Research Article in Scientific Thinking

Exploring Gender Differences in Scientific Reasoning Ability Among Junior Secondary Students by Mediation Analysis

Luoyi Zheng[1], Xingkai Chen[1], Shantong Huand[1], Yuhong Zhu[1,*]

[1] Hangzhou Normal University, Hangzhou, Zhejiang , China

[*] Corresponding Author: zyh@hznu.edu.cn

(Received: 05-10/2026; Accepted: 06/28/2026; Published: 08/15/2026)

DOI: https://doi.org/10.37906/real.2026.3

Abstract: This study examined gender differences in scientific reasoning ability among Chinese junior secondary school students and investigated the mediating mechanisms underlying these differences. A total of 453 students from four public schools in Hangzhou, China, completed the Lawson Classroom Test of Scientific Reasoning (LCTSR) and self-reported their subject preference (language vs. mathematics) and thinking style orientations (memory-oriented thinking vs. logic-oriented thinking). The results revealed a significant overall gender difference in scientific reasoning ability, with higher performance observed among male students, particularly in conservation, proportional, and hypothetical-deductive reasoning, although the effect sizes were small. However, when controlling for the proposed mediators, the direct effect of gender on scientific reasoning ability was no longer significant. Mediation analyses further indicated that gender influenced scientific reasoning indirectly through mathematics preference and logic-oriented thinking, which jointly accounted for a substantial proportion (approximately 92%) of the total effect. Further analyses at the sub-dimension level showed that these mediating effects were particularly pronounced in higher-order reasoning components, such as proportional and hypothetical-deductive reasoning, indicating dimension-specific mediation patterns. These findings suggest that observed gender disparities in scientific reasoning ability may be partially shaped by socialization processes associated with stereotyped academic preferences rather than reflecting innate cognitive differences.

Keywords: gender differences; Lawson Classroom Test of Scientific Reasoning (LCTSR); mediation analysis; scientific reasoning ability

1. Introduction

Against the backdrop of global educational reform, scientific literacy has become one of the core competencies expected of individuals in the twenty-first century. Although this concept is widely discussed and applied within science education, there remains no internationally agreed-upon definition, with diverse and influential interpretations proposed by international organizations and academic traditions. The Organization for Economic Co-operation and Development (OECD) defines scientific literacy as encompassing the competencies to explain phenomena scientifically, evaluate and design scientific enquiry, and interpret data and evidence scientifically—capacities that together enable individuals to engage in reasoned discourse about science and technology (OECD, 2017). Within academic scholarship, Roberts offers a widely cited distinction between two visions of scientific literacy(Roberts, 2013). Vision I is concerned with the internal aspects of science, emphasizing mastery of core scientific knowledge, theoretical frameworks, and scientific inquiry processes. Vision II, by contrast, foregrounds the application of science in everyday and societal contexts, conceptualizing scientific literacy as citizens' capacity to use scientific knowledge when making decisions related to real-life situations and socially relevant scientific issues. Building upon Vision II, Feinstein conceptualizes scientifically literate individuals as "competent outsiders" who are able to recognize when science is relevant, evaluate scientific claims, and participate meaningfully in public discourse within everyday contexts(Feinstein, 2011). Achieving this advanced form of scientific literacy, however, requires more than the recall of scientific knowledge. It necessitates a core competency that enables individuals to critically evaluate evidence, construct plausible explanations, and establish logical connections between theories and empirical data. This competency is commonly conceptualized as scientific reasoning ability (Kind & Osborne, 2017).

Although scientific reasoning ability is widely regarded as a core component of scientific literacy and a key element of higher-order thinking, its development has not been sufficiently emphasized or systematically cultivated within many educational systems worldwide. Traditional science instruction often prioritizes the transmission of factual knowledge and rote memorization, rather than explicit engagement with critical thinking and scientific reasoning processes (Forawi, 2016). This imbalance may result in students possessing substantial subject knowledge yet lacking the capacity to analyze evidence and solve complex real-world problems involving scientific reasoning. Therefore, cultivating thinking skills in science education is of paramount importance (García-Carmona, 2025). Against this backdrop, identifying factors that influence the development of scientific reasoning ability has become particularly crucial. Among various influencing factors, gender—as a fundamental demographic variable—has long attracted scholarly attention, particularly in relation to disparities in participation and performance within science, technology, engineering, and mathematics (STEM) fields (Wang & Degol, 2017). However, existing research has primarily concentrated on gender differences in academic achievement, subject interest, or career aspirations, and findings remain inconclusive regarding whether gender-based disparities extend to deeper, cross-disciplinary scientific reasoning ability.

1.1. Theoretical Background

1.1.1. Scientific Reasoning Ability

Scientific reasoning ability, as a key dimension of scientific literacy, refers to an individual's capacity for logical judgement, control of variables, and causal inference within scientific contexts. Within science education research, a relatively mature set of assessment instruments has been developed to measure scientific reasoning ability. Early studies, grounded in Piaget's theory of cognitive development, conceptualized students' reasoning development through formal operational tasks that delineated distinct developmental stages (Piaget & Inhelder, 1999). Building on this foundation, scholars such as Lawson further conceptualized and operationalized the core components of scientific reasoning and developed the widely used Lawson Classroom Test of Scientific Reasoning (LCTSR). Its two-tier item format was designed to distinguish more precisely between students' selected answers and the underlying reasoning that supports them (Lawson et al., 2000). In parallel, the Group Assessment of Logical Thinking (GALT) and the Test of Logical Thinking (TOLT), both derived from Piagetian task development, were applied at secondary and tertiary levels to systematically assess key reasoning dimensions, including control of variables, proportional reasoning, and probabilistic reasoning (Roadrangka, 1983; Tobin. & Capie., 1981). More recently, the assessment paradigm for scientific reasoning has continued to evolve, shifting from the measurement of discrete and structured reasoning skills toward more nuanced evaluations of higher-order cognitive processes. These approaches emphasize how students utilize evidence, interpret data, and construct causal explanations—such as mechanistic reasoning—within more complex and open-ended contexts (Martin & Graulich, 2023). This evolution has fostered a complementary relationship between structured psychometric tests and contextualized assessments, thereby providing diversified evidence for learning diagnostics and instructional improvement.

Existing research has extensively examined the relationship between scientific reasoning ability and factors such as students' academic performance and cultural background. For example, Bao et al. employed the LCTSR to conduct a cross-cultural comparison between Chinese and American university students (Bao et al., 2009). The results indicated that although Chinese students significantly outperformed their American counterparts on physics concept inventories (e.g., FCI and BEMA)—a difference often attributed to longer and more intensive STEM instruction—no significant differences were observed in scientific reasoning ability between the two groups. These findings suggest that the accumulation of subject-specific knowledge does not necessarily lead to the development of scientific reasoning ability. And other studies demonstrating the predictive validity of scientific reasoning ability for later academic success. For instance, Sapia et al. reported that higher LCTSR scores were consistently associated with stronger academic performance in higher education, underscoring the instrument's utility in identifying students with high learning potential (Sapia et al., 2022). Taken together, while these studies support the relative independence and predictive validity of scientific reasoning ability, they provide limited insight into why such abilities vary across different groups.

1.1.2. Gender Differences Research

Gender differences in scientific literacy remain a subject of sustained attention in international educational assessments and academic research, with studies primarily focusing on gender disparities in academic achievement, scientific learning, and scientific cognition. Some empirical research has reported that males tend to demonstrate higher performance in certain domains of science-related learning (Balducci et al., 2024). In contrast, other studies have found no significant differences between males and females in general intelligence (Giofrè et al., 2022; Reilly et al., 2022). Extensive research indicates that performance disparities between genders in learning do not stem from a single innate factor, but from the dynamic interplay of cognitive, motivational, and socio-environmental influences. For instance, performance gaps observed at younger ages have been shown to be closely associated with sociocognitive factors rather than biological determinants (Martinot et al., 2025). Moreover, gender differences in motivational characteristics play a direct role in shaping students' academic achievement and educational aspirations (Olive et al., 2022). Students' perceptions of gender norms within their social environments further influence their academic self-concept, thereby challenging simplistic explanations that attribute gender disparities to inherent brain differences alone (Fine, 2010; Sicard et al., 2021). Taken together, the gender differences in STEM fields constitute a complex phenomenon shaped by interacting cognitive, motivational, and sociocultural factors, which operate jointly across educational contexts (Wang & Degol, 2017).

At the international level, research specifically examining the relationship between scientific reasoning ability and gender remains relatively limited. Some studies have reported that males tend to outperform females on scientific reasoning assessments (Luo et al., 2021), a finding that aligns with several large-scale investigations suggesting an average male advantage in tasks involving scientific and quantitative reasoning (Strand et al., 2006; Wai et al., 2010). Conversely, other studies indicate that males do not outperform females in all domains, with no discernible gender differences observed in scientific reasoning ability (Bezci & Sungur, 2021; Piraksa et al., 2014). Given the lack of consensus in existing empirical findings and the well-documented influence of social and educational contexts on gender-related outcomes, further examination of gender differences in scientific reasoning across diverse contexts is warranted. Accordingly, situated within the cultural and educational context of China, the present study focuses on junior secondary school students to empirically investigate gender differences in scientific reasoning ability and to explore the potential mechanisms underlying these differences.

1.2. Research Questions

The core issue of this study lies in investigating gender differences in scientific reasoning ability among secondary school students. Based on data obtained from the LCTSR, this study further examines the extent to which students' self-assessment tendencies are associated with scientific reasoning ability across genders. Accordingly, this research aims to address the following two key questions:

1. Do gender differences exist in secondary school pupils' scientific reasoning abilities?

2. If such differences exist, through what mechanisms do gender differences shape variations in students' scientific reasoning abilities?

2. Methods

2.1. Participants

This study involved students from four state-funded secondary schools located in Hangzhou, Zhejiang Province, China. To enhance representation across year groups, students were randomly selected from different districts and year levels within Hangzhou, with each school including students from at least three different year groups.

The assessment was conducted during the first term of the 2024-2025 academic year, lasting one month. It took the form of a paper-and-pencil test, with each sitting lasting 35 minutes and administered under teacher supervision on a class-by-class basis. Prior to administration, students were informed that their participation was voluntary, their responses would not affect their course grades, and they were encouraged to respond honestly.

A total of 457 students participated in the study. After data screening and the exclusion of incomplete or invalid responses, 452 valid questionnaires were retained for analysis, yielding a usable response rate of 98.91%. The number of students and gender distribution across each year group are presented in Table 1.

Table 1. Sample Size and Gender Distribution by Grade Level.

Grade	N	Male	Female	Male/female ratio (%)
7	159	81	78	50.9/49.1
8	155	85	70	54.8/45.2
9	138	73	65	52.9/47.1

2.2. Instruments

2.2.1. Self-Assessment

This study employed two self-designed dichotomous self-assessment items to assess students' subject preferences (Language vs. Mathematics) and their thinking style orientations (memory-oriented thinking vs. logic-oriented thinking). These items were intended to provide descriptive data for group-level analyses and to offer an exploratory perspective for examining potential mechanisms through which gender differences may be associated with scientific reasoning ability in subsequent analyses.

The self-assessment items were as follows:

1. Do you prefer Language or Mathematics?

2. Do you consider yourself better at memory-oriented thinking or logic-oriented thinking?

2.2.2. Lawson Classroom Test of Scientific Reasoning (LCTSR)

The scientific reasoning ability of junior secondary school pupils was assessed using the LCTSR. The LCTSR has been widely employed in physics education research to examine the relationship between scientific reasoning and conceptual learning in physics (Coletta & Phillips, 2005). Evidence for the reliability of the instrument has been reported in prior studies (Bao et al., 2018). The LCTSR is a 24-item, two-tier multiple-choice instrument consisting of 12 item pairs designed to measure scientific reasoning. It is a revised version of the original Classroom Test of Formal Reasoning developed by Lawson (1978), with the revision completed in 2000. The test assesses six reasoning dimensions: conservation reasoning, proportional reasoning, control of variables reasoning, probability reasoning, correlational reasoning, and hypothetical-deductive reasoning. These dimensions are widely recognized within academia, and their results substantially reflect students' scientific reasoning abilities. For each item pair, students first predict an outcome based on a presented scenario and then select the corresponding reasoning. A score of one point is awarded only when both responses in the pair are correct. The total test score is 24 points, with the question numbers corresponding to each dimension shown in Table 2. Reliability analysis of the 24 items using SPSS yielded a Cronbach's α coefficient of 0.82, which exceeds the conventional threshold of 0.70 commonly interpreted as acceptable internal consistency in science education research. This suggests satisfactory internal consistency for the instrument in the context of this study (Taber, 2018).

Table 2. LCTSR Dimensions and Corresponding Item Numbers.

Dimensions	Item numbers
conservation reasoning	1-4
proportional reasoning	5-8
control of variables reasoning	9-14
probability reasoning	15-18
correlational reasoning	19-20
hypothetical-deductive reasoning	21-24

2.2.3. Generalized Mediated Effects Model

Upon obtaining LCTSR data, this study formulated hypotheses regarding potential mediating effects to investigate whether subject preferences and thinking style orientations mediate the relationship between students 'gender and scientific reasoning ability. These hypotheses were tested using a generalized mediation effect model. Traditional mediation models typically assume that all variables are continuous (Baron & Kenny, 1986). Consequently, directly applying conventional models to dichotomous mediating variables violates model assumptions and may produce biased estimates. Prior research has proposed a more general causal mediation analysis framework. This framework is model agnostic, allowing researchers to specify appropriate models for estimating effects and making inferences based on the types of mediating and outcome variables (e.g., logistic regression for dichotomous mediators and linear regression for continuous outcomes). Consequently, it offers methodological advantages when handling mixed-type variables (Imai et al., 2010). In addition, the model employed in the present study incorporates nonparametric bootstrap methods to estimate confidence intervals for indirect effects, without assuming normality in the distribution of these effects. Relevant simulation studies have shown that this approach provides more reliable statistical inference than traditional Sobel tests, particularly in finite samples. (Hayes, 2022; Preacher & Hayes, 2008).

Based on the types of the mediating and outcome variables, the generalized mediation effect model employed in this study comprises the following two link equations:

1.Mediator model (logistic regression)

$$logit(P(M = 1|X)) = a_0 + aX$$

Where X denotes students ‘gender, M represents a dichotomous mediator variable (e.g., preference for language or mathematics), and a is the regression coefficient of X on M.

2.Outcome model (linear regression)

$$Y = \beta_0 + \beta X + \gamma M + \epsilon$$

Where Y denotes the total score for scientific reasoning ability, γ represents the effect of M on Y after controlling for X, and β represents the direct effect of X on Y after controlling for M.

Based on the aforementioned model, the mediating effect is estimated by the product of path coefficients $\alpha \times \gamma$. For binary mediating variables, the estimation method for mediating effects has been theoretically substantiated and empirically validated in methodological research, demonstrating that the product approach provides consistent estimates of indirect effects under this specification (Li et al., 2007). The analysis was conducted using the mediation package in R, employing a nonparametric bootstrap with 5000 resamples to estimate the confidence interval for the average causal mediation effect (ACME) and test its statistical significance. In mediation analysis, if the 95% confidence interval for the ACME does not include zero, the effect is considered statistically significant at $p < 0.05$. This confidence interval-based inference method has been widely adopted and recommended (Hayes, 2022; MacKinnon et al., 2004).

3. Results

3.1. Description statistics

Descriptive statistics were first computed for the key study variables (Table 3). The participating students exhibited an approximately balanced gender distribution (52.2% male, 47.8% female). Regarding self-assessment tendencies, 48.8% of students favored mathematics as their preferred subject, while around 49.8% reported an orientation for logic-oriented thinking. LCTSR total scores ranged from 2 to 24 points, with a mean of 14.08 (SD = 4.65), indicating that students' overall scientific reasoning performance was moderate, with considerable variability observed among individuals.

Table 3. Descriptive Statistics and Correlation Matrix.

Variables	M	SD	Min	Max	1	2	3	4
Gender (1 = male, 0 = female)	0.52	0.50	0	1	1	-	-	-
Mathematics preference (1 = yes, 0 = no)	0.49	0.50	0	1	0.28***	1	-	-
Logic-oriented thinking (1 = yes, 0 = no)	0.50	0.50	0	1	0.34***	0.43***	1	-
LCTSR test total score	14.08	4.65	2	24	0.16***	0.30***	0.25***	1

Note. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

To examine the relationships among the key study variables, Pearson product-moment correlation coefficients were first calculated for each pair of variables, and two-tailed significance tests were conducted at $\alpha = 0.05$. The correlation matrix on the right-hand side of Table 3 illustrates the correlations between the various variables. Results indicate that gender was significantly positively correlated with both mathematics preference ($r = 0.28$, $p < 0.001$) and logic-oriented thinking ($r = 0.34$, $p < 0.001$), indicating that male students tend to report higher preferences for mathematics and logic-oriented thinking than female students. Furthermore, mathematics preference ($r = 0.30$, $p < 0.001$) and logic-oriented thinking ($r = 0.25$, $p < 0.001$) were positively correlated with scientific reasoning ability, suggesting that students who self-report a preference for mathematics or logic-oriented thinking tend to score higher on the LCTSR test.

From another perspective, the analysis revealed that students who self-reported an orientation for memory-oriented thinking tended to prefer Language, whereas those who self-reported an orientation for logic-oriented thinking tended to prefer Mathematics. The distribution of subject preferences and the number of students in each thinking style orientations category are presented in Table 4.

Table 4. Distribution of Subject Preferences and Thinking Style Orientations Categories.

	Memory-oriented thinking	**Logic-oriented thinking**
Language preference	152	61
Mathematics preference	57	146

Note. Values represent the number of students.

3.2. Gender Differences in Sub-dimensions of Scientific Reasoning Ability

A multivariate analysis of variance (MANOVA) was conducted to examine gender differences across six dimensions of scientific reasoning. The multivariate test indicated a significant overall effect of gender, Wilks' $\Lambda = 0.94$, $F(6, 409) = 4.25$, $p < 0.001$. Follow-up univariate analyses with corrected significance levels revealed significant gender differences in three dimensions: conservation reasoning, $F = 7.47$, $p = 0.040$, proportional reasoning, $F = 18.97$, $p < 0.001$, and hypothetical-deductive reasoning, $F = 8.35$, $p = 0.025$. In all three dimensions, male students scored higher on average than female students. No significant gender differences were observed in control of variable reasoning, probability reasoning, or correlational reasoning. Effect size analyses indicated that the magnitude of the observed gender differences was small, with Cohen's d values ranging from 0.27 to 0.43 across the significant dimensions. This suggests that although the gender differences reached statistical significance, their practical significance was limited.

Table 5. Results of Multivariate Analysis of Variance.

Dimensions	Male (M)	Male (SD)	Female (M)	Female (SD)	F	p (corrected)	Cohen's d[1]
Conservation reasoning	3.45	0.94	3.17	1.16	7.47	0.040*	0.27
Proportional reasoning	3.10	1.21	2.51	1.54	18.97	< 0.001***	0.43
Control of variables reasoning	3.35	1.40	3.15	1.48	2.19	0.836	0.14
Probability reasoning	1.58	1.67	1.67	1.61	0.30	1.000	-0.05
Correlational reasoning	0.85	0.95	0.67	0.92	3.83	0.307	0.19
Hypothetical-deductive reasoning	2.48	1.32	2.12	1.23	8.35	0.025*	0.28
Multivariate analysis of variance (population)	Wilks' Λ =0.94, F(6,409) = 4.25, p = 0.0004						

[1] According to Cohen's guidelines, effect sizes of |d| ≥ 0.2, 0.5, and 0.8 are generally regarded as small, medium, and large, respectively. In this study, the d values for sub-dimensions showing significant gender differences ranged from 0.27 to 0.43, which fall within the small effect size range. p values were adjusted using the Bonferroni correction. Significance levels are indicated as: *p < 0.05, **p < .001, ***p < 0.001. (Cohen, 1988)

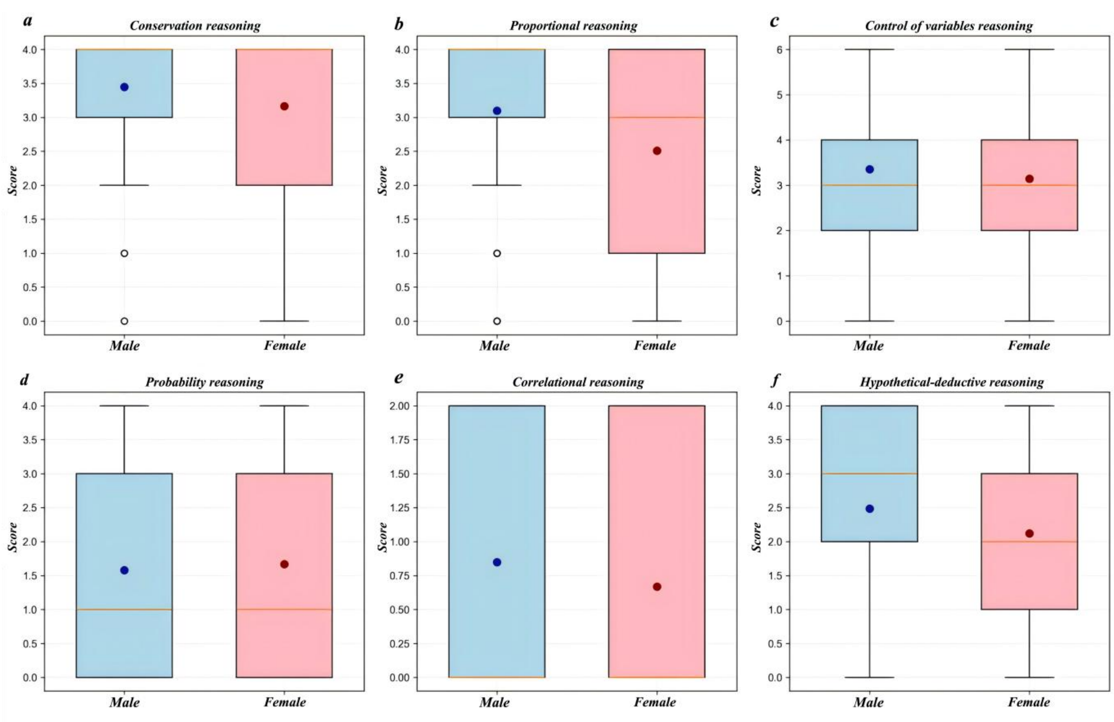


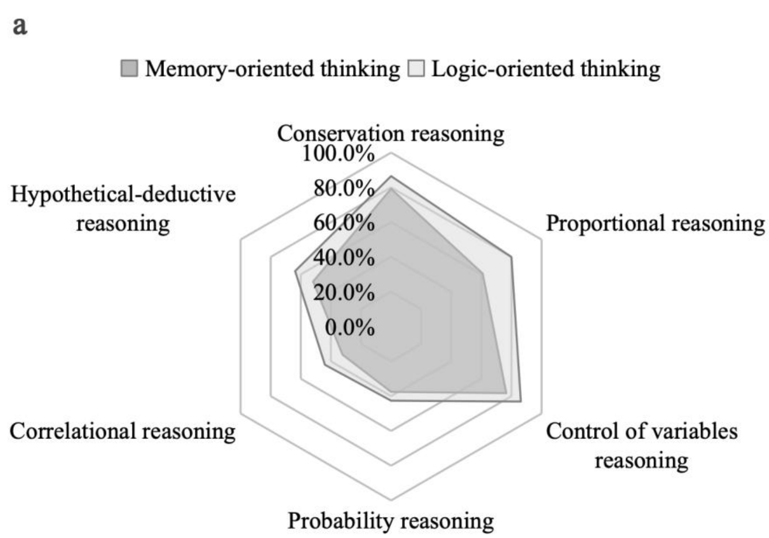
Figure 1. Gender Differences in LCTSR Scores Across Six Dimensions: A Boxplot.

The remaining three LCTSR dimensions, which showed no significant gender differences, can be interpreted from a cognitive development perspective. Control of variables reasoning and probability reasoning are generally considered early or mid-stage components of the formal operational stage, emerging relatively early, with which observable in most secondary school students. Research indicates that control of variable reasoning primarily involves the systematic manipulation of single variables. For example, Chen and Klahr demonstrated that children acquire the ability to design single-factor experiments through training, laying the foundation for mastering control of variables strategies (Chen & Klahr, 1999). Similarly, in probability reasoning, individuals often rely on simple frequency judgments rather than complex probability calculations (Gigerenzer & Hoffrage, 1995). As a result, at the secondary school level, no significant gender differences are typically observed in control of variable reasoning or probabilistic reasoning (Sarıbaş & Ay, 2023; Van Vo et al., 2023). In contrast, correlational reasoning, which involves understanding relationships between variables and statistical dependencies, represents a later stage of formal operational thinking (Piaget & Inhelder, 1999). Most secondary school students have yet to develop systematic analytical skills for correlational relationships, exhibiting low overall proficiency and little individual variation (Lawson et al., 1979). Consequently, gender differences in correlational reasoning also remain non-significant at this developmental stage. Overall, the observed patterns of gender differences across these three reasoning abilities align with their cognitive developmental trajectories: early-stage components are generally well-developed, leaving limited scope for differentiation, whereas later-stage components remain underdeveloped, preventing the emergence of significant gender gaps.

Overall, genuine gender differences were observed in three dimensions of scientific reasoning: conservation reasoning, proportional reasoning, and hypothetical-deductive reasoning, with male students scoring higher than female students.

3.3. Gender Differences within Student Groups with the Same Subject Preferences and Thinking Style Orientations

As noted in the preceding descriptive statistical analysis, both subject preferences and thinking style orientations are positively associated with scientific reasoning ability. The figure below further illustrates that students with different subject preferences and thinking style orientations exhibit clear differences in their scientific reasoning performance.



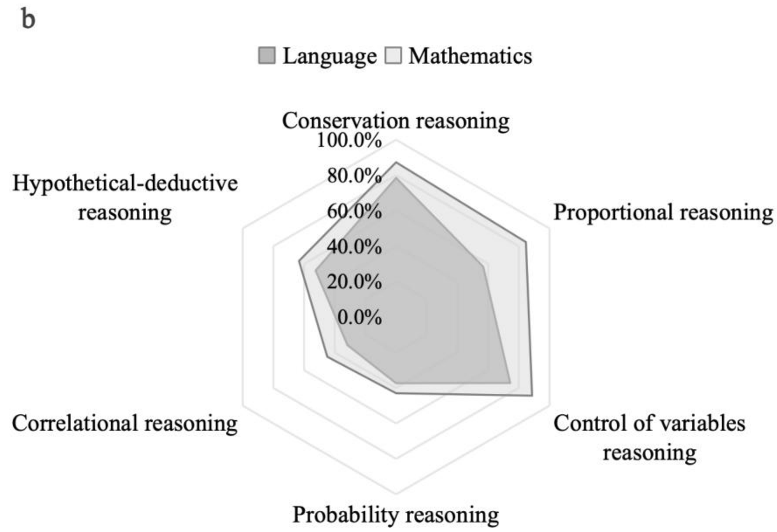


Figure 2. Module Score Comparison of Scientific Reasoning Ability by Subject Preference and Thinking Style Orientations. (a) Subject Preference; (b) Thinking Style Orientations.

In terms of subject preference, students who preferred mathematics scored significantly higher on average than those who preferred language. With respect to thinking style orientations, students who self-identified as logical thinkers obtained significantly higher mean scores than those who identified as memory-oriented thinkers. Across the six dimensions of scientific reasoning ability, mathematics-preferring students outperformed language-preferring students in conservation reasoning, proportional reasoning, control of variables reasoning, and correlational reasoning, whereas no significant difference was observed in probability reasoning. Similarly, students with a logic-oriented thinking demonstrated significantly higher performance than memory-oriented students in conservation reasoning, proportional reasoning, control of variables reasoning, correlational reasoning, and hypothetical-deductive reasoning, with no significant difference found in probability reasoning.

Further analyses indicated that no significant gender differences in scientific reasoning ability were found among students with the same subject preferences or thinking style orientations, as indicated by independent-samples t tests (see Table 6).

Table 6. Results of Multivariate Analysis of Variance.

		Female			**Male**			**significance**
		N	**M**	**SD**	**N**	**M**	**SD**	
Subject preferences	Language	131	12.38	4.8	82	13.21	5.44	t = 1.13 p = 0.262
	Mathematics	68	15.0	4.07	135	15.79	3.93	t = 1.33 p = 0.186
Thinking style orientations	Memory-oriented thinking	135	12.85	4.57	74	12.99	5.12	t = 0.20 p = 0.846
	Logic-oriented thinking	64	14.17	4.93	143	15.76	4.21	t = 2.37 p = 0.019

As shown in Table 6, subject preference was significantly associated with students' scientific reasoning ability. Within the group of students with a preference for language, no statistically significant gender differences in overall scientific reasoning ability were observed. Similarly, no significant gender differences were identified among students with a preference for mathematics. With respect to thinking style orientations, no significant gender differences were found in the self-reported memory-oriented thinking group. In contrast, within the self-reported logic-oriented thinking group, differences in mean scores between male and female students were observed. Further independent-samples t-tests conducted on the sub-dimensions of scientific reasoning ability indicated a statistically significant gender difference in conservation reasoning ($p = 0.0351$). Proportional reasoning exhibited a marginal trend toward statistical significance ($p = 0.0517$), whereas no significant gender differences were detected in the remaining sub-dimensions.

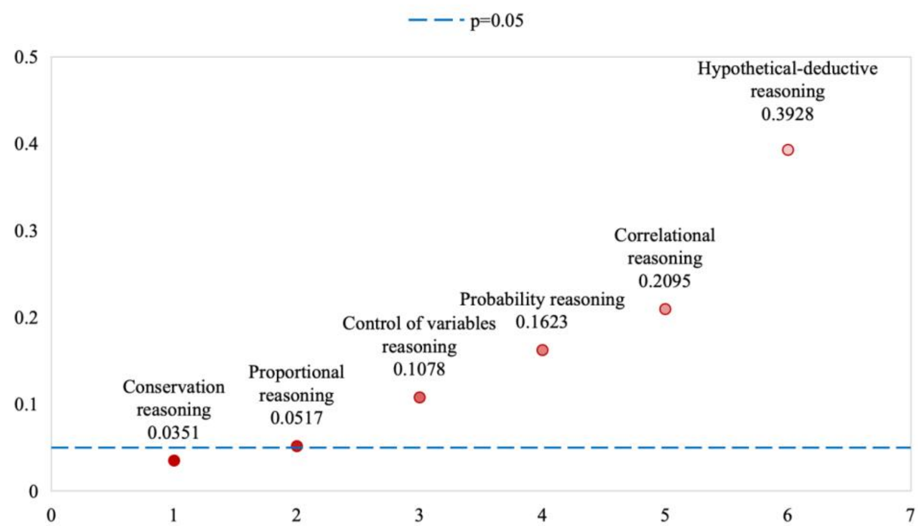


Figure 3. Distribution of p-values Across Sub-dimensions.

3.4. The Mediating Role of Subject Preference and Thinking Style Tendencies Using Scientific Reasoning Ability as the Outcome Variable

Given that no significant gender differences were observed among student cohorts with similar subject preferences and thinking style orientations, the question arises: through which intrinsic mechanisms do gender influence scientific reasoning ability, thereby leading to the observed gender differences in this capacity? To address this question, the present study proposes a dual mediation model (Figure 4). This model posits that gender not only directly affects scientific reasoning ability but also indirectly influences it through two mediating pathways: mathematics preference and logic-oriented thinking.

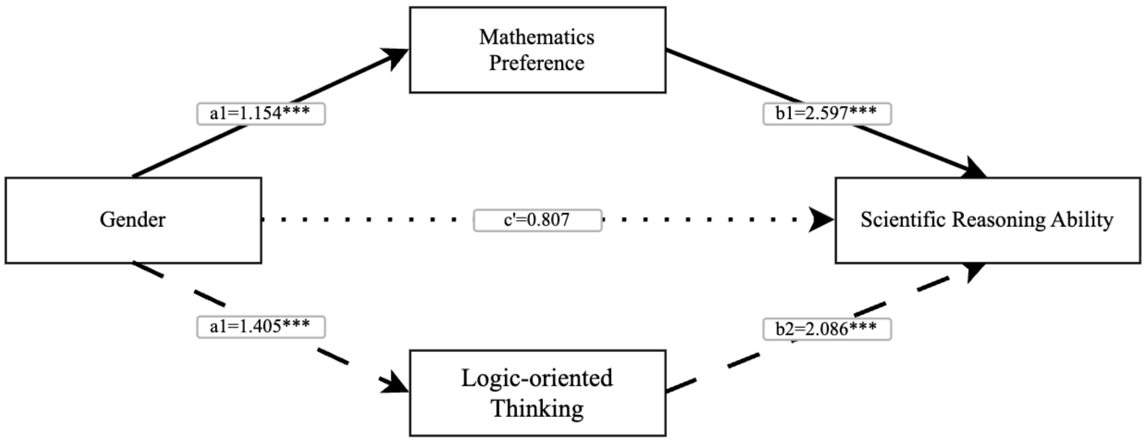


Figure 4. Mediation Model of Gender Effects on Scientific Reasoning: Exploring the Roles of Mathematics Preference and Logic-oriented Thinking.

Note. The figure depicts standardized path coefficients. a_1 and a_2 represent the effects of gender on the mediators (mathematics preference and logic-oriented thinking, respectively); b_1 and b_2 represent the effects of the mediators on scientific reasoning ability; c′ denotes the direct effect of gender on scientific reasoning ability. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

This study hypothesizes that gender influences scientific reasoning ability through its effects on mathematics preference and logic-oriented thinking. Figure 4 illustrates the theoretical model and corresponding path coefficients. As shown, gender exerts a significant positive effect on both mediators ($a_1 = 1.154$, $p < 0.001$; $a_2 = 1.405$, $p < 0.001$), indicating that males are more likely to demonstrate a preference for mathematics and logic-oriented thinking. Concurrently, both mathematics preference ($b_1 = 2.597$, $p < 0.001$) and logic-oriented thinking ($b_2 = 2.086$, $p < 0.001$) positively predict scientific reasoning ability. In contrast, the direct effect of gender on scientific reasoning ability ($c' = 0.807$, $p = 0.084$) was not statistically significant, consistent with full mediation. That is, gender appears to influence scientific reasoning primarily indirectly via these two mediators. To examine the mediation effects more rigorously, a bootstrap analysis with 5,000 resamples was conducted, The mediating effect values and their confidence intervals were obtained through direct Bootstrap estimation, with the results presented in Table 7.

Table 7. Results of Bootstrap Mediation Analysis for the Effects of Gender on Scientific Reasoning Ability.

Path	Coefficient (β)	SE	p	95%CI
Indirect effect via Mathematics Preference				
Gender → Mathematics Preference (a_1)	1.154	0.205	<0.001***	[0.754, 1.555]
Mathematics Preference → Scientific Reasoning Ability (b_1)	2.597	0.465	<0.001***	[1.683, 3.511]
Indirect effect ($a_1 \times b_1$)	0.693	-	<0.001***	[0.360, 1.161]
Direct effect (c′)	0.807	0.466	0.084	[-0.111, 1.725]
Indirect effect via Logic-oriented Thinking				

Gender → Logic-oriented Thinking (a_2)	1.405	0.217	<0.001***	[0.980, 1.830]
Logic-oriented Thinking → Scientific Reasoning Ability (b_2)	2.086	0.456	<0.001***	[1.192, 2.981]
Indirect effect ($a_2 \times b_2$)	0.727	-	<0.001***	[0.341, 1.160]
Direct effect (c′)	0.831	0.505	0.101	[-0.165, 1.827]

Note. N = 416. Unstandardized coefficients are reported. CI = confidence interval. Confidence intervals for indirect effects are derived from bias-corrected bootstrap tests with 5,000 resamples.

As shown in Table 7, the indirect effect of gender on scientific reasoning ability through mathematics preference was significant, ACME = 0.693, 95% CI [0.360, 1.161], $p < 0.001$. The direct effect was not significant, ADE = 0.807, 95% CI [-0.111, 1.725], $p = 0.084$, supporting a model of full mediation. The proportion mediated was 0.46, 95% CI [0.22, 1.15]. Similarly, the indirect effect through logic-oriented thinking was also significant, ACME = 0.727, 95% CI [0.341, 1.160], $p < .001$, with a non-significant direct effect, ADE = 0.831, 95% CI [-0.165, 1.827], $p = 0.101$, which also supports full mediation. The proportion mediated for this path was 0.47, 95% CI [0.19, 1.23].

A comparison of the two mediation models indicates that the mediating effects of mathematics preference and logic-oriented thinking are comparable in magnitude (0.693 vs. 0.727), with each accounting for approximately 46% of the total effect. This result suggests that gender differences in scientific reasoning ability are largely transmitted through students' subject preferences and thinking style orientations, rather than arising from a direct effect of gender itself.

3.5. Using Sub-dimensions of Scientific Reasoning Ability as Outcome Variables

We further specified the sub-dimensions of scientific reasoning ability as outcome variables and examined the mediating effects of mathematics preference and logic-oriented thinking on the association between gender and each sub-dimension. The results are presented in Table 8.

Table 8. Results of Bootstrap Mediation Analysis for Gender Effects on Scientific Reasoning Sub-dimensions.

Outcome Variable	Mediating variable	Indirect Effect (ACME)	Direct Effect (ADE)	Total Effect
Conservation reasoning	Mathematics preference	0.079* [0.020, 0.160]	0.198 [-0.015, 0.418]	0.277** [0.078, 0.490]
	Logic-oriented thinking	0.077* [0.002, 0.162]	0.206 [-0.019, 0.430]	0.284** [0.079, 0.484]
Proportional reasoning	Mathematics preference	0.291** [0.163, 0.440]	0.301* [0.047, 0.561]	0.593** [0.324, 0.860]
	Logic-oriented thinking	0.202** [0.103, 0.347]	0.374** [0.087, 0.650]	0.576** [0.317, 0.854]

Control of variables reasoning	Mathematics preference	0.164** [0.063, 0.268]	0.053 [-0.245, 0.351]	0.218 [-0.075, 0.491]
	Logic-oriented thinking	0.126* [0.020, 0.241]	0.089 [-0.209, 0.385]	0.215 [-0.067, 0.485]
Probability reasoning	Mathematics preference	0.039 [-0.057, 0.138]	-0.123 [-0.462, 0.207]	-0.084 [-0.406, 0.219]
	Logic-oriented thinking	0.089 [-0.018, 0.220]	-0.180 [-0.517, 0.144]	-0.091 [-0.407, 0.218]
Correlational reasoning	Mathematics preference	0.073* [0.010, 0.132]	0.115 [-0.076, 0.299]	0.188 [-0.007, 0.357]
	Logic-oriented thinking	0.079* [0.002, 0.138]	0.114 [-0.081, 0.295]	0.193 [-0.007, 0.357]
Hypothetical-deductive reasoning	Mathematics preference	0.110** [0.025, 0.189]	0.262* [0.015, 0.519]	0.372** [0.120, 0.611]
	Logic-oriented thinking	0.145** [0.046, 0.243]	0.228 [-0.033, 0.481]	0.373** [0.115, 0.609]

Note. ACME = Average Causal Mediation Effect; ADE = Average Direct Effect. Unstandardized coefficients are reported. Values in brackets are 95% bias-corrected bootstrap confidence intervals based on 5,000 resamples. *p < 0.05, **p < .001, ***p < 0.001.

The table above shows distinct mediating mechanisms through which gender is associated with different sub-dimensions of scientific reasoning ability. Bootstrap analyses indicate that gender differences primarily operate through two mediating pathways—mathematics preference and logic-oriented thinking—although substantial variation was observed across specific dimensions.

The mediating effect of mathematics preference was consistent and statistically significant. With the exception of probability reasoning, significant indirect effects ($p < 0.05$) were observed across the remaining five dimensions: conservation reasoning, proportional reasoning, control of variables reasoning, correlational reasoning, hypothetical-deductive reasoning, and overall score. Notably, for control of variable reasoning, the mediation proportion reached 75.5%, indicating that most of the observed gender differences were transmitted through mathematics preference. In proportional reasoning, the mediation proportion reached 49.2%, underscoring the importance of mathematics-related preference for performance in this dimension.

In parallel, logic-oriented thinking constituted another effective mediating pathway. Its operational pattern closely resembled that of mathematics preference, with significant indirect effects observed across the same six dimensions and with comparable effect sizes. This suggests that subject preferences (mathematics preference) and thinking style orientations (logic-oriented thinking) tend to function synergistically, rather than independently, in accounting for gender differences in scientific reasoning ability.

For probability reasoning, neither mediation pathway reached statistical significance (all $p > 0.05$), and the total effect of gender was also non-significant. This indicates that no meaningful gender differences were observed for this dimension, with individual variation likely attributable to other factors, such as prior knowledge or contextual experience.

The analysis further differentiated between full and partial mediation patterns. For conservation reasoning, control of variables reasoning, and Correlational reasoning, significant indirect effects coupled with non-significant direct effects supported a full mediation model, in which gender influences were entirely transmitted through subject preferences and thinking style orientations. In contrast, for proportional reasoning and hypothetical-deductive reasoning, both direct and indirect effects were significant, indicating partial mediation and suggesting that gender may exert direct influences beyond the mediated pathways.

Overall, these findings indicate that gender differences in scientific reasoning ability do not arise from a single uniform mechanism but vary according to the cognitive demands of specific reasoning tasks. When tasks are closely aligned with mathematics preference or logic-oriented thinking, preferences and modes of thinking—likely shaped by broader sociocultural influences—appear to play a central mediating role.

To further examine the mediating relationships identified above, this study estimated a path model depicting the associations between gender and each sub-dimension of scientific reasoning ability, as shown in Figure 5. The model illustrates the respective mediating roles of mathematics preference and logic-oriented thinking in linking gender to different reasoning dimensions.

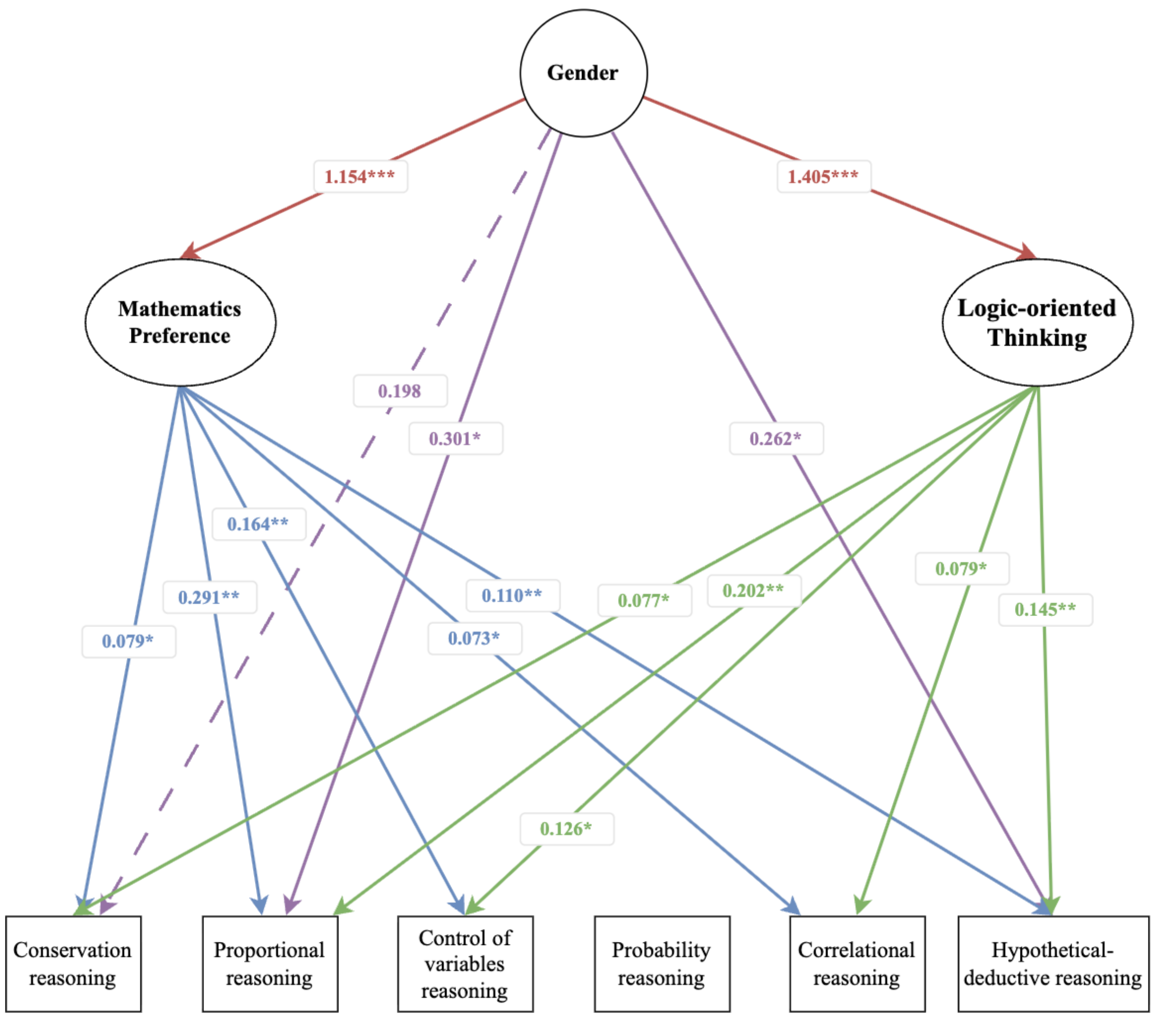


Figure 5. Path Diagram of Mediation Effects for Sub-dimensions.

4. Discussion and Synthesis

4.1. Discussion

This study demonstrates that gender differences exist in the scientific reasoning abilities of secondary school pupils, consistent with prior research on Finnish upper secondary students, where disparities were observed prior to any instructional intervention (Nieminen et al., 2013). However, some studies report no significant gender differences in scientific reasoning (Bezci & Sungur, 2021; Piraksa et al., 2014), likely reflecting sample characteristics such as national education systems or gender ratios that influence statistical detection. Overall, the present findings indicate that gender differences are relatively small and exhibit sub-dimensional specificity: conservation reasoning, proportional reasoning, and hypothetical-deductive reasoning show more pronounced differentiation, whereas control variable reasoning, probabilistic reasoning, and correlational reasoning display minimal or non-significant differences. This suggests that gender differences in scientific reasoning do not manifest as a unified construct but rather emerge across distinct cognitive components.

When controlling for aligned subject preferences and thinking style orientations, gender differences become negligible among students sharing identical preferences. This indicates that observed disparities are primarily mediated by differences in interests, subject preferences, and cognitive styles rather than inherent ability. Motivation, interest, and cognitive styles are known to influence students' engagement, strategy selection, and problem representation, thereby affecting performance in higher-order cognitive tasks (Eccles & Wigfield, 2002; Krapp, 2005; Pintrich, 2003). In STEM learning, interest structures and cognitive preferences often exhibit pronounced gendered patterns (Su & Rounds, 2015; Wang & Degol, 2017), which translate into performance differences through their impact on learning commitment, persistence, and strategy use (Eccles, 2009; OECD, 2019). Therefore, after controlling for these variables, gender differences are substantially attenuated or even eliminated, suggesting the presence of highly plastic mediating mechanisms.

Further analysis indicates that the influence of gender on scientific reasoning is fully mediated by subject preferences and thinking style orientations. A critical question arises concerning the origins of gender differences in these preferences. Stereotypes provide a plausible explanation: evaluative experiences in academic settings can heighten self-awareness regarding group-based expectations, and negative stereotype effects may impair performance on cognitive tasks (Steele, 1997; Steele et al., 2002; Walton. & Spencer., 2009). In the present context, the widely held belief in China that "male excel in science and female excel in language" reinforces gendered patterns in subject preferences and thinking styles. Expectancy-value theory posits that perceptions of subject value and ability beliefs, shaped by socialization, foster gendered motivations and choices (Eccles & Wigfield, 2002). Importantly, these differences are not attributable to innate abilities; meta-analyses confirm high similarity between genders in core cognitive competencies for science and mathematics (Hyde, 2014). Research consistently shows that sociocultural factors, including stereotypes and environmental contexts, are pivotal in shaping gender gaps in STEM interest and participation (Wang & Degol, 2017). These patterns reflect the internalization of social stereotypes, which influence students' self-perceptions, interests, motivations, and strategic engagement, ultimately affecting scientific reasoning performance (Eccles, 2009; Su & Rounds, 2015).

Large-scale longitudinal evidence supports this mechanism. Tracking over 2.65 million French children, studies reveal that gender gaps in mathematical ability emerge and widen early in schooling, primarily reflecting internalization of sociocultural stereotypes regarding females' math performance (Martinot et al., 2025). Consequently, the gender differences observed in the present study likely stem from social and motivational factors rather than innate ability. Interventions, such as value-affirming activities and supportive educational environments, may mitigate these influences and promote equitable outcomes (Cohen et al., 2006; Martens et al., 2006).

Encouragingly, despite statistically significant differences in certain sub-dimensions, all effect sizes fall within the "small" range, indicating that male and female pupils demonstrate comparable core scientific reasoning abilities. This suggests that targeted educational interventions and environment shaping can effectively reduce minor disparities. Educational focus should therefore prioritize fostering equal opportunities that fully develop all students' reasoning potential, rather than reinforcing gender-based ability assumptions.

4.2. Limitations

As an empirical study, the present research has several limitations. First, data were collected from students of different year groups at approximately the same time points, without conducting longitudinal repeated measurements on the same cohort. This cross-sectional design imposes inherent constraints on causal inference. Although the analyses support the pathway in which gender influences scientific reasoning ability through subject preferences and thinking style orientations, the possibility of reverse causation cannot be ruled out (e.g., early success in scientific reasoning may reinforce preferences for mathematics and logic-oriented thinking).Second, the present study examined only two mediating variables: subject preferences and thinking style orientations. Future research should consider incorporating additional factors, such as parental educational attainment (Osterhaus & Koerber, 2024), scientific motivation (Van Vo & Csapó, 2023), and spatial reasoning ability (Hodgkiss et al., 2018), to construct a more comprehensive explanatory model. Finally, subject preference and thinking style orientations were assessed using self-report measures, which rely on students' self-assessment and are therefore subject to potential biases. Nevertheless, preference-based variables are inherently shaped by implicit and contextual factors, making fully objective measurement challenging. Consequently, self-report instruments retain reasonable validity and utility. Future research may employ complementary multi-method approaches to further enhance measurement accuracy and robustness.

5. Conclusions

This study shows that gender differences in junior secondary students' scientific reasoning exist but are relatively small and selective, primarily appearing in conservation reasoning, proportional reasoning, and hypothetical-deductive reasoning. Importantly, these differences are largely mediated by subject preferences and thinking style orientations. That is, gender influences students' subject and thinking style preferences, which in turn affect their performance on scientific reasoning tasks. Once these mediating factors are accounted for, direct gender differences are greatly reduced or disappear, indicating that sociocultural and motivational factors, rather than inherent ability, drive the observed gaps. These findings highlight the plasticity of gender differences and suggest that targeted interventions in interest development, thinking style cultivation, and stereotype mitigation can effectively promote equitable reasoning skills in STEM education.

References:

Balducci, M., Larose, M. P., Stoet, G., & Geary, D. C. (2024). The Gender-Equality Paradox in Intraindividual Academic Strengths: A Cross-Temporal Analysis. Psychol Sci, 35(11), 1246–1259. https://doi.org/10.1177/09567976241271330

Bao, L., Cai, T., Koenig, K., Fang, K., Han, J., Wang, J., Liu, Q., Ding, L., Cui, L., Luo, Y., Wang, Y., Li, L., & Wu, N. (2009). Physics. Learning and scientific reasoning. Science, 323(5914), 586–587. https://doi.org/10.1126/science.1167740

Bao, L., Xiao, Y., Koenig, K., & Han, J. (2018). Validity evaluation of the Lawson classroom test of scientific reasoning. Physical Review Physics Education Research, 14(2), Article 020106. https://doi.org/10.1103/PhysRevPhysEducRes.14.020106

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol, 51(6), 1173–1182. https://doi.org/10.1037//0022-3514.51.6.1173

Bezci, F., & Sungur, S. (2021). How is Middle School Students' Scientific Reasoning Ability Associated with Gender and Learning Environment? [Article]. Science Education International, 32(2), 96–106. https://doi.org/10.33828/sei.v32.i2.2

Chen, Z., & Klahr, D. (1999). All other things being equal: acquisition and transfer of the control of variables strategy. Child Dev, 70(5), 1098–1120. https://doi.org/10.1111/1467-8624.00081

Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: a social-psychological intervention. Science, 313(5791), 1307–1310. https://doi.org/10.1126/science.1128317

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences. Taylor & Francis Group.

Coletta, V. P., & Phillips, J. A. (2005). Interpreting FCI scores: normalized gain, preinstruction scores, and scientific reasoning ability. American Journal of Physics, 73(12), 1172–1182. https://doi.org/10.1119/1.2117109

Eccles, J. (2009). Who Am I and What Am I Going to Do With My Life? Personal and Collective Identities as Motivators of Action. Educational Psychologist, 44(2), 78–89, Article Pii 910596492. https://doi.org/10.1080/00461520902832368

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals [Article]. Annual Review of Psychology, 53, 109–132. https://doi.org/10.1146/annurev.psych.53.100901.135153

Feinstein, N. (2011). Salvaging science literacy [Article]. Science Education, 95(1), 168–185. https://doi.org/10.1002/sce.20414

Fine, C. (2010). Delusions of gender : how our minds, society, and neurosexism create difference.

Forawi, S. A. (2016). Standard-based science education and critical thinking. Thinking Skills & Creativity, 20, 52–62. https://doi.org/10.1016/j.tsc.2016.02.005

García-Carmona, A. (2025). Scientific Thinking and Critical Thinking in Science Education : Two Distinct but Symbiotically Related Intellectual Processes [Article]. Science and Education, 34(1), 227–245. https://doi.org/10.1007/s11191-023-00460-5

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: frequency formats. Psychological Review, 102, 684–704. https://doi.org/10.1037/0033-295X.102.4.684

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

Hayes, A. F. (2022). Introduction to Mediation, Moderation, and Conditional Process Analysis : A Regression-Based Approach. Guilford Publications.

Hodgkiss, A., Gilligan, K. A., Tolmie, A. K., Thomas, M. S. C., & Farran, E. K. (2018). Spatial cognition and science achievement: The contribution of intrinsic and extrinsic spatial skills from 7 to 11 years. Br J Educ Psychol, 88(4), 675–697. https://doi.org/10.1111/bjep.12211

Hyde, J. S. (2014). Gender Similarities and Differences. Annual Review of Psychology, 65(Volume 65, 2014), 373–398. https://doi.org/https://doi.org/10.1146/annurev-psych-010213-115057

Imai, K., Keele, L., & Tingley, D. (2010). A General Approach to Causal Mediation Analysis [Article]. Psychological Methods, 15(4), 309–334. https://doi.org/10.1037/a0020761

Kind, P. E. R., & Osborne, J. (2017). Styles of Scientific Reasoning: A Cultural Rationale for Science Education? Science Education, 101(1), 8–31. https://doi.org/10.1002/sce.21251

Krapp, A. (2005). Basic needs and the development of interest and intrinsic motivational orientations. Learning & Instruction, 15(5), 381–395. https://doi.org/10.1016/j.learninstruc.2005.07.007

Lawson, A. E., Adi, H., & Karplus, R. (1979). Development of Correlational Reasoning in Secondary Schools: Do Biology Courses Make a Difference? The American Biology Teacher, 41(7), 420–430. https://doi.org/10.2307/4446678

Lawson, A. E., Clark, B., Cramer-Meldrum, E., Falconer, K. A., Sequist, J. M., & Kwon, Y.-J. (2000). Development of Scientific Reasoning in College Biology: Do Two Levels of General Hypothesis-Testing Skills Exist? Journal of Research in Science Teaching, 37(1), 81–101. https://doi.org/10.1002/(sici)1098-2736(200001)37:1<81::Aid-tea6>3.0.Co;2-i

Li, Y., Schneider, J. A., & Bennett, D. A. (2007). Estimation of the mediation effect with a binary mediator. Stat Med, 26(18), 3398–3414. https://doi.org/10.1002/sim.2730

Luo, M., Sun, D., Zhu, L., & Yang, Y. (2021). Evaluating scientific reasoning ability: Student performance and the interaction effects between grade level, gender, and academic achievement level [Article]. Thinking Skills and Creativity, 41, Article 100899. https://doi.org/10.1016/j.tsc.2021.100899

MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods [Article]. Multivariate Behavioral Research, 39(1), 99–128. https://doi.org/10.1207/s15327906mbr3901_4

Martens, A., Johns, M., Greenberg, J., & Schimel, J. (2006). Combating stereotype threat: The effect of self-affirmation on women's intellectual performance [Article]. Journal of Experimental Social Psychology, 42(2), 236–243. https://doi.org/10.1016/j.jesp.2005.04.010

Martin, P. P., & Graulich, N. (2023). When a machine detects student reasoning: a review of machine learning-based formative assessment of mechanistic reasoning. Chemistry Education Research and Practice, 24(2), 407–427. https://doi.org/10.1039/d2rp00287f

Martinot, P., Colnet, B., Breda, T., Sultan, J., Touitou, L., Huguet, P., Spelke, E., Dehaene-Lambertz, G., Bressoux, P., & Dehaene, S. (2025). Rapid emergence of a maths gender gap in first grade. Nature, 643(8073), 1020–1029. https://doi.org/10.1038/s41586-025-09126-4

Nieminen, P., Savinainen, A., & Viiri, J. (2013). GENDER DIFFERENCES IN LEARNING OF THE CONCEPT OF FORCE, REPRESENTATIONAL CONSISTENCY, AND SCIENTIFIC REASONING. International Journal of Science and Mathematics Education, 11(5), 1137–1156. https://doi.org/10.1007/s10763-012-9363-y

OECD. (2017). PISA 2015 Assessment and Analytical Framework. https://doi.org/10.1787/9789264281820-en

OECD. (2019). PISA 2018 Results (Volume I). https://doi.org/10.1787/5f07c754-en

Olive, K., Tang, X., Loukomies, A., Juuti, K., & Salmela-Aro, K. (2022). Gendered difference in motivational profiles, achievement, and STEM aspiration of elementary school students [Original Research]. Frontiers in Psychology, Volume 13 - 2022. https://doi.org/10.3389/fpsyg.2022.954325

Osterhaus, C., & Koerber, S. (2024). The personal epistemology of parents predicts the development of scientific reasoning in children aged 6-10 years. Dev Sci, 27(3), e13474. https://doi.org/10.1111/desc.13474

Piaget, J., & Inhelder, B. (1999). The growth of logical thinking from childhood to adolescence : an essay on the construction of formal operational structures. Routledge.

Pintrich, P. R. (2003). A Motivational Science Perspective on the Role of Student Motivation in Learning and Teaching Contexts. Journal of Educational Psychology, 95(4), 667–686. https://doi.org/10.1037/0022-0663.95.4.667

Piraksa, C., Srisawasdi, N., & Koul, R. (2014). Effect of Gender on Student's Scientific Reasoning Ability: A Case Study in Thailand. Procedia - Social and Behavioral Sciences, 116, 486–491. https://doi.org/https://doi.org/10.1016/j.sbspro.2014.01.245

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behavior Research Methods, 40(3), 879–891. https://doi.org/10.3758/BRM.40.3.879

Reilly, D., Neumann, D. L., & Andrews, G. (2022). Gender Differences in Self-Estimated Intelligence: Exploring the Male Hubris, Female Humility Problem [Original Research]. Frontiers in Psychology, Volume 13 - 2022. https://doi.org/10.3389/fpsyg.2022.812483

Roadrangka, V., Yeany, R. H., & Padilla, M. J. (1983). The Construction and Validation of Group Assessment of Logical Thinking (GALT) Annual Meeting of the National Association for Research in Science Teaching (NARST), Dallas, TX.

Roberts, D. A. (2013). Scientific literacy/science literacy. In Handbook of research on science education (pp. 729–780). Routledge.

Sapia, P., Napoli, F., & Bozzo, G. (2022). The Lawson's Test for Scientific Reasoning as a Predictor for University Formative Success: A Prospective Study. Education Sciences, 12(11). https://doi.org/10.3390/educsci12110814

Sarıbaş, H. M., & Ay, Z. S. (2023). Investigation of Middle School Students' Probabilistic Reasoning Levels in Terms of Some Variables. Cumhuriyet International Journal of Education, 12(2), 292–304. https://doi.org/10.30703/cije.1152810

Sicard, A., Martinot, D., Toczek, M.-C., Pironom, J., & Darnon, C. (2021). Turning the gender tables: evidence of students' awareness of a reversal in gender status between academic and occupational contexts. Social Psychology of Education, 24(1), 247–272. https://doi.org/10.1007/s11218-021-09607-0

Steele, C. M. (1997). A threat in the air: how stereotypes shape intellectual identity and performance. American Psychologist, 52, 613–629. https://doi.org/10.1037/0003-066X.52.6.613

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In Advances in Experimental Social Psychology (Vol. 34, pp. 379–440). Academic Press. https://doi.org/https://doi.org/10.1016/S0065-2601(02)80009-0

Strand, S., Deary, I. J., & Smith, P. (2006). Sex differences in cognitive abilities test scores: a UK national picture. Br J Educ Psychol, 76(Pt 3), 463–480. https://doi.org/10.1348/000709905X50906

Su, R., & Rounds, J. (2015). All STEM fields are not created equal: People and things interests explain gender disparities across STEM fields [Original Research]. Frontiers in Psychology, Volume 6 - 2015. https://doi.org/10.3389/fpsyg.2015.00189

Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. Research in Science Education, 48(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Tobin., -. K. G., & Capie., -. W. (1981). The Development and Validation of a Group Test of Logical Thinking. Educational and Psychological Measurement, 41(1981)(2), 413–423. https://doi.org/doi: 10.1177/001316448104100220

Van Vo, D., & Csapó, B. (2023). Exploring Inductive Reasoning, Scientific Reasoning and Science Motivation, and Their Role in Predicting STEM Achievement Across Grade Levels. International Journal of Science and Mathematics Education, 21(8), 2375–2398. https://doi.org/10.1007/s10763-022-10349-4

Van Vo, D., Csapó, B., & Greiff, S. (2023). Development of the control of variables strategy in physics among secondary school students [Article]. Thinking Skills and Creativity, 49, Article 101371. https://doi.org/10.1016/j.tsc.2023.101371

Wai, J., Cacchio, M., Putallaz, M., & Makel, M. C. (2010). Sex differences in the right tail of cognitive abilities: A 30year examination. Intelligence, 38(4), 412–423. https://doi.org/https://doi.org/10.1016/j.intell.2010.04.006

Walton., -. G. M., & Spencer., -. S. J. (2009). Latent Ability. Psychological Science, 20(2009)(9), 1132–1139. https://doi.org/doi: 10.1111/j.1467-9280.2009.02417.x

Wang, M.-T., & Degol, J. L. (2017). Gender Gap in Science, Technology, Engineering, and Mathematics (STEM): Current Knowledge, Implications for Practice, Policy, and Future Directions. EDUCATIONAL PSYCHOLOGY REVIEW, 29(1), 119–140. https://doi.org/10.1007/s10648-015-9355-x