety levels associated with schooling typically increase with age (Hill et al., 2016; Robson et al., 2023). Therefore, it is possible that the relation between anxiety with academic achievement is more pronounced in older students. Moreover, the cause-effect relation between test anxiety (and other forms of academic anxiety) and achievement are not fully understood (Geary et al., 2023; Ma & Xu, 2004). With increased exposure to different subjects in school, including mathematics, anxiety could become more differentiated. According to this account, anxiety in young children might initially manifest as a broad, undifferentiated emotional state. However, as they move through schooling and experience stressful events, such as exams, and more complex material (Embse & von der Embse & Hasson, 2012; Putwain, 2008), anxiety could become more common in some domains than others (Geary et al., 2023). During this period, the student-teacher relationship also becomes critically linked to levels of anxiety (Semeraro et al., 2020). In other words, our general test anxiety measure might underestimate the relation between specific forms of academic anxiety, such as math anxiety, and achievement in the associated domain.

Echoing the calls from both Robson et al. (2023) and von der Embse et al. (2018), our study highlights the need for longitudinal research to unravel the dynamic interplay between test anxiety, gender, and academic performance over time. The insights from Seipp (1991) stress the importance of considering various moderators and mediators that that may help explain this complex relationship. There are only a few longitudinal studies investigating the impact of anxiety and other attitudes on achievement (e.g., Putwain et al., 2018; Putwain & Wood, 2023). It is essential to undertake such studies to track the development of distinct forms of anxiety (e.g., test anxiety, mathematics anxiety) and their relation to academic achievement. These studies should incorporate a broad assessment of additional factors that may influence outcomes, such as self-esteem, and evaluate the role of contextual elements like the student-teacher relationship.

5. Conclusions

To conclude, we provided robust and reliable estimates of the relationship between test anxiety and both mathematics and reading performance using large and representative samples of Italian children. Our meta-analytic approach was particularly important because it allowed us to demonstrate that the effect was robust across our samples and when combined with other meta-analytic results overcame the important limits of previous studies. Test anxiety seems to be an important contributor to gender differences in both mathematics and reading test performance, although as noted cause-effect relations cannot be established due to the correlational nature of our data. Whatever the cause-effect relations, about two-thirds of the effects are probably explained by other factors underscoring the need for future research to explore the additional variables that might be related to academic performance disparities.

CRediT authorship contribution statement

Sara Caviola: Writing – review & editing, Writing – original draft, Supervision, Conceptualization. **David Giofrè:** Writing – review & editing, Writing – original draft, Supervision, Conceptualization. **Tommaso Feraco:** Writing – review & editing, Formal analysis, Data curation, Conceptualization. **Enrico Toffalini:** Writing – review & editing, Formal analysis, Data curation. **Katie Allen:** Writing – review & editing, Writing – original draft. **David C. Geary:** Writing – review & editing, Writing – original draft, Conceptualization.

Declaration of ethical approval

All relevant ethical guidelines and principles were carefully considered in the preparation of this scientific article. The research, relying on secondary data analyses, strictly adhered to ethical standards to ensure that potential impacts on humans and the environment were minimized.

Ethical statement

The study used anonymized secondary data, not requiring institutional approval, in line with Italian rules. It adheres to ethical standards (COPE, APA, Helsinki), ensuring privacy and confidentiality.

Funding

This work was carried out within the framework of the project "Cognitive & emotional pathways to hidden STEM: longitudinal and intervention studies fostering higher educational choices" (Prot. 2022FZ595K; CUP: D53D23009660006), funded under the PRIN 2022 call—Progetti di Ricerca di Rilevante Interesse Nazionale (Decreto Direttoriale n. 104 del 02-02-2022), by the Italian Ministry of University and Research (MUR), and co-funded by the European Union—NextGenerationEU).

Declaration of competing interest

Dr. Sara Caviola had no involvement in the peer-review of this article and has no access to information regarding its peer-review. Full responsibility for the editorial process for this article was taken by Dr.

Greiff (Editor in Chief).

Acknowledgements

This work was carried out within the framework of the project "Cognitive & emotional pathways to hidden STEM: longitudinal and intervention studies fostering higher educational choices" (Prot. 2022FZ595K; CUP: D53D23009660006), funded under the PRIN 2022 call—Progetti di Ricerca di Rilevante Interesse Nazionale (Decreto Direttoriale n. 104 del 02-02-2022), by the Italian Ministry of University and Research (MUR), and co-funded by the European Union—NextGenerationEU. The authors would like to thank the Ufficio Statistico INVALSI for granting access to the data. This paper was written while David Giofrè was spending a period of research at the University of Missouri with David Geary, under the auspices of a Fullbring.

Appendix A. Supplementary information

Supplementary information to this article can be found online at https://doi.org/10.1016/j.lindif.2025.102726.

Data availability

The data that support the findings of this study are available from INVALSI. Restrictions apply to the availability of these data, which were used under license for this study. Data are available https://invalsi-serviziostatistico.cineca.it/ with the permission of INVALSI.

References

Ahmed, W., Minnaert, A., Kuyper, H., & van der Werf, G. (2012). Reciprocal relationships between math self-concept and math anxiety. *Learning and Individual Differences, 22*(3), 385–389. https://doi.org/10.1016/j.lindif.2011.12.004

Ang, R. P., & Huan, V. S. (2006). Academic expectations stress inventory. *Educational and Psychological Measurement, 66*(3), 522–539. https://doi.org/10.1177/0013164405282461

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

Barroso, C., Ganley, C. M., McGraw, A. L., Geer, E. A., Hart, S. A., & Daucourt, M. C. (2021). A meta-analysis of the relation between math anxiety and math achievement. *Psychological Bulletin, 147*(2), 134–168. https://doi.org/10.1037/bul0000307

Bosshard, M., & Gomez, P. (2024). Effectiveness of stress arousal reappraisal and stress-is-enhancing mindset interventions on task performance outcomes: A meta-analysis of randomized controlled trials. *Scientific Reports, 14*(1), 7923. https://doi.org/10.1038/s41598-024-58408-w

Camacho-Morles, J., Slemp, G. R., Pekrun, R., Loderer, K., Hou, H., & Oades, L. G. (2021). Activity achievement emotions and academic performance: A Meta-analysis. *Educational Psychology Review, 33*(3), 1051–1095. https://doi.org/10.1007/s10648-020-09585-3

Cassady, J. C., & Johnson, R. E. (2002). Cognitive test anxiety and academic performance. *Contemporary Educational Psychology, 27*(2), 270–295. https://doi.org/10.1006/ceps.2001.1094

Cassady, J., Finch, W. H., & Heath, J. A. (2022). Early assessment of cognitive skills, self-regulated learning skills, and attitudes toward education predict university success at graduation. *Journal of Postsecondary Student Success, 1*(4). https://doi.org/10.33009/fsop_jpss129806

Caviola, S., Toffalini, E., Giofrè, D., Ruiz, J. M., Szűcs, D., & Mammarella, I. C. (2022). Math performance and academic anxiety forms, from sociodemographic to cognitive aspects: A meta-analysis on 906,311 participants. *Educational Psychology Review, 34*(1), 363–399. https://doi.org/10.1007/s10648-021-09618-5

Cheng, B. H., & McCarthy, J. M. (2018). Understanding the dark and bright sides of anxiety: A theory of workplace anxiety. *Journal of Applied Psychology, 103*(5), 537–560. https://doi.org/10.1037/apl0000266

Choe, K. W., Jenifer, J. B., Rozek, C. S., Berman, M. G., & Beilock, S. L. (2019). Calculated avoidance: Math anxiety predicts math avoidance in effort-based decision-making. *Science Advances, 5*(11), Article eaay1062. https://doi.org/10.1126/sciadv.aay1062

Cizek, G. J., & Burg, S. S. (2006). *Addressing test anxiety in a high-stakes environment: Strategies for classroom and schools*. Corwin Press.

Deieso, D., & Fraser, B. J. (2018). In i.e. Progress (Ed.), *52. Mathematics anxiety: Its assessment, determinants and remedies* (pp. 113–137).

Dekhtyar, S., Weber, D., Helgertz, J., & Herlitz, A. (2018). Sex differences in academic strengths contribute to gender segregation in education and occupation: A longitudinal examination of 167,776 individuals. *Intelligence, 67*, 84–92. https://doi.org/10.1016/j.intell.2017.11.007

Derakshan, N., & Eysenck, M. W. (2009). Anxiety, processing efficiency, and cognitive performance. *European Psychologist, 14*(2), 168–176. https://doi.org/10.1027/1016-9040.14.2.168

Donolato, E., Giofrè, D., & Mammarella, I. C. (2019). Working memory, negative affect and personal assets: How do they relate to mathematics and reading literacy? *PLoS One, 14*(6), Article e0218921. https://doi.org/10.1371/journal.pone.0218921

Donolato, E., Toffalini, E., Giofrè, D., Caviola, S., & Mammarella, I. C. (2020). Going beyond mathematics anxiety in primary and middle school students: The role of ego-resiliency in mathematics. *Mind, Brain, and Education, 14*(3), 255–266. https://doi.org/10.1111/mbe.12251

Duchesne, S., & Larose, S. (2018). Academic competence and achievement goals: Self-pressure and disruptive behaviors as mediators. *Learning and Individual Differences, 68*, 41–50. https://doi.org/10.1016/j.lindif.2018.09.008

Durik, A. M., Vida, M., & Eccles, J. S. (2006). Task values and ability beliefs as predictors of high school literacy choices: A developmental analysis. *Journal of Educational Psychology, 98*(2), 382–393. https://doi.org/10.1037/0022-0663.98.2.382

Efklides, A., & Volet, S. (2005). Emotional experiences during learning: Multiple, situated and dynamic. *Learning and Instruction, 15*(5), 377–380. https://doi.org/10.1016/j.learninstruc.2005.07.006

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

Eysenck, M. W., & Calvo, M. G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition & Emotion, 6*(6), 409–434. https://doi.org/10.1080/02699939208409696

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion, 7*(2), 336–353. https://doi.org/10.1037/1528-3542.7.2.336

Foley, A. E., Herts, J. B., Borgonovi, F., Guerriero, S., Levine, S. C., & Beilock, S. L. (2017). The math anxiety-performance link. *Current Directions in Psychological Science, 26*(1), 52–58. https://doi.org/10.1177/0963721416672463

Galla, B. M., & Wood, J. J. (2012). Emotional self-efficacy moderates anxiety-related impairments in math performance in elementary school-age youth. *Personality and Individual Differences, 52*(2), 118–122. https://doi.org/10.1016/j.paid.2011.09.012

Ganley, C. M., & McGraw, A. L. (2016). The development and validation of a revised version of the math anxiety scale for young children. *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.01181

Geary, D. C., Hoard, M. K., Nugent, L., Chu, F., Scofield, J. E., & Ferguson Hibbard, D. (2019). Sex differences in mathematics anxiety and attitudes: Concurrent and longitudinal relations to mathematical competence. *Journal of Educational Psychology, 111*(8), 1447–1461. https://doi.org/10.1037/edu0000355

Geary, D. C., Hoard, M. K., Nugent, L., & Ünal, Z. E. (2023). Sex differences in developmental pathways to mathematical competence. *Journal of Educational Psychology, 115*(2), 212–228. https://doi.org/10.1037/edu0000763

Giofrè, D., Allen, K., Toffalini, E., & Caviola, S. (2022). The impasse on gender differences in intelligence: A meta-analysis on WISC batteries. *Educational Psychology Review, 34*(4), 2543–2568. https://doi.org/10.1007/s10648-022-09705-1

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*, Article 101467. https://doi.org/10.1016/j.intell.2020.101467

Goetz, T., & Hall, N. C. (2013). Emotion and achievement in the classroom. In *International guide to student achievement* (pp. 192–195). https://psycnet.apa.org/record/2013-13643-059.

Guo, J., Marsh, H. W., Parker, P. D., Morin, A. J. S., & Yeung, A. S. (2015). Expectancy-value in mathematics, gender and socioeconomic background as predictors of achievement and aspirations: A multi-cohort study. *Learning and Individual Differences, 37*, 161–168. https://doi.org/10.1016/j.lindif.2015.01.008

Halpern, D. F., & LaMay, M. L. (2000). Smarter sex: A critical review of sex differences in intelligence. *Educational Psychology Review, 12*(2), 229. https://doi.org/10.1023/A:1009027516424

Hasty, L. M., Malanchini, M., Shakeshaft, N., Schofield, K., Malanchini, M., & Wang, Z. (2021). When anxiety becomes my propeller: Mental toughness moderates the relation between academic anxiety and academic avoidance. *British Journal of Educational Psychology, 91*(1), 368–390. https://doi.org/10.1111/bjep.12366

Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science, 269*(5220), 41–45. https://doi.org/10.1126/science.760427

Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research, 58*(1), 47–77.

Hill, F., Mammarella, I. C., Devine, A., Caviola, S., Passolunghi, M. C., & Szűcs, D. (2016). Maths anxiety in primary and secondary school students: Gender differences, developmental changes and anxiety specificity. *Learning and Individual Differences, 48*, 45–53. https://doi.org/10.1016/j.lindif.2016.02.006

Huberty, T. J. (2012). The developmental psychopathology of depression. In T. J. Huberty (Ed.), *Anxiety and depression in children and adolescents* (pp. 55–75). New York: Springer.

Humphreys, L. G., Lubinski, D., & Yao, G. (1993). Utility of predicting group membership and the role of spatial visualization in becoming an engineer, physical scientist, or artist. *Journal of Applied Psychology, 78*, 250–261. https://doi.org/10.1037/0021-9010.78.2.250

Hutchison, J. E., Lyons, I. M., & Ansari, D. (2019). More similar than different: Gender differences in children's basic numerical skills are the exception not the rule. *Child Development, 90*(1). https://doi.org/10.1111/cdev.13044

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science, 321*(5888), 494–495. https://doi.org/10.1126/science.116036

Jeanneteau, F., & Coutellier, L. (2022). The glucocorticoid footprint on the memory engram. *Current Opinion in Endocrine and Metabolic Research, 25*, Article 100378. https://doi.org/10.1016/j.coemr.2022.100378

Kovačević, M., Kuvač Kraljević, J., & Cepanec, M. (2007). Sex differences in lexical and grammatical development in Croatian. In M. Eriksson (Ed.), *Proceedings from the first European network meeting on the communicative development inventories* (pp. 5–15). Gävle, Sweden: Gävle University Press.

Liebert, R. M., & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological Reports, 20*(3), 975–978. https://doi.org/10.2466/pr0.1967.20.3.975

Linnenbrink, E. A. (2006). Emotion research in education: Theoretical and methodological perspectives on the integration of affect, motivation, and cognition. *Educational Psychology Review, 18*(4), 307–314. https://doi.org/10.1007/s10648-006-9028-x

Linnenbrink-Garcia, L., & Pekrun, R. (2011). Students' emotions and academic engagement: Introduction to the special issue. *Contemporary Educational Psychology, 36*(1), 1–3. https://doi.org/10.1016/j.cedpsych.2010.11.004

Loderer, K., Pekrun, R., & Lester, J. C. (2020). Beyond cold technology: A systematic review and meta-analysis on emotions in technology-based learning environments. *Learning and Instruction, 70*, Article 101162. https://doi.org/10.1016/j.learninstruc.2018.08.002

Lowe, P. A. (2014). The test anxiety measure for adolescents (TAMA): Examination of the reliability and validity of the scores of a new multidimensional measure of test anxiety for middle and high school students. *Journal of Psychoeducational Assessment, 32*(5), 404–416. https://doi.org/10.1177/0734282913520595

Lupien, S. J., Maheu, F., Tu, M., Fiocco, A., & Schramek, T. E. (2007). The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition. *Brain and Cognition, 65*(3), 209–237. https://doi.org/10.1016/j.bandc.2007.02.007

Lynn, R., & Mikk, J. (2009). National IQs predict educational attainment in math, reading and science across 56 nations. *Intelligence, 37*(3), 305–310. https://doi.org/10.1016/j.intell.2009.01.002

Ma, X., & Xu, J. (2004). The causal ordering of mathematics anxiety and mathematics achievement: A longitudinal panel analysis. *Journal of Adolescence, 27*(2), 165–179. https://doi.org/10.1016/j.adolescence.2003.11.003

Malanchini, M., Wang, Z., Voronin, I., Schenker, V. J., Plomin, R., Petrill, S. A., & Kovas, Y. (2017). Reading self-perceived ability, enjoyment and achievement: A genetically informative study of their reciprocal links over time. *Developmental Psychology, 53*(4), 698–712. https://doi.org/10.1037/dev0000209

Mandler, G., & Sarason, S. B. (1952). A study of anxiety and learning. *The Journal of Abnormal and Social Psychology, 47*(2), 166–173. https://doi.org/10.1037/h0062855

McGeown, S., Goodwin, H., Henderson, N., & Wright, P. (2012). Gender differences in reading motivation: Does sex or gender identity provide a better account? *Journal of Research in Reading, 35*(3), 328–336. https://doi.org/10.1111/j.1467-9817.2010.01481.x

Nie, Y., Lau, S., & Liau, A. K. (2011). Role of academic self-efficacy in moderating the relation between task importance and test anxiety. *Learning and Individual Differences, 21*(6), 736–741. https://doi.org/10.1016/j.lindif.2011.09.005

Pekrun, R. (2000). A social cognitive, control-value theory of achievement emotions. In Heckhausen, J. (Ed.), Motivational psychology of human development. Oxford, UK: Elsevier.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315–341. https://doi.org/10.1007/s10648-006-9029-9

Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., & Goetz, T. (2017a). Achievement emotions and academic performance: Longitudinal models of reciprocal effects. *Child Development, 88*(5), 1653–1670. https://doi.org/10.1111/cdev.12704

Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., & Goetz, T. (2017b). Achievement emotions and academic performance: Longitudinal models of reciprocal effects. *Child Development, 88*(5), 1653–1670. https://doi.org/10.1111/cdev.12704

Pekrun, R., & Linnenbrink-Garcia, L. (2014). Introduction to emotions in education. In *International handbook of emotions in education* (pp. 1–10). https://doi.org/10.4324/9780203148211

Pekrun, R., & Marsh, H. W. (2022). Research on situated motivation and emotion: Progress and open problems. *Learning and Instruction, 81*, Article 101664. https://doi.org/10.1016/j.learninstruc.2022.101664

Pekrun, R., Marsh, H. W., Suessenbach, F., Frenzel, A. C., & Goetz, T. (2023a). School grades and students' emotions: Longitudinal models of within-person reciprocal effects. *Learning and Instruction, 83*(May 2022), 101626. https://doi.org/10.1016/j.learninstruc.2022.101626

Pekrun, R., Marsh, H. W., Suessenbach, F., Frenzel, A. C., & Goetz, T. (2023b). School grades and students' emotions: Longitudinal models of within-person reciprocal effects. *Learning and Instruction, 83*, Article 101626.

Pekrun, R., & Stephens, E. J. (2010). Achievement emotions: A control-value approach. *Social and Personality Psychology Compass, 4*(4), 238–255. https://doi.org/10.1111/j.1751-9004.2010.00259.x

Putwain, D. W. (2008). Deconstructing test anxiety. *Emotional and Behavioural Difficulties, 13*(2), 141–155. https://doi.org/10.1080/13632750802027713

Putwain, D. W., Becker, S., Symes, W., & Pekrun, R. (2018). Reciprocal relations between students' academic enjoyment, boredom, and achievement over time. *Learning and Instruction, 54*, 73–81. https://doi.org/10.1016/j.learninstruc.2017.08.004

Putwain, D. W., Connors, L., & Symes, W. (2010). Do cognitive distortions mediate the test anxiety–examination performance relationship? *Educational Psychology, 30*(1), 11–26. https://doi.org/10.1080/01443410903328866

Putwain, D. W., Gallard, D., Beaumont, J., Loderer, K., & von der Embse, N. P. (2021). Does test anxiety predispose poor school-related wellbeing and enhanced risk of emotional disorders? *Cognitive Therapy and Research, 45*(6), 1150–1162. https://doi.org/10.1007/s10608-021-10211-x

Putwain, D. W., Stockinger, K., von der Embse, N. P., Suldo, S. M., & Daumiller, M. (2021). Test anxiety, anxiety disorders, and school-related wellbeing: Manifestations of the same or different constructs? *Journal of School Psychology, 88*, 47–67. https://doi.org/10.1016/j.jsp.2021.08.001

Putwain, D. W., von der Embse, N. P., Rainbird, E. C., & West, G. (2021). The development and validation of a new multidimensional test anxiety scale (MTAS). *European Journal of Psychological Assessment, 37*(3), 236–246. https://doi.org/10.1027/1015-5759/a000604

Putwain, D. W., & Wood, P. (2023). Anxiety in the mathematics classroom: Reciprocal relations with control and value, and relations with subsequent achievement. *ZDM Mathematics Education, 55*(2), 285–298. https://doi.org/10.1007/s11858-022-01390-2

R Core Team. (2023). *R: A language and environment for statistical computing*. Vienna, Austria, Austria: R Foundation for Statistical Computing. http://www.r-project.org/.

Ramirez, G., Hooper, S. Y., Kersting, N. B., Ferguson, R., & Yeager, D. (2018). Teacher math anxiety relates to adolescent students' math achievement. *AERA Open, 4*(1), Article 233285841875605. https://doi.org/10.1177/2332858418756052

Reilly, D. (2012). Gender, culture, and sex-typed cognitive abilities. *PLoS One, 7*(7), Article e39904. https://doi.org/10.1371/journal.pone.0039904

Reilly, D., Neumann, D. L., & Andrews, G. (2015). Sex differences in mathematics and science achievement: A meta-analysis of National Assessment of educational Progress assessments. *Journal of Educational Psychology, 107*(3), 645–662. https://doi.org/10.1037/edu0000012

Reilly, D., Neumann, D. L., & Andrews, G. (2016). Sex and sex-role differences in specific cognitive abilities. *Intelligence, 54*, 147–158. https://doi.org/10.1016/j.intell.2015.12.004

Reilly, D., Neumann, D. L., & Andrews, G. (2019). Gender differences in reading and writing achievement: Evidence from the National Assessment of educational Progress (NAEP). *American Psychologist, 74*(4), 445–458. https://doi.org/10.1037/amp0000356

Robson, D. A., Johnstone, S. J., Putwain, D. W., & Howard, S. (2023). Test anxiety in primary school children: A 20-year systematic review and meta-analysis. *Journal of School Psychology, 98*, 39–60. https://doi.org/10.1016/j.jsp.2023.02.003

Schutz, P. A., & Pekrun, R. (2007). *Emotion in education*. Elsevier Academic Press.

Seipp, B. (1991). Anxiety and academic performance: A meta-analysis of findings. *Anxiety Research, 4*(1), 27–41. https://doi.org/10.1080/08917779108248762

Semeraro, C., Giofrè, D., Coppola, G., Lucangeli, D., & Cassibba, R. (2020). The role of cognitive and non-cognitive factors in mathematics achievement: The importance of the quality of the student-teacher relationship in middle school. *PLoS One, 15*(4), Article e0231381. https://doi.org/10.1371/journal.pone.0231381

Sorvo, R., Koponen, T., Viholainen, H., Aro, T., Räikkönen, E., Peura, P., Dowker, A., & Aro, M. (2017). Math anxiety and its relationship with basic arithmetic skills among primary school children. *British Journal of Educational Psychology, 87*(3), 309–327. https://doi.org/10.1111/bjep.12151

Spielberger, C. D., Anton, W. D., & Bedell, J. (1976). The nature and treatment of test anxiety. In M. Zuckerman, & C. D. Spielberger (Eds.), *Emotions and anxiety: New concepts, methods, and applications* (pp. 317–345). Erlbaum.

Stanley, J. C., Benbow, C. P., Brody, L. E., Dauber, S., & Lupwowski, A. E. (1992). Gender differences on eighty- six nationally standardized aptitude and achievement tests. In N. Colangelo, S. G. Assouline, & D. L. Ambroson (Eds.), *Talent development* (pp. 42–65). Unionville, New York: Trillium Press.

Stoet, G., Bailey, D. H., Moore, A. M., & Geary, D. C. (2016). Countries with higher levels of gender equality show larger national sex differences in mathematics anxiety and relatively lower parental mathematics valuation for girls. *PLoS One, 11*(4), 1–24. https://doi.org/10.1371/journal.pone.0153857

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), Article e57988. 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). Sex-specific academic ability and attitude patterns in students across developed nations. *Intelligence, 81*, Article 101453. https://doi.org/10.1016/j.intell.2020.101453

Szczygieł, M., & Pieronkiewicz, B. (2022). Exploring the nature of math anxiety in young children: Intensity, prevalence, reasons. *Mathematical Thinking and Learning, 24*(3), 248–266. https://doi.org/10.1080/10986065.2021.1882363

Talbot, L. (2016). Test anxiety: Prevalence, effects, and interventions for elementary school students. *James Madison Undergraduate Research Journal, 3*(1), 41–50.

von der Embse, N., & Hasson, R. (2012). Test anxiety and high-stakes test performance between school settings: Implications for educators. *Preventing School Failure, 56*(3), 180–187. https://doi.org/10.1080/1045988X.2011.633285

von der Embse, N., Jester, D., Roy, D., & Post, J. (2018). Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review. *Journal of Affective Disorders, 227*, 483–493. https://doi.org/10.1016/j.jad.2017.11.048

von der Embse, N. P., & Witmer, S. E. (2014). High-stakes accountability: Student anxiety and large-scale testing. *Journal of Applied School Psychology, 30*(2), 132–156. https://doi.org/10.1080/15377903.2014.888529

Voyer, D., Voyer Susan, D. D, & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin, 140*(4), 1174–1204. https://doi.org/10.1037/a0036620

Wigfield, A., Muenks, K., & Rosenzweig, E. Q. (2015). Children's achievement motivation in school. In *The routledge international handbook of social psychology of the classroom* (pp. 9–20). https://doi.org/10.4324/9781315716923

World Economic Forum. (2023). Global gender gap report. https://www3.weforum.org/docs/WEF_GGGR_2023.pdf.

Wu, S. S., Barth, M., Amin, H., Malcarne, V., & Menon, V. (2012). Math anxiety in second and third graders and its relation to mathematics achievement. *Frontiers in Psychology, 3*. https://doi.org/10.3389/fpsyg.2012.00162

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology, 18*(5), 459–482. https://doi.org/10.1002/cne.920180503

Zeidner, M. (1998). An introduction to the domain of test anxiety. Test anxiety: The state of the art, Plenum, New York.