Article

"It's Different for Girls!" The Role of Anxiety, Physiological Arousal, and Subject Preferences in Primary School Children's Math and Mental Rotation Performance

Michelle Lennon-Maslin * and Claudia Michaela Quaiser-Pohl

Department of Developmental Psychology and Psychological Assessment, Faculty of Educational Sciences, Institute of Psychology, University of Koblenz, 56070 Koblenz, Germany; quaiser@uni-koblenz.de
* Correspondence: mlennonm@uni-koblenz.de

Abstract: (1) Background: This study examines the role of subjective anxiety (mathematics and spatial anxiety), along with physiological responses, in mathematics or math and mental rotation performance in 131 German primary school students (65 girls, 66 boys; *Mean age* = 8.73 years). (2) Method: Students' preference for math vs. German and their subjective anxiety were assessed using standardized questionnaires. Emotional reactivity was measured using the Galvanic Skin Response (GSR). Math performance was evaluated via percentage scored and completion times on number line estimation, word problems, and missing terms tasks. Spatial skills were assessed using a novel mental rotation task (nMRT) incorporating gender-congruent and -neutral stimuli. (3) Results: Girls outperformed boys on percentage scored on the math task but took longer to complete this. No gender differences were found in performance on the nMRT. Girls demonstrated higher math anxiety and were less likely to prefer math over German. Math anxiety predicted math scores and accuracy on the nMRT while gender predicted math performance and mental rotation response time. Subject preference was associated with longer completion times and emotional reactivity with longer response times. Girls' preference for math and lower emotional reactivity was linked to shorter completion times, while lower math anxiety predicted higher scores. In contrast, these factors did not affect boys' math performance. Additionally, subjective anxiety, emotional reactivity, or subject preference did not impact spatial performance for either gender. (4) Conclusions: Supporting mathematical self-efficacy and emotional regulation, especially in girls, is crucial for enhancing STEM outcomes in primary education. Gender-fair assessment in mental rotation reveals equitable spatial performance and reduces the impact of anxiety.

Keywords: math anxiety; physiological arousal; math performance; mental rotation skills; gender differences; primary education

1. Introduction

The recent social media phenomenon, "Girl Math", which began as a light-hearted viral joke about women's alleged shopping and bargaining habits, has also reignited a persistent stereotype that women have a negative relationship with mathematics [1]. In a recent interview, British mathematician Hannah Fry, who describes herself as a "numberphile", discussed the low number of women pursuing careers in math-related fields. When asked about the reasons behind this gender disparity, Fry stated that it cannot be attributed to a single cause. Instead, she described it as "death by a thousand cuts" [2]. Fry elaborated that these 'cuts' collectively reinforce the notion that women do not belong in the field, highlighting a significant impact of stereotype threat. The article further outlines evidence from research which consistently demonstrates that stress and math anxiety severely affect girls' focus, undermining their working memory and limiting their ability to perform well in the subject [3–5]. Poor female math performance perpetuates the gender stereotype that "girls are just not good at math" a belief which, despite their desire to excel, many

girls internalize hindering their success in the field [2]. Therefore, the aim of this study is to investigate the role of emotion, particularly anxiety, in the performance of primary school children on mathematical and spatial tasks. By measuring specific forms of subjective anxiety as well as physiological arousal, this study seeks to unravel the complex relationship between these psychological factors and their influence on performance in early education.

1.1. Gender Differences in Mathematics Performance in Primary School Students

Evidence for gender differences in math performance in primary school is mixed and context-dependent. Historically, studies indicate no significant differences between boys and girls in early education settings. Meta-analyses and longitudinal studies have generally found that any observed gender differences are small and often not statistically significant. For instance, Hyde et al. (2008) [6] conducted a meta-analysis in the USA and found gender differences in math performance to be negligible. However, gender differences can vary significantly by country and cultural context. Cross-national studies, such as those by the Program for International Student Assessment (PISA), show varying gender gaps across different educational systems [7]. The 2018 PISA study found that in Germany, the gender gap in math performance among adolescents was wider compared to the OECD average [7]. Research suggests that stereotypes and peer influence contribute to this outcome. Wolff (2021) [8] found that in German secondary schools, stereotypes like "boys are better at math than girls" negatively impacted girls' self-concept but not boys', even after controlling for grades and age. Gender stereotypes regarding math ability can emerge as early as primary school, affecting confidence and interest in the subject [9,10]. Additionally, negative cultural stereotypes such as the "STEM-nerd" influence the attractiveness of Science, Technology, Engineering, and Mathematics (STEM) and related careers, often limiting girls' identification with and choice of related fields [11]. The effects of societal stereotypes and associated thinking can also affect STEM subject preferences as early as primary and secondary school, leading to gender differences in the pursuit of mathematically related studies and STEM careers going forward [12–14].

1.2. Mental Rotation and the Role of Stereotypes in Its Assessment

Mental rotation (MR), the ability to mentally manipulate two- and three-dimensional objects, is a critical spatial skill extensively examined in psychology and education [15,16]. MR is typically assessed using psychometric instruments such as the mental rotation test (MRT), which usually consists of items with cubical stimuli (Vandenberg and Kuse, 1978 [17]). These tasks consistently show large and reliable performance differences favoring males [18]. Traditionally, MR tasks have shown a male performance advantage, potentially due to boys' greater engagement in spatially oriented play [19,20]. However, recent studies indicate that the design of MR tasks, including the use of gender-stereotyped or -neutral stimuli, can influence these gender differences by either reinforcing or mitigating stereotype threat [16,21]. Findings indicate that the familiarity or gender congruence of stimulus objects significantly influences performance differences. There is a positive correlation between the stereotyped content of the stimuli and children's MR performance [22–24].

1.3. The Association between Mental Rotation and Mathematical Performance

Numerous studies confirm the association between spatial and mathematical abilities [25–27]. Mental rotation (MR), like other spatial reasoning skills, can be improved through targeted practice and training, which often translates to better mathematical performance [28,29]. MR skills are linked to advanced mathematical skills and are strong predictors of mathematical competence in young children [30]. Moreover, MR performance is a significant indicator of academic success and career choices in STEM fields [31]. Similar to mathematical ability, gender differences in MR performance are influenced by stereotype threat, self-concept, and emotional arousal. These factors tend to negatively affect women's task performance due to a fear of confirming negative stereotypes about female weaknesses in visuo-spatial cognition [32–34].

Gender differences in performance on math and spatial tasks may be particularly evident in processing speed. Limited working memory capacity constrains speed, impacting completion, reaction, and response times for cognitive tasks like math and MR. Girls may be more sensitive to stress from time pressure on a math or spatial test or fear making mistakes, which diverts their cognitive resources, slowing the mental operations needed for rapid and accurate responses on timed tasks [34,35]. Boys, on the other hand, tend to demonstrate higher math self-concept and confidence in their mathematical ability [13]. They often overestimate their aptitude and are more influenced by peer dynamics, with social comparisons shaping their confidence in math [36]. Additionally, boys' tendency to prioritize speed over accuracy which can result in shorter processing times on cognitive tasks, potentially sacrificing thoroughness and leading to errors [37,38]. There is evidence from research that girls' performance on mathematical and spatial tasks is slower than boys', which in turn leads to a disadvantage in their overall performance in these areas [39,40]. This may be due to a variety of factors such as the effects of stereotype threat, higher anxiety levels during the tasks, and diverse processing strategies [34,41–43].

1.4. Interplay of Psychological Factors and Their Role in Mathematical and Spatial Performance

Mathematics performance is a fundamental component of academic achievement and cognitive development in primary school children [44]. Understanding mathematical concepts is crucial for success in various educational domains, including STEM fields [45,46]. However, individual differences in math performance are influenced by factors beyond cognitive abilities, such as self-concept and emotional arousal [47,48].

Mathematics anxiety, characterized by fear or apprehension during mathematical tasks, significantly hinders learning and achievement by depleting working memory and leading to poorer cognitive performance [49,50]. High math anxiety often results in reduced motivation and avoidance behaviors, contributing to underperformance [42,51,52]. Compared to boys, girls' appear to experience more negative effects of math anxiety on performance [53,54]. Similarly, spatial anxiety, which involves discomfort with spatial reasoning tasks, is linked to deficits in spatial abilities, particularly in girls and women. It can disrupt performance on mental rotation tests, especially when gender stereotypes are at play [41,55,56]. Hence, anxiety in both math and spatial tasks can hinder children's cognitive performance.

Recent studies highlight the interconnectedness of psychological factors in gender differences in mathematical and spatial performance. Spatial anxiety and spatial self-efficacy have been found to mediate gender differences in mental rotation performance, particularly in challenging tasks [57]. Additionally, mental rotation performance and perceived competence mediate the relationship between gender and math anxiety in younger populations [58]. Research also links statistics anxiety with spatial processing and performance, and recent findings suggest that lower math self-concept in girls may affect accuracy in mental rotation, with spatial anxiety playing a mediating role [13,59]. These studies highlight the interconnectedness of mathematical and spatial skills and associated anxieties.

1.5. Physiological Markers of Emotion Regulation and Their Role in Spatial and Mathematical Ability

Arousal, representing the level of central nervous system activity, plays a critical role in regulating the human stress response [60]. Transactional models of stress suggest that various events, tasks, and conditions, as well as the individual's assessment of these, trigger physiological, cognitive, behavioral, and emotional responses [61]. Measuring physiological markers, such as systolic blood pressure and electrodermal activity (EDA), offers valuable insights into emotional reactivity and cognitive engagement during learning tasks [62,63].

Galvanic Skin Responses (GSRs) serve as biomarkers of autonomic nervous system (ANS) arousal and are well established for assessing psychophysiological functioning in humans [64,65]. The GSR effectively captures physiological changes associated with

emotional expressions such as anxiety and stress in children [66,67]. Wearable devices for GSR measurement facilitate the examination of physiological responses in children engaged in cognitive tasks under real-life conditions [68,69]. Galvanic Skin Responses (GSRs) in children vary depending on the context in which they are measured. These variations are influenced by the emotional and stressful nature of tasks, reflecting differences in emotional arousal and stress responses [67,70]. For example, anxiety linked to math tasks is associated with increased vigilance and heightened amygdala activity, leading to greater physiological reactivity [48,71]. Mental rotation tasks, due to their demand for spatial reasoning, can trigger significant stress and emotional arousal, especially when new to children [41]. In contrast, familiar math tasks practiced regularly in school may reduce anxiety and physiological arousal [72].

The ability to regulate emotions impacts performance on spatial tasks, with gender differences in skin conductance responses observed during mental rotation tasks [73,74]. A pilot study by Lennon-Maslin and colleagues (2023) highlighted that increased physiological arousal, measured by GSR, influenced primary school children's response times on specific items in a novel mental rotation task [75]. These physiological responses are context-dependent, influenced by factors such as stimulus type, measurement context, emotion regulation strategies, and sociocultural influences [76]. Understanding these variables is crucial for interpreting GSR differences among children in educational settings.

1.6. Aims of the Current Study

Despite growing interest in the role of anxiety and physiological arousal in mathematics and spatial performance [48,62], limited research has investigated their combined effects in primary school children. Moreover, only one pilot study to date has examined the role of emotional reactivity measured by the GSR in mental rotation performance in the same target group [75]. The present study, therefore, represents a novel contribution to research in this field by examining the simultaneous impact of math and spatial anxiety, along with physiological arousal measured through skin conductance, on both mathematical performance and mental rotation skills in primary school children. By utilizing standardized measures for anxiety and physiological assessment, this study aims to provide an unprecedented, comprehensive understanding of the interplay between emotional reactivity and anxiety-related psychological factors in shaping early cognitive performance. This integrated approach not only addresses a critical gap in the literature but also offers valuable insights which could inform educational interventions designed to enhance learning outcomes in young children.

1.7. Research Questions

Based on an extensive review of the literature, our study set out to explore some key questions about gender differences in performance on math and spatial ability among primary school children. Specifically, we sought to determine if there are noticeable differences between boys and girls in their performance on mathematical and spatial tasks and their association with task-related subjective anxiety, emotional reactivity, and their preferences for school subjects.

Our first research question seeks to answer whether there are gender differences in the performance of our sample on cognitive tasks, specifically in math and mental rotation, in subjective anxiety (math and spatial anxiety), and in the preference for math vs. German. This question addresses a critical gap in the literature, as previous studies have shown mixed and context-dependent findings regarding gender differences in these areas [6–8]. By examining these variables in primary school children within a specific cultural context, our study aims to clarify these inconsistencies.

Our second research question delves into the interplay between several factors—gender, subject preference, emotional reactivity (as measured by the Galvanic Skin Response or GSR), and subjective anxiety (both math and spatial anxiety)—and how they collectively influence performance on cognitive tasks, specifically in math and mental rotation. While

previous research has extensively explored these variables individually, few studies have examined how they interact to influence cognitive performance in young children, particularly in a primary school setting [42,48,77]. This research question seeks to address this gap by providing a more holistic view of how these factors combine to affect performance.

1.8. Hypotheses

H1. *We hypothesized that there would be significant gender differences in both the percentage scored and the completion times on the math task. We further hypothesized that girls would outperform boys in terms of the percentage scored but would take longer to complete the task.*

Rationale: This hypothesis explores the idea that girls often perform just as well or better than boys in math accuracy but often at the cost of speed, potentially due to increased anxiety and diverse problem-solving strategies [40,43].

H2. *We expected no significant gender differences in terms of accuracy and response time on the mental rotation task.*

Rationale: Recent research challenges traditional beliefs about gender differences in spatial abilities, particularly when tasks are designed to minimize stereotype threat [22–24,75]. This hypothesis also tests whether these findings hold in a primary school context.

H3. *We anticipated that girls would report higher levels of subjective anxiety (both math and spatial) compared to boys.*

Rationale: This expectation aligns with the existing literature, which frequently documents higher anxiety levels among females in math-related contexts [53,54] and also on spatial tasks [13,41,55]. Studies exploring gender differences in spatial anxiety in primary school children are particularly limited, as are studies examining its effects on math performance [13,78].

H4. *We hypothesized that girls in primary education would be less likely than boys to choose math as their preferred school subject over German.*

Rationale: Although some studies have examined how stereotypical thinking can lead to gender differences in the preference for mathematics and STEM as a subject at secondary school and beyond, few studies have examined this in the primary school context [12–14]. This hypothesis addresses a gap in understanding how early subject preferences play a role in performance on math and spatial tasks in primary education. This may have implications for the pursuit of STEM studies and careers, a field in which gender disparities are most pronounced [79,80].

H5. *We hypothesized that there would be a significant association between gender, subject preference, and performance on the math task when controlling for emotional reactivity and math and spatial anxiety. For girls, we expected that a preference for math, lower emotional reactivity, and lower math and spatial anxiety would predict higher percentages scored and shorter completion times on the math task compared to boys.*

Rationale: This hypothesis aims to identify potential protective factors, such as subject preference, reduced subjective anxiety and physiological arousal that might mitigate potential negative impacts on performance [13,59,81]. This area remains underexplored in the literature.

H6. *We hypothesized that there would be a significant association between gender, subject preference, and performance on the mental rotation task when controlling for math and spatial anxiety and*

emotional reactivity. However, we expected no significant gender differences in how these factors affect performance on the mental rotation task, that is, the effects would be similar for both girls and boys.

Rationale: Anxiety, emotional reactivity, and physiological arousal are known to affect performance on complex cognitive tasks such as mental manipulation which may trigger stereotype effects known to be associated with spatial ability [48,67,73]. This hypothesis examines whether commonly observed gender differences in emotional and physiological factors persist on a mental rotation task consisting of gender-congruent and -neutral stimuli.

2. Materials and Methods

A total of one hundred and thirty-one students were recruited from primary schools in the state of Rhineland Palatinate in Germany (*N* = 131). All of the students belonged to second-, third-, and fourth-grade classes, respectively, and were enrolled in five local primary schools in Rhineland Palatinate, Germany. Schools were situated in both urban and rural settings. Although all of the children were fluent German speakers, some were also bilingual and had diverse ethnic backgrounds. Unfortunately, specific data regarding ethnicity and socio-economic status was not collected from parents and guardians. These observations were made by the researchers and through information volunteered by children prior to testing.

All students completed the math task, the mental rotation task, and various questionnaires. Included in this sample were 66 students who identified as boys and 65 as girls. There were no students who identified as non-binary in this sample. The average age of the students was *M* = 8.73 (*SD* = 1.02) years old. Mean grade level distribution by gender is shown in the bar chart in Table 1.

Table 1. Class, sex crosstabulation, and grade in primary school.

| | | Sex | | Total |
		Boys	Girls	
	2nd Grade	16	21	37
Class	3rd Grade	23	22	45
	4th Grade	27	22	49
	Total	66	65	131

2.1. Skin Conductance

Shimmer3 GSR+ Unit® (https://shimmersensing.com/product/consensyspro-software/, accessed on 1 January 2024) was used to measure galvanic skin response (GSR). These devices were synchronized with ConsensysBasic® multi-sensor management software (https://shimmersensing.com/product/consensyspro-software/, accessed on 1 January 2024) where they were calibrated for recording and a sampling rate of 1 Hz was set. A baseline recording of 2 min per participant was planned in order to compare this with GSR during the math task and during the MRT.

Shimmer devices were attached to the index and middle fingers of the non-dominant hand. This placement allows for optimal signal detection while minimizing interference with the tasks performed using their dominant hand. Each child was asked to try and keep the monitored hand as still as possible throughout the measurement period. This was to ensure consistent and accurate readings by minimizing movement artifacts. The GSR measurement began with a 2 min baseline recording. During this period, the children were asked to sit quietly and relax. This baseline data served as a control to compare physiological arousal during the subsequent tasks. After the baseline measurement, the students performed a series of math problems. The GSR device continued to record physiological responses throughout this task. Following the completion of the math task, students proceeded to the mental rotation task. The GSR measurement continued during this period to capture physiological responses associated with this cognitive activity.

Once the tasks were completed, the devices were redocked to stop the recording. This was completed before administering the self-report questionnaires to ensure that the GSR data were isolated to the periods of baseline, math, and mental rotation tasks. The measurements were conducted in a controlled environment, specifically a separate classroom with adequate lighting and individual seating arrangements. This setup aimed to minimize external distractions and create a comfortable setting for the children to perform the tasks. The room temperature and other ambient conditions were kept consistent to avoid any environmental factors that might influence physiological responses. This careful administration and controlled environment ensured the reliability and validity of the GSR measurements, allowing for an accurate assessment of the children's physiological arousal in response to the cognitive tasks.

Skin conductivity is measured in units referred to as microsiemens (S). "Micro" is a prefix meaning millionths, so 1 microsiemen (1 S) is a unit of time in the International System of Units (SI) equal to one millionth of a siemen [82].

2.2. Math Task

Students were given a paper-and-pencil math task consisting of 9 items of varying difficulty levels: A, B, and C. Each item had 5 possible solutions, with only one correct answer. A-level tasks were the easiest, worth 3 points each, B-level tasks were more difficult, worth 4 points each, and C-level tasks were the most difficult, worth 5 points each. Second- and third-grade students had to solve three A-level, four B-level, and two C-level problems and could score a maximum of 35 points, while fourth-grade students could score 36 points due to an additional C-level problem. The problems were drawn from the Mathematical Kangaroo Competition an international math challenge in over 77 countries, the purpose of which is to show students that math can be interesting, beneficial, and even fun [83]. The Kangaroo math problems have been translated into many languages; therefore, for our task, we drew the problems in the German language from the "Mathematikwettbewerb Känguru" website [84]. To ensure unfamiliarity, problems from past competitions were chosen. Our math task included three types of math problems: number line estimation, word problem representation, and missing terms. Math performance was measured based on the percentage scored and the completion time, that is, the interval from when the task is presented to when the participant finishes solving the problems and submits their answers.

Students received written instructions to work individually, circle or mark answers, and refrain from using a calculator or other electronic aid. Zero points were given for unanswered questions. The task aimed to engage students in math and minimize fatigue. An example of each type of math problem used in the task is listed in Appendix A.

2.3. Mental Rotation Task

A novel, computerized mental rotation task (nMRT) based on Vandeberg and Kuse's original MRT but consisting of more varied stimuli was programmed in PsychoPy® software (www.psychopy.org/download.html, accessed on 6 April 2023) and installed on Microsoft Pro 8 Surface tablets, each with a keyboard and a mouse. The task was programmed to record both the number of correct responses (accuracy) as well as the duration in seconds from when the stimulus is presented to when the participant provides the answer (response time). Items included mental rotation stimuli for younger children, i.e., animals and letters [85], as well as other concrete stimuli and abstract stimuli, all either rotated in-depth or in picture plane. Abstract items consisted of stimuli such as cubes, pellets [22], and polyhedral figures [24]. Other concrete items consisted of male and female gender-stereotyped stimuli [21,23] and gender-neutral stimuli [24]. Examples of some of the items used are shown in Figure 1 and are also appended to this report.

The MRT was divided into two parts, each with a time limit. Part one (MRT 1) consisted of 6 abstract and 10 concrete items which were rotated in picture plane only. It was limited to 5 min. Part two (MRT 2) contained 6 abstract and 6 concrete items rotated in-depth and was limited to 8 min. Items were presented randomly in each part of the task

with one target stimulus on the left and four comparison stimuli on the right. Participants were instructed to identify two out of four stimuli on the right which, although rotated, were identical to that on the left. A reliability analysis of MRT conducted on the data from our sample yielded a Cronbach's Alpha of $\alpha = 0.86$.
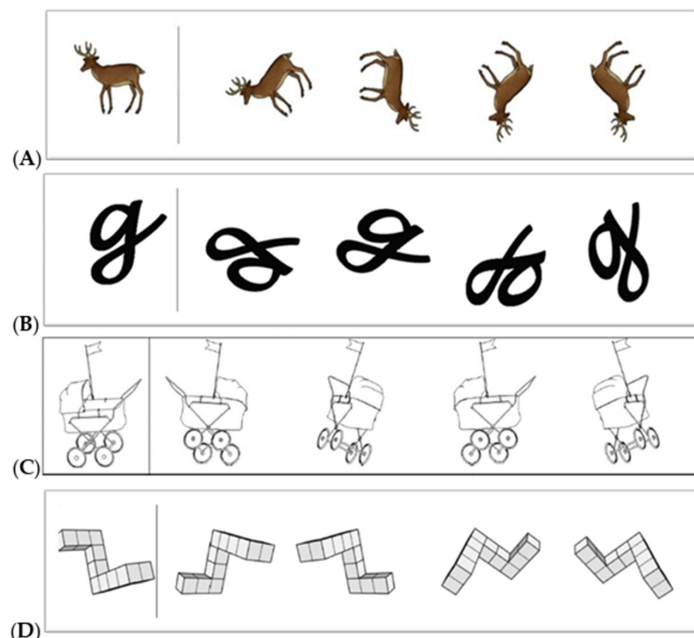

Figure 1. Examples of MRT items used with concrete objects: (**A**) animal stimulus and (**B**) letter stimulus [85]; (**C**) female-stereotyped stimulus "Pram" [21,23]; and with an abstract object (**D**) "Cube" figure [22].

2.4. Demographic Data

An online questionnaire was presented to each participant following the MRT to collect data relating to participants' age and gender. The position of this and the other self-report questionnaires in the procedural sequence, i.e., being placed after the math and mental rotation tasks, aimed to avoid priming or eliciting stereotype threat in male and female participants.

2.5. Self-Report Questionnaires

Three further self-report questionnaires were also presented in PsychoPy®, the purpose of which was to assess the students' preference for math vs. German, their spatial anxiety, and their math anxiety. All the questionnaires were presented to students in the German language.

2.5.1. Math versus German Preference Survey

Students were surveyed to indicate their preferred school subject among math, German, or no preference for either subject. Their responses were coded in SPSS for statistical analysis as follows: math = 1; German = 2; no preference = 3. This question is part of the Academic Self-Concept in Primary School Children (ASKG) questionnaire, a German self-report measure of reading, writing, and math self-concept for children [86]. We used only the Math vs. German survey section, which is appended to this article.

2.5.2. Child Spatial Anxiety Questionnaire

The Child Spatial Anxiety Questionnaire (CSAQ) [41] consists of 8 items in which students are asked to rate how anxiety-provoking they find a particular task. All tasks described in the questionnaire require spatial ability and skills. Researchers explained to students in advance that this questionnaire was about feelings and gave an example of

how children might experience nervousness and anxiety, i.e., heart racing, rapid breathing, hands trembling. An example of one item from the CSAQ is as follows: "How would you feel if your teacher asked you to measure something with a ruler?" Participants are then asked to rate on an emoji scale of 1 to 5, 1 being "not nervous at all, calm", 3 being "neither calm nor nervous", and 5 being "very, very nervous", how nervous they would feel about completing this task. A reliability analysis of the CSAQ conducted on the data from our sample yielded a Cronbach's Alpha of $\alpha = 0.69$.

2.5.3. Modified Abbreviated Math Anxiety Scale (mAMAS)

The Modified Abbreviated Math Anxiety Scale (mAMAS) is based on the Abbreviated Math Anxiety Scale, an American self-report instrument developed by Hopko and colleagues (2003) [87]. It was modified by [88] to be used with British children aged 8–13 year-old and is designed to measure levels of mathematics anxiety. The scale consists of 9 items, which are rated on a 5-point Likert scale ranging from 0 (strongly disagree) to 4 (strongly agree). Respondents indicate the extent to which they experience anxiety-related thoughts and feelings when faced with mathematical tasks or situations. The mAMAS items cover various aspects of math anxiety, including worries about math classes, fear of performing poorly in math-related activities, and discomfort with mathematical problem solving. The scale aims to provide a brief yet reliable measure of math anxiety, making it suitable for use in research studies where brevity and efficiency are a priority. We used a translation of this scale previously validated for its psychometric properties by native German-speaking colleagues. Students rated their level of math anxiety on a scale of 1 to 5 with 1 being "low anxiety" and 5 being "high anxiety" based on statements such as the following: "Finding out that you are going to have a surprise math test when you start your math lesson?" Total scores on the mAMAS can range from 0 to 36, with higher scores indicating greater levels of mathematics anxiety. The reported internal consistency of this questionnaire is high. We found a Cronbach's Alpha consistent with high reliability ($\alpha = 0.87$) when we tested this on our sample data.

2.6. Procedure

Approval for the pilot study was provided by the Ethics Committee of the University of Koblenz and the state authorities in Rhineland Palatinate overseeing schools. Informed consent was sought and provided by parents and guardians of all students involved in this study. The class teacher and the principal also gave permission for this study to be conducted in the school. Students provided verbal assent prior to commencing the experiment and were informed that they could withdraw their participation with no consequences at any point. Students were tested by two female researchers in a separate classroom with access to their teacher, if required. Consistent supervision of the data collection by teaching staff could not be provided in any of the schools where data were collected and, although this arrangement was approved by the ethics committee, it may have inadvertently evoked increased anxiety in some students, an issue discussed in the limitations of this article. The room had adequate lighting and individual seating arrangements. The researchers explained the mental rotation task to the students by rotating objects such as a pair of scissors, a toy, and a wooden object, it does not change its features. Students were then asked to imagine the object in their mind then try to rotate it mentally. The purpose of this study and the significance of mental rotation in everyday life and for schoolwork was also explained. The researchers also checked in advance that students were familiar and comfortable with the use of a keyboard and a mouse. Furthermore, any student who required eye-glasses was reminded to wear these while viewing the tablet screen.

2.7. Data Analysis

Quantitative data analyses were conducted using SPSS® 29 software. We employed a range of statistical tests to explore various aspects of our research questions. Raw scores

on the math task were standardized by converting them to percentages for comparability during the statistical analyses.

The accuracy scores and response time on the mental rotation task and the spatial anxiety score variables exhibited positive skewness, as evidenced by the skewness statistic and visual inspection of the histogram and Q-Q plot. To address this, a square root transformation was applied to accuracy (Sqr_Acc) and response time (Sqr_RT) and a logarithmic transformation was applied to the spatial anxiety variable (log_CSA). This transformation improved the normality of the distribution, reducing skewness and bringing the distribution of these variables closer to a normal shape.

A Chi-Square test of association was used to test whether girls or boys were more likely to choose math or German as their preferred school subject. The assumptions for the application of this test were met, that is, the sample size was large enough, the nominal variables were categorical, participants were randomly and independently sampled, the data consisted of raw frequencies and the expected frequency (Fe) within each cell was greater than 1, and no more than 20% of the cells should have less than 5.

A multivariate analysis of covariance (MANCOVA) was used to assess gender differences in math and mental rotation performance and math and spatial anxiety. Assumptions for the MANCOVA, such as multivariate normality of the dependent variables, homogeneity of variances (verified with Levene's test) and covariances (verified with Box's test), linearity, homogeneity of slopes (verified by checking interaction terms), the independence of observations, and the absence of multicollinearity and outliers, were satisfied for all variables [89].

Prior to data analyses, raw skin conductance data were pre-processed by removing artefacts and noise [82]. The mean of the GSR for the three experimental conditions of baseline and during the math and mental rotation tasks was calculated for each participant in order to facilitate analysis. GSR variability can lead to skewed distributions within the data which require further processing and robust statistical methods which are not sensitive to non-normality. Following the removal of outliers and checks for normality on the skin conductance variables, the data were found to be skewed. This is common with physiological data due to the nature of measurements, which often do not follow a normal distribution [82]. Pre-processing and data transformation such as square-root and logarithmic transformation did not improve this. However, the test of normality within the MANCOVA did not reveal any significant effects. Additionally, as skin conductance can vary widely between individuals due to differences in baseline arousal levels, skin properties, and other factors, raw GSR data were standardized by performing a z-score transformation. Z-scores help normalize differences, making individual responses more comparable [90].

3. Results

3.1. Gender Differences in Math and Mental Rotation Performance, Math and Spatial Anxiety, and Subject Preference

A multivariate analysis of variance (MANOVA) was conducted to examine the effects of the fixed factor gender on the combined dependent variables of math performance (percentage scored and completion times), mental rotation performance (accuracy and response time), math anxiety, and spatial anxiety. The multivariate effect of gender was significant, $\lambda = 0.80$, $F(1, 129) = 5.22$, $p < 0.001$, with a small effect size of $\eta^2 = 0.20$, indicating significant differences between boys and girls across the dependent variables (see Table 2).

Table 2. Multivariate tests: effect of gender (sex) on performance on math and mental rotation tasks, math, and spatial anxiety in primary school children.

	Effect	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Sex	Wilks' Lambda	0.798	5.219	6.000	124.000	<0.001	0.202

Design: intercept + sex.

A series of between-subjects effects tests were conducted to examine gender differences in percentage scored and completion times on the math task, accuracy and response time on the mental rotation task, math anxiety, and spatial anxiety. The analyses yielded the following results.

For the math percentage scored, there was a statistically significant effect of gender, $F(1, 129) = 4.63$, $p = 0.033$, $\eta^2 = 0.03$. This indicates that gender accounts for 3% of the variance in percentage scores. Similarly, for completion times on the math task, the effect of gender was also statistically significant, $F(1, 129) = 7.20$, $p = 0.008$, $\eta^2 = 0.05$, indicating that gender accounts for 5% of the variance in completion times. In contrast, gender did not have a significant effect on accuracy in the mental rotation task, $F(1, 129) = 0.03$, $p = 0.872$, $\eta^2 = 0.00$, nor on response time, $F(1, 129) = 3.20$, $p = 0.076$, $\eta^2 = 0.02$. The effect of gender on math anxiety was significant, $F(1, 129) = 10.50$, $p = 0.002$, $\eta^2 = 0.08$. This indicates that gender explains 8% of the variance in math anxiety scores. For spatial anxiety, the effect of gender was not statistically significant, $F(1, 129) = 2.28$, $p = 0.134$, $\eta^2 = 0.02$.

Overall, these results indicate significant gender differences in percentage scores and completion times on the math task, as well as in math anxiety. The effect of gender on accuracy, response time, and spatial anxiety was not significant (see Table 3).

Table 3. Tests of between-subjects: effects of gender (sex) on percentage scored (Math_PT) and completion time (Math_CT) on the math task, and math anxiety (MA) in primary school children.

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
	Math_PT	1643.461	1	1643.461	4.630	0.033	0.035
Sex	Math_CT	195.693	1	195.693	7.200	0.008	0.053
	MA	6.015	1	6.015	10.501	0.002	0.075

Pairwise comparisons were conducted to examine the differences between boys and girls on various dependent variables, including math percentage scores, math completion times, accuracy, response time, math anxiety, and spatial anxiety.

For math percentage scores, girls scored significantly higher than boys, with a mean difference of 7.08, *SE* = 3.30, $p = 0.033$, 95% *CI* [0.57, 13.56]. Regarding math completion times, girls took significantly longer to complete the tasks compared to boys, *mean difference* = 2.44, *SE* = 0.91, $p = 0.008$, 95% *CI* [0.64, 4.25]. In terms of accuracy on the nMRT, there were no significant differences between boys and girls, *mean difference* = 0.005, *SE* = 0.03, $p = 0.872$, 95% *CI* [−0.05, 0.06], nor in response time, *mean difference* = −0.19, *SE* = 0.12, $p = 0.076$, 95% *CI* [−0.40, 0.02]. Regarding math anxiety, girls reported significantly higher levels of math anxiety than boys, *mean difference* = 0.43, *SE* = 0.13, $p = 0.002$, 95% *CI* [0.17, 0.69]. For spatial anxiety, there were no significant differences between boys and girls, *mean difference* = 0.005, *SE* = 0.03, $p = 0.872$, 95% *CI* [−0.05, 0.06].

Overall, the results suggest that girls outperform boys in terms of math percentage scores but take longer to complete the math task. The error bars in Figure 2 represent the standard error of the mean (SEM). The difference in the width of the error bars for percentage scored and completion times indicates that there is more variability in the scores compared to the completion times. Scores might be more sensitive to individual differences in mathematical ability, understanding, and problem-solving strategies, leading to greater variability in scores. Completion times might be less variable because they are influenced by factors such as time constraints and task pacing, which could be more consistent across participants. There are no significant differences between boys and girls in accuracy nor response time on the mental rotation task or in spatial anxiety.
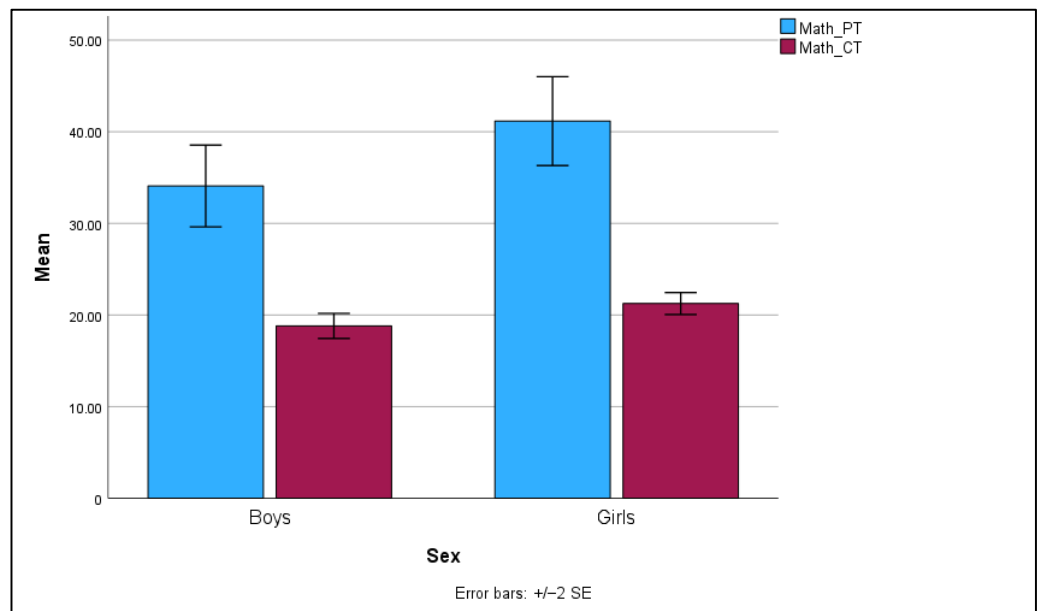
Figure 2. Gender (sex) differences in percentage scored (blue bar = Math_PT) and completion times (red bar = Math_CT) on the math task. Error bars represent the standard error of the mean (SEM). The error bars indicate similar variability in scores and completion times between boys and girls.

A Chi-Square test for association was performed to examine the relationship between gender and preference for math vs. German. The results indicated a strong significant association between gender and subject preference, χ^2 (2, N = 131) = 13.50, p < 0.001. Specifically, girls were significantly less likely than boys to indicate a preference for math and were more likely to choose German or indicate no preference for either subject (see Figure 3). A crosstabulation analysis was conducted to examine the relationship between gender (sex) and subject preference (math vs. German). The results are as follows. Out of the 66 boys, 44 (65.7%) preferred math, 8 (28.6%) preferred German, and 14 (38.9%) had no preference. Out of the 65 girls, 23 (34.3%) preferred math, 20 (71.4%) preferred German, and 22 (61.1%) had no preference. The overall distribution across both sexes showed that 67 participants (51.1%) preferred math, 28 participants (21.4%) preferred German, and 36 participants (27.5%) had no preference.
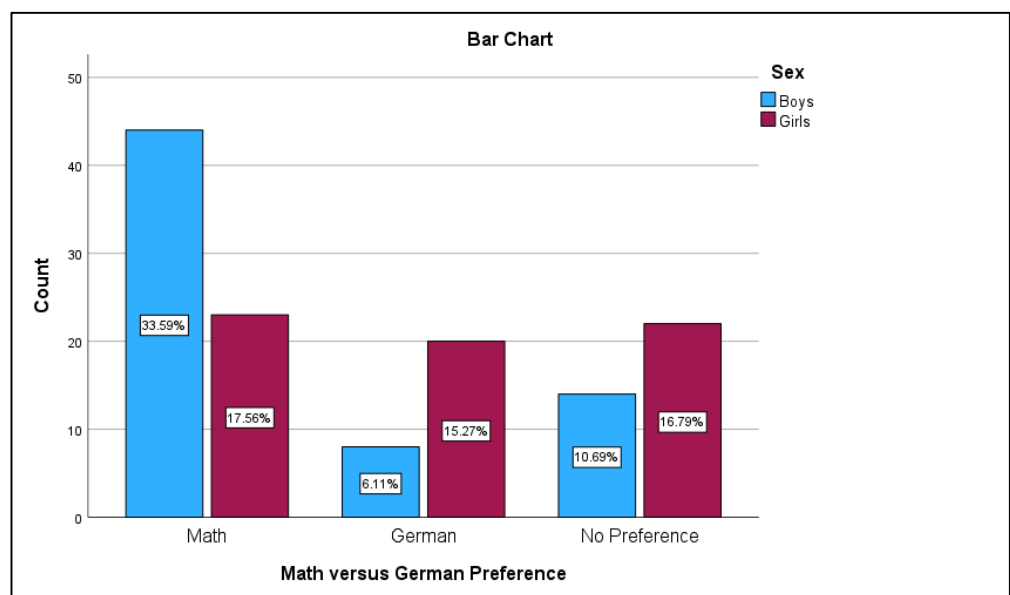

Figure 3. Percentage of math vs. German as the preferred school subject by gender (sex).

3.2. Association between Gender, Subject Preference, Emotional Reactivity (Measured by GSR), Subjective Anxiety (Math and Spatial Anxiety), and Performance on the Math and Mental Rotation Tasks

3.2.1. Effects on Math Performance

A multivariate analysis of covariance (MANCOVA) was conducted to examine the effects of various predictors on two dependent variables: percentage scored and completion times on the math task. The predictors included two fixed factors, gender and subject preference, and three covariates, emotional reactivity measured by GSR, math anxiety, and spatial anxiety.

The results of the multivariate tests, as indicated the main effect of gender was significant, *Wilks'* $\Lambda = 0.91$, $F(2, 121) = 5.60$, $p = 0.004$, $\eta^2 = 0.09$, indicating a significant difference in the combined dependent variables based on gender. Additionally, Wilks' Lambda showed a significant effect of math anxiety on the combined dependent variables, *Wilks'* $\Lambda = 0.910$, $F(2, 121) = 5.96$, $p = 0.003$, $\eta^2 = 0.09$, suggesting that math anxiety significantly impacts the combined dependent variables. The main effect of subject preference was also significant, *Wilks'* $\Lambda = 0.89$, $F(4, 242) = 3.43$, $p = 0.009$, $\eta^2 = 0.05$, suggesting that subject preference significantly impacts the combined dependent variables. The interaction effect between gender and subject preference was not significant, *Wilks'* $\Lambda = 0.98$, $F(4, 242) = 0.52$, $p = 0.718$, $\eta^2 = 0.01$, indicating that this interaction does not significantly affect the combined dependent variables. The individual effects of emotional reactivity measured by GSR and spatial anxiety on the combined dependent variables were not significant, with *Wilks'* $\Lambda = 0.97$, $F(2, 121) = 1.78$, $p = 0.172$, $\eta^2 = 0.03$ for emotional reactivity measured by GSR and *Wilks'* $\Lambda = 0.97$, $F(2, 121) = 1.93$, $p = 0.15$, $\eta^2 = 0.03$ for spatial anxiety.

Overall, these multivariate tests demonstrate that gender has a significant impact on the dependent variables, with notable contributions from subject preference and math anxiety. The effect of emotional reactivity measured by GSR and spatial anxiety were not significant. The results are summarized in Table 4.

Table 4. Multivariate tests: effects of math anxiety (MA), gender (sex), and subject preference on performance on the math task in primary school children.

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
MA	Wilks' Lambda	0.910	5.965	2.000	121.000	0.003	0.090
Sex	Wilks' Lambda	0.913	5.797	2.000	121.000	0.004	0.087
Subject Preference	Wilks' Lambda	0.895	3.433	4.000	242.000	0.009	0.054

Design: intercept + MA + sex + subject preference.

A series of tests of between-subjects effects were conducted to examine the influence of gender, subject preference on percentage scored, and completion times on the math task while controlling for emotional reactivity measured by GSR and math anxiety and spatial anxiety. The analyses yielded the following results.

For the dependent variable percentage scored, the between-subjects effects indicated several significant results. There was a significant effect of math anxiety on percentage scored, $F(1, 122) = 11.86$, $p < 0.001$, $\eta^2 = 0.09$. The effect of gender on percentage scored was also significant, $F(1, 122) = 6.53$, $p = 0.012$, $\eta^2 = 0.051$.

For the dependent variable completion times, the between-subjects effects showed the following significant results. The effect of gender on completion times was significant, $F(1, 122) = 4.44$, $p = 0.037$, $\eta^2 = 0.03$. There was also a significant effect of subject preference on completion times, $F(2, 122) = 5.94$, $p = 0.003$, $\eta^2 = 0.09$. The effect of emotional reactivity measured by GSR on completion times was not significant, $F(1, 122) = 3.06$, $p = 0.083$, $\eta^2 = 0.02$, nor was the effect of spatial anxiety, $F(1, 122) = 3.79$, $p = 0.054$, $\eta^2 = 0.03$. The interaction effect between gender and subject preference was not significant for either percentage scored, $F(2, 122) = 1.01$, $p = 0.367$, $\eta^2 = 0.02$, or completion times, $F(2, 122) = 0.08$, $p = 0.921$, $\eta^2 = 0.00$.

These results indicate that, while gender and math anxiety have significant effects on math percentage scored, gender and subject preference are significant for completion times. The interaction between gender and subject preference does not significantly impact either dependent variable. Additionally, emotional reactivity measured by GSR and spatial anxiety do not show significant effects on math completion times (see Table 5).

Table 5. Tests of between-subjects effects of math anxiety (MA), gender (sex), and subject preference on percentage scored (Math_PT) and completion times (Math_CT) on the math task.

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
MA	Math_PT	3884.253	1	3884.253	11.856	<0.001	0.089
Sex	Math_PT	2140.623	1	2140.623	6.534	0.012	0.051
	Math_CT	108.920	1	108.920	4.437	0.037	0.035
Subject_Preference	Math_CT	291.391	2	145.696	5.936	0.003	0.089

3.2.2. Gender-Specific Effects on Math Performance

A multivariate analysis of covariance (MANCOVA) was conducted to examine the effects of subject preference, emotional reactivity measured by GSR, math anxiety, and spatial anxiety on math percentage scored and completion times separately for boys and girls.

For boys:

Multivariate tests showed that the effects of emotional reactivity measured by GSR on boys' performance on the math task were not significant (*Wilks'* $\Lambda = 0.10$, $F(2, 59) = 0.09$, $p = 0.915$, $\eta^2 = 0.00$), neither was math anxiety (*Wilks'* $\Lambda = 0.95$, $F(2, 59) = 1.55$, $p = 0.220$, $\eta^2 = 0.05$), nor spatial anxiety (*Wilks'* $\Lambda = 0.98$, $F(2, 59) = 0.72$, $p = 0.489$, $\eta^2 = 0.02$), nor subject preference (*Wilks'* $\Lambda = 0.90$, $F(4, 118) = 1.62$, $p = 0.173$, $\eta^2 = 0.05$). Tests of between-subjects effects showed no significant effect of emotional reactivity on boys' percentage scored ($F(1, 60) = 0.01$, p $= 0.929$, $\eta^2 = 0.00$). Neither math anxiety ($F(1, 60) = 3.14$, $p = 0.081$ $\eta^2 = 0.05$), nor spatial anxiety ($F(1, 60) = 0.02$, $p = 0.876$, $\eta^2 = 0.00$), nor subject preference ($F(2, 60) = 1.04$, $p = 0.359$, $\eta^2 = 0.03$) showed significant effects on percentage scored on the math task.

For boys' completion times on the math task, neither emotional reactivity measured by GSR ($F(1, 60) = 0.172$, $p = 0.680$, $\eta^2 = 0.00$), nor math anxiety ($F(1, 60) = 0.02$, $p = 0.878$, $\eta^2 = 0.00$), nor spatial anxiety ($F(1, 60) = 1.44$, $p = 0.234$, $\eta^2 = 0.02$), nor subject preference ($F(2, 60) = 2.31$, $p = 0.108$, $\eta^2 = 0.07$) showed significant effects.

For girls:

Multivariate tests showed significant effects of emotional reactivity measured by GSR (*Wilks'* $\Lambda = 0.84$, $F(2, 58) = 5.53$, $p = 0.006$, $\eta^2 = 0.16$) on girls' performance on the math task. Math anxiety (*Wilks'* $\Lambda = 0.83$, $F(2, 58) = 5.93$, $p = 0.005$, $\eta^2 = 0.17$) and subject preference (*Wilks'* $\Lambda = 80$, $F(4, 116) = 3.45$, $p = 0.011$, $\eta^2 = 0.11$) also had significant effects, but spatial anxiety did not (*Wilks'* $\Lambda = 0.95$, $F(2, 58) = 1.67$, $p = 0.197$, $\eta^2 = 0.05$) (see Table 6).

Table 6. Multivariate tests: gender-specific effects of emotional reactivity measured by GSR (ZGSR_Math), math anxiety (MA), and subject preference and performance on the math task.

Sex	Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
	ZGSR_Math	Wilks' Lambda	0.840	5.532	2.000	58.000	0.006	0.160
Girls	MA	Wilks' Lambda	0.170	5.929	2.000	58.000	0.005	0.170
	Subject_Preference	Wilks' Lambda	0.799	3.450	4.000	116.000	0.011	0.106

Design: intercept + ZGSR_Math + MA + Subject_Preference.

Tests of between-subjects effects showed that emotional reactivity measured by GSR had no significant effects on girls' percentage scored, $F(1, 59) = 1.34$, $p = 0.251$, $\eta^2 = 0.02$.

Math anxiety showed significant effects ($F(1, 59) = 10.22$, $p = 0.002$, $\eta^2 = 0.145$), but spatial anxiety ($F(1, 59) = 1.13$, $p = 0.292$, $\eta^2 = 0.02$) and subject preference ($F(2, 59) = 1.45$, $p = 0.242$, $\eta^2 = 0.05$) did not (see Table 7).

Table 7. Tests of between-subjects gender-specific effects of math anxiety (MA), emotional reactivity measured by GSR (ZGSR_Math), and subject preferences on percentage scored (Math_PT) and completion times (Math_CT) on the math task.

Sex	Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
	ZGSR_Math	Math_CT	158.365	1	158.365	8.480	0.005	0.126
Girls	MA	Math_PT	3394.999	1	3394.999	10.223	0.002	0.148
	Subject_Preference	Math_CT	202.922	2	101.461	5.433	0.007	0.156

For girls' completion times, emotional reactivity measured by GSR showed significant effects ($F(1, 59) = 8.48$, $p = 0.005$, $\eta^2 = 0.13$), as did subject preference ($F(2, 59) = 5.43$, $p = 0.007$, $\eta^2 = 0.16$), but math anxiety ($F(1, 59) = 0.65$, $p = 0.423$, $\eta^2 = 0.01$) and spatial anxiety ($F(1, 59) = 2.76$, $p = 0.102$, $\eta^2 = 0.04$) did not (see Figure 4).
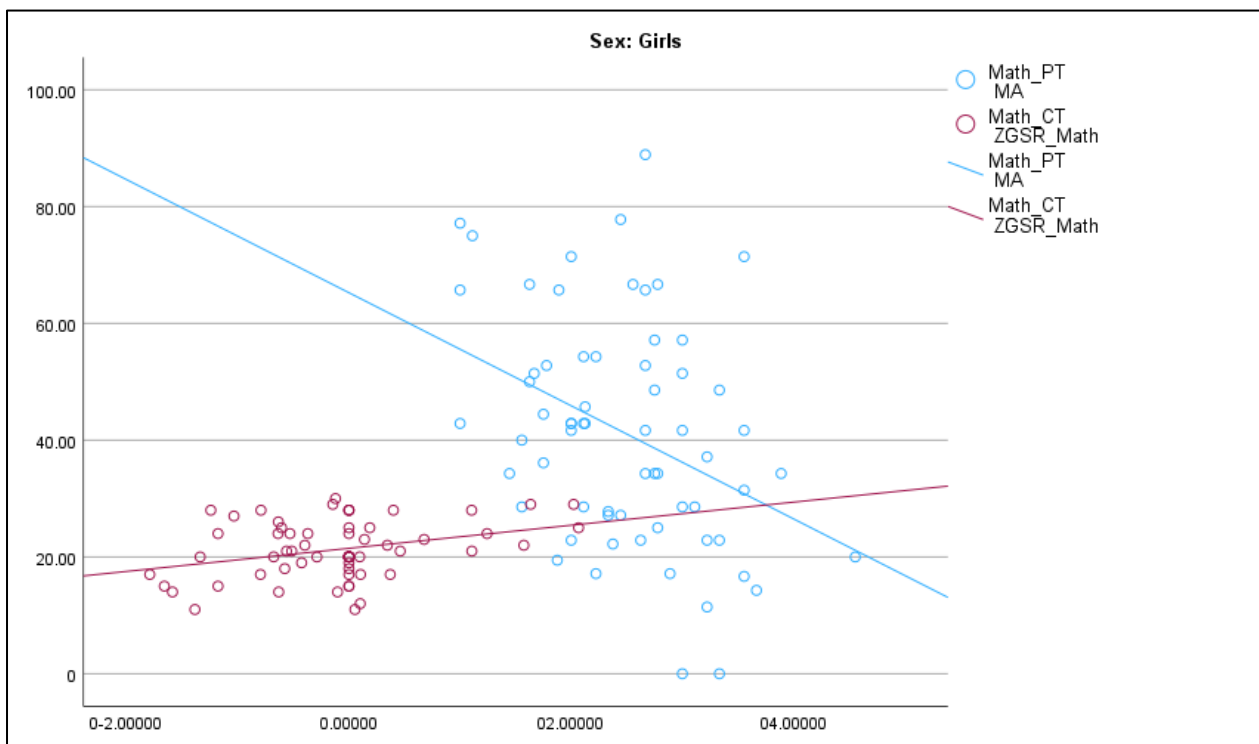
Figure 4. Effects of math anxiety on percentage scores and emotional reactivity (GSR) on completion times for girls. X-Axis = standardized scores; Y-Axis = percentage scores and completion times; blue circles: percentage scores (Math_PT) vs. math anxiety (MA); blue line: trend line for percentage scores (Math_PT); red circles: completion times (Math_CT) vs. emotional reactivity (ZGSR_Math); red line: trend line for completion times (Math_CT).

For boys, none of the predictors were significant for performance on the math task. For girls, math anxiety had a significant effect on percentage scored, and both emotional reactivity and subject preference had significant effects on completion times. For girls, lower math anxiety is significantly associated with higher math scores, and reduced emotional reactivity is significantly associated with shorter completion times on the math task. Subject preference does not significantly affect their math scores but girls who preferred math also had significantly shorter completion times.

3.2.3. Effects on Mental Rotation Performance

A multivariate analysis of covariance (MANCOVA) was conducted to examine the associations between the fixed factors gender, subject preference, and performance on the combined dependent variables accuracy and response time on the mental rotation task when controlling for covariates of emotional reactivity measured by GSR and math and spatial anxiety. The results of the multivariate tests, as indicated by Wilks' Lambda, showed no significant effect of emotional reactivity measured by GSR on mental rotation performance, *Wilks'* Λ = 0.96, $F(2, 121)$ = 2.70, p = 0.071, η^2 = 0.04. Neither math anxiety, *Wilks'* Λ = 0.96, $F(2, 121)$ = 2.64, p = 0.075, η^2 = 0.04, nor spatial anxiety, *Wilks'* Λ = 0.10, $F(2, 121)$ = 0.23, p = 0.798, η^2 = 0.004, had a significant effect. A significant effect of gender, *Wilks'* Λ = 0.94, $F(2, 121)$ = 3.73, p = 0.027, η^2 = 0.06, but no significant effect of subject preference, *Wilks'* Λ = 0.96, $F(4, 242)$ = 1.11, p = 0.350, η^2 = 0.02, was found. There was no significant interaction effect for gender and subject preference, *Wilks'* Λ = 0.98, $F(4, 242)$ = 0.71, p = 0.582, η^2 = 0.01 (see Table 8).

Table 8. Multivariate tests: effects of gender (sex) on performance on the mental rotation task.

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Sex	Wilks' Lambda	0.942	3.731	2.000	121.000	0.027	0.058

Design: intercept + ZGSR_MRT + MA + sex.

The results of multivariate tests further indicate that gender had a significant effect on accuracy and response time in the mental rotation task, while emotional reactivity measured by GSR and math anxiety both showed no significant effects.

A series of between-subjects effect tests were conducted to examine the influence of gender, emotional reactivity measured by GSR, math anxiety, and spatial anxiety on accuracy and response time on the mental rotation task. The analyses yielded the following results.

For accuracy, the between-subjects effects indicated the following results. The effect of emotional reactivity measured by GSR was not significant, $F(1, 122)$ = 0.27, p = 0.603, η^2 = 0.00, but the effect of math anxiety was significant, $F(1, 122)$ = 5.26, p = 0.023, η^2 = 0.04. The effects of spatial anxiety, $F(1, 122)$ = 0.310, p = 0.579, η^2 = 0.003, gender, $F(1, 122)$ = 0.02, p = 0.885, η^2 = 0.00, and subject preference, $F(2, 122)$ = 0.138, p = 0.871, η^2 = 0.00, on accuracy were not significant.

For response time, the between-subjects effects showed the following results. The effect of emotional reactivity measured by GSR was significant, $F(1, 122)$ = 5.43, p = 0.021, η^2 = 0.04. The effects of math anxiety, $F(1, 122)$ = 0.80, p = 0.373, η^2 = 0.01, and spatial anxiety, $F(1, 122)$ = 0.274, p = 0.602, η^2 = 0.00, were not significant. The effect of gender was significant, $F(1, 122)$ = 7.12, p = 0.009, η^2 = 0.06, but subject preference was not, $F(2, 122)$ = 2.06, p = 0.131, η^2 = 0.03 (see Table 9).

Table 9. Tests of between-subjects effects of emotional reactivity measured by GSR (ZGSR_MRT), math anxiety (MA), and gender (sex) on accuracy (Sqr_Acc) and response times (Sqr_RT) on the mental rotation task.

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
ZGSR_MRT	Sqr_RT	1.929	1	1.929	5.433	0.021	0.043
MA	Sqr_Acc	0.133	1	0.133	5.263	0.023	0.041
Sex	Sqr_RT	2.527	1	2.527	7.120	0.009	0.055

These results indicate that math anxiety has a significant effect on accuracy, while emotional reactivity measured by GSR and gender have significant effects on response

time. However, neither spatial anxiety, nor subject preference, nor their interaction have significant effects on mental rotation performance.

3.2.4. Gender-Specific Effects on Mental Rotation Performance

The results of a MANCOVA examining the gender-specific effects of emotional reactivity, math anxiety, and spatial anxiety, as well as subject preference on accuracy and response time on the mental rotation task, are presented below.

For boys:

Multivariate tests showed no significant effect of emotional reactivity measured by GSR on boys' mental rotation performance (*Wilks'* Λ = 0.95, $F(2, 59)$ = 1.57, p = 0.217, η^2 = 0.050). Neither math anxiety (*Wilks'* Λ = 0.94, $F(2, 59)$ = 1.74, p = 0.185, η^2 = 0.06), spatial anxiety (*Wilks'* Λ = 0.99, $F(2, 59)$ = 0.36, p = 0.699, η^2 = 0.01), nor subject preference (*Wilks'* Λ = 0.912, $F(4, 118)$ = 1.40, p = 0.239, η^2 = 0.04) show significant effects.

Tests of between-subjects effects showed no significant effects of emotional reactivity on boys' accuracy ($F(1, 60)$ = 0.61, p = 0.437, η^2 = 0.01). Neither math anxiety ($F(1, 60)$ = 3.35, p = 0.072, η^2 = 0.05), nor spatial anxiety ($F(1, 60)$ = 0.003, p = 0.954, η^2 = 0.00), nor subject preference ($F(2, 60)$ = 0.31, p = 0.734, η^2 = 0.01) showed significant effects.

For boys' response time, neither emotional reactivity measured by GSR ($F(1, 60)$ = 3.17, p = 0.080, η^2 = 0.05), nor math anxiety ($F(1, 60)$ = 0.08, p = 0.783, η^2 = 0.00), nor spatial anxiety, ($F(1, 60)$ = 0.60, p = 0.442, η^2 = 0.01) nor subject preference ($F(2, 60)$ = 1.91, p = 0.156, η^2 = 0.06) showed significant effects.

For girls:

Multivariate tests showed no significant effect of emotional reactivity measured by GSR on girls' mental rotation performance (*Wilks'* Λ = 0.94, $F(2, 58)$ = 1.70, p = 0.191, η^2 = 0.05). Neither math anxiety (*Wilks'* Λ = 0.97, $F(2, 58)$ = 0.974, p = 0.384, η^2 = 0.03), nor spatial anxiety (*Wilks'* Λ = 0.98, $F(2, 58)$ = 0.72, p = 0.490, η^2 = 0.02), nor subject preference (*Wilks'* Λ = 0.98, $F(4, 116)$ = 0.28, p = 0.893, η^2 = 0.01) showed significant effects.

Tests of between-subjects effects showed no significant effects of emotional reactivity on girls' accuracy, $F(1, 59)$ = 0.34, p = 0.562, η^2 = 0.01), and neither did math anxiety ($F(1, 59)$ = 1.37, p = 0.246, η^2 = 0.02), nor spatial anxiety ($F(1, 59)$ = 1.28, p = 0.262, η^2 = 0.02), nor subject preference ($F(2, 59)$ = 0.12, p = 0.891, η^2 = 0.00).

For girls' response times, neither emotional reactivity measured by GSR ($F(1, 59)$ = 2.57, p = 0.114, η^2 = 0.04), nor math anxiety ($F(1, 59)$ = 1.02, p = 0.316, η^2 = 0.02), nor spatial anxiety ($F(1, 59)$ = 0.03, p = 0.854, η^2 = 0.00), nor subject preference ($F(2, 59)$ = 0.478, p = 0.623, η^2 = 0.00) showed significant effects.

In summary, none of the factors showed significant effects on spatial performance for either gender.

4. Discussion

4.1. Overview of Findings

This study explored gender differences in mathematics and mental rotation performance, focusing on the roles of subjective anxiety (math and spatial anxiety), subject preference math versus German, and emotional reactivity as measured by the Galvanic Skin Response (GSR). The findings indicate that girls outperformed boys in percentage scores on the math task but took longer to complete it. This aligns with previous research suggesting that girls often perform just as well or better than boys in terms of scores but take more time, possibly due to higher levels of anxiety [35,40]. Lower math anxiety in girls was associated with higher percentage scores, highlighting the critical role of anxiety in cognitive performance [43,91]. However, no significant gender differences were found in mental rotation accuracy nor response times, suggesting that spatial ability might be more evenly distributed among primary school children when tasks are designed to mitigate stereotype threat [22–24,75].

4.2. Gender Differences in Math and Mental Rotation Performance, Subjective Anxiety, and Subject Preference

The MANOVA results demonstrated significant gender differences in math performance, with girls scoring higher but taking longer to complete the tasks. This may be due to the cognitive load imposed by higher anxiety levels, which can limit working memory capacity and slow down processing speed [34,35]. Girls' higher math anxiety, consistent with previous findings [77,91], likely contributed to their longer completion times. This anxiety may cause girls to adopt a more meticulous approach to problem solving, resulting in higher accuracy but slower performance [36,38]. Girls may have taken extra time to thoroughly understand and accurately solve each problem, leading to fewer mistakes and higher overall scores. This strategy contrasts with potentially quicker, but less precise, approaches and has been found to produce better results in children [38]. Boys were faster but scored lower on the math task, which might be explained by a tendency to prioritize speed over accuracy. This approach may reflect social validation sought by boys from peers regarding their math ability [36]. This involves rushing through problems to complete the task quickly, which can lead to more mistakes and lower overall scores. Additionally, boys might have been more confident in their initial responses and less likely to double-check their work, resulting in a higher rate of errors. This contrasts with a more thorough approach, where taking additional time to carefully solve each problem can lead to higher accuracy and better performance, a strategy also found to be effective on spatial tasks [37].

Interestingly, no significant gender differences were found in spatial anxiety, contrasting with some studies that report higher spatial anxiety in girls [41,55]. The superior math performance of girls in this study may have mitigated the expected differences in spatial anxiety, suggesting that mathematical self-efficacy might mitigate anxiety levels across related cognitive domains [13]. Furthermore, some studies show that girls excel in the Kangaroo Challenge, in which the math problems involved are known to correlate with spatial abilities [92,93]. Additionally, the order of the math task, which preceded the spatial task, demonstrates that practice on number line estimation, word problems, and missing term tasks, known to correlate with mental rotation, may also have reduced students' spatial anxiety [94,95]. From a statistical perspective, the spatial anxiety variable in the current study, although improved through log transformation, deviated marginally from normality. The potential effects of this are discussed further in the limitations.

Contrary to traditional findings that boys outperform girls on mental rotation tasks [96], this study found no significant gender differences in accuracy or response times on the nMRT. This result supports recent research suggesting that, when tasks are designed to minimize stereotype threat, gender differences in spatial abilities may diminish [20,23]. The use of gender-congruent and -neutral stimuli and the inclusion of a balanced task design likely contributed to this outcome [22,24]. As observed by the researchers, a preference expressed by children for the tablet-based task over the paper-and-pencil math task also highlights the potential impact of task format on engagement and performance.

The significant association between gender and subject preference, with girls less likely to choose math as their preferred subject, reflects enduring societal stereotypes and their influence on academic choices [12–14]. This preference for German over math among girls may stem from their higher math anxiety and the internalization of societal beliefs about gender and mathematical ability [9,10].

4.3. Association between Gender, Subject Preference, Emotional Reactivity Measured by GSR, Subjective Anxiety, and Performance on the Math and Mental Rotation Tasks

A unique aspect of the present study is that it explored the association between gender and subject preference, but also multiple related psychological as well as physiological factors and their effects on performance on math and mental rotation tasks among primary school children. The findings provide insights into the role of these factors in task performance, with notable differences observed across gender and task type.

4.3.1. Subject Preference

Subject preference significantly impacted math completion times but not scores. Students who preferred math completed tasks quicker, highlighting the importance of interest and engagement in efficient performance. This supports the hypothesis and aligns with research suggesting that subject interest, societal stereotypes, and intrinsic motivation significantly influence academic performance [12,14,97]. Additionally, this result provides further evidence for the reciprocal effects of engagement and interest on math self-concept, which in turn leads to greater achievement, and vice versa [98,99].

The finding that a preference for math over German had a protective effect on completion times for the math task, particularly among girls, is noteworthy. This suggests that girls who prefer math may experience reduced cognitive and emotional barriers when engaging with mathematical tasks. One possible explanation is that these girls may be less susceptible to stereotype threat. By identifying more strongly with math, these girls may be less affected by societal stereotypes that suggest girls are less capable in math, thereby mitigating the anxiety that can deplete working memory and hinder task performance [34]. Girls with a preference for math might view the challenge of the task as an opportunity to improve rather than as threats to their self-concept, which could further reduce performance anxiety and increase task efficiency. Additionally, a genuine interest and enjoyment in math could create a positive feedback loop where pleasure and fun derived from engaging in the subject counteract the stress typically associated with challenging tasks [97,100]. When girls enjoy math, they are likely to approach the task with greater confidence and motivation, which can enhance focus and efficiency, leading to faster completion times. Another potential factor could be the development of more effective problem-solving strategies. Girls who prefer math might have more practice and experience with mathematical thinking, allowing them to navigate the tasks more quickly eliciting less cognitive load [43]. These factors work together to create a more favorable psychological environment for girls, enabling them to perform math tasks more efficiently.

4.3.2. Emotional Reactivity

The results indicate that emotional reactivity, as measured by GSR, plays a nuanced role in mathematical and spatial task performance. For the math task, higher emotional reactivity was a significant predictor of longer completion times but did not significantly impact the percentage scored. This suggests that, while physiological arousal may slow down task completion, it does not necessarily detract from the precision of performance. Similarly, emotional reactivity was a significant predictor of response time on the mental rotation task, indicating that higher physiological arousal was associated with slower performance. These findings align with the Yerkes–Dodson law, which posits that higher levels of arousal can hamper performance on complex tasks by decreasing focus and cognitive resources [60,101]. This also provides evidence to support the theory that executive resources such as working memory are limited by emotional reactivity, leading to interference with cognitive focus, slowing down responses on both tasks [34,95].

A particular significance of this study, therefore, lies in its contribution to understanding the physiological underpinnings of how negative emotions and stereotype thinking impact cognitive performance. This is particularly relevant for females, who are often subjected to stereotypes about their abilities in mathematics and spatial tasks and related fields [102]. The findings highlight that the physiological arousal associated with stereotype threat not only affects how quickly tasks are completed but also suggests a potential mechanism by which stereotype threat undermines cognitive performance. When females face tasks in domains where they are stereotypically expected to perform poorly, the resulting emotional arousal may deplete the cognitive resources needed for optimal performance, thereby reinforcing the stereotype [34,103].

This study adds important empirical evidence to the literature on gender differences in cognitive performance by demonstrating that emotional reactivity, an often-overlooked factor, plays a significant role in these differences. By measuring physiological arousal

through GSR, this study provides concrete data showing that the emotional and physiological responses elicited by stereotype threat can slow performance, even if accuracy remains unaffected. This finding is critical for educators and policymakers aiming to address gender disparities in STEM fields, as it suggests that interventions designed to reduce stereotype threat and manage emotional reactivity could be key to improving performance outcomes for girls in mathematics and spatial tasks.

Additionally, the findings underscore the complex interplay between emotional states and cognitive performance, particularly under conditions of stereotype threat. This study not only broadens the understanding of how gender differences in performance manifest, but also points to the importance of considering physiological and emotional factors in educational interventions. By emphasizing the role of emotional reactivity, this research makes a vital contribution to knowledge on how gender disparities in STEM fields can be mitigated.

4.3.3. Subjective Anxiety

Math anxiety emerged as a significant predictor of performance on both the math and the mental rotation tasks, with higher levels of anxiety associated with lower math percentage scores and less mental rotation accuracy. The former is consistent with the existing literature highlighting the detrimental effects of math anxiety on math performance, likely due to increased cognitive load and interference with working memory [34,91,95]. Similar to research on the effects of psychological factors such as math self-concept on accuracy on a mental rotation task [13], the findings of the current study suggest that individuals with higher anxiety about math also tend to struggle more with spatial tasks like mental rotation.

Interestingly, in this study, math anxiety did not significantly affect math completion times nor mental rotation response time. This suggests that the primary impact of math anxiety may be on the precision and accuracy of math and spatial task performance rather than on the speed of processing. This distinction is important because it indicates that, while anxious individuals may not necessarily take longer to complete tasks, their heightened anxiety could lead to more errors or less accurate responses. This could be due to the fact that anxiety diverts cognitive resources away from the task at hand, leading to a focus on avoiding mistakes rather than on efficient problem solving [37,39].

In contrast, spatial anxiety did not emerge as a significant predictor of either percentage scores or completion times on the math task nor accuracy and response times on the spatial task. This finding is intriguing, particularly in light of the literature discussed in the introduction, which suggests that spatial anxiety can disrupt performance on tasks that require spatial reasoning, such as mental rotation [41,55,56]. One possible explanation for this result is that the math problems, although correlating with mental rotation, were mathematical in nature, thus diminishing the influence of spatial anxiety on performance outcomes. Additionally, the prior engagement with these kinds of math problems may have served as a warm-up that reduced the impact of spatial anxiety on subsequent spatial performance. The lack of a significant relationship between spatial anxiety and performance in this study could also reflect the influence of other moderating factors, such as task familiarity or the presence of stereotype threat, which were mentioned in the introduction as important considerations in understanding the effects of anxiety on cognitive performance.

Overall, these findings contribute to a nuanced understanding of the role of anxiety in educational contexts, highlighting that, while math anxiety clearly impairs performance accuracy, its effects on speed are less straightforward. This insight underscores the importance of considering the specific cognitive demands of tasks when assessing the effects of different types of anxiety on performance, as well as the potential for targeted interventions to mitigate these effects.

4.4. Gender-Specific Effects of Subject Preference, Emotional Reactivity, and Subjective Anxiety on Math and Mental Rotation Performance

This study revealed significant gender differences in how subject preferences, emotional reactivity, and subjective anxiety impact task performance. For boys, none of the predictors played a significant role in performance outcomes on the math tasks. These findings may indicate that boys' performance is less affected by anxiety and physiological arousal, or it may reflect different coping mechanisms or cognitive strategies employed by boys [54,73]. Additionally, boys may also be aware of stereotypes suggesting male superiority in math, which could increase pressure to excel in math tasks. This pressure to uphold the stereotype and compete with male peers may lead to faster performance but result in more errors [37]. Additionally, an element of bravado along with peer pressure may lead boys to under-report subjective anxiety [36,104]. On the mental rotation task, none of the factors were significantly associated with boys' spatial performance.

Girls, on the other hand, performed better on the math task when they reported less subjective anxiety, that is, they had better percentage scores. Additionally, their preference for math, and an ability to keep physiological arousal in check and regulate emotions, was associated with shorter completion times on the same task. None of the factors played a role in girls' performance on the mental rotation task. These findings suggest that girls' math performance is more sensitive to the effects of subject preference, anxiety, and physiological arousal. There are several possible explanations for this difference, which have been previously identified in the literature. Firstly, females are known to employ an analytic, piecemeal processing strategy when attempting to solve items on a task with spatial elements [102], such as the Kangaroo problems. This approach is known to be more time-consuming than the alternative holistic approach employed by males. It is well-known that having sufficient time is an important factor for females in performing spatial tasks [105]. Also, higher sensitivity to emotional and physiological states could be attributed to gender differences in emotion regulation strategies and stress responses [73,74]. Stereotypes associated with female math inferiority, in other words, stereotype threat, may have had a more powerful effect for girls' performance on the math task than on the less familiar mental rotation task [9,51,54]. Moreover, evidence from research has found that individuals with high levels of math anxiety and increased physiological arousal showed decreases in math accuracy compared to those who had lower math-related anxiety [48].

Girls who indicated a preference for math had significantly higher percentage scores and lower completion times on the math task. This preference for math reflects a characteristic of the sample and provides further evidence that reciprocal effects exist between engagement, practice, a more positive math attitude, and performance on both mathematical tasks [13,99,106].

As mentioned, neither boys' nor girls' mental rotation performance was significantly influenced by any of the factors measured in this study, including emotional reactivity, math anxiety, or spatial anxiety. This finding is noteworthy, particularly in light of the well-documented gender differences in mental rotation tasks, where boys have traditionally outperformed girls [96]. However, recent research suggests that, when the design of such tasks is adjusted to reduce gender bias and stereotype threat, these differences may diminish or even disappear [20,22–24]. By carefully selecting stimuli that were gender-congruent or gender-neutral and ensuring that the task was novel and free from culturally ingrained biases, this study likely reduced the activation of any stereotype-related concerns. This might explain why none of the measured factors, including those typically associated with stereotype vulnerability, such as math and spatial anxiety, had a significant impact on performance [34,41,51]. Additionally, the extended time limits may have allowed for more thoughtful and strategic processing, further reducing the influence of anxiety or emotional reactivity. The opportunity to approach the task without the pressure of a strict time constraint might have enabled both boys and girls to employ their optimal cognitive strategies, leading to similar levels of performance across genders. This outcome supports the idea that time constraints and task pressure are critical factors that can exacerbate or

mitigate gender differences in spatial tasks [35]. This highlights the importance of task design in educational assessments and suggests that gender differences in spatial abilities might be more context-dependent than previously assumed. Overall, the lack of significant effects on children's spatial performance in this study provides valuable insights into how stereotype threat can be mitigated through thoughtful task design, and it challenges the notion that gender differences in spatial abilities are innate or unchangeable.

4.5. Limitations and Future Research

Generally, there is a need to replicate this study in future research due to a number of limitations outlined in this section. The log-transformed spatial anxiety variable was approximately normally distributed, with minor deviations that did not significantly impact the analyses' validity. However, caution is advised regarding these deviations as they can lead to the incorrect interpretation of related findings.

Physiological measurements, such as skin conductance, inherently vary and can be skewed. Despite standardization and normality tests, some limitations persist due to individual differences in baseline levels and potential measurement artifacts [82]. Thus, while our findings offer valuable insights, further research is needed to confirm these results and explore their generalizability across different populations. Other physiological measurements such as pupillometry, eye-tracking measurement, or photoplethysmography (PPG) could capture autonomic nervous system activity during cognitive tasks, which may provide more insight into the effects of physiological arousal.

The math task used in this study, derived from the Kangaroo Challenge, differs from the standardized problems children typically encounter in the German primary school curriculum. Although designed to engage and encourage enjoyment of math [83], the unfamiliar and creative nature of these problems might have caused stress or led to rushing through the task to engage with the computerized spatial task.

Future studies could benefit from utilizing a computerized math task. Such a task, which also records the response times of individual items, would allow for a full investigation of the effects of gender, subject preferences, subjective anxiety, and physiological arousal on performance on more difficult tasks, for example. This could facilitate an analysis of whether gender differences in scores and processing times persist on more difficult math tasks. An analysis of the current paper-and-pencil data allowed for only a limited insight into the effects of physiological arousal and the math completion times as a whole rather than individual items.

A notable limitation is the reliance on quantitative data, which does not capture the nuanced experiences and thoughts of the children during the tasks. Including qualitative data could provide deeper insights into the cognitive and emotional processes children experience while performing math and mental rotation tasks. Mixed methods allow for the integration of both quantitative (physiological measures, test scores) and qualitative (e.g., interviews, observations) data, providing a more comprehensive understanding of the phenomena under study. This approach helps to capture the complexity of children's experiences with anxiety and cognitive tasks, offering insights that might be missed with a single-method approach.

The sample may be biased because the children who agreed to participate likely had more interest and enjoyment in math than the general school population. Furthermore, their parents may have consented for children with a more positive attitude towards math after discussing participation in this study with them. Additionally, the sample was taken from one region in a single European WEIRD (Western, Educated, Industrialized, Rich, and Democratic) state. Replicating this study with a more diverse socio-economic and cultural population may contribute to a higher generalizability of the findings to primary school contexts globally.

A factor which may have led to increased anxiety levels in students is testing by unfamiliar individuals which did not take place in their own classroom. Although they had access to their teachers at any stage, this may have caused some children to feel more anx-

ious. Moreover, the effects of teacher and educator attitudes and anxieties regarding math and spatial activities were not explored in this study, but could have valuable educational implications. Information about parent's socio-economic status and cultural background could also enhance findings in future research.

Future research should investigate the associations between anxiety and physiological arousal with larger, more diverse samples to confirm and elaborate on these findings. Longitudinal studies could offer insights into how these relationships evolve over time and across different educational stages. Additionally, examining the effectiveness of specific interventions such as mindfulness or relaxation in primary education aimed at reducing anxiety and managing physiological arousal could offer practical solutions for improving academic performance.

4.6. Educational Implications of This Study

The results of this study suggest several implications for teachers and educators. Given that girls demonstrate higher math performance but also higher anxiety in mathematics, interventions should focus on reducing math anxiety and providing support to build confidence. Encouraging a positive math identity in girls from a young age could help balance subject preferences and potentially increase the number of girls pursuing STEM fields in the future [97]. Furthermore, fostering interest and enthusiasm for math in girls can have significant benefits for their performance [100]. This could involve providing positive role models, offering engaging math activities, and addressing anxiety through targeted interventions [107]. Engaging and relevant math problems that connect to real-world applications can make the subject more relatable and interesting [92,93]. Interventions such as stress management training, mindfulness practices, and anxiety reduction techniques can help students manage physiological arousal and negative emotions, thereby enhancing performance. For boys, strategies to improve focus and attention during tasks may be more effective. Also, techniques such as positive reinforcement and encouraging a growth mindset can also be beneficial [51,108]. Educators could counteract gender stereotypes in math and science by promoting positive role models and creating an inclusive classroom environment. Encouraging all students to see themselves as capable mathematicians can improve self-concept and reduce anxiety [8,100]. Overall, these educational strategies could contribute to more balanced academic outcomes and encourage greater participation of girls in STEM disciplines.

Additionally, as spatial ability is critical in many STEM-related tasks, such as geometry, engineering design, architecture, and physics, the spill-over effects of math-related psychological and emotional factors such as math self-concept, math anxiety, and physiological arousal negatively impact spatial abilities. Students with high math anxiety and high emotional reactivity may struggle in these areas, even if their spatial reasoning skills would otherwise be strong. This could result in students avoiding STEM subjects or careers, perpetuating under-representation in fields that rely heavily on spatial reasoning. Curriculums could integrate more spatial reasoning activities in non-math contexts, which would allow students to develop spatial skills without the pressure of math tasks, potentially reducing the impact of math anxiety. For example, including more visual arts, design, and hands-on learning can improve spatial reasoning in less anxiety-inducing environments [109].

5. Conclusions

This study highlights significant gender differences in mathematics performance, subjective anxiety, subject preference, and emotional reactivity among primary school children. Girls outperformed boys in percentage scores on the math task but took longer to complete it, suggesting a more meticulous approach with higher anxiety levels playing a significant role. Boys, on the other hand, prioritized speed over precision, resulting in lower scores but faster completion times. Math anxiety significantly impacted students' performance, with higher anxiety associated with lower scores for both genders, though its effect was more pronounced in girls. Emotional reactivity, as measured by GSR, influenced

completion times for the math and the spatial tasks, indicating that higher physiological arousal can slow performance on both. Girls who preferred math performed better in this domain and completed the task faster, highlighting the importance of fostering a positive attitude towards the subject at this important educational stage in order to enhance performance. No significant gender differences were found in mental rotation accuracy nor response times, suggesting that spatial ability may be more evenly distributed among primary school children or that educational practices in Germany effectively mitigate these differences. These findings underscore the need for educational interventions designed to address the unique needs of boys and girls.

Furthermore, addressing societal influences and systemic issues such as gender stereotypes, poor attitudes towards mathematics, and gender inequity in STEM fields is vital for creating a more equitable and supportive environment for all students. Several interventions can be implemented; for example, training teachers in gender-sensitive practices and diverse teaching methods [110]; developing educational materials that promote gender equality and feature diverse STEM role models [111]; engaging parents and communities through workshops and resources to support positive attitudes towards math and STEM [112]; launching media campaigns to change societal perceptions and highlight successful women in STEM [113]; advocating for policies that promote gender equity in education, including training for educators and anti-discrimination policies [114]; and conducting research to continuously assess and improve these interventions. These steps aim to create a more inclusive and supportive environment, benefiting all students and fostering a culture of respect and equity in education.

Appendix A. Sample Problems from Math Task

Sample number line estimation problem for second-grade students.
Difficulty level: high.

C-2

Die Mädchen Lia und Mia und die Jungen Leo, Ole und Ali stehen in einer Reihe.

Lia steht zwischen Ole und Leo.
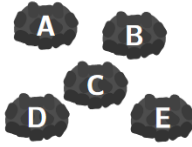Leo steht am Rand.
Ali hat einen Pulli mit einem Fußball.

Welches Kind ist Mia?

A	B	C	D	E

Translation from German: "The girls Lia and Mia and the boys Leo, Ole and Ali are standing in a row. Lia is standing between Ole and Leo. Leo is standing on the edge. Ali has a sweater with a soccer ball on it. Which child is Mia?"

Sample word problem for third-grade students.
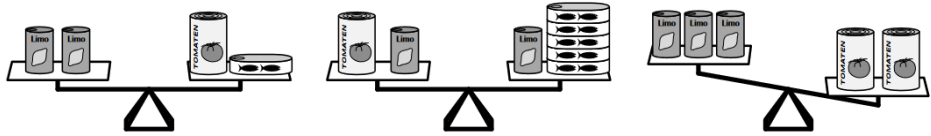Difficulty level: high.

C4 Egon, the mole, is working on his burrow. He already made 5 molehills, 2 of them today. Only one of the hills A, C and D was made today. C or E was made today, but not both. A and D were both made yesterday. Which 2 molehills were made today?

(**A**) A and E (**B**) B and C (**C**) B and D (**D**) C and E (**E**) B and E

Sample missing term problem for fourth-grade students.
Difficulty level: high.

C7 Using 3 types of cans, 2 scales were brought into balance. What has to be added to the left side of the third scale so that it is also balanced?
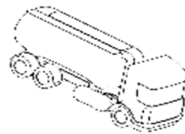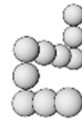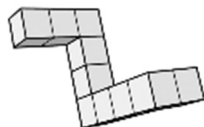
(A) (B) (C) (D) (E)

All problems are available in several languages from the German Kangaroo and Mini-Kangaroo competition website, but were presented to participants in German.

Appendix B. Examples of Stimuli Used in the Two Subtests of the Mental Rotation Test

MRT 1

Letter_g

Lorry_2D

Goggles_2D

Cubes_2D_#1

MRT 2

Hairbrush_3D

Cubes_#5_3D

Pellets_3D_#5

Polyhedron_3D_#1

Appendix C. Math versus German Survey

Children were asked to indicate whether they preferred math or German or preferred both subjects equally well by ticking one of the following boxes:

☐ Math
☐ German

I like both subjects the same.

Appendix D. Child Spatial Anxiety Questionnaire

Children rated their anxiety on a sliding 5-point emoji scale based on the following questions:

1. How do you feel being asked to say which direction is right or left?
2. How would you feel if your teacher asked you to build this house out of these blocks in 5 min? [Show child card with picture of a Lego house].

Sample picture shown to students:

3. How would you feel if you were given this problem: John is taller than Mary, and Mary is taller than Chris? Who is shorter, John or Chris?
4. How do you feel when you are asked to point to a certain place on a map, like this one? [Show card with image of a map of Germany].
5. How do you feel when your teacher asks you whether these shapes are rectangles and why? [Show child card with similar shapes].
6. How do you feel when you have to solve a maze like this in one minute? [Show child card with maze].
7. How do you feel if you are asked to measure something with a ruler?
8. How do you feel when a friend asks you how to get from school to your house?

The scale was translated from the original English version and presented to participants in German.

Appendix E. Modified Abbreviated Math Anxiety Scale (mAMAS)

Children rated their level of math anxiety on a scale of 1 to 5 based on the following statements:

1. Having to use the tables in the back of a math book
2. Thinking about a math test the day before you take it
3. Watching the teacher work out a math problem on the board
4. Taking a math test
5. Being given math homework with lots of difficult questions that you have to hand in the next day
6. Listening to the teacher talk for a long time in math
7. Listening to another child in your class explain a math problem
8. Finding out that you are going to have a surprise math test when you start your math lesson
9. Starting a new topic in math

The scale was translated by a native German speaker at the university and presented to participants in German.

References

1. Thomson, C. A Feminist Formulae?: What "Girl Math" Really Means. *The Glasgow Guardian*, 23 November 2023.
2. Satnoianu, A. More Gender Balance in STEAM? Not without More Empowerment for Girls! *Medium*, 9 March 2023.
3. Pelegrina, S.; Justicia-Galiano, M.J.; Martín-Puga, M.E.; Linares, R. Math Anxiety and Working Memory Updating: Difficulties in Retrieving Numerical Information From Working Memory. *Front. Psychol.* **2020**, *11*, 669. [CrossRef] [PubMed]
4. Pellizzoni, S.; Cargnelutti, E.; Cuder, A.; Passolunghi, M.C. The Interplay between Math Anxiety and Working Memory on Math Performance: A Longitudinal Study. *Ann. N. Y. Acad. Sci.* **2022**, *1510*, 132–144. [CrossRef] [PubMed]
5. Skagerlund, K.; Östergren, R.; Västfjäll, D.; Träff, U. How Does Mathematics Anxiety Impair Mathematical Abilities? Investigating the Link between Math Anxiety, Working Memory, and Number Processing. *PLoS ONE* **2019**, *14*, e0211283. [CrossRef] [PubMed]
6. Hyde, J.S.; Lindberg, S.M.; Linn, M.C.; Ellis, A.B.; Williams, C.C. Gender Similarities Characterize Math Performance. *Science* **2008**, *321*, 494–495. [CrossRef]
7. OECD. *PISA 2018 Results (Volume I): What Students Know and Can Do*; PISA; OECD: Paris, France, 2019; ISBN 978-92-64-46038-6.
8. Wolff, F. How Classmates' Gender Stereotypes Affect Students' Math Self-Concepts: A Multilevel Analysis. *Front. Psychol.* **2021**, *12*, 599199. [CrossRef]
9. Cvencek, D.; Meltzoff, A.N.; Greenwald, A.G. Math–Gender Stereotypes in Elementary School Children. *Child Dev.* **2011**, *82*, 766–779. [CrossRef]
10. Passolunghi, M.C.; Rueda Ferreira, T.I.; Tomasetto, C. Math–Gender Stereotypes and Math-Related Beliefs in Childhood and Early Adolescence. *Learn. Individ. Differ.* **2014**, *34*, 70–76. [CrossRef]
11. Cheryan, S.; Plaut, V.C.; Handron, C.; Hudson, L. The Stereotypical Computer Scientist: Gendered Media Representations as a Barrier to Inclusion for Women. *Sex Roles* **2013**, *69*, 58–71. [CrossRef]
12. Aeschlimann, B.; Herzog, W.; Makarova, E. Studienpräferenzen von Gymnasiastinnen und Gymnasiasten: Wer entscheidet sich aus welchen Gründen für ein MINT-Studium? *Swiss J. Educ. Res.* **2015**, *37*, 285–300. [CrossRef]
13. Lennon-Maslin, M.; Quaiser-Pohl, C.M.; Wickord, L.-C. Beyond Numbers: The Role of Mathematics Self-Concept and Spatial Anxiety: For Mental Rotation Performance and STEM Preferences. *Front. Educ.* **2024**, *9*, 1300598. [CrossRef]
14. Lesperance, K.; Munk, S.; Holzmeier, Y.; Braun, M.; Holzberger, D. *Geschlechterunterschiede im Bildungskontext: Von Wissenschaftlichen Studien zu Impulsen für die Unterrichtspraxis*; Pedocs: Münster, Germany, 2022; ISBN 978-3-8309-9534-0.
15. Buckley, J.; Seery, N.; Canty, D. A Heuristic Framework of Spatial Ability: A Review and Synthesis of Spatial Factor Literature to Support Its Translation into STEM Education. *Educ. Psychol. Rev.* **2018**, *30*, 947–972. [CrossRef]
16. Neuburger, S.; Ruthsatz, V.; Jansen, P.; Quaiser-Pohl, C. Can Girls Think Spatially? Influence of Implicit Gender Stereotype Activation and Rotational Axis on Fourth Graders' Mental-Rotation Performance. *Learn. Individ. Differ.* **2015**, *37*, 169–175. [CrossRef]
17. Vandenberg, S.G.; Kuse, A.R. Mental Rotations, a Group Test of Three-Dimensional Spatial Visualization. *Percept. Mot. Ski.* **1978**, *47*, 599–604. [CrossRef] [PubMed]
18. Wraga, M.; Duncan, L.; Jacobs, E.C.; Helt, M.; Church, J. Stereotype Susceptibility Narrows the Gender Gap in Imagined Self-Rotation Performance. *Psychon. Bull. Rev.* **2006**, *13*, 813–819. [CrossRef] [PubMed]
19. Kersh, J.E.; Casey, M.B.; Mercer Young, J. Research on Spatial Skills and Block Building in Boys and Girls: The Relationship to Later Mathematics Learning. In *Contemporary Perspectives on Mathematics in Early Childhood Education*; Saracho, O.N., Spodek, B., Eds.; Information Age Publishing Inc.: Charlotte, NC, USA, 2008; pp. 233–252.
20. Ruthsatz, V.; Rahe, M.; Schurmann, L.; Quaiser-Pohl, C. Girls' Stuff, Boys' Stuff and Mental Rotation: Fourth Graders Rotate Faster with Gender-Congruent Stimuli. *J. Cogn. Psychol.* **2019**, *31*, 225–239. [CrossRef]
21. Rahe, M.; Quaiser-Pohl, C. Cubes or Pellets in Mental-Rotation Tests: Effects on Gender Differences and on the Performance in a Subsequent Math Test. *Behav. Sci.* **2019**, *10*, 12. [CrossRef]
22. Ruthsatz, V.; Neuburger, S.; Jansen, P.; Quaiser-Pohl, C. Pellet Figures, the Feminine Answer to Cube Figures? Influence of Stimulus Features and Rotational Axis on the Mental-Rotation Performance of Fourth-Grade Boys and Girls. In *Spatial Cognition IX*; Freksa, C., Nebel, B., Hegarty, M., Barkowsky, T., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2014; Volume 8684, pp. 370–382; ISBN 978-3-319-11214-5.
23. Ruthsatz, V.; Neuburger, S.; Jansen, P.; Quaiser-Pohl, C. Cars or Dolls? Influence of the Stereotyped Nature of the Items on Children's Mental-Rotation Performance. *Learn. Individ. Differ.* **2015**, *43*, 75–82. [CrossRef]
24. Ruthsatz, V.; Saunders, M.; Lennon-Maslin, M.; Quaiser-Poh, C.M. Male? Female? Neutral! Using Novel Polyhedral Figures as Gender-Neutral Stimuli in a Mental-Rotation Test. *J. Learn. Individ. Differ.* **2024**; *under review*.
25. Atit, K.; Power, J.R.; Pigott, T.; Lee, J.; Geer, E.A.; Uttal, D.H.; Ganley, C.M.; Sorby, S.A. Examining the Relations between Spatial Skills and Mathematical Performance: A Meta-Analysis. *Psychon. Bull. Rev.* **2022**, *29*, 699–720. [CrossRef]
26. Gilligan, K.A.; Hodgkiss, A.; Thomas, M.S.C.; Farran, E.K. The Developmental Relations between Spatial Cognition and Mathematics in Primary School Children. *Dev. Sci.* **2019**, *22*, e12786. [CrossRef]
27. Mix, K. Why Are Spatial Skill and Mathematics Related? *Child Dev. Perspect.* **2019**, *13*, 121–126. [CrossRef]
28. Hawes, Z.; Moss, J.; Caswell, B.; Poliszczuk, D. Effects of Mental Rotation Training on Children's Spatial and Mathematics Performance: A Randomized Controlled Study. *Trends Neurosci. Educ.* **2015**, *3*, 60–68. [CrossRef]
29. Uttal, D.H.; Meadow, N.G.; Tipton, E.; Hand, L.L.; Alden, A.R.; Warren, C.; Newcombe, N.S. The Malleability of Spatial Skills: A Meta-Analysis of Training Studies. *Psychol. Bull.* **2013**, *139*, 352–402. [CrossRef] [PubMed]

30. Moe, A. Mental Rotation and Mathematics: Gender-Stereotyped Beliefs and Relationships in Primary School Children. *Learn. Individ. Differ.* **2018**, *61*, 172–180. [CrossRef]

31. Moe, A. Does Experience with Spatial School Subjects Favour Girls' Mental Rotation Performance? *Learn. Individ. Differ.* **2016**, *47*, 11–16. [CrossRef]

32. Moe, A.; Pazzaglia, F. Beyond Genetics in Mental Rotation Test Performance The Power of Effort Attribution. *Learn. Individ. Differ.* **2010**, *20*, 464–468. [CrossRef]

33. Neuburger, S.; Jansen, P.; Heil, M.; Quaiser-Pohl, C. A Threat in the Classroom: Gender Stereotype Activation and Mental-Rotation Performance in Elementary-School Children. *Z. Psychol.* **2012**, *220*, 61–69. [CrossRef]

34. Schmader, T.; Johns, M.; Forbes, C. An Integrated Process Model of Stereotype Threat Effects on Performance. *Psychol. Rev.* **2008**, *115*, 336–356. [CrossRef]

35. Goodmon, L.B.; Brown, K.; Edwards, L.; Hurley, K.; Hartzell, K.; Powell, T. Ready, Set, Rotate: The Relationship Between Working Memory Capacity and Mental Rotation Speed. *J. Sci. Psychol.* **2019**, 1–21.

36. Raabe, I.J.; Block, P. The Gendered Maths Confidence Gap, Social Influence and Social Integration. *Eur. Soc.* **2024**, 1–36. [CrossRef]

37. Liesefeld, H.; Fu, X.; Zimmer, H. Fast and Careless or Careful and Slow? Apparent Holistic Processing in Mental Rotation Is Explained by Speed-Accuracy Trade-Offs. *J. Exp. Psychology. Learn. Mem. Cogn.* **2014**, *41*, 1140–1151. [CrossRef]

38. Tancoš, M.; Chvojka, E.; Jabůrek, M.; Portešová, Š. Faster ≠ Smarter: Children with Higher Levels of Ability Take Longer to Give Incorrect Answers, Especially When the Task Matches Their Ability. *J. Intell.* **2023**, *11*, 63. [CrossRef] [PubMed]

39. Liesefeld, H.R.; Janczyk, M. Combining Speed and Accuracy to Control for Speed-Accuracy Trade-Offs(?). *Behav. Res.* **2019**, *51*, 40–60. [CrossRef] [PubMed]

40. Stoevenbelt, A.H.; Wicherts, J.M.; Flore, P.C.; Phillips, L.A.T.; Pietschnig, J.; Verschuere, B.; Voracek, M.; Schwabe, I. Are Speeded Tests Unfair? Modeling the Impact of Time Limits on the Gender Gap in Mathematics. *Educ. Psychol. Meas.* **2023**, *83*, 684–709. [CrossRef] [PubMed]

41. Ramirez, G.; Gunderson, E.A.; Levine, S.C.; Beilock, S.L. Spatial Anxiety Relates to Spatial Abilities as a Function of Working Memory in Children. *Q. J. Exp. Psychol.* **2012**, *65*, 474–487. [CrossRef]

42. Ramirez, G.; Shaw, S.T.; Maloney, E.A. Math Anxiety: Past Research, Promising Interventions, and a New Interpretation Framework. *Educ. Psychol.* **2018**, *53*, 145–164. [CrossRef]

43. Zhu, Z. Gender Differences in Mathematical Problem Solving Patterns: A Review of Literature. *Int. Educ. J.* **2007**, *8*, 187–203.

44. Kliziene, I.; Paskovske, A.; Cizauskas, G.; Augustiniene, A.; Simonaitiene, B.; Kubiliunas, R. The Impact of Achievements in Mathematics on Cognitive Ability in Primary School. *Brain Sci.* **2022**, *12*, 736. [CrossRef]

45. Goos, M.; Carreira, S.; Namukasa, I.K. Mathematics and Interdisciplinary STEM Education: Recent Developments and Future Directions. *ZDM Math. Educ.* **2023**, *55*, 1199–1217. [CrossRef]

46. Maass, K.; Geiger, V.; Ariza, M.R.; Goos, M. The Role of Mathematics in Interdisciplinary STEM Education. *ZDM Math. Educ.* **2019**, *51*, 869–884. [CrossRef]

47. Cohen, L.D.; Korem, N.; Rubinsten, O. Math Anxiety Is Related to Math Difficulties and Composed of Emotion Regulation and Anxiety Predisposition: A Network Analysis Study. *Brain Sci.* **2021**, *11*, 1609. [CrossRef]

48. Pizzie, R.G.; Kraemer, D.J.M. The Association Between Emotion Regulation, Physiological Arousal, and Performance in Math Anxiety. *Front. Psychol.* **2021**, *12*, 639448. [CrossRef] [PubMed]

49. Ashcraft, M.H. Math Anxiety: Personal, Educational, and Cognitive Consequences. *Curr. Dir. Psychol. Sci.* **2002**, *11*, 181–185. [CrossRef]

50. Hembree, R. The Nature, Effects, and Relief of Mathematics Anxiety. *J. Res. Math. Educ.* **1990**, *21*, 33–46. [CrossRef]

51. Dowker, A.; Sarkar, A.; Looi, C.Y. Mathematics Anxiety: What Have We Learned in 60 Years? *Front. Psychol.* **2016**, *7*, 508. [CrossRef]

52. Li, D.; Liew, J.; Raymond, D.; Hammond, T. Math Anxiety and Math Motivation in Online Learning during Stress: The Role of Fearful and Avoidance Temperament and Implications for STEM Education. *PLoS ONE* **2023**, *18*, e0292844. [CrossRef]

53. Devine, A.; Fawcett, K.; Szűcs, D.; Dowker, A. Gender Differences in Mathematics Anxiety and the Relation to Mathematics Performance While Controlling for Test Anxiety. *Behav. Brain Funct.* **2012**, *8*, 33. [CrossRef]

54. Van Mier, H.I.; Schleepen, T.M.J.; Van den Berg, F.C.G. Gender Differences Regarding the Impact of Math Anxiety on Arithmetic Performance in Second and Fourth Graders. *Front. Psychol.* **2019**, *9*, 2690. [CrossRef]

55. Alvarez-Vargas, D.; Abad, C.; Pruden, S.M. Spatial Anxiety Mediates the Sex Difference in Adult Mental Rotation Test Performance. *Cogn. Res.* **2020**, *5*, 31. [CrossRef]

56. Campbell, S.; Collaer, M. Stereotype Threat and Gender Differences in Performance on a Novel Visuospatial Task. *Psychol. Women Q.* **2009**, *33*, 437–444. [CrossRef]

57. Arrighi, L.; Hausmann, M. Spatial Anxiety and Self-Confidence Mediate Sex/Gender Differences in Mental Rotation. *Learn. Mem.* **2022**, *29*, 312–320. [CrossRef]

58. Rahe, M.; Quaiser-Pohl, C. Can (Perceived) Mental-Rotation Performance Mediate Gender Differences in Math Anxiety in Adolescents and Young Adults? *Math. Ed. Res. J.* **2021**, *35*, 255–279. [CrossRef]

59. Gibeau, R.-M.; Maloney, E.A.; Béland, S.; Lalande, D.; Cantinotti, M.; Williot, A.; Chanquoy, L.; Simon, J.; Boislard-Pépin, M.-A.; Cousineau, D. The Correlates of Statistics Anxiety: Relationships With Spatial Anxiety, Mathematics Anxiety and Gender. *J. Numer. Cogn.* **2023**, *9*, 16–43. [CrossRef]

60. Stokes, A.F.; Kite, K. *Flight Stress: Stress, Fatigue and Performance in Aviation*; Routledge: London, UK, 2017; ISBN 978-1-315-25520-0.

61. Lazarus, R.S.; Folkman, S. *Stress, Appraisal, and Coping*, 1st ed.; Springer: New York, NY, USA, 1984; ISBN 978-0-8261-4191-0.

62. Hunt, T.E.; Bhardwa, J.; Sheffield, D. Mental Arithmetic Performance, Physiological Reactivity and Mathematics Anxiety amongst U.K. Primary School Children. *Learn. Individ. Differ.* **2017**, *57*, 129–132. [CrossRef]

63. Nourbakhsh, N.; Wang, Y.; Chen, F.; Calvo, R.A. Using Galvanic Skin Response for Cognitive Load Measurement in Arithmetic and Reading Tasks. In Proceedings of the 24th Australian Computer-Human Interaction Conference, Association for Computing Machinery, New York, NY, USA, 26 November 2012; pp. 420–423.

64. Christopoulos, G.I.; Uy, M.A.; Yap, W.J. The Body and the Brain: Measuring Skin Conductance Responses to Understand the Emotional Experience. *Organ. Res. Methods* **2019**, *22*, 394–420. [CrossRef]

65. Fowles, D.C. The Measurement of Electrodermal Activity in Children. In *Developmental Psychophysiology: Theory, Systems, and Methods*; Schmidt, L.A., Segalowitz, S.J., Eds.; Cambridge University Press: Cambridge, UK, 2007; pp. 286–316; ISBN 978-0-521-82106-3.

66. Horvers, A.; Tombeng, N.; Bosse, T.; Lazonder, A.W.; Molenaar, I. Detecting Emotions through Electrodermal Activity in Learning Contexts: A Systematic Review. *Sensors* **2021**, *21*, 7869. [CrossRef]

67. Najafpour, E.; Asl-Aminabadi, N.; Nuroloyuni, S.; Jamali, Z.; Shirazi, S. Can Galvanic Skin Conductance Be Used as an Objective Indicator of Children's Anxiety in the Dental Setting? *J. Clin. Exp. Dent.* **2017**, *9*, e377–e383. [CrossRef]

68. Geršak, V.; Vitulić, H.S.; Prosen, S.; Starc, G.; Humar, I.; Geršak, G. Use of Wearable Devices to Study Activity of Children in Classroom; Case Study—Learning Geometry Using Movement. *Comput. Commun.* **2020**, *150*, 581–588. [CrossRef]

69. Thammasan, N.; Stuldreher, I.V.; Schreuders, E.; Giletta, M.; Brouwer, A.-M. A Usability Study of Physiological Measurement in School Using Wearable Sensors. *Sensors* **2020**, *20*, 5380. [CrossRef]

70. Pizzie, R.G.; Kraemer, D.J.M. Avoiding Math on a Rapid Timescale: Emotional Responsivity and Anxious Attention in Math Anxiety. *Brain Cogn.* **2017**, *118*, 100–107. [CrossRef]

71. Tyng, C.M.; Amin, H.U.; Saad, M.N.M.; Malik, A.S. The Influences of Emotion on Learning and Memory. *Front. Psychol.* **2017**, *8*, 1454. [CrossRef]

72. Ng, C.; Chen, Y.; Wu, C.; Chang, T. Evaluation of Math Anxiety and Its Remediation through a Digital Training Program in Mathematics for First and Second Graders. *Brain Behav.* **2022**, *12*, e2557. [CrossRef] [PubMed]

73. Fladung, A.-K.; Kiefer, M. Keep Calm! Gender Differences in Mental Rotation Performance Are Modulated by Habitual Expressive Suppression. *Psychol. Res.* **2016**, *80*, 985–996. [CrossRef] [PubMed]

74. Nolen-Hoeksema, S.; Aldao, A. Gender and Age Differences in Emotion Regulation Strategies and Their Relationship to Depressive Symptoms. *Personal. Individ. Differ.* **2011**, *51*, 704–708. [CrossRef]

75. Lennon-Maslin, M.; Quaiser-Pohl, C.M.; Ruthsatz, V.; Saunders, M. Under My Skin: Reducing Bias in STEM through New Approaches to Assessment of Spatial Abilities Considering the Role of Emotional Regulation. *Soc. Sci.* **2023**, *12*, 356. [CrossRef]

76. Murphy, M.C.; Steele, C.M.; Gross, J.J. Signaling Threat: How Situational Cues Affect Women in Math, Science, and Engineering Settings. *Psychol. Sci.* **2007**, *18*, 879–885. [CrossRef]

77. Szczygiel, M. Gender, General Anxiety, Math Anxiety and Math Achievement in Early School-Age Children. *Issues Educ. Res.* **2020**, *30*, 1126–1142.

78. Lauer, J.E.; Esposito, A.G.; Bauer, P.J. Domain-Specific Anxiety Relates to Children's Math and Spatial Performance. *Dev. Psychol.* **2018**, *54*, 2126–2138. [CrossRef]

79. Ceci, S.J.; Ginther, D.K.; Kahn, S.; Williams, W.M. Women in Academic Science: A Changing Landscape. *Psychol. Sci. Public Interest* **2014**, *15*, 75–141. [CrossRef]

80. Makarova, E.; Aeschlimann, B.; Herzog, W. The Gender Gap in STEM Fields: The Impact of the Gender Stereotype of Math and Science on Secondary Students' Career Aspirations. *Front. Educ.* **2019**, *4*, 60. [CrossRef]

81. Rahe, M.; Schürmann, L.; Jansen, P. Self-Concept Explains Gender Differences in Mental Rotation Performance after Stereotype Activation. *Front. Psychol.* **2023**, *14*, 1168267. [CrossRef]

82. Braithwaite, J.; Watson, D.; Jones, R.; Rowe, M. A Guide for Analysing Electrodermal Activity & Skin Conductance Responses for Psychological Experiments. *Psychophysiology* **2013**, *49*, 1017–1034.

83. Akveld, M.; Caceres-Duque, L.F.; Nieto Said, J.H.; Sánchez Lamoneda, R. The Math Kangaroo Competition. *Espac. Math.* **2020**, *1*, 74–91. [CrossRef]

84. e.V. Mathematikwettbewerb Känguru Känguru der Mathematik e.V. | Startseite. Available online: https://www.mathe-kaenguru.de/ (accessed on 4 June 2024).

85. Quaiser-Pohl, C. The Mental Cutting Test "Schnitte" and the Picture Rotation Test-Two New Measures to Assess Spatial Ability. *Int. J. Test.* **2003**, *3*, 219–231. [CrossRef]

86. Ehm, J.-H. *Akademisches Selbstkonzept im Grundschulalter. Entwicklungsanalyse Dimensionaler Vergleiche und Exploration Differenzieller Unterschiede*; Pedocs: Münster, Germany, 2014.

87. Hopko, D.R.; Mahadevan, R.; Bare, R.L.; Hunt, M.K. The Abbreviated Math Anxiety Scale (AMAS): Construction, Validity, and Reliability. *Assessment* **2003**, *10*, 178–182. [CrossRef]

88. Carey, E.; Hill, F.; Devine, A.; Szűcs, D. The Modified Abbreviated Math Anxiety Scale: A Valid and Reliable Instrument for Use with Children. *Front. Psychol.* **2017**, *8*, 11. [CrossRef]

89. Tabachnick, B.G.; Fidell, L.S. *Using Multivariate Statistics: Pearson New International Edition*, 6th ed.; Microsoft Press: Harlow, UK, 2013; ISBN 978-1-292-02131-7.

90. Society for Psychophysiological Research Ad Hoc Committee on Electrodermal Measures. Publication Recommendations for Electrodermal Measurements. *Psychophysiology* **2012**, *49*, 1017–1034. [CrossRef]

91. Ashcraft, M.H.; Moore, A.M. Mathematics Anxiety and the Affective Drop in Performance. *J. Psychoeduc. Assess.* **2009**, *27*, 197–205. [CrossRef]

92. Applebaum, M. Spatial Abilities as a Predictor to Success in the Kangaroo Contest. *J. Math. Syst. Sci.* **2017**, *7*, 154–163. [CrossRef]

93. Applebaum, M. Gender Issues in Solving Problems in the Kangaroo Contest. *Mediterr. J. Res. Math. Educ.* **2019**, *16*, 19–31.

94. Siegler, R.S.; Opfer, J.E. The Development of Numerical Estimation: Evidence for Multiple Representations of Numerical Quantity. *Psychol. Sci.* **2003**, *14*, 237–250. [CrossRef]

95. Yang, X.; Yu, X. The Relationship between Mental Rotation and Arithmetic: Do Number Line Estimation, Working Memory, or Place-value Concept Matter? *Br. J. Educ. Psychol.* **2021**, *91*, 793–810. [CrossRef]

96. Linn, M.C.; Petersen, A.C. Emergence and Characterization of Sex Differences in Spatial Ability: A Meta-Analysis. *Child Dev.* **1985**, *56*, 1479–1498. [CrossRef] [PubMed]

97. Cvencek, D.; Paz-Albo, J.; Master, A.; Herranz Llácer, C.V.; Hervás-Escobar, A.; Meltzoff, A.N. Math Is for Me: A Field Intervention to Strengthen Math Self-Concepts in Spanish-Speaking 3rd Grade Children. *Front. Psychol.* **2020**, *11*, 593995. [CrossRef] [PubMed]

98. Gunderson, E.; Park, D.; Maloney, E.; Beilock, S.; Levine, S. Reciprocal Relations among Motivational Frameworks, Math Anxiety, and Math Achievement in Early Elementary School. *J. Cogn. Dev.* **2018**, *19*, 21–46. [CrossRef]

99. Marsh, H.W.; Craven, R.G. Reciprocal Effects of Self-Concept and Performance From a Multidimensional Perspective: Beyond Seductive Pleasure and Unidimensional Perspectives. *Perspect. Psychol. Sci.* **2006**, *1*, 133–163. [CrossRef] [PubMed]

100. Boaler, J. *The Elephant in the Classroom: Helping Children Learn and Love Maths*; Souvenir Press: London, UK, 2015; ISBN 978-0-285-64319-2.

101. Teigen, K.H. Yerkes-Dodson: A Law for All Seasons. *Theory Psychol.* **1994**, *4*, 525–547. [CrossRef]

102. Bartlett, K.A.; Camba, J.D. Gender Differences in Spatial Ability: A Critical Review. *Educ. Psychol. Rev.* **2023**, *35*, 8. [CrossRef]

103. Johns, M.; Inzlicht, M.; Schmader, T. Stereotype Threat and Executive Resource Depletion: Examining the Influence of Emotion Regulation. *J. Exp. Psychol. Gen.* **2008**, *137*, 691–705. [CrossRef]

104. Heyder, A.; Kessels, U. Boys Don't Work? On the Psychological Benefits of Showing Low Effort in High School. *Sex Roles* **2017**, *77*, 72–85. [CrossRef]

105. Taragin, D.; Tzuriel, D.; Vakil, E. Mental Rotation: The Effects of Processing Strategy, Gender and Task Characteristics on Children's Accuracy, Reaction Time and Eye Movements' Pattern. *J. Eye Mov. Res.* **2019**, *12*. [CrossRef]

106. Sella, F.; Cohen Kadosh, R. What Expertise Can Tell About Mathematical Learning and Cognition. *Mind Brain Educ.* **2018**, *12*, 186–192. [CrossRef]

107. Bauer, R.; Jansen, P. A Short Mindfulness Induction Might Increase Women's Mental Rotation Performance. *Conscious. Cogn.* **2024**, *123*, 103721. [CrossRef] [PubMed]

108. Živković, M.; Pellizzoni, S.; Doz, E.; Cuder, A.; Mammarella, I.; Passolunghi, M.C. Math Self-Efficacy or Anxiety? The Role of Emotional and Motivational Contribution in Math Performance. *Soc. Psychol. Educ.* **2023**, *26*, 579–601. [CrossRef]

109. Zhu, C.; Klapwijk, R. The Spatial Aspect of Designing: Opportunities, Challenges, and Conjectures on Engaging Pupils in Spatial Thinking Through Design Education. In Proceedings of the Spatial Cognition XIII, Dublin, Ireland, 25–28 June 2024; Živković, M., Buckley, J., Pagkratidou, M., Duffy, G., Eds.; Springer Nature: Cham, Switzerland, 2024; pp. 97–113.

110. Gamarra, E.; Tenbrink, T.; Mills, D. Gender in Teacher-Student Interactions: Another Factor in Spatial Ability Development and STEM Affiliation. In Proceedings of the Spatial Cognition XIII, Dublin, Ireland, 25–28 June 2024; Živković, M., Buckley, J., Pagkratidou, M., Duffy, G., Eds.; Springer Nature: Cham, Switzerland, 2024; pp. 51–65.

111. Cruz Neri, N.; Schwenzer, F.; Sprenger, S.; Retelsdorf, J. Geschlechterdarstellungen in Geographie-Schulbüchern: Von mangelnder Repräsentation weiblicher Figuren und Reproduktion von Geschlechterstereotypen. *Z. Erzieh.* **2024**, *27*, 977–994. [CrossRef]

112. Olive, K.; Tang, X.; Loukomies, A.; Juuti, K.; Salmela-Aro, K. Gendered Difference in Motivational Profiles, Achievement, and STEM Aspiration of Elementary School Students. *Front. Psychol.* **2022**, *13*, 954325. [CrossRef]

113. Alkhammash, R. 'It Is Time to Operate Like a Woman': Representation of Women in STEM Fields in Social Media: A Corpus-Based Study. *Int. J. Engl. Linguist.* **2019**, *9*, 217–226. [CrossRef]

114. Global Education Monitoring Report Team. *Global Education Monitoring Report 2020: Gender Report, A New Generation: 25 Years of Efforts for Gender Equality in Education—UNESCO Digital Library*; United Nations Educational, Scientific and Cultural Organization: Paris, France, 2020; ISBN 978-92-3-100411-7.