Bridging Gaps!

Spatial Ability Development in the Primary School Years and Its Relation to Affective and Academic Measures in STEM Learning in Boys and Girls.

Kumulative Dissertation

zur Erlangung des akademischen Grades

Doktor der Philosophie

Institut für Psychologie

Fachbereich 1: Bildungswissenschaften

Universität Koblenz

Erstgutachterin: Prof. Dr. Claudia Quaiser-Pohl

Zweitgutachterin: apl. Prof. Dr. Martina Endepohls-Ulpe

vorgelegt von Michelle Lennon-Maslin, MSc.

am 12.09.2024

Table of Contents

Abstract: ..... **v**

I. Acknowledgements ..... **vi**

II. Publication Overview ..... **vii**

i. Submitted Publications vii

ii. Conference and Congress Contributions vii

III. Summary of Research ..... **viii**

1 Chapter 1: Introduction **1**

1.1 Background and Context 1

1.3 Research Questions 3

1.3.1 Emotional Reactivity and Performance: 3

1.3.2 Item Difficulty and Stereotype Effects: 3

1.3.3 Gender and Age Differences: 3

1.3.4 Predictive Factors: 4

1.4 Significance of the Research 4

2 Chapter 2: Literature Review **4**

2.1 Gendered Career Choices and STEM Engagement 5

2.2 Spatial Ability and Mental Rotation (MR) 5

2.3 Emotional Reactivity and Cognitive Performance 6

2.4 Computer Adaptive Testing (CAT) in MR 7

2.5 Gender Differences in Maths Self-Concept and STEM Preferences 7

2.6 Developmental Aspects and Childhood Brain Development 8

2.7 Physiological Measures of Emotional Arousal 9

2.8 Stereotype Threat and Working Memory 10

2.9 Gaps in the Literature 11

2.9.1 Gender Disparities in Spatial and Mathematical Abilities: 11

2.9.2 Impact of Emotional Arousal on Cognitive Performance: 12

2.9.3 Developmental Differences in Cognitive and Emotional Responses: 12

2.9.4 Effectiveness of Adaptive Testing: 12

2.9.5 Interplay Between Self-Concept, Anxiety, and Performance: 12

2.9.6 Influence of Societal Stereotypes on STEM Engagement: 13

2.9.7 Methodological Advances in Measuring Cognition and Emotion: 13

3 Chapter 3: Methodology **13**

3.1 Research Design 13

3.2 Participants 14

3.3 Instruments Used to Measure Constructs 14

3.3.1 Cognitive Tasks 14
3.3.2 Self-Report Questionnaires 15
3.3.3 Physiological Measure of Galvanic Skin Response (GSR) 16
3.4 Data Analysis 17
3.4.1 Statistical Techniques: 17
3.5 Ethical Considerations 18
4 Chapter 4: Summary of Methodology and Results of the Submitted Publications .. 18
4.1 Publication 1 19
4.1.1 Methodology: 19
4.1.2 Results: 20
4.2 Publication 2 21
4.2.1 Adjustments to the Research Design Based on Pilot Study Outcomes 21
4.2.2 Methodology: 22
4.2.3 Results: 23
4.3 Publication 3 24
4.3.1 Adjustments to the Research Design Based on Outcomes of Study 2 24
4.3.2 Methodology: 24
4.3.3 Results: 25
5 Chapter 5: Synthesis and General Discussion 29
5.1 Integration of Findings 29
5.1.1 Emotional Reactivity and Performance: 29
5.1.2 Item Difficulty and Stereotype Effects 30
5.1.3 Gender and Age Differences 30
5.1.4 Predictive Factors 32
5.2 Theoretical Contributions 33
5.3 Limitations and Future Research 35
6 Chapter 6: Conclusion 36
6.1 Summary of Key Findings 36
6.2 Overall Contributions 36
6.3 Conclusion 37
7 References 38
8 Appendices 56
A.1 Measures of Maths and Spatial Constructs 56
A.2 Legal Statement 59
A.3 Publications 60

List of Figures

Figure 1 *Under-Representation of Women Among New Entrants in STEM Fields in Higher Education (OECD, 2017).* 6

Figure 2 *Structural Changes During Early Adolescence in Brain Regions in the Superior Parietal Cortex Support Improved Spatial Manipulation (Modroño et al., 2018).* 9

Figure 3 *Integrated Model of Stereotype Threat Effects on Cognitive Performance (Schmader et al., 2008).* 11

Figure 4 *Flowchart for Research Questions and Progression of the Three Studies within the Research Project.* 19

Figure 5 *Summary of Commonalities and Differences in Research Design across Three Studies* 27

Abstract:

Gender disparities in STEM (Science, Technology, Engineering, and Mathematics) fields remain a significant challenge, with women often underrepresented. Spatial abilities, particularly mental rotation (MR), are crucial for success in STEM, yet significant gender differences in these skills persist. This research aims to explore the factors contributing to these differences, focusing on emotional reactivity, self-concept, anxiety, and their impact on performance in mathematical and spatial tasks among primary school children.

This research synthesizes findings from three related studies involving *N*=303 primary school students, consisting of 146 girls and 155 boys with a *mean age*=8.70 (*SD*=1.11) years. Data were collected through standardized questionnaires assessing self-concept, spatial and maths anxiety, and preferences for STEM subjects. Cognitive performance was evaluated using a computerized, novel Mental Rotation Task (nMRT) incorporating gender-congruent and neutral stimuli and various maths tasks correlating with mental rotation. Physiological responses were measured using galvanic skin response (GSR) to assess the impact of emotional reactivity on task performance. All data were collected in the classroom environment to increase ecological validity and generalizability of findings.

Across studies, girls demonstrated higher maths and spatial anxiety, lower maths self-concept, and a lower preference for maths as a STEM subject compared to boys. These factors were significantly associated with performance differences in both maths and MR tasks. Higher emotional reactivity, as evidenced by GSR, and increased response time were associated with better scores on difficult items, that is, abstract stimuli rotated in-depth. Emotional reactivity also affected maths task completion times, with girls demonstrating lower physiological arousal linked to shorter processing time. Gender, subject preference, math self-concept and anxiety levels emerged as significant predictors of task performance on both maths and spatial tasks.

The results underscore the influence of self-concept, anxiety and physiological responses on cognitive performance, highlighting significant gender differences. Girls demonstrated higher subjective anxiety and physiological arousal during maths tasks. However, in the same group, lower emotional reactivity and maths anxiety served as protective influences, leading to improved scores and shorter completion times. Moreover, girls and tweens demonstrated lower maths self-concept and preference for maths, indicating that stereotype effects are already impacting their interest during primary school. These findings suggest that psychological factors play a crucial role in learning outcomes, particularly in STEM subjects.

This integrated research contributes to a deeper understanding of how psychological factors such as self-concept, subjective anxiety but also physiological arousal and subject preferences affect mathematical and spatial performance in primary school children. The findings have practical implications for educators and policymakers, advocating for the development of strategies to enhance self-concept, manage anxiety and support emotional regulation, particularly in girls, fostering a supportive learning environment that mitigates the impact of stereotype threat. Enhanced self-efficacy and reduced anxiety thereby increase the likelihood of their engagement with maths, subsequently improving their performance and expanding their future career options in STEM fields.

Keywords: Gender Differences; Self-Concept; Anxiety; Spatial Skills; STEM Subjects; Emotional Reactivity; Childhood Development; Primary Education**.**

I. Acknowledgements

I would like to express my deepest gratitude to those who have supported me throughout the journey of completing this research. First and foremost, to my lovely hubba, Paul Maslin, whose unwavering love and encouragement sustained me through the most challenging moments. Your belief in me has been a constant source of strength. I am immensely grateful to my supervisor, Prof. Dr. Claudia Quaiser-Pohl, for your invaluable guidance, wisdom, and insight. Your dedication to my academic and personal growth has been instrumental in shaping this work, and I very much appreciate your mentorship throughout this process. A heartfelt thank you to my research network, SellSTEM (Spatially Enhanced Learning linked to Science, Technology, Engineering and Mathematics), in particular my fellow PhD colleagues, for fostering an inspiring, collaborative and enjoyable environment providing rich opportunities for intellectual and personal exchange (our coffee hour will be missed). The support and feedback from this community have been vital for refining my ideas. I look forward to following your journey and remain in awe of your amazing talent. I would also like to extend my sincere appreciation to the students, principals and teachers of the primary schools who participated in this research. Your cooperation and enthusiasm were essential to the success of this project, and I am truly grateful for your willingness to engage in the study. Without your valuable contributions, this research would not have been possible. Thank you as well to our team, for your willingness to share your knowledge and experience from which I have benefitted greatly. I am truly blessed to have found two wonderful research assistants and I thank you for your hard work, dedication, and patience throughout the project. Your assistance in data collection, analysis, and relationship management made the completion of this project possible. Finally, to my family and friends, thank you for your constant support, understanding, and encouragement throughout this journey. This dissertation would not have been possible without the support of all these incredible people, and for that, I am very grateful.

II. Publication Overview

i. Submitted Publications

Lennon-Maslin, M., Quaiser-Pohl, C. M., Ruthsatz, V., & Saunders, M. (2023). Under My Skin: Reducing Bias in STEM through New Approaches to Assessment of Spatial Abilities Considering the Role of Emotional Regulation. *Social Sciences*, 12(6), Article 6. https://doi.org/10.3390/socsci12060356

Lennon-Maslin, M., Quaiser-Pohl, C. M., & Wickord, L.-C. (2024). Beyond Numbers: The Role of Mathematics Self-Concept and Spatial Anxiety in Mental Rotation Performance and STEM Preferences in Primary Education. *Frontiers in Education*, 9. https://doi.org/10.3389/feduc.2024.1300598

Lennon-Maslin, M., & Quaiser-Pohl, C. M. (Accepted for Publication on 11th Sept 2024). "It's Different for Girls!" The Role of Anxiety, Physiological Arousal and Subject Preferences in Primary School Children's Math and Mental-Rotation Performance. *Behavioral Sciences.*

ii. Conference and Congress Contributions

Lennon-Maslin, M., & Quaiser-Pohl, C. (2024). Mindset Matters: The Role of Mathematics Self-concept and Age in Mental Rotation Performance Among Primary School Children. In M. Živković, J. Buckley, M. Pagkratidou, & G. Duffy (Eds.), Spatial Cognition XIII (pp. 19–31). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-63115-

Lennon-Maslin, M., & Quaiser-Pohl, C. M. (2024). Be an Enchantress of Numbers! How Mathematics Self-Concept and Spatial Anxiety Influence Mental Rotation and STEM Choices as Early as Primary School. *Network Gender & STEM Conference 2024, Heidelberg, 18th to 20th July 2024.*

Lennon-Maslin, M., & Quaiser-Pohl, C. M. (2022). Closing the gap: Reducing gender bias in STEM through the development of new approaches to assessment of spatial abilities. *8th Conference of the International Association for Computerized Adaptive Testing (IACAT) 2022, Frankfurt, 20th to 22nd Sept. 2022.*

III. Summary of Research

This research conducted as part of the Marie Sklodowska-Curie Actions doctoral network "SellSTEM" (Spatially Enhanced Learning linked to Science, Technology, Engineering and Mathematics) provides a comprehensive understanding of the interplay between self-concept, emotion, physiological arousal, and educational performance in primary school children. It emphasizes the significant gender differences in maths self-concept, maths and spatial anxiety and their role in mathematical and spatial performance, as well as the critical role of emotional reactivity, providing evidence for the potential benefits of targeted interventions to support emotional well-being and cognitive performance in STEM education. These insights are central for developing effective educational strategies and policies to enhance learning outcomes and promote gender equity in STEM fields.

1 Chapter 1: Introduction

1.1 Background and Context

Spatial skills, which involve the ability to mentally manipulate objects and visualize spatial relationships, are critical for success in STEM (Science, Technology, Engineering, Mathematics) disciplines (Linn & Petersen, 1985; Newcombe, 2017). Among these skills, mental rotation stands out as a key component of spatial ability. Mental rotation refers to the capacity to rotate two- or three-dimensional objects mentally and is essential for understanding and solving problems in various scientific and technical fields (Shepard & Metzler, 1988).

In the context of primary education, developing strong mental rotation skills can significantly enhance children's spatial reasoning and overall cognitive development (Newcombe & Frick, 2010). Early proficiency in mental rotation not only supports learning in subjects such as mathematics and science but also fosters problem-solving abilities (Frick et al., 2013). Recognizing the importance of these skills, educators are increasingly focusing on incorporating activities and curricula that promote spatial thinking and mental rotation in early education, laying a solid foundation for future success in STEM disciplines (Zhu et al., 2023).

Research has consistently found gender differences in mental rotation performance, with boys typically outperforming girls (Linn & Petersen, 1985; Voyer et al., 1995). These differences are often attributed to a variety of factors, including socialization processes, stereotype threat, and differences in experience with spatial tasks with many researchers emphasising social over cognitive influences (Newcombe, 2017). Understanding these gender disparities is crucial for developing targeted interventions that can help close this gap and ensure that all students have the opportunity to develop strong spatial skills, which are vital for success in STEM fields (Newcombe, 2017; Uttal & Cohen, 2012).

Mathematics self-concept, or an individual's perception of their own abilities in mathematics, plays a crucial role in shaping their academic preferences and performance, particularly in STEM subjects (Cvencek et al., 2020; Marsh et al., 2018). A positive maths self-concept is strongly associated with a higher preference for and engagement in STEM subjects, fostering a pathway towards careers in these fields (Goldman & Penner, 2016). This self-perception not only influences students' motivation and interest in pursuing STEM-related courses but due to the link between mathematics skills and spatial ability, it also plays a role in their spatial performance (Lennon-Maslin & Quaiser-Pohl, 2024). Research has shown that school children with a strong competence and confidence in spatial tasks tend to exhibit more motivation for mathematics, further enhancing their overall academic achievement and likelihood of pursuing STEM studies (Atit et al., 2022). Therefore, understanding the interplay between maths self-concept, STEM preferences, and spatial performance is essential for developing educational strategies that support and encourage students, particularly those who might otherwise be deterred by low self-efficacy in mathematics.

Mathematics and spatial anxiety are prevalent issues in educational settings, significantly affecting students' academic performance and STEM self-efficacy (Živković et al., 2023). Maths anxiety refers to the feeling of tension and fear that interferes with the manipulation of numbers and the solving of mathematical problems in various academic and everyday situations (Ashcraft, 2002; Hembree, 1990). Spatial anxiety involves a similar apprehension related to tasks that require spatial thinking, such as mental rotation and navigation (Alvarez-Vargas et al., 2020; Ramirez et al., 2012). Both types of anxiety can hinder students' ability to perform well in STEM (Science, Technology, Engineering, Mathematics) subjects, which are crucial for many career paths in the modern world (Delage et al., 2021)

Research has consistently shown that anxiety negatively impacts cognitive performance by occupying working memory and cognitive resources needed for problem-solving and task execution (Johns et al., 2008; Schmader et al., 2008). This is particularly significant in educational settings, where high levels of anxiety can lead to reduced participation, lower achievement, and a negative self-concept regarding one's abilities in maths and spatial tasks (Delage et al., 2021; Szczygiel, 2020; Živković et al., 2023). Despite extensive research in this area, gaps remain, particularly concerning the early identification of anxiety's impact on young children's performance and the role of physiological responses in this dynamic.

1.2 Problem Statement

The primary aim of this research is to examine the role of psychological factors such as self-concept and anxiety in primary school children's performance on spatial tasks and how this relates to performance in STEM subjects, in particular mathematics. By understanding these relationships, we can identify key factors that contribute to performance disparities and develop strategies to mitigate these effects. The research focuses on gender differences in psychological constructs associated with maths and spatial tasks and performance, exploring how boys and girls differ in their experiences and outcomes. Additionally, the studies investigate emotional reactivity, as measured through physiological responses, and children's preference for STEM subjects and their role of in influencing task performance. Understanding these dynamics is important for developing effective educational interventions and policies.

1.3 Research Questions

Overall, this research sought to answer the following questions on these topics:

1.3.1 Emotional Reactivity and Performance:

This research explored how emotional reactivity, as measured by galvanic skin responses (GSR), changes during a mental rotation (MR) task and is influenced by task difficulty. It aimed to uncover the relationship between emotional responses and performance metrics such as accuracy and response time on a MR task. The study also examined whether participants who invest more time in the task demonstrate greater accuracy, providing insights into the dynamics of emotional reactivity and its impact on cognitive performance in primary school children. The research also investigated how emotional reactivity, alongside subjective anxiety (both maths and spatial), influences performance on cognitive tasks such as maths and mental rotation. It explored whether these influences differ between girls and boys, and also how these factors along with subject preference impact task performance. This comprehensive approach aimed to provide a nuanced understanding of the interplay between cognitive and emotional factors in primary education.

1.3.2 Item Difficulty and Stereotype Effects:

The research investigated how patterns in speed and accuracy during a mental rotation (MR) task are affected by item difficulty, which varies based on stimulus type and rotational axis. It also explored the relationship between participants' perceptions of stereotyped and stimulus difficulty and their actual performance, assessing the influence of stereotype effects on spatial-cognitive outcomes.

1.3.3 Gender and Age Differences:

This research examined whether gender and age differences exist in primary school children's maths self-concept, spatial anxiety, perceived difficulty of spatial tasks, mental-rotation performance, and preference for STEM subjects. It aimed to understand how these differences manifest in cognitive task performance, particularly in maths and mental rotation, as well as in subject preferences, comparing their preference for maths versus German. It also focused on developmental differences based on childhood stages.

1.3.4 Predictive Factors:

The research sought to determine whether students' preference for STEM or non-STEM subjects predicts their maths self-concept. It examined the interplay between maths self-concept, spatial anxiety, and perceived task difficulty to understand how these factors collectively predict mental rotation performance, highlighting the underlying psychological dynamics. Additionally, it explored whether anxiety acts as a mediator in the relationship between math self-concept and mental rotation and whether gender, subject preference, emotional reactivity, and subjective anxiety are associated with to impact math and spatial task performance. This comprehensive approach aimed to provide a nuanced understanding of the interplay between cognitive and emotional factors in primary education.

1.4 Significance of the Research

The significance of this research lies in its potential to inform educational interventions and policy-making. By identifying the factors that contribute to deficits in self-efficacy and to task-related anxiety and its role in performance, educators can develop targeted strategies to support students, particularly those who are most affected by these issues. The research findings can help in designing curricula and teaching methods that increase self-efficacy, reduce anxiety, enhance emotional regulation, and improve overall academic performance in maths and spatial tasks. Furthermore, understanding the role of physiological responses in educational performance can lead to innovative approaches in assessing and supporting students' emotional well-being. This research aims to contribute to a more inclusive and effective educational environment, mitigating stereotyped attitudes and promoting better learning outcomes for all students, regardless of gender.

2 Chapter 2: Literature Review

Gender disparities in STEM fields have been a persistent issue worldwide. Despite significant efforts to promote gender equality, women remain underrepresented in many STEM disciplines, particularly those requiring advanced mathematical and spatial skills (Makarova et al., 2019; World Economic Forum, 2019). This integrated literature review synthesizes research from three key studies to explore the factors contributing to this disparity. Specifically, it examines gender differences in spatial ability, the role of emotional reactivity and stereotype threat, the benefits of computer adaptive testing (CAT), and the developmental aspects of spatial and mathematical skills in children.

2.1 Gendered Career Choices and STEM Engagement

Despite the advances in globalization, gendered career choices continue to persist globally. According to the World Economic Forum (2019), only one-third of female students pursue higher education or research careers in STEM fields. This underrepresentation is especially pronounced in mathematically intensive disciplines such as geoscience, engineering, and computer science (Ceci et al., 2014). A number of initiatives have been rolled out in the last decades in OECD countries to address gender issues in STEM subjects and careers. One such initiate was the National Pact for Women in MINT[1] Careers launched in Germany in 2008. This initiative, also known as “Go MINT” and instigated by the German Federal Ministry for Education and Research, brought together politics, business, science and the media with the goal of improving the image of STEM-related professions and careers in society (OECD, 2017). Additionally, the initiative aimed to increase young women's interest in scientific and technical degree courses. Another initiative, funded by the EU’s Horizon 2020, is the SellSTEM (Spatially Enhanced Learning linked to STEM) group. This research network, consisting of ten universities across Europe and nine partners in industry and technology education, aims to enhance spatial abilities in children across Europe. Due to consistent gender differences found in spatial ability in favour of males, SellSTEM has a particular focus on improving girls’ spatial ability to better prepare them for STEM education and careers (*SellSTEM MSCA ITN*, 2021).

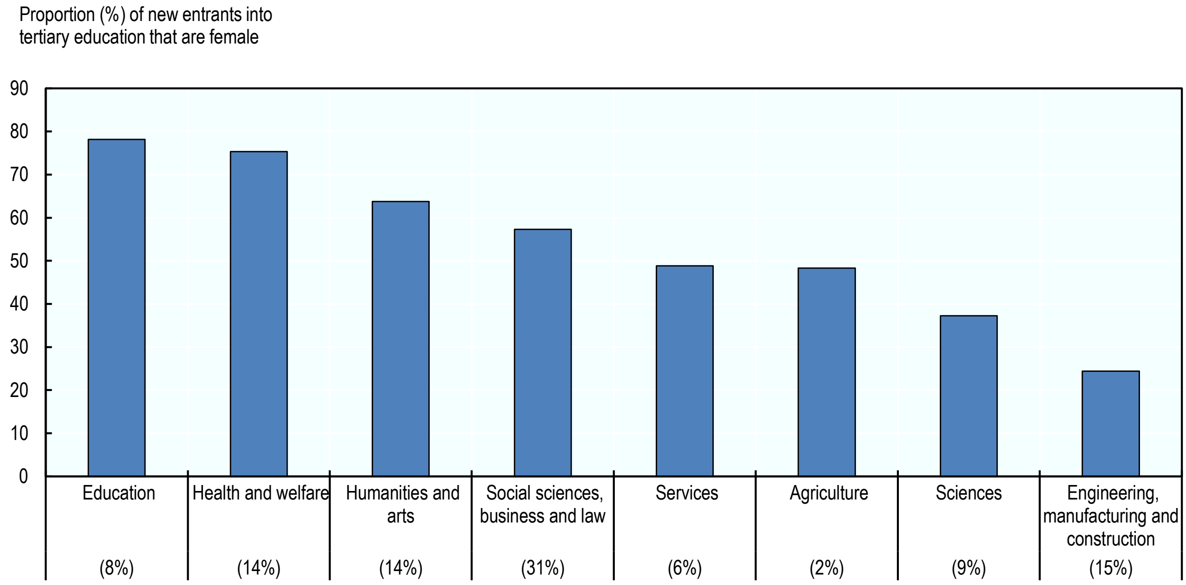
2.2 Spatial Ability and Mental Rotation (MR)

Spatial ability, particularly MR, is crucial for success in STEM fields. MR consistently shows significant gender differences, often favouring males, which are particularly evident in tasks involving three-dimensional objects (Linn & Petersen, 1985; Voyer et al., 1995). These differences. MR is traditionally assessed using psychometric instruments that have evolved to digital formats, allowing for more sophisticated analyses of performance (Schmand, 2019). Factors influencing MR performance include item difficulty, task-solving strategies, and

[1] MINT is the acronym for STEM in German and stands for Mathematik (Mathematics), Informatik (Computer Science), Naturwissenschaft (Natural Science) and Technologie (Technology).

Figure 1

Under-Representation of Women Among New Entrants in STEM Fields in Higher Education (OECD, 2017).



gender-stereotyped stimuli (Rahe & Quaiser-Pohl, 2021; Ruthsatz et al., 2015). Digital MR tests provide opportunities to explore these factors in detail, offering insights into designing assessments that minimize gender biases.

Recent studies have turned their attention to the impact of gender-stereotyped attributes of stimuli used in MR tests and how they might influence performance. Findings suggest that the degree to which stimulus objects are familiar or gender-congruent significantly affects the gender difference in MR performance (Neuburger et al., 2011). There is also a positive correlation between stereotyped stimulus content and children's performance on MR tests (Ruthsatz et al., 2014).

2.3 Emotional Reactivity and Cognitive Performance

Emotional reactivity significantly impacts cognitive performance on tasks like MR. Anxiety, stereotype threat, and other emotional responses can reduce the working memory and executive resources needed for successful task completion (Schmader et al., 2008). For instance, spatial anxiety (SA) can mediate gender differences in MR performance, particularly on more challenging spatial tasks, with girls often experiencing higher SA and lower self-confidence than boys (Arrighi & Hausmann, 2022; Ramirez et al., 2012). Effective emotional regulation strategies can mitigate these negative effects, improving performance on cognitive tasks (Fladung & Kiefer, 2016). Physiological measures, such as GSR, provide valuable insights into participants' emotional state during testing, helping to identify when and how emotional regulation strategies can be most effective (Deng et al., 2016). The significance of emotional responses is further underscored by research demonstrating that gender differences in spatial abilities are often influenced by emotional states and regulation strategies (Fladung & Kiefer, 2016). Identifying and addressing these emotional factors is therefore crucial for developing interventions that support equitable performance across genders.

2.4 Computer Adaptive Testing (CAT) in MR

Computer adaptive testing (CAT) represents a significant advancement over traditional fixed item tests (FIT). CAT adapts to the participant's performance, presenting items based on previous responses, potentially reducing anxiety and improving motivation, particularly among female students (Eggen & Verschoor, 2006; Martin & Lazendic, 2018). CAT's flexibility and efficiency make it a promising approach for assessing MR in educational settings (Chuesathuchon, 2008). Research indicates that CAT can lead to better performance and a more positive testing experience by reducing the pressure associated with fixed tests (Linden & Glas, 2000; Ling et al., 2017). By adapting to the test-taker's ability level, CAT can provide a more accurate assessment of their skills while maintaining engagement and reducing test-related anxiety (Fritts & Marszalek, 2010). This method also allows for shorter tests, reduced scheduling and supervision, and faster, more accurate scoring and reporting. CAT's advantages extend to educational practices by offering teachers greater convenience and flexibility (Chuesathuchon, 2008). The adaptive nature of CAT ensures that each student's abilities are accurately assessed, enabling more personalized and effective educational strategies.

2.5 Gender Differences in Maths Self-Concept and STEM Preferences

Gender disparities in maths self-concept and spatial anxiety emerge early in education, influenced by societal stereotypes and peer dynamics (Cvencek et al., 2011; Wolff, 2021). Girls often exhibit lower maths self-concept and higher spatial anxiety than boys, impacting their engagement and performance in STEM subjects (OECD, 2013; Raabe & Block, 2024). Longitudinal studies show that these factors significantly influence academic trajectories and career choices in STEM fields (Marsh et al., 2018; Parker et al., 2014).

Maths self-concept, the belief in one's ability to succeed in mathematics, is a critical factor influencing academic performance and subject preference (Cvencek et al., 2020). Research consistently shows that girls have lower maths self-concept than boys, leading to lower participation in maths and related subjects (Goldman & Penner, 2016). Addressing these disparities through targeted interventions can help improve girls' confidence and interest in STEM.

The Reciprocal Effects Model (REM) posits a mutually reinforcing relationship between self-concept and academic achievement over time (Marsh, 1990; Marsh & Craven, 2006). Encouraging active involvement and participation in mathematics is therefore anticipated to yield positive effects on students' self-concept. Studies conducted in the United States and China with pre- and primary school children found a reciprocal relationship between maths self-concept, interest, and achievement (Cai et al., 2018; Fisher et al., 2012).

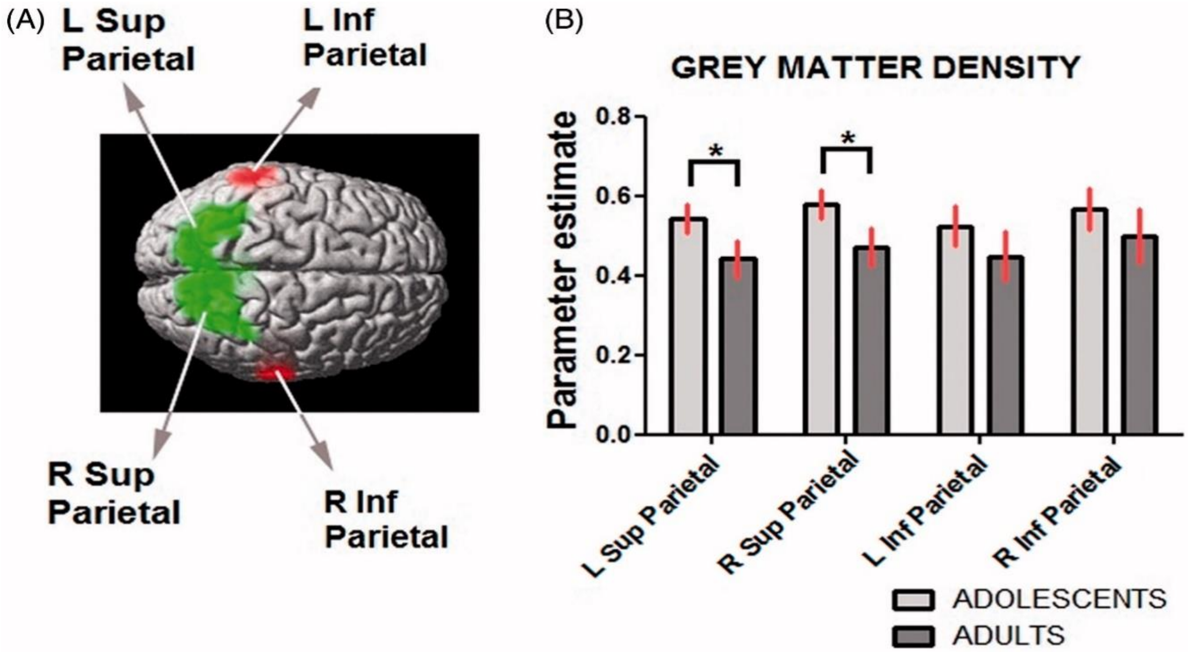
2.6 Developmental Aspects and Childhood Brain Development

Developmental aspects play a crucial role in shaping spatial and mathematical abilities, particularly during the pre-adolescent or tween stage (ages 9 to 12). This period is characterized by significant cognitive and brain development, including maturation of the parietal cortex, which is critical for spatial processing and cognitive control (Hodgkiss et al., 2021; Modroño et al., 2018). The development of these abilities during this stage can set the foundation for future success in STEM subjects.

Tweens are at a critical juncture where societal hierarchies, stereotypes, and peer influence become more salient. They are beginning to develop an awareness of gender identity and societal norms, which can impact their preferences and engagement with STEM subjects (McArthur et al., 2021; McGuire et al., 2022). Peer influence and teacher expectations play significant roles in shaping attitudes towards STEM (Heyder & Kessels, 2017; Wolff, 2021). Additionally, interventions such as peer mentoring during this period can be particularly effective in fostering positive attitudes and reducing the impact of negative stereotypes (Quaiser-Pohl, Endepohls-Ulpe, et al., 2014; Quaiser-Pohl & Endepohls-Ulpe, 2012; Space Science Institute, 2023; Starr, 2018).

Figure 2

Structural Changes During Early Adolescence in Brain Regions in the Superior Parietal Cortex Support Improved Spatial Manipulation (Modroño et al., 2018).



The development of mental rotation in pre-adolescence is associated with changes in the brain, particularly in regions related to spatial processing and cognitive control (s. Figure 2). Structural changes, such as increased grey matter density and synaptic pruning, contribute to improved spatial abilities (Modroño et al., 2018). This developmental period therefore presents a prime opportunity for the expansion and enhancement of cognitive abilities, which are essential for excelling in STEM subjects during secondary education. Subjects such as advanced mathematics, computer science, physics, and chemistry benefit greatly from development in spatial cognition. This developmental juncture sets the stage for not only cognitive enhancement but also potential challenges such as the emergence of negative attitudes and emotions related to spatial tasks. Hence, it is important that in girls and pre-adolescents, spatial skills are actively nurtured and interest in mathematics sustained.

2.7 Physiological Measures of Emotional Arousal

Physiological measures, such as galvanic skin response (GSR), provide valuable insights into emotional arousal and its impact on cognitive performance. GSR effectively captures physiological changes associated with anxiety and stress, offering a real-time measure of emotional reactivity during cognitive tasks (Christopoulos et al., 2019; Horvers et al., 2021; Nourbakhsh et al., 2012). Understanding these physiological responses can inform interventions aimed at reducing anxiety and improving cognitive outcomes in educational settings. Using wearable devices to measure GSR in real-life conditions, such as classrooms, can provide detailed data on how emotional arousal affects learning and performance (Geršak et al., 2020). This information is crucial for designing effective educational interventions that address the emotional and cognitive needs of students (Thammasan et al., 2020) .

Tests and exams trigger a range of emotions, including excitement, frustration, anxiety, and boredom (Tyng et al., 2017). Increased anxiety and negative thought patterns are linked to heightened electrodermal activity (EDA) (Pizzie & Kraemer, 2021). Maths anxiety, in particular, is connected to increased vigilance when dealing with mathematical tasks. Previous research has shown that biological factors, such as heightened amygdala activity, are associated with greater reactivity in individuals with high maths anxiety when they encounter mathematics (Pizzie & Kraemer, 2017). Additionally, mental-rotation tasks, which require significant spatial reasoning and mental visualization, can be particularly challenging for primary school children, leading to increased stress and emotional arousal (Ramirez et al., 2012). The novelty of these tasks might trigger interest and curiosity but also higher emotional reactivity due to their uncertainty. In contrast, repeated exposure to maths activities can lead to greater competence, comfort, and reduced anxiety (Ng et al., 2022). Familiar and straightforward maths tasks regularly practiced in school may be less stressful. However, lower self-efficacy in performing both maths and spatial tasks might heighten anxiety and physiological arousal (Rahe & Quaiser-Pohl, 2021).

2.8 Stereotype Threat and Working Memory

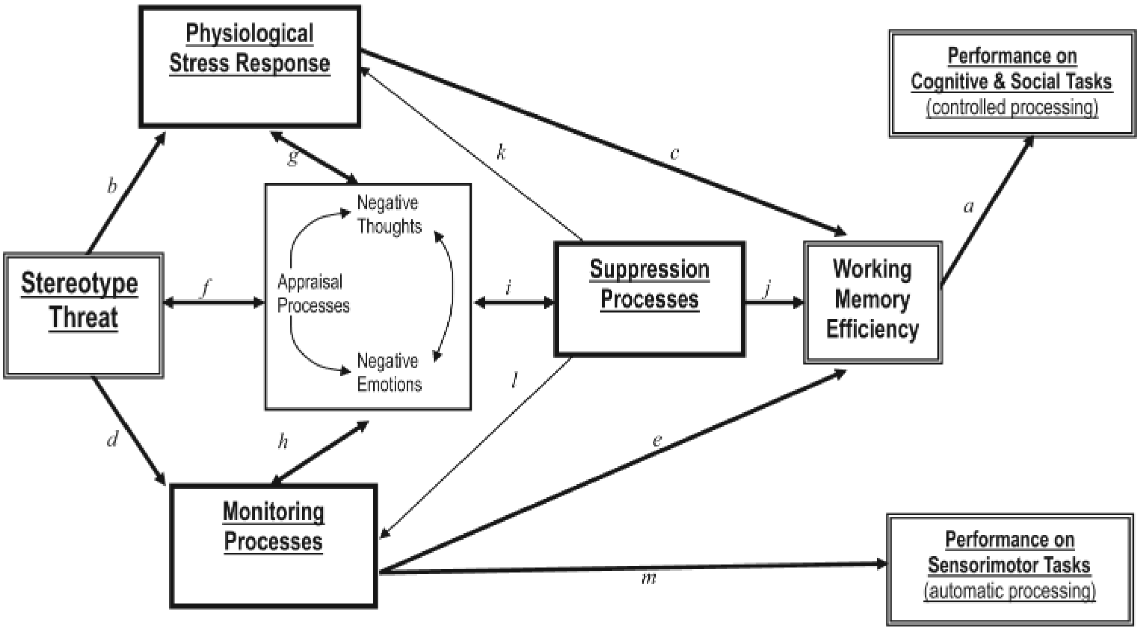
Stereotype threat (ST) occurs when individuals fear confirming negative stereotypes about their social group, which can impair their performance by consuming cognitive resources and increasing anxiety (Steele & Aronson, 1995). ST significantly affects cognitive performance by increasing anxiety and reducing working memory capacity, which is essential for complex cognitive tasks like MR and mathematics (Schmader, 2010; Schmader et al., 2008) (s. Figure 3). For girls, the stereotype that they are less capable in spatial and mathematical tasks can lead to increased anxiety and reduced working memory capacity, negatively impacting their performance (Neuburger et al., 2012; Schmader et al., 2008).

There is evidence that even very young children hold explicit stereotypes about spatial ability, associating it more with boys than girls (Ebert et al., 2024). Moreover, these stereotypes appear to be more pronounced in boys. This can be advantageous for boys due to stereotype lift effects, which enhance performance when individuals are aware that an outgroup is negatively stereotyped (Walton & Cohen, 2003). People can experience stereotype lift when the ability or value of an outgroup (boys vs girls for example) is questioned, even without specific reference to that outgroup, as long as the performance task is associated with a widely recognized negative stereotype.

Research has shown however that stereotype threat can be mitigated through interventions such as positive reinforcement, exposure to role models, and inclusive educational practices (Aronson et al., 2009). These strategies can help reduce the anxiety associated with stereotype threat, allowing students to perform to their full potential.

Figure 3

Integrated Model of Stereotype Threat Effects on Cognitive Performance (Schmader et al., 2008).



2.9 Gaps in the Literature

Based on the literature review, this research addresses several key gaps in the existing literature:

2.9.1 Gender Disparities in Spatial and Mathematical Abilities:

While there is a considerable body of research highlighting gender differences in spatial abilities, particularly mental rotation, there is a need for more focused studies on how these differences manifest in primary school children and how they are influenced by stereotype threat and emotional reactivity. Previous studies have established the presence of gender differences in spatial tasks, but there is limited understanding of the underlying cognitive and emotional mechanisms, particularly how subjective task-related anxiety and physiological arousal affect performance.

2.9.2 Impact of Emotional Arousal on Cognitive Performance:

Although the relationship between anxiety and cognitive performance is well-documented, there is a gap in understanding how physiological measures of emotional reactivity correlate with task performance in children. This research aims to fill this gap by using GSR to measure emotional responses during cognitive tasks. Most existing studies rely on self-reported measures of anxiety and emotional states. This research incorporates physiological data to provide a more objective assessment of how emotion impacts cognitive performance.

2.9.3 Developmental Differences in Cognitive and Emotional Responses:

There is a need for more research on how cognitive and emotional responses develop during the tween years (ages 9-12) and how these developmental changes impact performance on tasks requiring spatial and mathematical abilities. This research focuses on this critical developmental period to understand better how cognitive abilities and emotional reactivity evolve. While some studies have examined cognitive development in general, there is a lack of research specifically addressing how developmental changes influence susceptibility to stereotype threat and spatial anxiety during this stage.

2.9.4 Effectiveness of Adaptive Testing:

There is limited research on the use of computer adaptive testing (CAT) for assessing spatial abilities in children. This research provides some insights for further research into whether CAT can reduce anxiety and improve performance compared to traditional fixed item tests (FIT). Existing studies on CAT have primarily focused on academic subjects like mathematics. This research theoretically extends the application of CAT to spatial ability assessment, exploring its potential benefits in reducing test anxiety and providing a more tailored testing experience.

2.9.5 Interplay Between Self-Concept, Anxiety, and Performance:

While the relationship between self-concept and academic performance is well-documented, there is a gap in understanding how maths self-concept interacts with spatial anxiety and perceived task difficulty to influence performance on spatial tasks. This research aims to clarify these interactions and their implications for educational interventions. Previous research has often treated self-concept, anxiety, and performance as separate constructs. This research seeks to integrate these factors, providing a more holistic understanding of how they collectively impact cognitive task performance.

2.9.6 Influence of Societal Stereotypes on STEM Engagement:

There is a need for more research on how societal stereotypes about gender and STEM influence children's interest and engagement with these subjects. This research explores how stereotype threat may affect primary school children's performance and interest in STEM, particularly focusing on the impact of negative cultural stereotypes and academic choices. While stereotypes are recognized as a barrier to STEM engagement, there is limited empirical evidence on how early these stereotypes begin to affect children's self-concept, subject preferences and task performance. This research addresses this gap by examining the influence of stereotypes during the critical developmental period of primary school.

2.9.7 Methodological Advances in Measuring Cognition and Emotion:

Traditional studies have relied heavily on self-report questionnaires to measure anxiety and self-concept. This research incorporates physiological measures (e.g., GSR) to provide a more nuanced understanding of cognitive and emotional responses. By combining subjective and objective measures, this research aims to develop more reliable and valid methods for assessing the impact of emotional regulation on cognitive performance, particularly in young children. By addressing these gaps, the research contributes to a deeper understanding of the factors influencing gender differences in spatial and mathematical abilities, the role of emotional reactivity in cognitive performance, and the development of effective interventions to support children's engagement and success in STEM fields.

3 Chapter 3: Methodology

3.1 Research Design

This research employs a combination of cross-sectional and experimental designs to investigate the impact of self-concept, anxiety and physiological responses on the mathematical and spatial performance of primary school children. The cross-sectional design allows for the assessment of differences between various groups (e.g., gender differences in anxiety and performance), while the experimental design facilitates the examination of causal relationships between variables, such as the impact of emotional reactivity on task performance (Vogt et al., 2012). Additionally, the setting in which this research was undertaken, namely the classroom, reinforces its external validity, specifically ecological validity. Ecological validity refers to the realism with which the study design matches participant's real-life context (Hartson & Pyla, 2012). This factor influences the generalisability of the findings to actual classroom settings (Laureati & Pagliarini, 2019).

3.2 Participants

The studies involved 303 (*N*=303) primary school students from Germany, consisting of 146 girls and 155 boys with a *mean age* = 8.70 (*SD*=1.11) years. Participants were recruited from local primary schools and were representative of the general population in terms of socio-economic status and educational background. Written consent from parents and guardians was obtained for all participants, as well as verbal assent from the students themselves and ethical guidelines were strictly followed throughout the study.

3.3 Instruments Used to Measure Constructs

3.3.1 Cognitive Tasks

3.3.1.1 Mathematical Task:

Participants completed a series of mathematical problems designed to assess various aspects of mathematical competence. These tasks included number-line estimation, word problem representation, and missing terms tasks and were borrowed from the Kangaroo Maths Challenge (Mathematikwettbewerb Känguru e.V, n.d.). Each task was selected to cover different mathematical skills, from basic arithmetic to problem-solving abilities which correlate with spatial skills (Applebaum, 2017). Performance was measured in terms of percentage scores and completion times. Some examples of the problems are appended to this report (s. Appendix.1).

3.3.1.2 Mental Rotation Task (MRT):

The mental rotation task involved participants viewing images of objects that had been rotated in space and identifying which objects were the same as a reference image. The tasks used novel, gender-fair Mental Rotation Task (nMRT) and included both abstract and concrete stimuli rotated in picture-plane and in-depth. The task is based on Vandenberg and Kuse's (1978) pioneering study, but instead of using cubical stimuli only, it also incorporates items for young children such as animals and letters (Quaiser-Pohl, 2003) and features gender-stereotyped (male and female) and neutral objects (Neuburger et al., 2012; Ruthsatz et al., 2014, 2015). Time limits were imposed but these were somewhat longer than on the original MRT to accommodate diverse processing strategies in girls and boys (Nolte et al., 2022; Voyer, 2011). Participants' accuracy and response time were recorded to assess their spatial skills. The nMRT were computerized and presented on tablet devices. Experimental programming was undertaken in Psychopy, an open-source software package designed for the creation and presentation of experiments in behavioural sciences, particularly in psychology (Peirce, 2007). For the full list of stimuli used in the mental rotation task, see the appendix of publication 1.

3.3.2 Self-Report Questionnaires

3.3.2.1 Stereotyped Nature of Stimuli Questionnaire:

The Stereotyped Nature of Stimuli Questionnaire is a computerised, abbreviated vesion of a questionnaire developed to assess the extent to which participants perceived items in the mental rotation task as gender-stereotyped (Ruthsatz et al., 2015). This questionnaire included six items that measured participants' beliefs about whether certain stimuli were more suitable for boys or girls. The aim was to understand how gender stereotypes might influence perceived difficulty and performance on the mental rotation task (s. Appendix Publication 1).

3.3.2.2 Perceived Difficulty of Stimuli Questionnaire:

This Questionnaire was designed to measure participants' perceptions of the difficulty of items on the mental rotation task. It was based on the Stereotyped Nature of Stimuli Questionnaire (Ruthsatz et al., 2015). Understanding perceived difficulty is crucial as it can influence anxiety levels and overall performance. This questionnaire, consisting of six items, was used to assess how difficult participants found each stimulus, to examine the relationship between perceived difficulty, self-concept, anxiety levels, and task performance and to identify whether perceived difficulty varies by gender and how it might contribute to performance differences (s. Appendix Publication 1).

3.3.2.3 Maths Self-Concept Questionnaire:

This is a validated instrument designed to measure students' perceptions of their abilities and self-efficacy in mathematics. This subscale is part of a broader self-concept questionnaire and focuses specifically on mathematics-related self-concept (Ehm, 2014). The subscale was included to evaluate how primary school children perceive their abilities in mathematics, to explore the relationship between maths self-concept and performance in the spatial task and to examine gender differences in maths self-concept and its role in task performance (s. Appendix A.1).

3.3.2.4 Maths Anxiety Questionnaire:

A standardized maths anxiety questionnaire was administered to assess the levels of anxiety participants felt towards maths. The questionnaire included items measuring cognitive, affective, and physiological components of anxiety. The primary purpose of the Modified Abbreviated Math Anxiety Questionnaire (mAMAS) (Carey et al., 2017) is to measure the levels of anxiety that primary school children experience when engaging in mathematical tasks. The mAMAS provides a reliable and concise tool for quantifying this anxiety. One of the key objectives of using the mAMAS is to identify gender differences in maths anxiety. Previous studies have shown that girls often report higher levels of maths anxiety than boys. This research aims to confirm these findings within the sample and examine how these differences influence performance in mathematical and spatial tasks (s. Appendix A.1).

3.3.2.5 Spatial Anxiety Questionnaire:

Similar to the maths anxiety questionnaire, the spatial anxiety questionnaire assessed participants' anxiety related to spatial tasks. This included items on feelings of tension and worry when engaging in activities that required spatial reasoning. The primary purpose of the Child Spatial Anxiety Questionnaire (Ramirez et al., 2012) is to measure the levels of anxiety experienced by primary school children when engaging in spatial tasks. Understanding these anxiety levels is crucial as they can significantly impact performance in tasks requiring spatial reasoning, such as mental rotation and spatial visualization. The questionnaire aims to identify any gender differences in spatial anxiety. Previous research suggests that girls often report higher levels of spatial anxiety compared to boys. This research seeks to confirm these findings and explore the implications for educational performance and intervention strategies (s. Appendix A.1).

3.3.3 Physiological Measure of Galvanic Skin Response (GSR)

Galvanic Skin Response (GSR) was used to measure physiological arousal during the tasks. GSR measures the electrical conductance of the skin, which varies with its moisture level. Since sweat gland activity is controlled by the sympathetic nervous system, GSR is an effective measure of physiological arousal related to stress and anxiety (Christopoulos et al., 2019). The Shimmer3 GSR+ Unit® was used for GSR measurement. This device was synchronized with ConsensysBasic® multi-sensor management software. A two-minute baseline GSR recording was taken while participants were at rest to establish a reference level of physiological arousal. GSR data were continuously recorded during the maths and mental rotation tasks. Electrodes were attached to the index and middle fingers of the participants' non-dominant hand. GSR data were analysed to identify peaks in arousal corresponding to task-related stress and anxiety. These peaks were then correlated with performance metrics to assess the impact of physiological arousal on task performance.

3.4 Data Analysis

3.4.1 Statistical Techniques:

Descriptive Statistics: Means, standard deviations, and frequencies were calculated to summarize the data (Howitt & Cramer, 2014).

Inferential Statistics: T-tests and ANOVAs were used to examine differences between groups (e.g., gender and age differences in anxiety and performance). Regression analyses were conducted to explore relationships between variables, such as the impact of anxiety and physiological arousal on performance (Agresti, 2018).

Multiple Regression Analysis: This was calculated to assess the strength and direction of relationships between variables, such as between GSR and task performance and to identify predictors of performance in maths and spatial tasks, incorporating variables such as subjective anxiety, self-concept, and physiological arousal (Agresti, 2018).

Multivariate Analysis of Variance: Multivariate analysis was employed in the three studies to comprehensively examine the complex relationships between various psychological and performance-related factors. This statistical approach allows for the simultaneous analysis of multiple variables, providing deeper insights into how these factors interact and influence each other (Tabachnick & Fidell, 2013). Across all three studies, multivariate analysis facilitated a holistic understanding of the interplay between maths self-concept, anxiety (both maths and spatial), physiological arousal, and cognitive performance. By allowing the simultaneous examination of multiple variables and their interactions, this approach provided robust insights into the factors influencing primary school children's academic outcomes. It enabled the identification of key predictors of performance, the validation of measurement instruments, and the exploration of complex causal relationships. The application of multivariate analysis thus played a pivotal role in advancing the research objectives and contributing to the field of educational psychology.

Mediation Analysis: Simple mediation analysis was utilized to investigate the pathways through which maths self-concept influences mental rotation performance. Specifically, the analysis examined whether the effects of maths self-concept on accuracy and response time in mental rotation tasks were mediated by spatial anxiety and perceived difficulty of the tasks (Hayes, 2022).

All statistical analyses were conducted using SPSS® version 29, with mediation analysis performed using the Process macro add-on, version 4.2 (Hayes, 2022). Moreover, an a-priori power analysis was used to calculate the minimum number of participants needed for the planned statistical analyses using G*Power, an open source statistical software tool (Faul et al., 2007).

3.5 Ethical Considerations

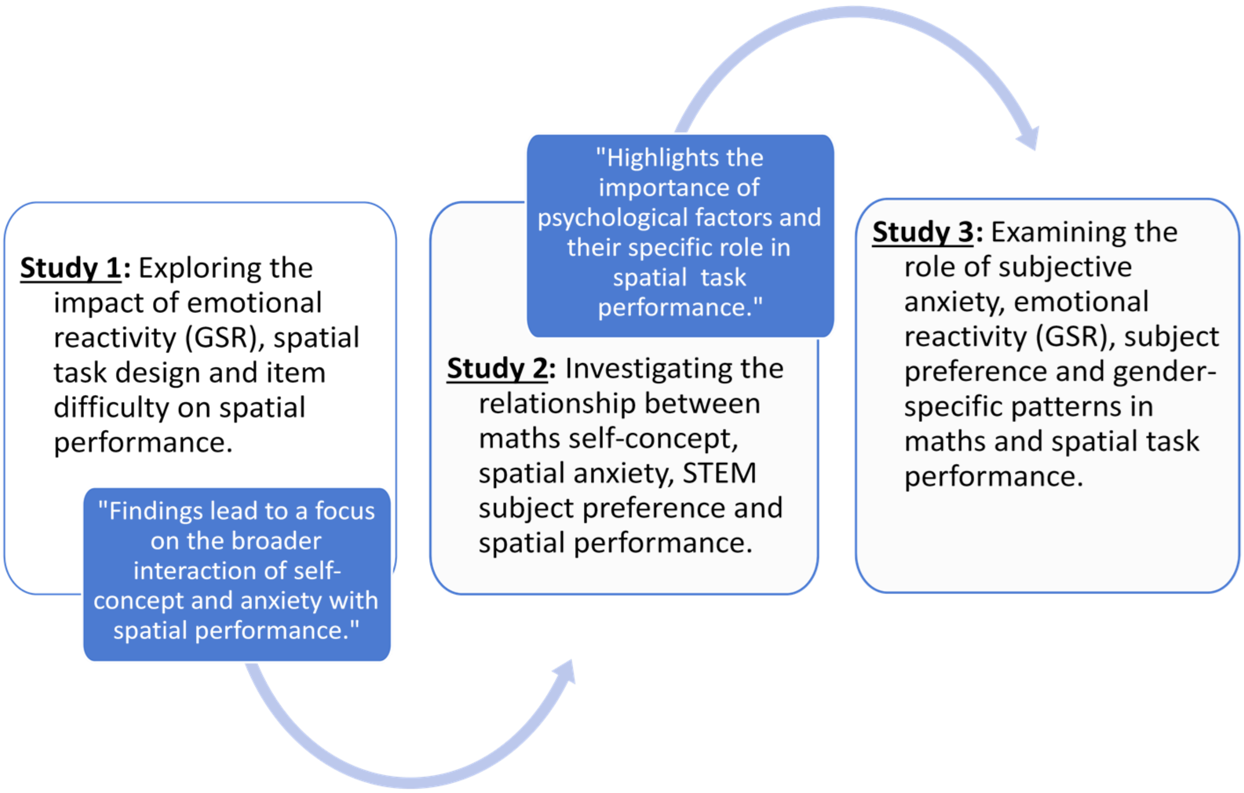
Ethical considerations were paramount in this study. The following measures were taken to ensure the ethical conduct of the research: Written informed consent was obtained from parents or guardians of all participants. The study's aims, procedures, and potential risks were clearly explained. Verbal assent was obtained from students prior to commencement of the experiment. Participant data were anonymized to protect privacy. Identifiable information was securely stored and only accessible to the research team. The tasks were designed to be age-appropriate and non-stressful. Breaks were provided to prevent fatigue, and participants could withdraw from the study at any time without consequence. The study was approved by the Ethics Committee of the University of Koblenz and adhered to the guidelines set by the Declaration of Helsinki (World Medical Association, 2013).

4 Chapter 4: Summary of Methodology and Results of the Submitted Publications

The following chapter outlines the methodology and findings of the three studies conducted within this research project and how the design evolved based on the findings of each study. The flowchart below illustrates this across the three studies (s. Figure 4).

Figure 4

Flowchart for Research Questions and Progression of the Three Studies within the Research Project.



4.1 Publication 1

Title: Under My Skin: Reducing Bias in STEM through New Approaches to Assessment of Spatial Abilities Considering the Role of Emotional Regulation (Lennon-Maslin et al., 2023).

4.1.1 Methodology:

A total of 29 third- and fourth-grade students from two primary schools in Koblenz, Germany, were recruited for this pilot study. After excluding one student who withdrew and three students whose data could not be retrieved, the final sample consisted of 25 participants (12 boys, 12 girls, and one student who did not specify gender) with an average age of 9.28 years. All parents and guardians provided written informed consent.

The MR task, programmed in PsychoPy® software and administered on Microsoft Pro 8 Surface tablets, recorded both accuracy and reaction time. The task included MR stimuli suitable for younger children, such as animals, letters, and cubes, as well as abstract and concrete stimuli rotated in picture-plane and in-depth. The task was divided into two parts: MRT 1 (easier, picture-plane rotation) and MRT 2 (more difficult, in-depth rotation), each with a time limit. MRT 1 consisted of 6 abstract and 10 concrete items rotated in picture-plane and allowed 5 minutes to complete and MRT 2 had 6 abstract and 6 concrete items rotated in-depth and with 8 minutes.

Two self-reported questionnaires were presented at the end of the MR task to assess the gender-stereotyped nature and perceived difficulty of the stimuli. The gender-stereotype questionnaire used a 5-point scale to rate the stimuli as more suitable for boys, girls, or neutral. The difficulty questionnaire used emojis on a 5-point scale to rate the perceived difficulty.

The Shimmer3 GSR+ Unit® was used to measure SCR. Baseline measurements were recorded for 2 minutes, and SCR was measured throughout the MR task. The data were categorized into low and high SCR groups based on mean SCR values.

Approval for the study was obtained from the Ethics Committee of the University of Koblenz and local school authorities. The MR task was explained to the students, who practiced with real objects before starting the computerized task. SCR was measured by attaching electrodes to the non-dominant hand, and baseline recordings were taken before the MR task began. The task included a practice run followed by two timed sections (MRT 1 and MRT 2).

The use of physiological measures, alongside traditional cognitive assessments, provides a robust framework for understanding how emotional factors impact spatial abilities. This approach aims to inform the development of more equitable and effective educational practices in STEM education.

4.1.2 Results:

A repeated-measures ANOVA with Greenhouse-Geisser correction showed significant differences in skin conductance levels across pre-test baseline, MRT 1 (picture-plane rotations), and MRT 2 (in-depth rotations), with 24% of the variation explained by the different test conditions. Pairwise comparisons indicated that skin conductance during MRT 2 was significantly higher than the baseline, and during MRT 1 was also significantly higher than the baseline. There was no significant difference between skin conductance in MRT 1 and MRT 2, suggesting that this increases during a mental rotation test regardless of task difficulty.

Pearson correlation analysis found a strong positive significant relationship between skin conductance levels and scores on items with abstract stimuli rotated in-depth, accounting for 67% of the variance. Additionally, there was a strong positive correlation between reaction time and scores on these items, accounting for 51% of the variation. Multiple regression analysis indicated that higher skin conductance levels predicted better scores on abstract objects rotated in-depth, but reaction time did not.

A one-way repeated-measures ANOVA found no significant differences in accuracy between skin conductance groups on concrete and abstract stimuli and stimuli rotated in picture-plane or in-depth. However, significant differences were found in accuracy scores based on stimulus type and rotational axis, with 26% of the variance in accuracy explained by these factors. Reaction time differences were also significant based on stimulus type and rotational axis, explaining 19% of the variance.

Descriptive analysis of the gender-stereotyped nature of stimuli questionnaire showed that items like the Car and Cube were rated as more masculine, while other items were rated as neutral. The perceived difficulty questionnaire indicated that items rotated in the picture-plane were rated as easier. Two-way repeated measures ANOVA found significant differences in accuracy and reaction time based on these questionnaire responses, explaining 21% and 20% of the variance, respectively.

This study highlights the critical role of emotional reactivity, as measured by skin conductance, in influencing MR task performance. The findings suggest that managing emotional responses can enhance cognitive functioning in spatial tasks, which has important implications for educational practices and assessment methods in STEM education. The study advocates for incorporating strategies to improve emotional regulation in educational settings to support the development of spatial skills necessary for success in STEM fields.

4.2 Publication 2

Title: Beyond numbers: the role of mathematics self-concept and spatial anxiety in shaping mental rotation performance and STEM preferences in primary education (Lennon-Maslin et al., 2024)**.**

4.2.1 Adjustments to the Research Design Based on Pilot Study Outcomes

Building on the findings from this pilot study (Publication 1), several key adjustments were made to the design of the next study (Publication 2) to enhance the robustness and applicability of the research. The study expanded the sample size and included a more comprehensive range of age groups (middle childhood and tweens) to explore age-related differences in emotional regulation and cognitive performance. Additionally, given the pilot study's findings that skin conductance levels were related to task difficulty but not significantly different between easier and more difficult mental rotation (MR) tasks, the second study retained the same MR tasks but added a focus on self-concept in mathematics and spatial anxiety. These new variables were introduced to explore how they might interact to influence MR performance and STEM preferences. The adjustments aimed to deepen the understanding of the cognitive and psychological factors that might impact STEM-related skills in a broader population of primary school children.

4.2.2 Methodology:

A total of 148 students from first- to fourth-grade primary schools in Germany, were recruited for the study. After excluding three students who were no longer in primary education and one student whose data could not be recorded due to a technical issue, the final sample consisted of 144 participants (74 boys, 70 girls). The average age of the students was 8.47 years. Two age categories were created based on stage of childhood development: A middle childhood group (6- to 8-year-olds) and a tween (pre-adolescent group) consisting of 9- to 11-year-olds.

An online questionnaire created in PsychoPy® was administered at the beginning of the experiment to collect data on participants' age and gender. Students named their favourite subject in a blank space provided. Using the Academic Self-Concept Questionnaire (ASKG) (Ehm, 2014) for primary school students', self-perceptions in mathematics, reading and writing were assessed using a 7-point scale. This study focused on the mathematics self-concept subscale, which demonstrated high reliability (Cronbach’s Alpha = 0.90). The Child Spatial Anxiety Questionnaire (CSAQ) (Ramirez et al., 2012) consists of 8 items assessed anxiety related to spatial tasks. The scale reliability was marginally below the acceptable range (Cronbach’s Alpha = 0.65) and the Perceived Difficulty of Stimuli (PDQ), adapted from the Stereotyped Nature of Stimuli questionnaire (Neuburger et al., 2015) used emojis on a 5-point scale to rate the difficulty of items in the MRT (Cronbach’s Alpha = 0.72).

The Mental Rotation Task (MRT) used was the same as that of the previous study. It was computerized MRT, based on Vandenberg & Kuse (1978) and was administered using Microsoft Pro 8 Surface tablets. The task recorded both accuracy and response time, including stimuli such as animals, letters, and cubes rotated in picture-plane and in-depth. The task was divided into two parts (Cronbach’s Alpha = 0.86; MRT 1: $\alpha = 0.71$, and MRT 2: $\alpha = 0.86$).

Approval was obtained from the Ethics Committee of the University of Koblenz and local authorities. Informed consent was obtained from parents and guardians. The study was conducted in a separate classroom with adequate lighting and individual seating. Researchers explained the MRT using physical objects to ensure understanding. Skin conductance was measured using Shimmer3 GSR+ Units. Quantitative analyses were conducted using SPSS® 29, employing various statistical tests.

The methodology ensured comprehensive data collection and analysis to explore the relationship between emotional regulation, cognitive performance, and spatial abilities in primary school children.

4.2.3 Results:

Using independent samples t-tests, the study found that girls had significantly lower maths self-concept compared to boys. Additionally, girls exhibited higher levels of spatial anxiety and perceived difficulty than boys. However, there were no significant gender differences in accuracy and response times on the mental rotation task. A Chi-Square test for association revealed that girls were less likely than boys to choose the STEM subject maths.

Independent samples t-tests indicated that tweens had lower maths self-concept compared to students in middle childhood. While there were no significant age differences in spatial anxiety and perceived difficulty, tweens performed better in accuracy on the mental rotation task. A Chi-Square test for association showed that tweens were also less likely than younger students to choose the STEM subject maths.

A one-way analysis of variance (ANOVA) was conducted to examine the relationship between students' preferred subject at school and their levels of maths self-concept. The results indicated that students who preferred the STEM subject maths had significantly higher maths self-concept compared to those who preferred non-STEM subjects.

Multiple regression analysis was used to investigate the association between maths self-concept and spatial anxiety, controlling for sex and age. The results indicated that higher maths self-concept was significantly associated with lower spatial anxiety. Another multiple regression analysis examined the association between maths self-concept and perceived difficulty of the mental rotation task, revealing that higher maths self-concept was significantly associated with lower perceived difficulty.

Mediation analysis using the PROCESS procedure in SPSS was performed to examine whether the relationship between maths self-concept and accuracy on the mental rotation task was mediated by spatial anxiety and perceived difficulty. The results indicated that the relationship between maths self-concept and accuracy was partially mediated by spatial anxiety, but not by perceived difficulty. For response time, mediation analysis showed that neither spatial anxiety nor perceived difficulty significantly mediated the same relationship.

4.3 Publication 3

Title: "It's Different for Girls!" The Role of Anxiety in Primary School Children's Mathematics and Mental Rotation Performance (Lennon-Maslin & Quaiser-Pohl, Under Review).

4.3.1 Adjustments to the Research Design Based on Outcomes of Study 2

The second study highlighted that while maths self-concept and spatial anxiety were significantly associated with MR performance, the role of maths anxiety emerged as a critical factor in shaping both maths and MR task outcomes. In response to these findings, the third study specifically targeted the measurement of maths anxiety and its interaction with emotional reactivity measured by Galvanic Skin Responses (GSR) to better understand its impact on both a maths and the mental rotation task. Considering the second study's finding that gender differences in STEM preferences and performance are influenced by self-concept and anxiety, the third study included a detailed examination of emotional reactivity during both maths and MR tasks, aiming to uncover any subtle gender-specific patterns that might inform more tailored educational interventions. These adjustments were designed to build on the prior study's insights and further clarify the complex interplay between cognitive, emotional, and gender-related factors in primary school students' STEM abilities.

4.3.2 Methodology:

The study recruited 131 students from second, third, and fourth-grade classes at five primary schools in Rhineland Palatinate, Germany. These schools were located in both urban and rural settings, with students from diverse socioeconomic and cultural backgrounds. The sample included 66 boys and 65 girls, with an average age of 8.73 years. Students completed a maths task, a mental rotation task, and several questionnaires.

The study utilized Shimmer3 GSR+ Units to measure galvanic skin response (GSR), synchronized with ConsensysBasic software. A baseline recording of 2 minutes was taken before participants engaged in the maths and mental rotation tasks. GSR data was pre-processed to remove artifacts and noise, and mean GSR was calculated for baseline, maths task, and mental rotation task periods. GSR devices were attached to the index and middle fingers of the non-dominant hand. Baseline measurements were followed by recordings during the maths and mental rotation tasks, conducted in a controlled environment to ensure reliability. The setup minimized external distractions and maintained consistent ambient conditions.

Students completed a paper-and-pencil maths task with items of varying difficulty, drawn from the Mathematical Kangaroo Competition. The task included number line estimation, word problem representation, and missing terms, with specific problems selected to be unfamiliar to the students. Written instructions were provided, and students worked individually without electronic aids.

A novel computerized mental rotation task (nMRT) was programmed using PsychoPy software, installed on Microsoft Pro 8 Surface tablets. The task recorded accuracy and response time, featuring stimuli such as animals, letters, and geometric shapes as well as gender-stereotyped and neutral objects. Items were presented randomly in two parts, with different time limits and types of rotation (in-picture-plane and in-depth).

An online questionnaire collected data on participants' age and gender after the cognitive tasks to avoid priming effects. Three questionnaires assessed students' preference for Maths vs German, spatial anxiety, and maths anxiety. These included a Maths-German Preference Survey which is part of the ASKG questionnaire, students indicated their preferred subject, the Child Spatial Anxiety Questionnaire (CSAQ) to assessed anxiety related to spatial tasks, using a 5-point emoji scale and the Modified Abbreviated Maths Anxiety Scale (mAMAS): Measured maths anxiety with a 9-item Likert scale, translated into German.

The study received ethical approval, and informed consent was obtained from parents and guardians. Testing was conducted in a separate classroom by two female researchers. Students were familiarized with the tasks and the significance of mental rotation. The setup ensured a comfortable environment, with reminders for students needing eyeglasses.

Data analyses were performed using SPSS 29 software. Maths task scores were standardized to percentages, and transformations were applied to address skewness in mental rotation accuracy, response times, and spatial anxiety scores. Skin conductance data was standardized using z-score transformations to account for individual variability.

4.3.3 Results:

A multivariate analysis of variance (MANOVA) was used to examine gender differences in maths performance, mental rotation performance, and anxiety measures. The analysis revealed significant gender differences across these variables. Between-subjects effects tests indicated that girls scored higher in maths but took longer to complete the tasks. Girls also reported higher maths anxiety compared to boys. However, there were no significant gender differences in accuracy or response times on the mental rotation task, nor in spatial anxiety. Pairwise comparisons confirmed these findings, highlighting that girls performed better in maths but experienced higher anxiety levels and took more time to complete the maths tasks.

A Chi-square test for association examined the relationship between gender and preference for maths versus German. The results showed a significant association, with girls being less likely to prefer maths as their school subject compared to boys.

A multivariate analysis of covariance (MANCOVA) was conducted to investigate the effects of gender, subject preference, emotional reactivity (measured by GSR), maths anxiety, and spatial anxiety on maths performance. The analysis indicated significant effects of gender and subject preference on maths performance, with maths anxiety also showing a significant impact. Emotional reactivity and spatial anxiety had marginally significant effects on completion times for the maths tasks. Gender-specific analyses further showed that maths anxiety significantly affected girls' maths scores, and emotional reactivity influenced their completion times. For boys, maths anxiety had a marginal effect on maths scores, while other factors did not show significant effects.

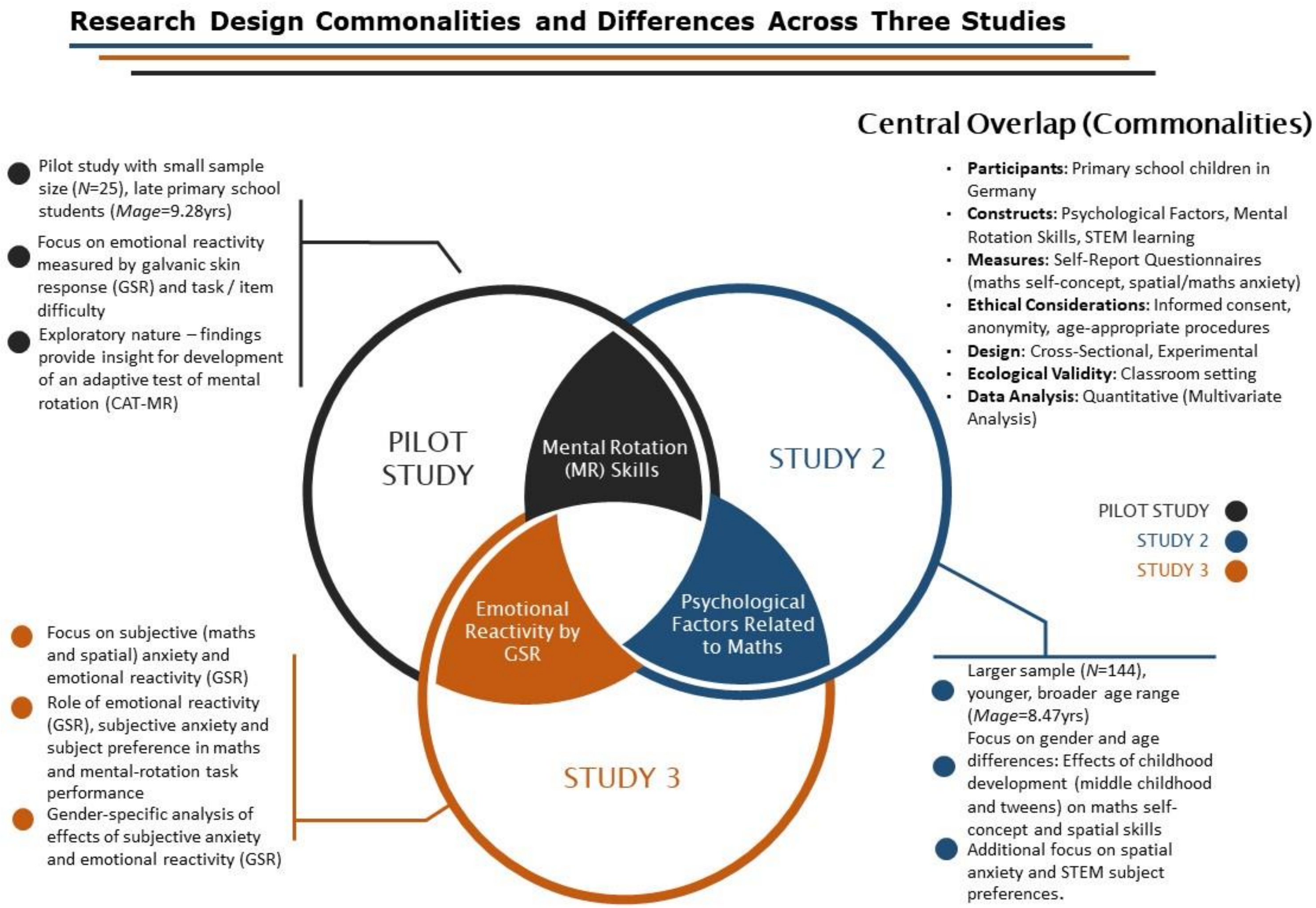
A MANCOVA was used to explore the associations between gender, subject preference, and performance on the mental rotation task, controlling for emotional reactivity, maths anxiety, and spatial anxiety. The results showed that gender significantly affected mental rotation task performance, with marginal effects from emotional reactivity and maths anxiety. Between-subjects effects tests revealed that maths anxiety significantly impacted accuracy, while emotional reactivity and gender influenced response times. Gender-specific analyses indicated that for boys, maths anxiety marginally affected accuracy, and emotional reactivity had a marginal effect on response times. For girls, none of the predictors significantly impacted accuracy or response time in the mental rotation task.

Overall, the study found significant gender differences in maths performance and anxiety levels, with girls outperforming boys in maths but experiencing higher anxiety and having longer completion times. Gender also influenced performance on the mental rotation task, with additional contributions from emotional reactivity and maths anxiety.

The Venn diagram below shows a summary of commonalities and differences in the research design across the three studies. The Venn diagram highlights shared elements such as participant characteristics, tasks, and measures, while also showcasing the unique aspects of each study. The flowchart traces the evolution of research focus from the exploration of emotional reactivity and task difficulty in Study 1, to the broader examination of mathsself-concept and spatial anxiety in Study 2, and finally to the targeted analysis of maths anxiety, physiological arousal and gender-specific patterns in Study 3 (s. Figure 5).

Figure 5

Summary of Commonalities and Differences in Research Design across Three Studies.



5 Chapter 5: Synthesis and General Discussion

5.1 Integration of Findings

The integrated findings from the three publications provide a comprehensive understanding of the interplay between emotional reactivity, gender differences, and task performance in primary school children.

5.1.1 Emotional Reactivity and Performance:

Emotional reactivity, as measured by skin conductance was found to significantly influence performance on mental rotation (MR) tasks. This increased significantly from baseline to the first part of the MR task, where stimuli were rotated in the picture-plane, and remained high throughout both parts of the task. However, no significant difference in emotional reactivity was observed between the easier first subscale and the more difficult second subscale of the task, suggesting that increased task difficulty did not significantly affect physiological arousal during the MR task.

Higher emotional reactivity, as indicated by skin conductance, and increased response time (RT) were associated with better scores on difficult items, specifically abstract stimuli rotated in-depth. This suggests that participants who managed their emotions effectively and invested more time in solving complex items achieved superior outcomes. Differences in accuracy and RT were observed between participants in "high" and "low" emotional reactivity groups, particularly concerning gender-stereotyped objects. Participants with higher emotional reactivity spent more time on gender-neutral objects, which were deemed more relevant and salient, leading to longer engagement and higher accuracy.

Emotional reactivity, also played a nuanced role in maths task performance. For this task, lower emotional reactivity was a significant as predictor of shorter completion times in girls but did not significantly impact their percentage scores. 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 in all students tested. Higher emotional reactivity may also indicate increased cognitive engagement which can also lead to cognitive overload, particularly in complex tasks like mental rotation. This aligns with the Yerkes-Dodson Law, which suggests that moderate levels of arousal improve performance, but excessive arousal can hinder it (Stokes & Kite, 2017; Teigen, 1994). This could also explain why participants with higher emotional reactivity performed better on difficult items but took longer to complete tasks overall.

5.1.2 Item Difficulty and Stereotype Effects

Accuracy and response time in mental rotation tasks are significantly influenced by both the type of stimuli and the rotational axis (Neuburger et al., 2012; Quaiser-Pohl, Neuburger, et al., 2014). Items rotated in-depth and abstract stimuli were more challenging, as reflected by greater variance in these performance metrics. Gender-stereotyped stimuli, such as the Car and Cube, were perceived as more masculine, while neutral items elicited less bias, suggesting stereotype effects in task perception. Furthermore, easier ratings were given to items rotated in the picture-plane which reflect existing findings (Neuburger et al., 2012). These results emphasize that both item difficulty and perceived stereotypes shape performance on MRTs, affecting both accuracy and speed (Neuburger et al., 2012). The study suggests that stereotype awareness and task difficulty perceptions may play crucial roles in cognitive outcomes, with important implications for reducing gender bias in spatial task performance in STEM education (Ruthsatz et al., 2019, Under Review).

Stereotype threat, especially in gender-stereotyped tasks such as the mental rotation, likely affected performance. For example, objects associated with masculine domains, such as hammers or cars, may have triggered stereotype-activated anxiety in girls, impacting their engagement or performance. However, when the tasks involved neutral or abstract objects, which are perceived as less gender-biased, this stereotype threat may have been reduced, leading to more equal engagement across genders (Ruthsatz et al., 2017, 2019). Gender-neutral objects, often more familiar to both boys and girls, were associated with longer response times but higher accuracy. This suggests that participants invested more effort in solving these items, likely due to their relevance and familiarity. This highlights the importance of careful task design in reducing gender biases in performance assessments, by including neutral stimuli that allow for fairer comparisons across genders (Ruthsatz et al., Under Review).

5.1.3 Gender and Age Differences

Significant gender disparities emerged in maths self-concept, spatial anxiety, and perceived difficulty of mental rotation tasks. Girls exhibited lower maths self-concept and higher levels of spatial anxiety and perceived task difficulty compared to boys. These discrepancies highlight variations in confidence and comfort levels between male and female students when engaging with mathematical and spatial concepts. Factors contributing to these gender differences include temperamental predispositions and socialization processes (Sanchis-Sanchis et al., 2020), exposure to spatially-related activities, and the influence of parental and teacher stereotypes (Garcia et al., 2021; Gunderson et al., 2012; Rocha et al., 2022).

Despite these challenges, no significant gender differences were found in mental rotation performance in terms of accuracy or response time. This lack of disparity may be attributed to effective educational practices in Germany (OECD, 2020), the nature of the mental rotation task used, and longer time limits that accommodated diverse processing strategies (Saunders & Quaiser-Pohl, 2020; Taragin et al., 2019). This suggests that spatial abilities may be more evenly distributed among genders when assessment techniques and educational practices mitigate stereotype threat. Moreover, ensuring equal access to resources may also benefit spatial ability (Ruthsatz et al., 2013).

In addition to gender differences, age-related patterns were also observed. Results revealed that tweens exhibited a lower maths self-concept compared to younger students in middle childhood. However, no significant age differences emerged in terms of spatial anxiety or perceived task difficulty. Interestingly, despite their lower math self-concept, tweens performed better in terms of accuracy on the mental rotation task underscoring how neurodevelopmentally and educationally tweens are better able to process complex spatial tasks despite appearing to lose interest in STEM subjects (Modroño et al., 2018; Ojose, 2008). Tweens were less likely than younger students to choose maths as their preferred subject, confirming a decline in interest or lack of identification in STEM subjects as children grow older (Kessels, 2005; Starr, 2018; Starr & Leaper, 2019). This addition highlights the nuanced relationship between age, self-concept, and task performance, while also emphasizing the decline in interest in maths among older students, contributing to a more comprehensive understanding of developmental differences.

In terms of subject preference, girls were also less likely to choose maths as their preferred subject. This persistent belief that maths is not a subject they identify with may result from socialization and environmental factors, influencing interest and achievement in STEM fields (Aeschlimann et al., 2015; Lesperance et al., 2022). Results also revealed a significant association between gender and subject preference and maths performance metrics. These finding highlights ongoing gender differences in subject interest and preference, influenced by higher reported maths anxiety among girls and societal stereotypes about gender and mathematical abilities (Cvencek et al., 2011; Passolunghi et al., 2014; Wolff, 2021).

5.1.4 Predictive Factors

The predictive influences of gender, age, and maths self-concept on spatial anxiety and perceived task difficulty were explored. Age and maths self-concept, but not gender, were significant predictors of spatial anxiety. Older children and those with higher maths self-concept exhibited lower levels of spatial anxiety. Similarly, maths self-concept accounted for a significant proportion of the variance in perceived task difficulty, whereas gender and age did not. These findings suggest that environmental circumstances, individual learning experiences, and socio-cultural influences may play a more prominent role in shaping spatial anxiety and perceived difficulty than gender alone (Quaiser-Pohl & Endepohls-Ulpe, 2012; Ruthsatz et al., 2013; Starr, 2018).

The mediation analysis revealed that higher levels of maths self-concept were associated with higher scores on the mental rotation task through reduced spatial anxiety. This underscores the importance of fostering a positive self-concept in maths from an early age. However, the relationship between maths self-concept and response time on the mental rotation task was not mediated by spatial anxiety or perceived difficulty, suggesting the involvement of other psychological constructs or variables such as motivation or task-specific strategies (Atit et al., 2022; Moe, 2016). High self-concept in maths among boys may act as a protective factor against the negative effects of spatial anxiety, allowing them to perform tasks with less interference. In contrast, girls, who reported lower maths self-concept, may have experienced more cognitive interference from anxiety, which likely affects spatial performance on timed tasks. This finding suggests that fostering self-concept in maths could be a critical intervention for reducing anxiety and improving performance in spatial tasks (Cvencek et al., 2020; Kaskens et al., 2020).

Maths anxiety emerged as a significant predictor of performance on the maths task, with higher anxiety levels associated with lower percentage scores. This is consistent with existing literature on the detrimental effects of maths anxiety on performance due to increased cognitive load and interference from working memory (Ashcraft & Moore, 2009; Schmader et al., 2008). However, maths anxiety did not significantly affect completion times, suggesting subjective anxiety determines the precision of task performance rather than speed.

The findings reveal significant gender differences in how subject preferences, emotional reactivity, and subjective anxiety impact maths an spatial task performance. For boys, none of the predictors played a significant role in performance outcomes on maths nor spatial tasks. This may reflect different coping mechanisms or cognitive strategies employed by boys

(Fladung & Kiefer, 2016; Van Mier et al., 2019). Girls performed better on the maths task when they reported less subjective anxiety, showing better percentage scores. Additionally, their preference for maths and the ability to regulate emotional arousal was associated with shorter completion times on the same task. Girls who indicated a preference for maths had significantly higher percentage scores and lower completion times on the maths task, reflecting reciprocal effects between engagement, practice, and a positive maths attitude (Gunderson et al., 2018; Marsh & Craven, 2006). None of the factors were associated with girls' performance on the spatial task underscoring the importance of equitable assessments methods which mitigate stereotype effects.

5.2 Theoretical Contributions

The integrated findings from the three publications provide several important theoretical contributions to the understanding of emotional reactivity, gender differences, and task performance in primary school children, particularly in the context of mathematical and spatial abilities.

The studies collectively highlight the crucial role of emotional regulation in influencing cognitive task performance, specifically in mental rotation (MR) tasks. The consistent finding that higher emotional reactivity, as measured by GSR, is associated with better performance on difficult items underscores the importance of managing emotional arousal during complex cognitive tasks. This contributes to the broader theoretical framework that posits emotional regulation as a key factor in cognitive functioning, particularly in tasks requiring significant working memory and attentional resources such as an MRT (Fladung & Kiefer, 2016; Ramirez et al., 2012; Schmader et al., 2008).

The research provides a nuanced understanding of gender differences in mathematical and spatial abilities among primary school children. The finding that girls exhibit lower maths self-concept and higher levels of spatial anxiety compared to boys, yet perform equally well on MR tasks, challenges traditional theories that suggest inherent gender differences in spatial abilities. Instead, these results support the theory that socialization processes and educational practices significantly influence cognitive performance (Garcia et al., 2021; Gunderson et al., 2012). The lack of significant gender differences in MR performance, despite differences in anxiety and self-concept, aligns with the perspective that educational interventions and supportive environments can mitigate stereotype threats and promote equity in cognitive abilities (OECD, 2015).

The studies underscore the impact of subject preference and motivation on task performance. The consistent association between a preference for maths and better performance on maths tasks highlights the role of intrinsic motivation and interest in STEM achievement (Atit et al., 2022). This finding reinforces the theoretical framework that posits subject preference as a critical factor in educational outcomes, suggesting that fostering positive attitudes towards subjects like maths can enhance performance and engagement (Cvencek et al., 2020).

The mediation analyses reveal the interconnectedness of various cognitive skills, particularly the role of maths self-concept in influencing spatial performance through reduced spatial anxiety. This supports the theoretical proposition that cognitive and affective factors are deeply intertwined, and interventions targeting one area (e.g., maths self-concept) can have spill over effects on related cognitive skills (Cai et al., 2018; Moe & Pazzaglia, 2010). This interconnectedness underscores the need for holistic educational approaches that address both cognitive and emotional development to optimize academic performance.

Additionally, the findings offer a valuable theoretical perspective on the interaction between task-specific strategies and emotional reactivity. The results suggest that while emotional reactivity, as measured by skin conductance, plays a crucial role in managing cognitive load during complex tasks, it is the combination of emotional regulation and effective problem-solving strategies that leads to superior outcomes, particularly on difficult items (Fladung & Kiefer, 2016). This contributes to the broader understanding that cognitive performance is not solely determined by intellectual capacity or emotional states, but by how well these elements are integrated during task engagement and aligns with theories of self-regulated learning, where learners who can manage their emotions and apply task-specific strategies effectively are better positioned to succeed, especially in challenging cognitive environments (Eysenck et al., 2007; Eysenck, 1998). The gender-specific findings, where girls' performance benefited from lower anxiety and emotional reactivity, also emphasize the protective effects of self-regulation and the need to consider individualized approaches in educational interventions that target both emotional and strategic components of learning. Age-related findings in which tweens demonstrated lower spatial anxiety and superior performance on spatial tasks support existing theories which show this period is characterized by significant cognitive and emotional development, including neurobiological maturation, which improves children's ability to manage anxiety related to tasks (Sanchis-Sanchis et al., 2020). This growth, along with increased exposure and experience, helps explain reductions in spatial anxiety and perceived task difficulty as children develop.

The findings contribute to the understanding of how educational practices and stereotype threat influence cognitive performance. The evidence that gender-neutral and balanced stimuli, along with extended time limits, mitigate gender differences in MR tasks supports theories that emphasize the importance of equitable educational practices in reducing stereotype threat (Passolunghi et al., 2014). This highlights the potential of thoughtfully designed educational interventions to create more inclusive and supportive learning environments that promote equity in cognitive abilities across genders.

5.3 Limitations and Future Research

A significant limitation across the studies was the sample size and diversity, particularly in the pilot study, which impacts the generalizability of the findings. Future research should replicate these studies with larger, culturally and socio-economically diverse samples to confirm and elaborate on the results.

The studies identified several measurement issues, such as the reliability of the Child Spatial Anxiety Questionnaire and potential biases introduced by using less familiar maths tasks like the Kangaroo Challenge. Future research should refine these measurement tools and consider using more standardized, computerized maths tasks to ensure reliability and validity. Additionally, incorporating qualitative data could provide deeper insights into the cognitive and emotional processes children experience during tasks.

While physiological measures like skin conductance provide valuable insights into emotional regulation, they are subject to individual variability and potential artifacts. Future research should continue to explore and validate these measures, possibly incorporating additional physiological methods such as heart rate or pupillometry to triangulate findings.

Given the cross-sectional nature of the studies, longitudinal research is needed to understand how cognitive and affective factors evolve over time. Following cohorts of students across their primary school years could also provide valuable insights into the long-term effects of interventions aimed at improving maths self-concept, reducing anxiety, and enhancing performance.

The role of stereotype threat in influencing cognitive performance was evident in the studies. Future research could explore more strategies to mitigate stereotype threat, such as promoting positive role models, creating inclusive curricula, and implementing training programs for teachers to adopt gender-sensitive practices. These steps can help create a more equitable educational environment that supports all students, regardless of gender.

By addressing these limitations and pursuing future research directions, we can deepen our understanding of the complex interplay between emotional regulation, cognitive skills, and educational practices, ultimately leading to more effective strategies for enhancing academic performance and equity in primary school education.

6 Chapter 6: Conclusion

6.1 Summary of Key Findings

This research has provided comprehensive insights into the complex interplay between maths self-concept, maths and spatial anxiety, and cognitive performance in primary school children. The key findings from the three studies are as follows:

High levels of maths anxiety significantly impair cognitive performance which is crucial for solving mathematical problems. Girls tend to exhibit higher levels of maths anxiety than boys, which negatively impacts their mathematical performance. Spatial skills, especially mental rotation, are critical for success in STEM disciplines. The studies reveal no significant gender differences in spatial performance when the instrument utilizes a broad range of stimuli including gender-congruent and neutral objects. Physiological arousal, measured through galvanic skin response (GSR), provides objective evidence of how emotional reactivity affects performance in maths and spatial tasks. In the spatial task, higher emotional reactivity was linked to longer response times and better performance on difficult items, indicating that effective management of emotions and investing more time can improve outcomes. Similarly, during the maths task, increased physiological arousal, especially in girls, resulted in longer completion times. Since both tasks were timed, difficulty in managing stress and anxiety may hinder task completion. These findings highlight the need for balancing time and precision, underscoring the critical role of emotional control in achieving success on challenging tasks. The interrelationship between maths self-concept, anxiety, and spatial performance highlights the need for integrated educational interventions. Positive self-concept enhances performance, while high anxiety levels, exacerbated by stereotype threats, may hinder it.

6.2 Overall Contributions

The research makes several significant contributions to the field of educational psychology and STEM education:

The integration of physiological measures with traditional psychological assessments extends existing theories on anxiety and educational performance. This approach provides a more nuanced understanding of how emotional and cognitive factors interact to influence learning outcomes. The findings inform the development of targeted educational interventions aimed at reducing anxiety and improving self-concept in mathematics. These strategies are crucial for fostering a positive learning environment and supporting students' academic success. The research highlights the importance of early interventions and inclusive teaching practices to address gender disparities in STEM education. Policymakers can use these insights to design educational policies that prioritize mental health and emotional well-being in schools. The use of GSR as a measure of physiological arousal in educational research provides a valuable tool for objectively assessing the impact of physiological arousal on cognitive performance. This methodological innovation can be applied in future studies to further explore these dynamics.

6.3 Conclusion

The integrated findings from the three publications underscore the importance of reducing disadvantages in cognitive testing to preserve the reliability and validity of psychometric instruments designed to measure spatial abilities, such as mental rotation (MR). The studies collectively highlight that mitigating stereotype threat in individuals exposed to this can restore executive resources, improve cognition, and thus enhance test performance (Johns et al., 2008). This can also lead to diminishing the gender difference in spatial test performance. Therefore, ongoing investigation into the effects of gender-congruent and neutral stimuli, emotional arousal during testing, item difficulty, and the chronological order of items on spatial ability performance is crucial. This research ensures that these factors do not compromise measurement accuracy nor contribute to increasing gender differences but instead serve to measure spatial ability accurately, regardless of gender.

The studies provide insights into critical factors influencing children's attitudes and performance in mathematics, spatial tasks, and STEM subjects. A persistent gender disparity in maths self-concept, spatial anxiety, perceived difficulty of spatial tasks, and preference for maths as a STEM subject was identified. This disparity highlights enduring negative stereotypes in these fields. Similarly, low maths self-concept and loss of interest and enthusiasm for maths in late primary school children warrant attention to prevent disengagement from STEM-related subjects in further education. Counteracting stereotypes early is essential to ensure children's academic choices are based on ability rather than attitude. Interventions aimed at improving maths self-concept may indirectly benefit spatial task performance by alleviating spatial anxiety. Understanding the complex interplay between cognitive, affective, and demographic factors influencing performance on spatial tasks is crucial for developing targeted interventions to support children's mathematical self-concept and spatial skills development in educational settings.

Interventions such as stress management training, mindfulness practices, and anxiety-reduction techniques can help students manage physiological arousal and negative emotions, thereby enhancing performance (Bauer & Jansen, 2024). For boys, strategies to improve focus and attention during tasks may be more effective. Techniques such as positive reinforcement and encouraging a growth mindset can also be beneficial (Cvencek et al., 2020). Educators could counteract gender stereotypes in maths 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 (Boaler, 2015).

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, such as training teachers in gender-sensitive practices and diverse teaching methods (Gamarra et al., 2024), developing educational materials that promote gender equality and feature diverse STEM role models (Cruz Neri et al., 2024), engaging parents and communities through workshops and resources to support positive attitudes towards maths and STEM (Olive et al., 2022), launching media campaigns to change societal perceptions and highlight successful women in STEM (Alkhammash, 2019), advocating for policies that promote gender equity in education, including training for educators and anti-discrimination policies (Global Education Monitoring Report Team, 2020), 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.

7 References

Aeschlimann, B., Herzog, W., & Makarova, E. (2015). Studienpräferenzen von Gymnasiastinnen und Gymnasiasten: Wer entscheidet sich aus welchen Gründen für ein MINT-Studium? *Swiss Journal of Educational Research*, *37*(2), Article 2. https://doi.org/10.24452/sjer.37.2.4954

Agresti, A. (2018). *Statistical Methods for the Social Sciences, Global Edition* (5th edition). Pearson.

Alkhammash, R. (2019). 'It Is Time to Operate Like a Woman': Representation of Women in STEM Fields in Social Media: A Corpus-Based Study. *International Journal of English Linguistics*, *9*, 217. https://doi.org/10.5539/ijel.v9n5p217

Alvarez-Vargas, D., Abad, C., & Pruden, S. M. (2020). Spatial anxiety mediates the sex difference in adult mental rotation test performance. *Cognitive Research: Principles and Implications*, *5*(1), 31. https://doi.org/10.1186/s41235-020-00231-8

Applebaum, M. (2017). Spatial Abilities as a Predictor to Success in the Kangaroo Contest. *Journal of Mathematics and System Science*, *7*, 154. https://doi.org/10.17265/2159-5291/2017.06.002

Aronson, J., Cohen, G., Mccolskey, W., & Montrosse-Moorhead, B. (2009). Reducing Stereotype Threat in Classrooms: A Review of Social-Psychological Intervention Studies on Improving the Achievement of Black Students. Summary. Issues & Answers. REL 2009-076. *Regional Educational Laboratory Southeast*.

Arrighi, L., & Hausmann, M. (2022). Spatial anxiety and self-confidence mediate sex/gender differences in mental rotation. *Learning & Memory*, *29*(9), 312–320. https://doi.org/10.1101/lm.053596.122

Ashcraft, M. H. (2002). Math Anxiety: Personal, Educational, and Cognitive Consequences. *Current Directions in Psychological Science*, *11*(5), 181–185. https://doi.org/10.1111/1467-8721.00196

Ashcraft, M. H., & Moore, A. M. (2009). Mathematics Anxiety and the Affective Drop in Performance. *Journal of Psychoeducational Assessment*, *27*(3), 197–205. https://doi.org/10.1177/0734282908330580

Atit, K., Power, J. R., Pigott, T., Lee, J., Geer, E. A., Uttal, D. H., Ganley, C. M., & Sorby, S. A. (2022). Examining the relations between spatial skills and mathematical performance: A meta-analysis. *Psychonomic Bulletin & Review*, *29*(3), 699–720. https://doi.org/10.3758/s13423-021-02012-w

Bauer, R., & Jansen, P. (2024). A short mindfulness induction might increase women's mental rotation performance. *Consciousness and Cognition*, *123*, 103721. https://doi.org/10.1016/j.concog.2024.103721

Boaler, J. (2015). *The Elephant in the Classroom: Helping Children Learn and Love Maths*. Souvenir Press.

Cai, D., Viljaranta, J., & Georgiou, G. K. (2018). Direct and indirect effects of self-concept of ability on math skills. *Learning and Individual Differences*, *61*, 51–58. https://doi.org/10.1016/j.lindif.2017.11.009

Carey, E., Hill, F., Devine, A., & Szűcs, D. (2017). The Modified Abbreviated Math Anxiety Scale: A Valid and Reliable Instrument for Use with Children. *Frontiers in Psychology*, *8*. https://www.frontiersin.org/article/10.3389/fpsyg.2017.00011

Ceci, S. J., Ginther, D. K., Kahn, S., & Williams, W. M. (2014). Women in Academic Science: A Changing Landscape. *Psychological Science in the Public Interest: A Journal of the American Psychological Society*, *15*(3), 75–141. https://doi.org/10.1177/1529100614541236

Christopoulos, G. I., Uy, M. A., & Yap, W. J. (2019). The Body and the Brain: Measuring Skin Conductance Responses to Understand the Emotional Experience. *Organizational Research Methods*, *22*(1), 394–420. https://doi.org/10.1177/1094428116681073

Chuesathuchon, C. (2008). Computerized adaptive testing in mathematics for primary schools in Thailand. *Theses: Doctorates and Masters*. https://ro.ecu.edu.au/theses/1591

Cruz Neri, N., Schwenzer, F., Sprenger, S., & Retelsdorf, J. (2024). Geschlechterdarstellungen in Geographie-Schulbüchern: Von mangelnder Repräsentation weiblicher Figuren und Reproduktion von Geschlechterstereotypen. *Zeitschrift für Erziehungswissenschaft*. https://doi.org/10.1007/s11618-024-01223-w

Cvencek, D., Meltzoff, A. N., & Greenwald, A. G. (2011). Math–Gender Stereotypes in Elementary School Children. *Child Development*, *82*(3), 766–779. https://doi.org/10.1111/j.1467-8624.2010.01529.x

Cvencek, D., Paz-Albo, J., Master, A., Herranz Llácer, C. V., Hervás-Escobar, A., & Meltzoff, A. N. (2020). Math Is for Me: A Field Intervention to Strengthen Math Self-Concepts in Spanish-Speaking 3rd Grade Children. *Frontiers in Psychology*, *11*, 593995. https://doi.org/10.3389/fpsyg.2020.593995

Delage, V., Trudel, G., Retanal, F., & Maloney, E. A. (2021). Spatial anxiety and spatial ability: Mediators of gender differences in math anxiety. *Journal of Experimental Psychology: General*. APA PsycArticles. https://doi.org/10.1037/xge0000884

Deng, Y., Chang, L., Yang, M., Huo, M., & Zhou, R. (2016). Gender Differences in Emotional Response: Inconsistency between Experience and Expressivity. *PloS One*, *11*(6), e0158666. https://doi.org/10.1371/journal.pone.0158666

Ebert, M., Jost, L., & Jansen, P. (2024). Gender stereotypes in preschoolers' mental rotation. *Frontiers in Psychology*, *15*, 1284314. https://doi.org/10.3389/fpsyg.2024.1284314

Eggen, T. J. H. M., & Verschoor, A. J. (2006). Optimal Testing With Easy or Difficult Items in Computerized Adaptive Testing. *Applied Psychological Measurement*, *30*(5), 379–393. https://doi.org/10.1177/0146621606288890

Ehm, J.-H. (2014). *Akademisches Selbstkonzept im Grundschulalter. Entwicklungsanalyse dimensionaler Vergleiche und Exploration differenzieller Unterschiede*. pedocs. http://nbn-resolving.de/urn:nbn:de:0111-opus-95657

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

Eysenck, N. D. M. W. (1998). Working Memory Capacity in High Trait-anxious and Repressor Groups. *Cognition and Emotion*, *12*(5), 697–713. https://doi.org/10.1080/026999398379501

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fisher, P. H., Dobbs-Oates, J., Doctoroff, G. L., & Arnold, D. H. (2012). Early math interest and the development of math skills. *Journal of Educational Psychology*, *104*(3), 673–681. https://doi.org/10.1037/a0027756

Fladung, A.-K., & Kiefer, M. (2016). Keep calm! Gender differences in mental rotation performance are modulated by habitual expressive suppression. *Psychological Research*, *80*(6), 985–996. https://doi.org/10.1007/s00426-015-0704-7

Frick, A., Hansen, M., & Newcombe, N. (2013). Development of mental rotation in 3-to 5-year-old children. *COGNITIVE DEVELOPMENT*, *28*(4), 386–399. https://doi.org/10.1016/j.cogdev.2013.06.002

Fritts, B., & Marszalek, J. (2010). Computerized adaptive testing, anxiety levels, and gender differences. *Social Psychology of Education*, *13*, 441–458. https://doi.org/10.1007/s11218-010-9113-3

Gamarra, E., Tenbrink, T., & Mills, D. (2024). Gender in Teacher-Student Interactions: Another Factor in Spatial Ability Development and STEM Affiliation. In M. Živković, J. Buckley, M. Pagkratidou, & G. Duffy (Eds.), *Spatial Cognition XIII* (pp. 51–65). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-63115-3_4

Garcia, N. L., Hall, L. V., & Pruden, S. M. (2021). *Individual Differences in Young Children's Spatial Ability: A Systematic Review*. PsyArXiv. https://doi.org/10.31234/osf.io/5mc2y

Geršak, V., Vitulić, H. S., Prosen, S., Starc, G., Humar, I., & Geršak, G. (2020). Use of wearable devices to study activity of children in classroom; Case study—Learning geometry using movement. *Computer Communications*, *150*, 581–588. https://doi.org/10.1016/j.comcom.2019.12.019

Global Education Monitoring Report Team. (2020). *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. https://unesdoc.unesco.org/ark:/48223/pf0000374514

Goldman, A. D., & Penner, A. M. (2016). Exploring international gender differences in mathematics self-concept. *International Journal of Adolescence and Youth*, *21*(4), 403–418. https://doi.org/10.1080/02673843.2013.847850

Gunderson, E., Park, D., Maloney, E., Beilock, S., & Levine, S. (2018). Reciprocal relations among motivational frameworks, math anxiety, and math achievement in early elementary school. *JOURNAL OF COGNITION AND DEVELOPMENT*, *19*(1), 21–46. https://doi.org/10.1080/15248372.2017.1421538

Gunderson, E., Ramirez, G., Beilock, S., & Levine, S. (2012). The Relation Between Spatial Skill and Early Number Knowledge: The Role of the Linear Number Line. *DEVELOPMENTAL PSYCHOLOGY*, *48*(5), 1229–1241. https://doi.org/10.1037/a0027433

Hartson, R., & Pyla, P. S. (2012). Chapter 14 - Rigorous Empirical Evaluation: Preparation. In R. Hartson & P. S. Pyla (Eds.), *The UX Book* (pp. 503–536). Morgan Kaufmann. https://doi.org/10.1016/B978-0-12-385241-0.00014-2

Hayes, A. F. (2022). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach* (3. edition). Taylor & Francis.

Hembree, R. (1990). The Nature, Effects, and Relief of Mathematics Anxiety. *Journal for Research in Mathematics Education*, *21*(1), 33–46. https://doi.org/10.2307/749455

Heyder, A., & Kessels, U. (2017). Boys Don't Work? On the Psychological Benefits of Showing Low Effort in High School. *Sex Roles*, *77*(1), 72–85. https://doi.org/10.1007/s11199-016-0683-1

Hodgkiss, A., Gilligan-Lee, K., Thomas, M., Tolmie, A., & Farran, E. (2021). The developmental trajectories of spatial skills in middle childhood. *BRITISH JOURNAL OF DEVELOPMENTAL PSYCHOLOGY*, *39*(4), 566–583. https://doi.org/10.1111/bjdp.12380

Horvers, A., Tombeng, N., Bosse, T., Lazonder, A. W., & Molenaar, I. (2021). Detecting Emotions through Electrodermal Activity in Learning Contexts: A Systematic Review. *Sensors (Basel, Switzerland)*, *21*(23). https://doi.org/10.3390/s21237869

Howitt, D., & Cramer, D. (2014). *Introduction to Statistics in Psychology* (6 edition). Pearson Education.

Johns, M., Inzlicht, M., & Schmader, T. (2008). Stereotype Threat and Executive Resource Depletion: Examining the Influence of Emotion Regulation. *Journal of Experimental Psychology. General*, *137*(4), 691–705. https://doi.org/10.1037/a0013834

Kaskens, J., Segers, E., Goei, S. L., van Luit, J. E. H., & Verhoeven, L. (2020). Impact of Children's math self-concept, math self-efficacy, math anxiety, and teacher competencies on math development. *Teaching and Teacher Education*, *94*, 103096. https://doi.org/10.1016/j.tate.2020.103096

Kessels, U. (2005). Fitting into the stereotype: How gender-stereotyped perceptions of prototypic peers relate to liking for school subjects. *European Journal of Psychology of Education*, *20*(3), 309–323. https://doi.org/10.1007/BF03173559

Laureati, M., & Pagliarini, E. (2019). 14—The effect of context on children's eating behavior. In H. L. Meiselman (Ed.), *Context* (pp. 287–305). Woodhead Publishing. https://doi.org/10.1016/B978-0-12-814495-4.00014-3

Lennon-Maslin, M., & Quaiser-Pohl, C. (2024). Mindset Matters: The Role of Mathematics Self-concept and Age in Mental Rotation Performance Among Primary School Children. In M. Živković, J. Buckley, M. Pagkratidou, & G. Duffy (Eds.), *Spatial Cognition XIII* (pp. 19–31). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-63115-3_2

Lennon-Maslin, M., & Quaiser-Pohl, C. M. (Under Review). "It's Different for Girls!" The Role of Anxiety in Primary School Children's Mathematics and Mental Rotation Performance. *Behavioral Sciences*, *14*(7). https://susy.mdpi.com/user/manuscripts/apc_info/764c1a785e8aba45579c770644acc0cb

Lennon-Maslin, M., Quaiser-Pohl, C. M., Ruthsatz, V., & Saunders, M. (2023). Under My Skin: Reducing Bias in STEM through New Approaches to Assessment of Spatial Abilities Considering the Role of Emotional Regulation. *Social Sciences*, *12*(6), Article 6. https://doi.org/10.3390/socsci12060356

Lennon-Maslin, M., Quaiser-Pohl, C. M., & Wickord, L.-C. (2024). Beyond Numbers: The role of mathematics self-concept and spatial anxiety: For Mental Rotation Performance and STEM Preferences. *Frontiers in Education*, *9*. https://doi.org/10.3389/feduc.2024.1300598

Lesperance, K., Munk, S., Holzmeier, Y., Braun, M., & Holzberger, D. (2022). *Geschlechterunterschiede im Bildungskontext. Von wissenschaftlichen Studien zu Impulsen für die Unterrichtspraxis*. pedocs. http://nbn-resolving.de/urn:nbn:de:0111-pedocs-253313

Linden, W. J. van der, & Glas, C. A. W. (2000). *Computer adaptive testing: Theory and practice.* Kluwer Academic Publishers. https://research.utwente.nl/en/publications/computer-adaptive-testing-theory-and-practice

Ling, G., Attali, Y., Finn, B., & Stone, E. A. (2017). Is a Computerized Adaptive Test More Motivating Than a Fixed-Item Test? *Applied Psychological Measurement*, *41*(7), 495–511. https://doi.org/10.1177/0146621617707556

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development, 56*(6), 1479–1498. https://doi.org/10.2307/1130467

Makarova, E., Aeschlimann, B., & Herzog, W. (2019). The Gender Gap in STEM Fields: The Impact of the Gender Stereotype of Math and Science on Secondary Students' Career Aspirations. *Frontiers in Education*, *4*. https://www.frontiersin.org/article/10.3389/feduc.2019.00060

Marsh, H. W. (1990). Causal ordering of academic self-concept and academic achievement: A multiwave, longitudinal panel analysis. *Journal of Educational Psychology*, *82*(4), 646–656. https://doi.org/10.1037/0022-0663.82.4.646

Marsh, H. W., & Craven, R. G. (2006). Reciprocal Effects of Self-Concept and Performance From a Multidimensional Perspective: Beyond Seductive Pleasure and Unidimensional Perspectives. *Perspectives on Psychological Science, 1*(2), 133–163. https://doi.org/10.1111/j.1745-6916.2006.00010.x

Marsh, H. W., Pekrun, R., Murayama, K., Arens, A. K., Parker, P. D., Guo, J., & Dicke, T. (2018). An integrated model of academic self-concept development: Academic self-concept, grades, test scores, and tracking over 6 years. *Developmental Psychology*, *54*(2), 263–280. https://doi.org/10.1037/dev0000393

Martin, A. J., & Lazendic, G. (2018). Computer-adaptive testing: Implications for students' achievement, motivation, engagement, and subjective test experience. *Journal of Educational Psychology*, *110*(1), 27–45. https://doi.org/10.1037/edu0000205

Mathematikwettbewerb Känguru e.V. (n.d.). *Känguru der Mathematik e.V. | Aufgaben*. Retrieved 4 June 2024, from https://www.mathe-kaenguru.de/chronik/aufgaben/index.html

McArthur, B. A., Madigan, S., & Korczak, D. J. (2021). Tweens are not teens: The problem of amalgamating broad age groups when making pandemic recommendations. *Canadian Journal of Public Health*, *112*(6), 984–987. https://doi.org/10.17269/s41997-021-00585-6

McGuire, L., Hoffman, A. J., Mulvey, K. L., Hartstone-Rose, A., Winterbottom, M., Joy, A., Law, F., Balkwill, F., Burns, K. P., Butler, L., Drews, M., Fields, G., Smith, H., & Rutland, A. (2022). Gender Stereotypes and Peer Selection in STEM Domains Among Children and Adolescents. *Sex Roles*, *87*(9), 455–470. https://doi.org/10.1007/s11199-022-01327-9

Modroño, C., Navarrete, G., Nicolle, A., González-Mora, J. L., Smith, K. W., Marling, M., & Goel, V. (2018). Developmental grey matter changes in superior parietal cortex accompany improved transitive reasoning. *Thinking & Reasoning*, *25*(2), 151–170. https://doi.org/10.1080/13546783.2018.1481144

Moe, A. (2016). Does experience with spatial school subjects favour girls' mental rotation performance? *LEARNING AND INDIVIDUAL DIFFERENCES*, *47*, 11–16. https://doi.org/10.1016/j.lindif.2015.12.007

Moe, A., & Pazzaglia, F. (2010). Beyond genetics in Mental Rotation Test performance The power of effort attribution. *LEARNING AND INDIVIDUAL DIFFERENCES*, *20*(5), 464–468. https://doi.org/10.1016/j.lindif.2010.03.004

Neuburger, S., Jansen, P., Heil, M., & Quaiser-Pohl, C. (2011). Gender differences in pre-adolescents' mental-rotation performance: Do they depend on grade and stimulus type? *Personality and Individual Differences*, *50*(8), 1238–1242. https://doi.org/10.1016/j.paid.2011.02.017

Neuburger, S., Jansen, P., Heil, M., & Quaiser-Pohl, C. (2012). A threat in the classroom: Gender stereotype activation and mental-rotation performance in elementary-school children. *Zeitschrift Für Psychologie*, *220*(2), 61–69. https://doi.org/10.1027/2151-2604/a000097

Neuburger, S., Ruthsatz, V., Jansen, P., & Quaiser-Pohl, C. (2012). Influence of Rotational Axis and Gender-Stereotypical Nature of Rotation Stimuli on the Mental-Rotation Performance of Male and Female Fifth Graders. In C. Stachniss, K. Schill, & D. Uttal (Eds.), *Spatial Cognition VIII* (pp. 220–229). Springer. https://doi.org/10.1007/978-3-642-32732-2_15

Neuburger, S., Ruthsatz, V., Jansen, P., & Quaiser-Pohl, C. (2015). Can girls think spatially? Influence of implicit gender stereotype activation and rotational axis on fourth graders' mental-rotation performance. *LEARNING AND INDIVIDUAL DIFFERENCES*, *37*, 169–175. https://doi.org/10.1016/j.lindif.2014.09.003

Newcombe, N. (2017). *Harnessing Spatial Thinking to Support Stem Learning*. OECD. https://doi.org/10.1787/7d5dcae6-en

Newcombe, N. S., & Frick, A. (2010). Early Education for Spatial Intelligence: Why, What, and How. *Mind, Brain, and Education*, *4*(3), 102–111. https://doi.org/10.1111/j.1751-228X.2010.01089.x

Ng, C., Chen, Y., Wu, C., & Chang, T. (2022). Evaluation of math anxiety and its remediation through a digital training program in mathematics for first and second graders. *Brain and Behavior*, *12*(5), e2557. https://doi.org/10.1002/brb3.2557

Nolte, N., Schmitz, F., Fleischer, J., Bungart, M., & Leutner, D. (2022). Rotational complexity in mental rotation tests: Cognitive processes in tasks requiring mental rotation around cardinal and skewed rotation axes. *Intelligence*, *91*, 101626. https://doi.org/10.1016/j.intell.2022.101626

Nourbakhsh, N., Wang, Y., Chen, F., & Calvo, R. A. (2012). Using galvanic skin response for cognitive load measurement in arithmetic and reading tasks. *Proceedings of the 24th Australian Computer-Human Interaction Conference*, 420–423. https://doi.org/10.1145/2414536.2414602

OECD. (2013). *Mathematics Self-Beliefs and Participation in Mathematics-Related Activities* (pp. 87–112). OECD. https://doi.org/10.1787/9789264201170-8-en

OECD. (2015). *The ABC of Gender Equality in Education: Aptitude, Behaviour, Confidence*. Organisation for Economic Co-operation and Development. https://www.oecd-ilibrary.org/education/the-abc-of-gender-equality-in-education_9789264229945-en

OECD. (2017). *The under-representation of women in STEM fields* (pp. 105–112). OECD. https://doi.org/10.1787/9789264281318-10-en

OECD. (2020). *Education Policy Outlook in Germany*. OECD. https://doi.org/10.1787/47b795b1-en

Ojose, B. (2008). Applying Piaget's Theory of Cognitive Development to Mathematics Instruction. *THE MATHEMATICS EDUCATOR*, *18*(1), Article 1. https://ojs01.galib.uga.edu/tme/article/view/1923

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

Parker, P. D., Marsh, H. W., Ciarrochi, J., Marshall, S., & Abduljabbar, A. S. (2014). Juxtaposing math self-efficacy and self-concept as predictors of long-term achievement outcomes. *Educational Psychology*, *34*(1), 29–48. https://doi.org/10.1080/01443410.2013.797339

Passolunghi, M. C., Rueda Ferreira, T. I., & Tomasetto, C. (2014). Math–gender stereotypes and math-related beliefs in childhood and early adolescence. *Learning and Individual Differences*, *34*, 70–76. https://doi.org/10.1016/j.lindif.2014.05.005

Passolunghi, M., Cargnelutti, E., & Pastore, M. (2014). The contribution of general cognitive abilities and approximate number system to early mathematics. *BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY*, *84*(4), 631–649. https://doi.org/10.1111/bjep.12054

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Pizzie, R. G., & Kraemer, D. J. M. (2017). Avoiding math on a rapid timescale: Emotional responsivity and anxious attention in math anxiety. *Brain and Cognition*, *118*, 100–107. https://doi.org/10.1016/j.bandc.2017.08.004

Pizzie, R. G., & Kraemer, D. J. M. (2021). The Association Between Emotion Regulation, Physiological Arousal, and Performance in Math Anxiety. *Frontiers in Psychology*, *12*, 639448. https://doi.org/10.3389/fpsyg.2021.639448

Quaiser-Pohl, C. (2003). The Mental Cutting Test 'Schnitte' and the Picture Rotation Test-Two New Measures to Assess Spatial Ability. *International Journal of Testing*, *3*(3), 219–231. https://doi.org/10.1207/S15327574IJT0303_2

Quaiser-Pohl, C., & Endepohls-Ulpe, M. (Eds.). (2012). *Women's choices in Europe. Influence of gender on education, occupational career and family development.* Waxmann.

Quaiser-Pohl, C., Endepohls-Ulpe, M., & Meyer, C. (2014, July 3). *How does mentoring with female pupils work? Results of multivariate analyses of the effects of a German* *mentoring program in the STEM field.* 2nd Gender and STEM Conference, Berlin, Germany.

Quaiser-Pohl, C., Neuburger, S., Heil, M., Jansen, P., & Schmelter, A. (2014). Is the Male Advantage in Mental-Rotation Performance Task Independent? On the Usability of Chronometric Tests and Paper-and-Pencil Tests in Children. *International Journal of Testing*, *14*(2), 122–142. https://doi.org/10.1080/15305058.2013.860148

Raabe, I. J., & Block, P. (2024). The gendered maths confidence gap, social influence and social integration. *European Societies*, *0*(0), 1–36. https://doi.org/10.1080/14616696.2024.2349217

Rahe, M., & Quaiser-Pohl, C. (2021). Can (perceived) mental-rotation performance mediate gender differences in math anxiety in adolescents and young adults? *Mathematics Education Research Journal*. https://doi.org/10.1007/s13394-021-00387-6

Ramirez, G., Gunderson, E. A., Levine, S. C., & Beilock, S. L. (2012). Spatial anxiety relates to spatial abilities as a function of working memory in children. *The Quarterly Journal of Experimental Psychology*, *65*(3), 474–487. https://doi.org/10.1080/17470218.2011.616214

Rocha, K., Lussier, C. M., & Atit, K. (2022). What makes online teaching spatial? Examining the connections between K-12 teachers' spatial skills, affect, and their use of spatial pedagogy during remote instruction. *Cognitive Research: Principles and Implications*, *7*. https://doi.org/10.1186/s41235-022-00377-7

Ruthsatz, V., Neuburger, S., Jansen, P., & Quaiser-Pohl, C. (2014). 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 C. Freksa, B. Nebel, M. Hegarty, & T. Barkowsky (Eds.), *Spatial Cognition IX* (Vol. 8684, pp. 370–382). Springer International Publishing. https://doi.org/10.1007/978-3-319-11215-2_26

Ruthsatz, V., Neuburger, S., Jansen, P., & Quaiser-Pohl, C. (2015). Cars or dolls? Influence of the stereotyped nature of the items on children’s mental-rotation performance. *Learning and Individual Differences*, *43*, 75–82. https://doi.org/10.1016/j.lindif.2015.08.016

Ruthsatz, V., Neuburger, S., Rahe, M., Jansen, P., & Quaiser-Pohl, C. (2017). The gender effect in 3D-Mental-rotation performance with familiar and gender-stereotyped objects—A study with elementary school children. *JOURNAL OF COGNITIVE PSYCHOLOGY*, *29*(6), 717–730. https://doi.org/10.1080/20445911.2017.1312689

Ruthsatz, V., Rahe, M., Schurmann, L., & Quaiser-Pohl, C. (2019). Girls’ Stuff, boys’ stuff and mental rotation: Fourth graders rotate faster with gender-congruent stimuli. *JOURNAL OF COGNITIVE PSYCHOLOGY*, *31*(2), 225–239. https://doi.org/10.1080/20445911.2019.1567518

Ruthsatz, V., S., N., & Quaiser-Pohl, C. (2013). The social relevance and the socio-cultural origins of gender differences in spatial abilities. *Folia Sociologica*, *43*, 17–32.

Ruthsatz, V., Saunders, M., Lennon-Maslin, M., & Quaiser-Poh, C. M. (Under Review). Male? Female? Neutral! Using Novel Polyhedral Figures as Gender-Neutral Stimuli in a Mental-Rotation Test. *Journal of Learning and Individual Differences*.

Sanchis-Sanchis, A., Grau, M. D., Moliner, A.-R., & Morales-Murillo, C. P. (2020). Effects of Age and Gender in Emotion Regulation of Children and Adolescents. *Frontiers in Psychology*, *11*, 946. https://doi.org/10.3389/fpsyg.2020.00946

Saunders, M., & Quaiser-Pohl, C. M. (2020). Identifying solution strategies in a mental-rotation test with gender-stereotyped objects by analyzing gaze patterns. *Journal of Eye Movement Research*, *13*(6), Article 6. https://doi.org/10.16910/jemr.13.6.5

Schmader, T. (2010). Stereotype Threat Deconstructed. *Current Directions in Psychological Science - CURR DIRECTIONS PSYCHOL SCI*, *19*, 14–18. https://doi.org/10.1177/0963721409359292

Schmader, T., Johns, M., & Forbes, C. (2008). An Integrated Process Model of Stereotype Threat Effects on Performance. *Psychological Review*, *115*(2), 336–356. https://doi.org/10.1037/0033-295X.115.2.336

Schmand, B. (2019). Why are neuropsychologists so reluctant to embrace modern assessment techniques? *The Clinical Neuropsychologist*, *33*(2), 209–219. https://doi.org/10.1080/13854046.2018.1523468

SellSTEM MSCA ITN. (2021). SellSTEM MSCA ITN. https://sellstem.eu/

Shepard, S., & Metzler, D. (1988). Mental rotation: Effects of dimensionality of objects and type of task. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(1), 3–11.

Space Science Institute. (2023). Tweens and STEM. *STAR Library Network*. https://www.starnetlibraries.org/resources/tweens-and-stem/

Starr, C. R. (2018). “I’m Not a Science Nerd!”: STEM Stereotypes, Identity, and Motivation Among Undergraduate Women. *Psychology of Women Quarterly*, *42*(4), 489–503. https://doi.org/10.1177/0361684318793848

Starr, C. R., & Leaper, C. (2019). Do adolescents’ self-concepts moderate the relationship between STEM stereotypes and motivation? *Social Psychology of Education*, *22*(5), 1109–1129. https://doi.org/10.1007/s11218-019-09515-4

Steele, C., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.69.5.797

Stokes, A. F., & Kite, K. (2017). *Flight Stress: Stress, Fatigue and Performance in Aviation*. Routledge. https://doi.org/10.4324/9781315255200

Szczygiel, M. (2020). Gender, General Anxiety, Math Anxiety and Math Achievement in Early School-Age Children. In *Issues in Educational Research* (Vol. 30, Issue 3, pp. 1126-1142 (17 Seiten)). Western Australian Institute for Educational Research Inc.

5/202 Coode Street, Como, Western Australia 6152, Australia. e-mail: editor@iier.org.au; Web site: http://www.waier.org.au.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics: Pearson New International Edition* (6th Edition). Microsoft Press.

Taragin, D., Tzuriel, D., & Vakil, E. (2019). Mental Rotation: The Effects of Processing Strategy, Gender and Task Characteristics on Children's Accuracy, Reaction Time and Eye Movements' Pattern. *Journal of Eye Movement Research*, *12*(8), 10.16910/jemr.12.8.2. https://doi.org/10.16910/jemr.12.8.2

Teigen, K. H. (1994). Yerkes-Dodson: A Law for all Seasons. *Theory & Psychology*, *4*(4), 525–547. https://doi.org/10.1177/0959354394044004

Thammasan, N., Stuldreher, I. V., Schreuders, E., Giletta, M., & Brouwer, A.-M. (2020). A Usability Study of Physiological Measurement in School Using Wearable Sensors. *Sensors (Basel, Switzerland)*, *20*(18), 5380. https://doi.org/10.3390/s20185380

Tyng, C. M., Amin, H. U., Saad, M. N. M., & Malik, A. S. (2017). The Influences of Emotion on Learning and Memory. *Frontiers in Psychology, 8*. https://www.frontiersin.org/articles/10.3389/fpsyg.2017.01454

Uttal, D., & Cohen, C. (2012). Spatial thinking and STEM education: When, why, and how? *Psychology of Learning and Motivation, 1*. https://doi.org/10.1016/B978-0-12-394293-7.00004-2

Van Mier, H. I., Schleepen, T. M. J., & Van den Berg, F. C. G. (2019). Gender Differences Regarding the Impact of Math Anxiety on Arithmetic Performance in Second and Fourth Graders. *Frontiers in Psychology*, *9*, 2690. https://doi.org/10.3389/fpsyg.2018.02690

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills, 47*(2), 599–604. https://doi.org/10.2466/pms.1978.47.2.599

Vogt, W. P., Gardner, D. C., & Haeffele, L. M. (2012). *When to Use What Research Design*. Guilford Press.

Voyer, D. (2011). Time limits and gender differences on paper-and-pencil tests of mental rotation: A meta-analysis. *Psychonomic Bulletin & Review*, *18*(2), 267–277. https://doi.org/10.3758/s13423-010-0042-0

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin*, *117*(2), 250–270.

Walton, G. M., & Cohen, G. L. (2003). Stereotype lift. *Journal of Experimental Social Psychology*, *39*(5), 456–467. https://doi.org/10.1016/S0022-1031(03)00019-2

Wolff, F. (2021). How Classmates' Gender Stereotypes Affect Students' Math Self-Concepts: A Multilevel Analysis. *Frontiers in Psychology*, *12*, 599199. https://doi.org/10.3389/fpsyg.2021.599199

World Economic Forum. (2019). *The Global Gender Gap Report 2020*. https://gdc.unicef.org/resource/global-gender-gap-report-2020

World Medical Association. (2013). World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA*, *310*(20), 2191–2194. https://doi.org/10.1001/jama.2013.281053

Zhu, C., Leung, C. O.-Y., Lagoudaki, E., Velho, M., Segura-Caballero, N., Jolles, D., Duffy, G., Maresch, G., Pagkratidou, M., & Klapwijk, R. (2023). Fostering spatial ability development in and for authentic STEM learning. *Frontiers in Education*, *8*. https://www.frontiersin.org/articles/10.3389/feduc.2023.1138607

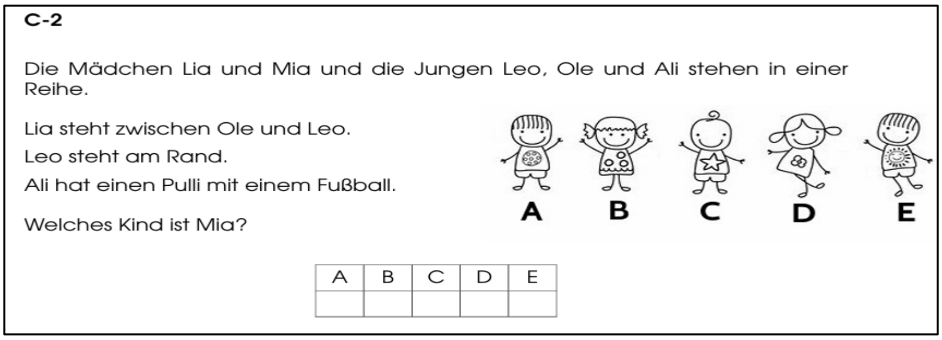
Živković, M., Pellizzoni, S., Doz, E., Cuder, A., Mammarella, I., & Passolunghi, M. C. (2023). Math self-efficacy or anxiety? The role of emotional and motivational contribution in math performance. *Social Psychology of Education*, *26*(3), 579–601. https://doi.org/10.1007/s11218-023-09760-8

8 Appendices

A.1 Measures of Maths and Spatial Constructs

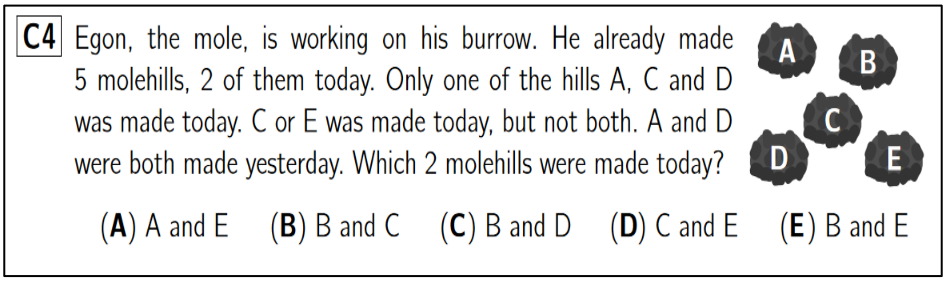
Sample Problems from Maths Task

Sample number line estimation problem for 2nd grade students.
Difficulty level: High

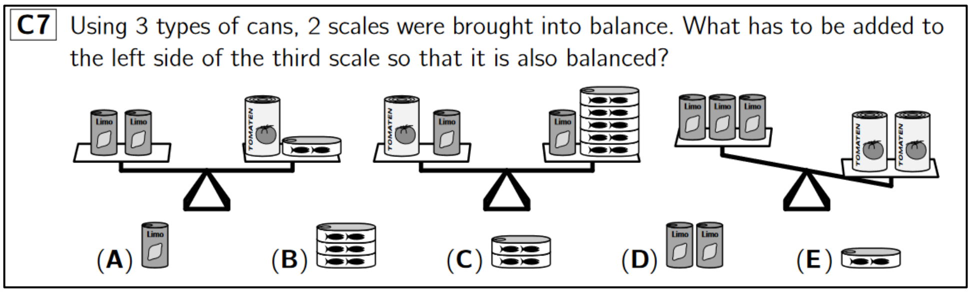


Translation from German: “The girls Lia and Mia and the boys Leo, Ole and Ali are standing in a row. Lia is standing between Ole and Leo. Leo is standing on the edge. Ali has a sweater with a soccer ball on it. Which child is Mia?”

Sample word problem for 3rd grade students.
Difficulty level: high



Sample missing term problem for 4th grade students.
Difficulty level high.



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

Mathematics Self-Concept Subscale (Translated from German).

My favourite/preferred subject is:
(Blank Space)

1. I find dealing with numbers....
   Very difficult.....Very easy
2. I like maths.....
   Not at all.....Very much
3. In arithmetic I am
   Very bad.....Very good
4. When calculating I make.....
   A lot of mistakes...No mistakes at all
5. Maths is.......
   No fun at all.....Lots of fun
6. In maths I am...
   Not talented at all .....Very talented

Children rated themselves on a sliding 7-point scale.

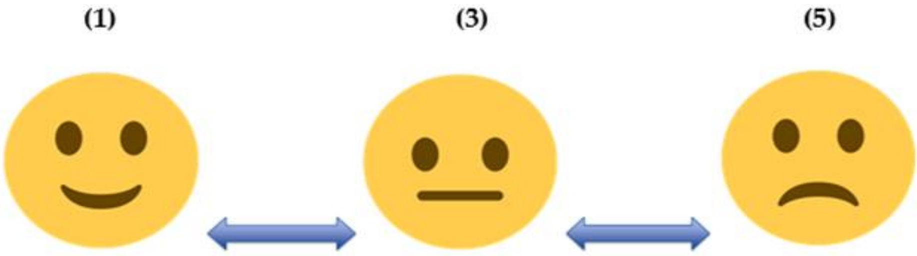
Child Spatial Anxiety Questionnaire.

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 minutes? [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.

Modified Abbreviated Math Anxiety Scale (mAMAS).

Children rated their level of math anxiety on a sliding 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.

A.2 Legal Statement

Hiermit erkläre ich, dass ich die vorliegende Dissertation selbstständig angefertigt, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und alle wörtlich oder inhaltlich entnommenen Stellen kenntlich gemacht habe. Diese Dissertation wurde noch in keinem anderen Verfahren zur Erlangung des Doktorgrades oder als Prüfungsarbeit für eine akademische oder staatliche Prüfung eingereicht.

Koblenz, den 12th Sept. 2024 Michelle Lennon-Maslin

Michelle Lennon-Maslin

A.3 Publications

social sciences

Article

Under My Skin: Reducing Bias in STEM through New Approaches to Assessment of Spatial Abilities Considering the Role of Emotional Regulation

Michelle Lennon-Maslin *, Claudia Michaela Quaiser-Pohl, Vera Ruthsatz and Mirko Saunders

Department of Developmental Psychology and Psychological Assessment, Faculty of Educational Sciences Institute of Psychology, University of Koblenz, 56070 Koblenz, Germany; quaiser@uni-koblenz.de (C.M.Q.-P.); ruthsatz@uni-koblenz.de (V.R.); mirkosaunders@uni-koblenz.de (M.S.)
* Correspondence: mlennonm@uni-koblenz.de

Abstract: Reducing gender bias in STEM is key to generating more equality and contributing to a more balanced workforce in this field. Spatial ability and its components are cognitive processes crucial to success in STEM education and careers. Significant gender differences have consistently been found in mental rotation (MR), the ability to mentally transform two- and three-dimensional objects. The aim of this pilot study is to examine factors in psychological assessment which may contribute to gender differences in MR performance. Moreover, findings will inform the development of the new approaches to assessment using computer adaptive testing (CAT). (1) Background: The study examines the impact of emotional regulation on MR performance in primary school children whose mean age was 9.28 years old. (2) Methods: Skin conductance was measured to assess the impact of emotional reactivity (ER) on performance during an MR task. (3) Results: Patterns of ER influence response time (RT) on specific items in the task. (4) Conclusions: Identifying the effects of emotional arousal and issues of test construction such as stereotyped stimuli and item difficulty in tests of spatial ability warrants ongoing investigation. It is vital to ensure that these factors do not compromise the accurate measurement of performance and inadvertently contribute to the gender gap in STEM.

Keywords: gender stereotypes; gender gap; STEM; spatial abilities; emotional regulation; skin conductance; CAT

check for updates

Citation: Lennon-Maslin, Michelle, Claudia Michaela Quaiser-Pohl, Vera Ruthsatz, and Mirko Saunders. 2023. Under My Skin: Reducing Bias in STEM through New Approaches to Assessment of Spatial Abilities Considering the Role of Emotional Regulation. *Social Sciences* 12: 356. https://doi.org/10.3390/socsci12060356

Academic Editors: Sabrina Sobieraj and Barbara Fawcett

Received: 6 April 2023
Revised: 12 June 2023
Accepted: 12 June 2023
Published: 16 June 2023

1. Introduction

Despite globalization, gendered career choices persist worldwide (Makarova et al. 2019). The World Economic Forum (2019) estimates that currently only one third of female students pursue higher education or research careers in the fields of science, technology, engineering and mathematics (STEM) (World Economic Forum 2019).

The current study, being undertaken at the University of Koblenz, Germany, is funded by the EU-financed (Horizon 2020) research network SellSTEM (Spatially Enhanced Learning linked to STEM). Spatial ability and its components are cognitive processes crucial to success in the STEM arena (Newcombe 2017). The SellSTEM network, consisting of 15 PhD candidates and their supervisors from 10 universities across Europe as well as 8 partners, is investigating the role of spatial ability in STEM learning.

The rationale for this research is based on three key findings: (1) too few young people are enrolling in STEM courses (Maselli and Beblavý 2014); (2) in scientific fields such as psychology, life and social sciences, women are present in much higher numbers; and (3) there isa widely-agreed-upon consensus that in undergraduate and postgraduate university programs, which are the most mathematically intensive, such as geoscience, engineering, economics, mathematics/computer science and the physical sciences, women are still underrepresented (Ceci et al. 2014). Furthermore, the percentage of women in some

STEM faculties, for example biomedical engineering, at the assistant, associate and full professor levels, remains low (Chesler et al. 2010).

Although gender differences on spatial tests on a broader scale have been found to be small or less consistent, there are subsets or components of spatial ability which continue to yield significant differences in favor of males (Voyer et al. 1995; Linn and Petersen 1985).

The aim of the four-year project is to raise spatial ability in children, and girls in particular, in a number of European countries, that is, Latvia, Norway, Ireland, Germany, The Netherlands, Sweden, Austria and the UK, so that they are better prepared for the cognitive demands of STEM education. Ultimately, the goal of the working group is to promote more successful STEM learning, triggering migration in larger numbers toward STEM careers and consequently generating a more gender-balanced ratio in the field (SellSTEM MSCA ITN 2021). A specific goal of the SellSTEM subproject based at the University of Koblenz is to provide primary school teachers across Europe with effective ways of assessing spatial ability development in children with the objective that such methods will assist them in identifying potential difficulties in STEM learning, thus facilitating early intervention.

1.1. Mental Rotation Tests (MRTs): Assessing a Spatial Ability

Mental rotation (MR), a component of spatial ability, has been studied extensively in psychology and education due to significant gender differences found in the ability to rotate mental representations of two- and three-dimensional objects (Neuburger et al. 2012b). Despite a small amount of evidence on cognitive disparity in spatial ability as a whole, MR tasks, particularly those incorporating three-dimensional objects, consistently yield large and reliable differences in performance favoring males, with no significant reduction (Wraga et al. 2006).

MR is traditionally assessed using psychometric instruments (often referred to as paper-and-pencil tests) such as the mental rotation test (VMRT; Vandenberg and Kuse 1978). In a bid to modernize psychometric assessment, many of these tests have become digitalized (Schmand 2019). During a computerized MRT, participants respond to each item or series of stimuli on a computer or touch-screen device (Monahan et al. 2008). Stimuli usually take the form of pictorial representations of abstract objects, such as cubical or pellet-shaped figures, and concrete objects such as animals or toys (Rahe and Quaiser-Pohl 2021). The stimuli are rotated in-depth or in picture-plane.

Recently, researchers examining MR testing have turned their attention to issues which may influence participants' performance. Due to the aforementioned gender differences, a number of studies have focused on how MR test characteristics, item and stimulus attributes might contribute to this phenomenon, e.g., time limitations (Rahe and Quaiser-Pohl 2021), item difficulty and rotational axis (Neuburger et al. 2012b), task-solving strategies (Saunders and Quaiser-Pohl 2020) and gender-stereotyped stimuli (Ruthsatz et al. 2015).

Performance on tests of spatial ability and other cognitive tasks is known to be influenced by the test-taker's experience and their emotional state (Shepard and Feng 1972; Johns et al. 2008). Moreover, findings demonstrate that emotion-regulating processes reduce working memory, and executive resources needed to perform well on tests of cognitive ability (Schmader et al. 2008). For instance, Ramirez and colleagues (2012) found that some younger children reported experiencing a phenomenon referred to as spatial anxiety (SA), that is, feeling anxious at the prospect of engaging in spatial activities. Furthermore, pronounced gender differences in SA and self-confidence have been found to mediate gender differences in MR performance, especially when task demands are high (Arrighi and Hausmann 2022).

Therefore, this pilot study is a first step in the development of new approaches to MR assessment in children. These approaches may reduce the negative effects of emotions such as anxiety and have a positive influence on self-confidence in participants. The study thereby examines factors beyond mental rotation itself, such as emotional regulation, stereotype threat and task and item difficulty, which are known to contribute to gender differences in performance on an MRT (Sanchis-Segura et al. 2018; Fladung and Kiefer 2016; Caissie et al. 2009).

1.2. Computer Adaptive Tests of MR (CAT-MR): An Alternative to Fixed Item Tests (FITs)

Computer adaptive testing (CAT) is an approach to psychometrics, which establishes a link between the participant's ability, their response to items and the underlying trait being measured (Veldkamp and Verschoor 2019). During CAT, a computer algorithm automatically presents items to participants and selects the next item based on their previous response. As opposed to fixed item testing (FIT), in which participants respond to the same set of items in the same order (Vispoel 1993), CAT adapts to their performance (Linden et al. 2000). Moreover, CAT has been found to elicit less participant anxiety (Fritts and Marszalek 2010; Ortner and Caspers 2011) and female primary school students achieved better results, reported a higher sense of motivation and a more positive subjective test experience after CAT (Martin and Lazendic 2018). In mathematics, CAT revealed promising results regarding the reduction in stress and anxiety, and a better overall performance among test candidates (Eggen and Verschoor 2006).

Moreover, findings demonstrate that CAT offers teachers many advantages such as convenience and flexibility, faster, more accurate scoring and reporting, potentially shorter tests, reduced scheduling and supervision, fewer test items needed to accurately estimate proficiency and less time needed for marking (Chuesathuchon 2008).

1.3. The Effects of Stereotype Threat on Cognitive Performance

More than a decade of research has now confirmed that experiencing negative stereotypes about one's social group can impair an individual's ability to achieve their potential (Schmader 2010). This effect is referred to as stereotype threat (ST) or situations in which individuals perceive themselves to be at risk of conforming to negative stereotypes about their ingroup (Steele and Aronson 1995). The beneficial effect of perceiving oneself as conforming to positive stereotypes about one's ingroup is known as stereotype lift (SL) (Walton and Cohen 2003).

ST has been found to arouse anxiety in targets, and in the process of emotional regulation, cognitive resources needed to successfully execute a task are thereby limited (Schmader et al. 2008). Consequently, performance on MR tests is susceptible to the effects of ST (Neuburger et al. 2012a). SL, on the other hand, may alleviate self-doubt, anxiety and fear of rejection that could otherwise hinder performance on MR and other cognitive tests (Walton and Cohen 2003).

Gender-stereotyped beliefs influence boys' and girls' preferences and behaviors early in life. This can often determine their choice of clothes and hairstyles and can also influence their preference for certain toys (Ruthsatz et al. 2019). Boys' preference for construction toys and other games that are related to object manipulation in space might lead to more practice in these skills than playing with stereotypically female toys, such as, dolls and cuddly toys (Newcombe and Frick 2010). This increased familiarity with construction-related play might be advantageous to male participants on an MR task consisting of items with the frequently used cubical figure (Kersh et al. 2008).

More recently, some studies have focused on the gender-stereotyped attributes of the stimuli used in MR tests and how they might impact a test-taker's performance (Rahe and Quaiser-Pohl 2019). Findings suggest that the degree to which stimulus objects are familiar to male versus female participants, that is, are gender-congruent, is a significant determinant of the gender difference (Neuburger et al. 2011). Moreover, a positive correlation between stereotyped stimulus content and children's performance on tests of MR has been established (Ruthsatz et al. 2014).

1.4. Assessing the Impact of Emotional Regulation on MR in Children

Performance on MR and other cognitive tasks is known to be heavily influenced by the participant's experience (Shepard and Feng 1972) and emotional state (Schmader et al.

2008). Various studies have examined the source of emotional arousal experienced by participants undertaking cognitive tests such as an MRT. As previously mentioned, the effects of stereotype threat can negatively impact cognitive performance (Johns et al. 2008), as can the mode of testing, e.g., FIT or CAT (Vispoel 1993). Additionally, anticipatory feelings, such as spatial anxiety, regarding the kinds of activities a test may involve are important factors which mediate performance (Ramirez et al. 2012).

Human cognitive functions such as perception, attention, memory and problem-solving skills are all significantly influenced by emotion (Tyng et al. 2017). Events or stimuli which evoke emotion are classified according to two main categories, valence and arousal (Costanzi et al. 2019). On a continuum ranging from negative to neutral to positive, the attractiveness or aversiveness of a stimulus or event is referred to as valence. An event or stimulus intensity is known as arousal when it ranges from being intensely calming to intensely exhilarating or agitating (Costanzi et al. 2019). Moreover, how an event or stimulus is evaluated plays a role in judgment and decision-making. Positive emotion surrounding an event or stimulus enhances thinking and reasoning and adversely negative emotion impairs these functions (Storbeck and Clore 2008). The relevance and salience of a stimulus or event to the observer also determines how it is prioritized. Emotional arousal enhances cognition for high-priority information and impairs it for that of low-priority (Turkileri et al. 2021).

Tests and examinations, for example, are often perceived as events which can elicit a variety of emotional states, including excitement, frustration, anxiety and even boredom (Tyng et al. 2017). As is evident from emotional experiences such as spatial anxiety, the emotional reaction to the subject matter itself can impair an individual's test performance (Tyng et al. 2017). Similarly, the stimuli which make up a test item can impact valence and arousal in ways that can either enhance or impair cognitive performance on that test (Costanzi et al. 2019). If these stimuli are not relevant to the observer, they will be assigned low priority; hence, cognition may be impaired by a lack of emotional arousal. Furthermore, individual differences in emotional regulation strategies have also been found to influence test performance. For example, in one study, the strategies used by males and females to regulate emotions during an MR test resulted in gender differences in their performance therein (Fladung and Kiefer 2016).

Investigating emotional experience arising in the course of an MR task could also be measured physiologically. Assessing somatic activity such as skin conductance and heart rate could provide valuable information with regard to differences in emotion expressive behavior, valence and arousal impacting performance on such tasks (Deng et al. 2016). Measurements of children's electrodermal and cardiovascular activity can be used as indicators of the autonomic nervous system's (ANS) activation during emotion-evoking events or when experiencing stimuli which may impact participants emotionally (Sohn et al. 2001).

Skin conductance responses (SCRs) are biomarkers of ANS arousal and are a well-established method for measuring psychophysiological functioning in humans (Christopoulos et al. 2019). These signals, stemming from the peripheral nervous system (PNS), have been long identified as important for mental functions, in particular emotions (James 1890). Furthermore, these somatic responses are likely an essential part of the emotional experience and act as cues based on which they are formed (Damasio 2001).

Galvanic skin response (GSR), a measure of skin conductance, has been successfully employed to measure physiological changes as a result of emotional expression such as anxiety and stress in children (Najafpour et al. 2017). Furthermore, the use of wearable devices to measure GSR can facilitate the analysis of physiological responses in children engaged in cognitive tasks in real-life conditions such as in the classroom (Geršak et al. 2019).

1.5. The Current Pilot Study

Using an experimental design, the goal of this study is to examine the impact of emotional regulation on spatial performance in primary school children during a computerized

MR test. It is anticipated that findings will also provide useful information about item characteristics and precedes the development of a computer adaptive test of mental rotation (CAT-MR).

1.5.1. Research Questions

Based on findings from previous research, this study aims to answer the following questions:

1. Will there be changes in participants' skin conductance in the course of a mental rotation task and will this be impacted by the difficulty of the task?
2. Will skin conductance levels influence participants' accuracy and reaction time on the MR task? Will participants who take more time demonstrate more accuracy on the MR task?
3. Will there be patterns in participants' speed and accuracy which relate to the item difficulty, i.e., stimulus type and rotational axis on a MR task?
4. Regarding responses on a stereotype nature of stimuli and a perceived difficulty questionnaire, will participants identify stereotyped and difficult stimuli and will their responses correspond with their performance on the task?

1.5.2. Hypotheses

Hypotheses 1. *Previous research has found that emotion regulation is an important factor for maintaining MR performance (Fladung and Kiefer 2016). Therefore, we hypothesize that there will be an increase in participants' skin conductance levels (SCL) measured by GSR during the MR task from baseline to subtest MRT1 and subtest MRT2.*

Hypotheses 2. *Item difficulty on an MR task is determined by stimulus type, that is, concrete or abstract objects, and rotational axis, that is, stimuli rotated in-depth or in picture-plane. Items containing abstract objects rotated in-depth are known to be more difficult in an MR task (Neuburger et al. 2012b). Furthermore, failure to effectively regulate negative emotional states, elicited, for example, by phenomena such as stereotype threat or spatial anxiety, can lead to a poorer performance on difficult cognitive tasks (Schmader 2010; Ramirez et al. 2012; Fladung and Kiefer 2016). Moreover, participants who spend more time and put more effort into solving difficult items on an MR test usually demonstrate better performance (Liesefeld et al. 2015). Therefore, we hypothesize that increased SCL measured by GSR and a longer reaction time (RT) predict higher scores on more difficult items, i.e., abstract stimuli rotated in-depth, on the MR task.*

Hypotheses 3. *Items with abstract objects rotated in-depth and gender-stereotyped objects can influence participants' accuracy and reaction time on an MR task. Mean accuracy scores are often higher on items with gender-congruent stimuli and lower on items with abstract stimuli rotated in-depth (Ruthsatz et al. 2019). Due to the difficulty level, reaction time is often longer on items with abstract stimuli rotated in-depth (Liesefeld et al. 2015). Furthermore, interaction between stimuli and emotional arousal and valence may account for some of these differences (Costanzi et al. 2019). Therefore, we hypothesize that there will be differences in accuracy scores and reaction time between two SCL groups on items with concrete and abstract stimuli rotated in picture-plane and in-depth and on three gendered objects, that is, male-stereotyped, female-stereotyped and gender-neutral objects, during the MR task.*

Hypotheses 4. *Long before an awareness of their own gender identity commences, children have already developed a schema of gender-associated traits and gender categories (Martin et al. 2002). Therefore, we hypothesize that participants will identify stimuli containing gender-stereotyped objects as such, that is, male-stereotyped objects as masculine and female-stereotyped objects as feminine.*

Hypotheses 5. *In line with previous findings, participants will identify stimuli containing concrete objects rotated in picture-plane as easier and abstract objects rotated in-depth as more difficult (Ruthsatz et al. 2017).*

2. Materials and Methods

2.1. Participants

For the pilot study, we recruited 29 students from third- and fourth-grade classes at two local primary schools in Koblenz, Rheinland Palatine, Germany. One student withdrew consent immediately prior to testing and three of the task data collected could not be retrieved from the devices. Therefore, the total number of participants whose data were used in the final analysis was 25 (N = 25). There were 12 children who identified as boys and 12 as girls and 1 participant who did not provide a response to the gender question. The average age of the students was 9.28 years old (M = 9.28). All 29 parents and guardians provided written informed consent.

2.2. Material and Instrumentation

2.2.1. Mental Rotation Task

A computerized MR task (Vandenberg and Kuse 1978) was programmed in PsychoPy® software (Supplementary Materials) 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 time taken to answer each item (reaction time).

Items included MR stimuli for younger children, i.e., animals, letters and cubes (Quaiser-Pohl 2003) as well as abstract and concrete stimuli rotated in-depth or in picture-plane. Abstract items consisted of stimuli such as cubes, pellets (Ruthsatz et al. 2014) and polyhedra (Ruthsatz et al., forthcoming). Concrete items consisted of male and female gender-stereotyped stimuli (Ruthsatz et al. 2015) and gender-neutral stimuli (Ruthsatz et al., forthcoming). Examples of some of the items used are shown in Figure 1.

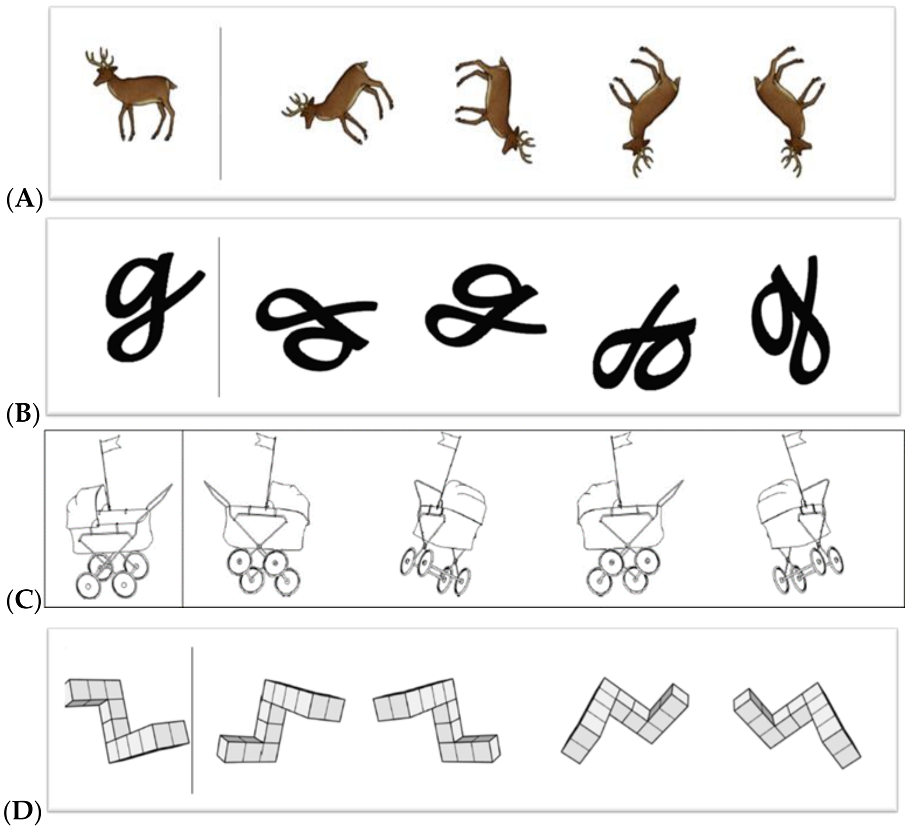


Figure 1. Examples of MRT items used with concrete objects: (**A**) animal stimulus and (**B**) letter stimulus (Hinze and Quaiser-Pohl 2003); (**C**) female-stereotyped stimulus "Pram" (Jansen et al. 2014; Ruthsatz et al. 2015) and with an abstract object (**D**) "Cube" figure (Ruthsatz et al. 2014).

The task was divided into two parts, each with a time limit: part one (MRT 1), considered an easier task due to stimuli rotated in picture-plane only, was limited to 5 min and part two (MRT 2), considered more difficult as items consisted of stimuli rotated in-depth, allowed participants 8 min to complete. MRT 1 had 6 abstract and 10 concrete items and MRT 2 had 6 abstract and 6 concrete items (See Appendix C). MRT 2 was also considered more difficult due to stimuli features, their complexity as well as rotational axis (Neuburger et al. 2015). 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.

2.2.2. Gender-Stereotype Nature and Perceived Difficulty of Stimuli Questionnaires

Two self-reported questionnaires were also created in PsychoPy® and presented at the end of the MRT to assess the gender-stereotyped nature of stimuli and the stimuli-perceived difficulty.

Both scales are adapted from the Stereotyped Nature of Stimuli questionnaire (Neuburger et al. 2015). The second questionnaire uses emojis on a sliding 5-point scale in order to assess perceived difficulty of stimuli. The first point on the scale (1) represents easy or happy face emoji, and the fifth (5) difficult or sad face emoji. The third point (3) represents neither easy nor difficult or neutral face emoji.

The first point (1) on the stereotype questionnaire represents the rating more for boys, and the fifth (5) more for girls. The third point (3) is a gender-neutral rating. Both scales are appended to this report.

2.2.3. Demographic Data

An online questionnaire was presented to each participant at the beginning of the experiment to collect data relating to participants' age and gender.

2.2.4. Skin Conductance

Shimmer3 GSR+ Unit® was used to measure galvanic skin response (GSR). These devices were synchronized with ConsensysBasic® multi-sensor management software where they were calibrated for recording and a sampling rate of 5 Hz was set. A baseline recording of 2 min per participant was planned in order to compare this with GSR during the MR task. Only 24 GSR datasets in total were analyzed as 1 dataset was missing from the Shimmer device. Furthermore, in three datasets at baseline was not recorded on the device; therefore, they could not be included in the analysis of SCL across the three conditions—baseline, MRT 1 and MRT 2.

For the purposes of analyses, two SCL-level groups were created—low and high GSR—during the MR task. SCL levels vary individually in humans; therefore, following checks for normality and removal of outliers on the SCL variable, descriptive statistics were calculated in SPSS as follows: minimum SCL during the MR task was 0.01 µS per minute, maximum was 12.64 µS and mean SCL was 4.69 µS per minute (M = 4.69, SD = 2.99). Therefore, all values < 4.69 µS were categorized as low SCL (N = 10) and all values > 4.69 µS per minute were categorized as high SCL (N = 11) during the MR task. Skin conductivity is measured in units referred to as microsiemens (µS). "Micro" is a prefix meaning millionths, so 1 microsiemen (1 µS) is a unit of time in the International System of Units (SI) equal to one millionth of a siemen (Braithwaite et al. 2013).

2.3. Procedure

Approval for the pilot study was provided by the Ethics Committee of the University of Koblenz and also by the state authorities in Rheinland Palatine overseeing schools. Informed consent was sought and provided by parents and guardians of all students involved in the study. The class teacher and the principal also permitted the study to be conducted in the school.

The students were tested by two female researchers in a separate classroom with access to their teacher, if required. The room had adequate lighting and individual seating arrangements.

The researchers explained the MR task to the students by rotating objects such as a pair of scissors, a toy and a wooden object, while explaining that by turning this object around, it does not change its features. Students were then asked to imagine the object in their mind and when they could see it, try to rotate the object mentally. The purpose of the study and the significance of mental rotation in everyday life and for school work was also explained to the students. 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.

Electrodes for measurement of GRS were attached to the fore- and index finder of the non-dominant hand of each participant. Shimmer devices were then switched on and synchronized with Consensys® software. Devices remained in their respective docks until the experiment could be initialized. Participants were advised to try not to move the hand to which the electrodes were attached.

Shimmer devices were then removed from each dock and a two-minute baseline measurement was recorded, after which the computerized experiment was initialized. Students were given time to read the on-screen instructions with the help of the researchers. After task-understanding was confirmed, an initial practice run followed, which contained three sample items providing feedback on accuracy.

When the practice run was complete, students were presented with further on-screen instructions, informing them that the first part of the test was about to begin and that there would be a 5 min time limit on this part of the MR task. They were also advised to try to work as fast and as accurately possible. In addition, students in each of the small groups were asked to wait until their peers could see a red stop sign which would appear at the end of MRT 1 before proceeding to MRT 2. The researchers kept track of the time throughout the experiment. The same procedure was repeated for the second part of the task.

At the end of the experiment, Shimmer devices were returned to their respective docks to end recording and electrodes were removed from participants' fingers. Data were labelled and imported to Consensys, then exported and saved on each respective table.

2.4. Data Analysis

Quantitative data analyses were performed on SPSS® 29 software for statistical analysis. A repeated-measures ANOVA was run to establish whether skin conductance level (SCL) differed significantly between the three test conditions—pre-test baseline, picture-plane and in-depth conditions. All participants took part in the three test conditions. The data in each of the groups were normally distributed. Sphericity was not violated and the Greenhouse–Geisser correction was reported as it is more stringent.

A Pearson coefficient correlation was used to determine whether there was a relationship between skin conductance, reaction time and scores on more difficult items, that is, items containing abstract objects rotated in-depth. Arising from this, multiple regression was run, entering those variables that produced a significant result to examine the way in which these variables relate to each other. The assumptions for using regression were checked and confirmed, i.e., the criterion variable was always continuous; the Mahalanobis distance values indicated that there were no substantial outliers; the residual scores were normally distributed and not related to the predicted values; and tolerance values did not exceed 0.2, indicating that there was no multi-collinearity.

One-way repeated-measures ANOVA were carried out to investigate the differences in skin conductance conditions across the MR task. Furthermore, accuracy and reaction time on items containing concrete and abstract stimuli rotated in-depth and in picture-plane and items containing gendered stimuli across the two SCl levels were analyzed. Observations were independent of one another and the sample was completely random. The independent variables were categorical on three and two levels and the dependent variables were continuous and scale variables. Significant outliers were removed from the dependent variables. Where sphericity was violated, the Greenhouse–Geisser correction was applied (Tabachnick and Fidell 2013).

3. Results

3.1. Skin Conductance Levels across the MR Task (H1)

A repeated-measures ANOVA using the Greenhouse–Geisser correction showed that skin conductance level (SCL) differed significantly between the three test conditions—pre-test baseline, MRT 1 (only rotations in picture-plane) and MRT 2 (only in-depth rotations) ($F(2, 40) = 9.46$, $p = 0.004$)—with a small to medium effect size ($\eta^2 = 0.241$). As a result, 24% of variation in skin conductance level can be explained by the different test conditions. See Table 1.

Table 1. Tests of within-subjects effects of skin conductance conditions (SCL levels) from pre-test baseline SCL to MRT1 and MRT2 on the MR task.

Measure: SCL_Levels							
Source		**Type III Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**	**Partial Eta Squared**
factor1	Sphericity Assumed	17.104	2	8.552	6.348	0.004	0.241
	Greenhouse–Geisser	17.104	1.752	9.764	6.348	0.006	0.241
error (factor1)	Sphericity Assumed	53.888	40	1.347			
	Greenhouse–Geisser	53.888	35.034	1.538			

Specifically, pairwise-comparisons-highlighted SCL during the MRT 2 condition was significantly higher than in the baseline condition ($M = 5.31$, $p = 0.018$, CI (95%) 0.182–2.233). Moreover, SCL in the MRT 1 condition was significantly higher than in the baseline condition ($M = 5.10$, $p = 0.009$, CI (95%) 0.233–1.701). There was no significant difference between SCL in the MRT 1 and SCL in the MRT 2 condition. Therefore, it can be concluded that SCL increases during an MR test. However, task difficulty did not have a statistically significant impact on SCL in this sample. See Table 2.

Table 2. Pairwise comparisons of skin conductance conditions (SCL levels) from pre-test baseline SCL to MRT1 and MRT2 on the MR task.

Measure: SCL_Levels						
(I) Factor1	**(J) Factor1**	**Mean Difference (I-J)**	**Std. Error**	**Sig. [a]**	**95% Confidence Interval for Difference [a]**	
					Lower Bound	**Upper Bound**
1	2	−0.962 *	0.283	0.009	−1.701	−0.223
	3	−1.207 *	0.393	0.018	−2.233	−0.182

* The mean difference is significant at the 0.05 level. [a]. Adjustment for multiple comparisons: Bonferroni.

Figure 2 graphically illustrates the mean SCL across three conditions from pre-test baseline SCL to SCL in MRT 1 and MRT 2. The error bars represent the standard error (SE), which indicates the variability of the estimated means within each condition. Those error bars representing MRT 1 and MRT 2 overlap, which suggests that the observed differences between the conditions are not statistically significant. This conclusion is supported by the results of the statistical analysis. See Figure 2.

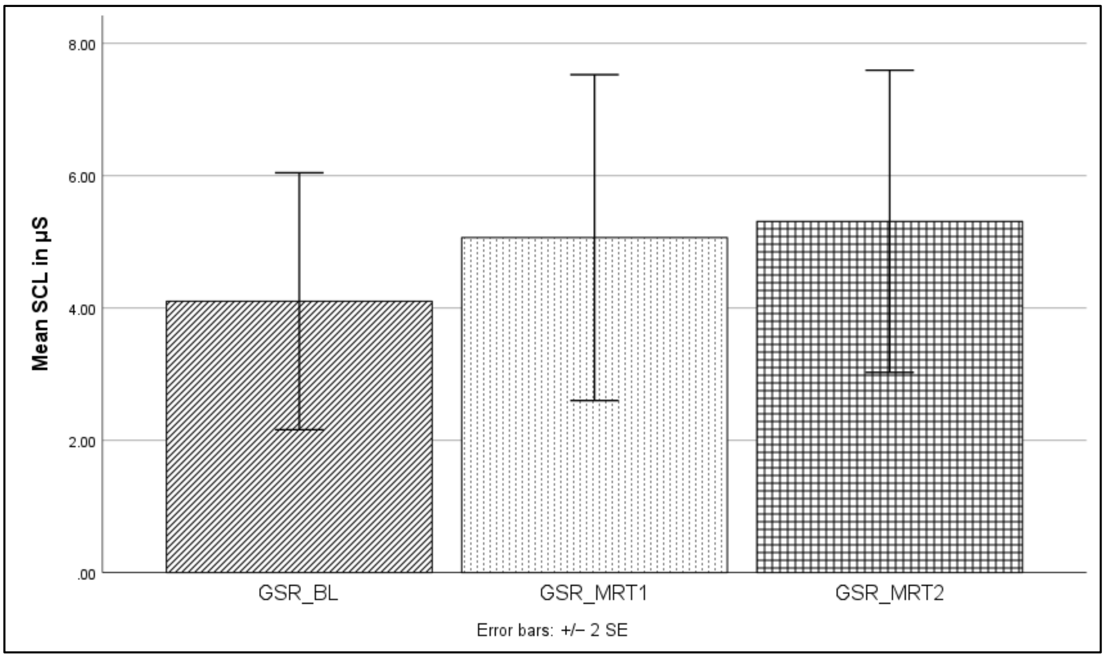


Figure 2. Differences in mean skin conductance level (SCL) across three test conditions: pre-test baseline (GSR_BL), first part of the test (GSR_MRT1) and second part of the test (GSR_MRT2).

3.2. The Relationship between Skin Conductance, Accuracy, Reaction Time and Item Difficulty (H2)

A Pearson correlation coefficient found that there was a strong positive significant relationship between SCL measured by GSR (M = 7.23, SD = 7.74) and scores on items with abstract stimuli rotated in-depth (M = 1.23, SD = 3.95) ($r(21) = 0.67$, $p < 0.001$). This relationship can account for 67% of variation in scores. Moreover, there was a strong positive correlation between participant reaction time (RT) (M = 16.83, SD = 6.50) and scores on items with abstract stimuli rotated in-depth (M = 1.23, SD = 3.95) ($r(21) = 0.512$, $p = 0.009$). This relationship can account for 51% of the variation in scores. See Table 3.

Table 3. Pearson coefficient correlations between scores on items with abstract stimuli rotated in-depth (Mean_Diff), reaction time (Mean_RT) and skin conductance (GSR_Test) during the MR task.

		Mean_Diff	**Mean_RT**	**GSR_Test**
Pearson Correlation	Mean_Diff	1.000	0.510	0.667
	Mean_RT	0.510	1.000	0.439
	GSR_Test	0.667	0.439	1.000
Sig. (1-tailed)	Mean_Diff		0.005	<0.001
	Mean_RT	0.005		0.016
	GSR_Test	0.000	0.016	
N	Mean_Diff	24	24	24
	Mean_RT	24	24	24
	GSR_Test	24	24	24

Both variables were then entered into a multiple regression, which was used to test whether participant SCL measured by GSR and reaction time (RT) were predictors of scores on items containing abstract stimuli rotated in-depth. The results of the regression indicated that the two predictors explained 43% of the variance ($R^2 = 0.46$, $F(2, 21) = 10.64$, $p < 0.001$). It was found that SCL measured by GSR predicted scores on abstract objects rotated in-depth ($\beta = 0.55$, $p = 0.004$, 95% CI = 0.10–0.47) but no statistically significant linear dependence of the mean of scores on abstract objects rotated in-depth on RT was detected. See Table 4.

Table 4. Model summary: skin conductance (GSR_Test) and reaction time (Mean_RT) predict scores on items with abstract stimuli rotated in-depth (Mean_Diff).

ANOVA [a]						
Model		**Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**
1	Regression	187.825	2	93.912	10.638	<0.001 [b]
	Residual	185.396	21	8.828		
	Total	373.221	23			

[a] Dependent variable: Mean_Diff. [b] Predictors: (constant), GSR_Test, Mean_RT.

3.3. Differences in Accuracy and Reaction Time between SCL Groups on Rotational Axis, Stimulus Type and on Gendered Objects (H3)

A one-way repeated-measures ANOVA using the Greenhouse–Geiser correction found that there was no significant differences in accuracy scores between SCL groups on concrete and abstract stimuli and stimuli rotated in picture-plane or in-depth ($F(3, 57) = 0.952$, $p = 0.444$, $\eta^2 = 0.04$). In relation to the main effect, there was a significant difference in accuracy scores on concrete and abstract stimuli and stimuli rotated in-depth and in picture-plane (F $(3, 57) = 6.55$, $p = 0.007$) with a small effect size ($\eta^2 = 0.26$). Therefore, rotational axis and stimulus type explain 26% of the variance in accuracy scores on the MR task. GSR level did not significantly explain variance in accuracy scores in this context. See Table 5.

Table 5. Tests of within-subjects effects of stimulus type and rotational axis (Rotation_Type) across two SCL groups (GSR_Levels) on accuracy on the MR task.

Measure: Accuracy							
Source		**Type III Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**	**Partial Eta Squared**
Rotation_Type	Sphericity Assumed	0.172	3	0.057	6.546	<0.001	0.256
	Greenhouse–Geisser	0.172	1.579	0.109	6.546	0.007	0.256
Rotation_Type * GSR_Levels	Sphericity Assumed	0.020	3	0.007	0.768	0.517	0.039
	Greenhouse–Geisser	0.020	1.579	0.013	0.768	0.444	0.039
Error (Rotation_Type)	Sphericity Assumed	0.500	57	0.009			
	Greenhouse–Geisser	0.500	30.005	0.017			

* Indicates all main effects and interactions among the variables Rotational Axis and Stimulus Type (Rotation_Type) and Skin Conductance Levels (GSR_Levels).

More specifically, pairwise comparisons highlighted accuracy in stimuli rotated in-depth was significantly lower than in stimuli rotated in picture-plane ($M = -0.127$, $p = 0.054$, CI (95%) −0.251–−0.002). Furthermore, there was a tendency for accuracy on abstract stimuli to be higher than on stimuli rotated in-depth ($M = 0.067$, $p = 0.057$, CI (95%) −0.001–0.135). See Figure 3.

A one-way repeated-measures ANOVA using the Greenhouse–Geiser correction found that there was no significant differences in reaction time between SCL groups on concrete and abstract stimuli and stimuli rotated in the picture-plane or in-depth ($F(3, 57) = 0.006$, $p = 0.983$, $\eta^2 = 0.00$). In relation to the main effect, there was a significant difference in reaction time on concrete and abstract stimuli and stimuli rotated in-depth and in picture-plane ($F(3, 57) = 4.37$, $p = 0.030$) with a small effect size ($\eta^2 = 0.19$). Therefore, stimulus type and rotational axis explain 19% of the variance in reaction time on the MR task.

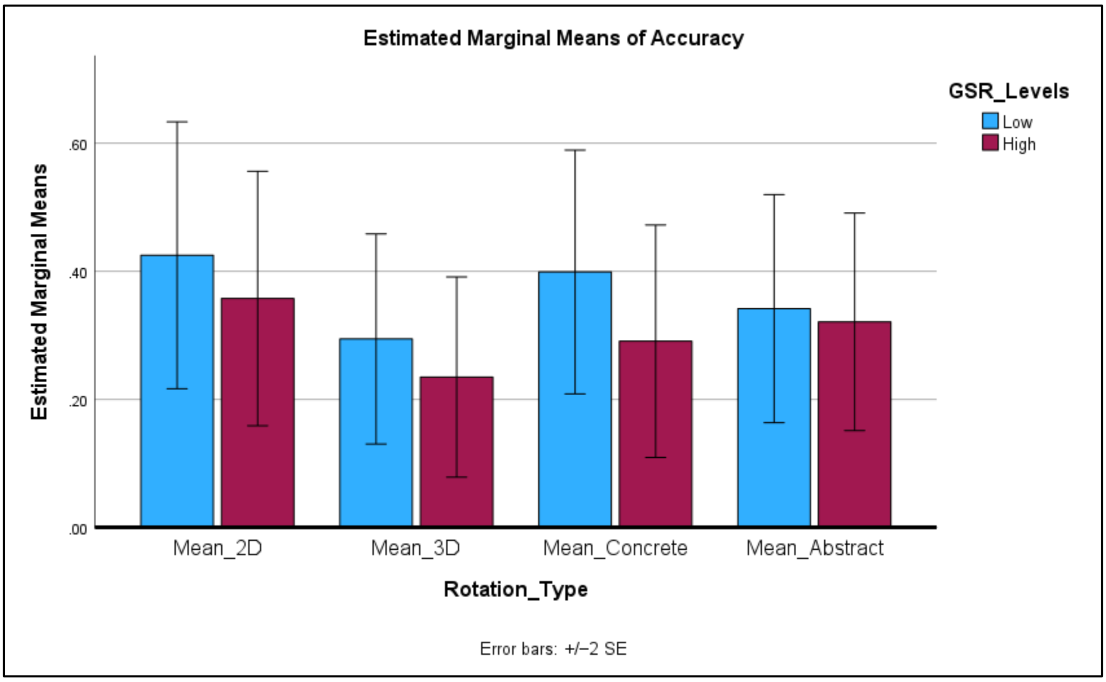


Figure 3. Differences in accuracy on stimuli rotated in-depth and in picture-plane and on concrete and abstract stimuli during the MR task dependent on GSR.

More specifically, pairwise comparisons highlighted that reaction time on abstract stimuli was significantly higher than on concrete stimuli ($M = -1.21$, $p = 0.041$, CI (95%) 0.103–0.7.22), but there was no significant difference in reaction time between stimuli rotated in picture-plane and stimuli rotated in-depth. See Figure 4.

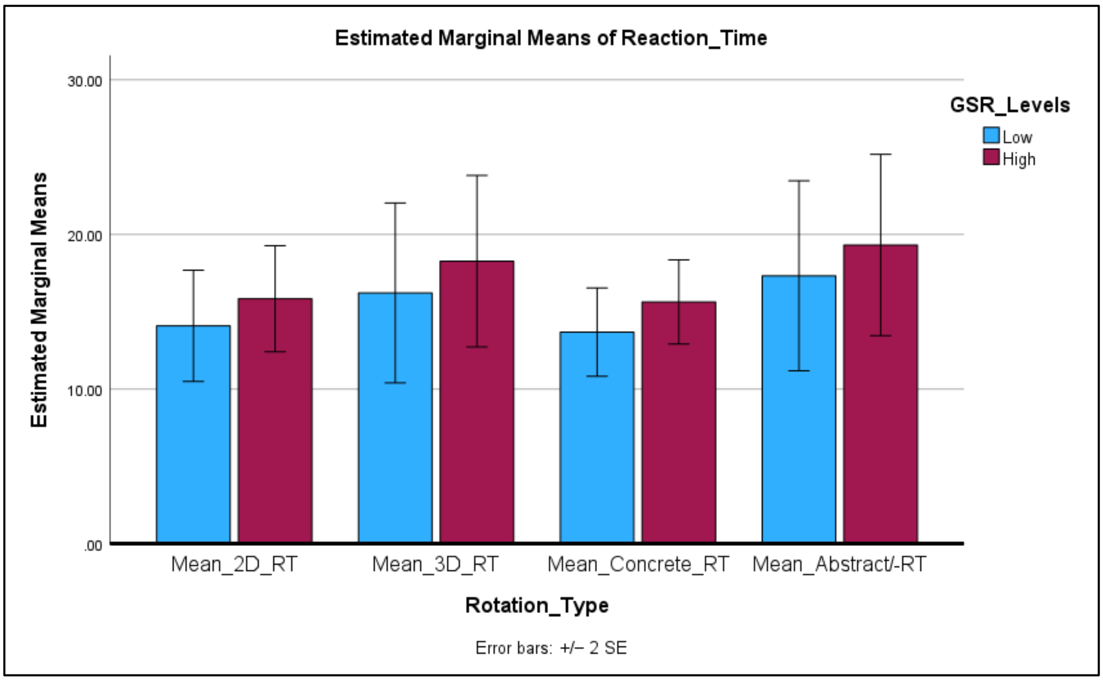


Figure 4. Differences in reaction time on stimuli rotated in-depth and in picture-plane and on concrete and abstract stimuli during the MR task dependent on GSR.

A one-way repeated-measures ANOVA using the Greenhouse–Geiser correction found that there was no significant differences in accuracy between SCL groups on the three gendered objects ($F(2, 38) = 0.585$, $p = 0.553$, $\eta^2 = 0.03$). In relation to the main effect, there was no significant difference in accuracy on the three gendered objects (F (2, 38) = 0.584, $p = 0.553$, $\eta^2 = 0.03$).

A one-way repeated-measures ANOVA using the Greenhouse–Geiser correction found that there was a tendency toward significant differences in reaction time between SCL groups on gender-stereotyped objects ($F(2, 38) = 2.91$, $p = 0.092$, $\eta^2 = 0.13$). In relation to the main effect, there was a significant difference in reaction time on the three gendered objects (F (2, 38) = 6.19, $p = 0.014$) with a small effect size ($\eta^2 = 0.25$). Therefore, the three gendered objects explain 25% of the variance in reaction time on the MR task. See Table 6.

Table 6. Tests of within-subjects effects of gendered objects (Gendered_Obj) across two SCL groups (GSR_Levels) on reaction time on the MR task.

Measure: Reaction_Time							
Source		**Type III Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**	**Partial Eta Squared**
Gendered_Obj	Sphericity Assumed	148.545	2	74.272	6.192	0.005	0.246
	Greenhouse–Geisser	148.545	1.293	114.848	6.192	0.014	0.246
Gendered_Obj * GSR_Levels	Sphericity Assumed	69.729	2	34.864	2.906	0.067	0.133
	Greenhouse–Geisser	69.729	1.293	53.911	2.906	0.092	0.133
Error (Gendered_Obj)	Sphericity Assumed	455.829	38	11.995			
	Greenhouse–Geisser	455.829	24.575	18.549			

* Indicates all main effects and interactions among the variables Gendered Objects (Gendered_Obj) and Skin Conductance Levels (GSR_Levels).

Moreover, pairwise comparisons highlighted reaction time (RT) on neutral objects was significantly higher than on female-stereotyped objects ($M = 3.73$, $p = 0.004$, CI (95%) 1.15–6.30). However, RT on neutral objects was not significantly higher than on the male-stereotyped objects ($M = 2.33$, $p = 0.334$, CI (95%) −1.34–6.00) See Table 7 and Figure 5.

Table 7. Pairwise comparisons with interactions between skin conductance conditions (SCL levels) and reaction time (RT) on items with gender-stereotyped stimuli: male-, female- and neutral objects.

Measure: Reaction_Time						
(I) Gendered_Obj	**(J) Gendered_Obj**	**Mean Difference (I-J)**	**Std. Error**	**Sig. [a]**	**95% Confidence Interval for Difference [a]**	
					Lower Bound	**Upper Bound**
1	2	1.394	0.722	0.205	−0.500	3.289
	3	−2.332	1.397	0.334	−6.000	1.336
2	1	−1.394	0.722	0.205	−3.289	0.500
	3	−3.726 *	0.981	0.004	−6.302	−1.151
3	1	2.332	1.397	0.334	−1.336	6.000
	2	3.726 *	0.981	0.004	1.151	6.302

Based on estimated marginal means. * The mean difference is significant at the 0.05 level. [a] Adjustment for multiple comparisons: Bonferroni.

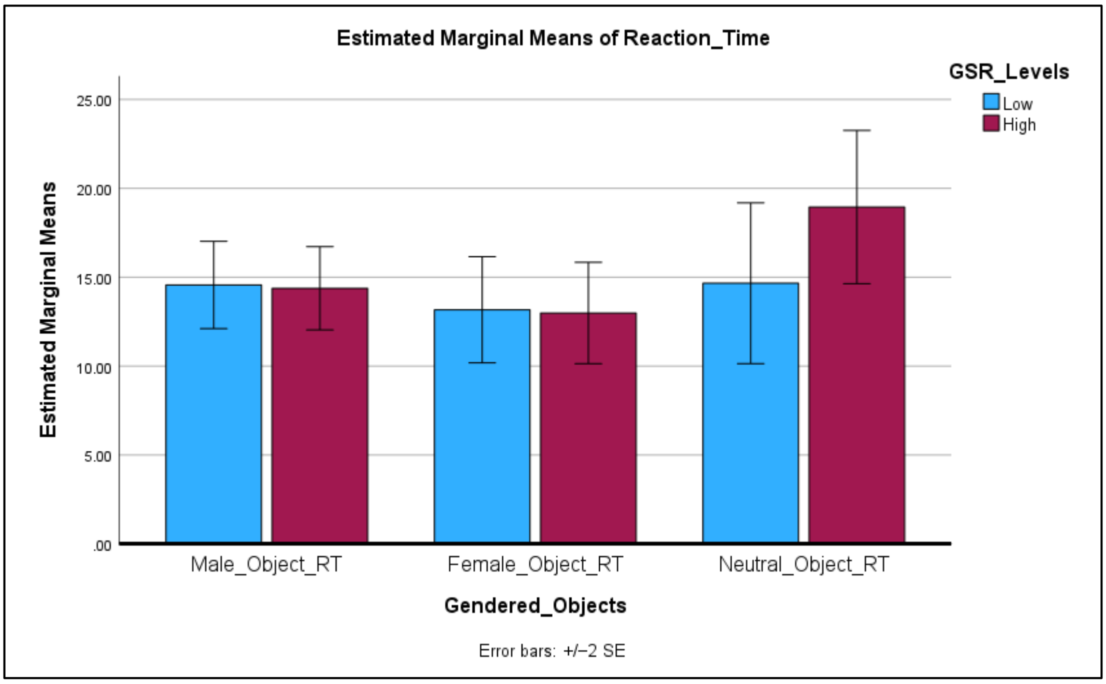


Figure 5. Differences in reaction time across two SCL levels (GSR_Levels) on gendered objects during the MR task.

3.4. Gender Stereotyped Nature of Stimuli and Difficulty of Stimuli Questionnaire Differences (H4 and H5)

Responses on the gender-stereotyped nature of stimuli questionnaire descriptively show that more students rated the *Car* (*M* = 1.88, *SD* = 0.74) and the *Cube figure* (*M* = 2.52, *SD* = 1.05) stimuli as more masculine. Other stimuli, that is, the *Letter_g*, the *Stag*, the *Pellet* and the *Polyhedron* figure presented in the questionnaire were rated as more neutral (*M* = 3.05, *SD* = 0.87).

Responses on the perceived difficulty of stimuli questionnaire descriptively show that more students rated the *Letter_F* (*M* = 1.58, *SD* = 0.83), the *Crocodile* (*M* = 2.45, *SD* = 0.71) and the *Hammer* (*M* = 2.50, *SD* = 1.02) stimuli, all of which were rotated in the picture-plane, as easier. As expected, concrete stimuli (*M* = 2.17, *SD* = 0.85) were rated easier than abstract stimuli (*M* = 3.03, *SD* = 0.84).

A two-way repeated measures ANOVA using the Greenhouse Geiser correction found that there were significant differences in accuracy ($F(4, 68) = 4.61$, $p = 0.012$, $\eta^2 = 0.21$) and reaction time ($F(4, 68) = 4.30$, $p = 0.029$, $\eta^2 = 0.20$) on the items identified as masculine in the stereotyped nature of stimuli questionnaire and those items identified as easier in the perceived difficulty questionnaire, explaining 21% of the variance in accuracy and 20% of the variance in reaction time on these items. See Table 8.

Upon inspection of the means for accuracy and reaction time on these stimuli, it was found that the *Hammer* had the highest mean accuracy (*M* = 0.56, *SD* = 0.511) and the highest reaction time (*M* = 16.12, *SD* = 5.37). The *Car* had the lowest mean accuracy (*M* = 0.17, *SD* = 0.383) and a high reaction time (*M* = 13.94, *SD* = 0.383). The *Crocodile* had a high mean accuracy (*M* = 0.50, *SD* = 0.514) and the lowest reaction time (*M* = 9.74, *SD* = 4.01). The *Letter_F* had mean accuracy (*M* = 0.44, *SD* = 0.511) and mean reaction time (*M* = 11.83, *SD* = 4.17). The *Cube figure* had mean accuracy (*M* = 0.39, *SD* = 0.502) and mean reaction time (*M* = 14.91, *SD* = 7.35). See Figure 6.

Table 8. Differences in accuracy and reaction time (RT) on items identified in the stereotyped nature of stimuli and perceived difficulty questionnaire.

Source		Measure	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
factor1	Accuracy	Sphericity Assumed	1.622	4	0.406	4.613	0.002	0.213
		Greenhouse–Geisser	1.622	2.372	0.684	4.613	0.012	0.213
	RT	Sphericity Assumed	472.286	4	118.072	4.301	0.004	0.202
		Greenhouse–Geisser	472.286	2.678	176.377	4.301	0.012	0.202
error (factor1)	Accuracy	Sphericity Assumed	5.978	68	0.088			
		Greenhouse–Geisser	5.978	40.318	0.148			
	RT	Sphericity Assumed	1866.817	68	27.453			
		Greenhouse–Geisser	1866.817	45.521	41.010			

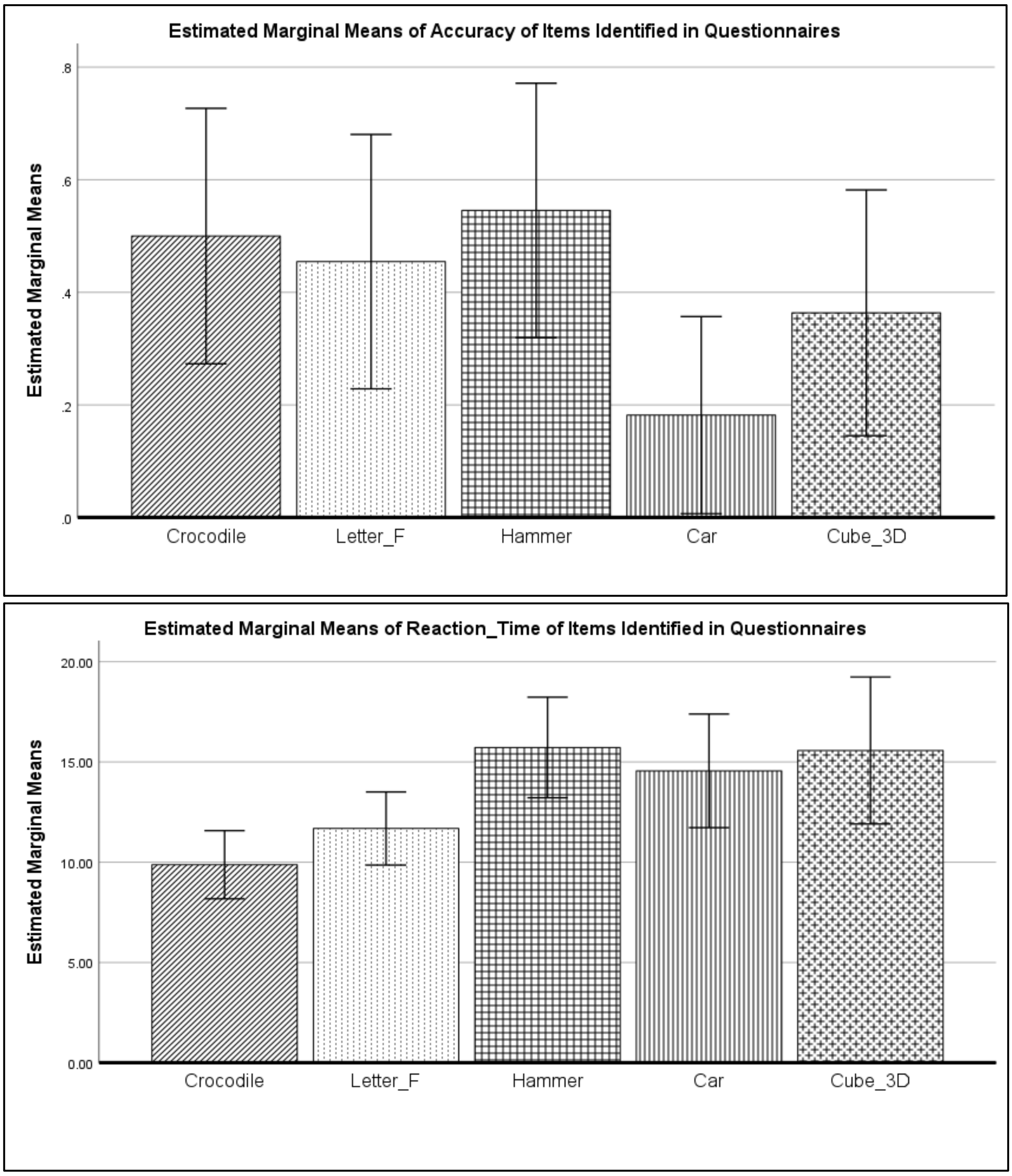


Figure 6. Differences in accuracy and reaction time on objects identified in the stereotyped nature of stimuli and perceived difficulty questionnaires.

4. Discussion

This pilot study set out to investigate whether emotional regulation measured by skin conductance has an impact on primary school children's performance on an MR task with a view to using the results to support the development of a computer adaptive test of mental rotation.

4.1. Skin Conductance Levels across the MR Task

Firstly, results demonstrate that skin conductance levels (SCL) did indeed change as soon as the MR task began. There was a significant increase in SCL from baseline measurements to part one of the task (MRT1), which was considered the easier part as stimuli were rotated in picture-plane only. Moreover, SCL on both parts of the task was significantly higher than in the baseline condition. Contrary to expectations, there was, however, no significant difference in SCL between the first part of the task (MRT 1) and the second, which was considered the more difficult part due to stimuli being rotated in-depth (MRT 2). Therefore, in this sample, an increased task difficulty had no significant effect on emotional reactivity measured by SCL in the second part of the MR task. There were, however, clear groups of participants who experienced "high" and "low" SCL levels during the MR task and these states influenced their performance on some items. These results will now be discussed.

4.2. The Relationship between Skin Conductance, Accuracy, Reaction Time and Item Difficulty

It is known from previous research that items containing abstract stimuli rotated in-depth are more difficult to solve on an MR task than items with concrete stimuli rotated in picture-plane. Therefore, this study also investigated the relationship between item difficulty, emotional regulation measured by SCL and reaction time (RT) on the MR task. The results demonstrate that there was a positive relationship between emotional regulation measured by SCL, RT and scores on difficult items (considered all items across the full MRT containing abstract stimuli rotated in-depth) in this study. Indeed, as emotional regulation measured by SCL and RT increased, so did scores on these items. Therefore, it appears that performance on the difficult items benefitted from increases in participant effort and time spent on these items. Furthermore, findings of this study also show that emotional regulation measured by SCL, in particular along with RT, were significant predictors of scores on difficult items. Thus, in this sample, participants who were able to regulate emotions effectively and who spent more time trying to determine the correct solution to more difficult items benefitted in that they achieved better outcomes.

4.3. Differences in Accuracy and Reaction Time between SCL Groups on Rotational Axis, Stimulus Type and on Gendered Objects

Individual SCL varied, and likewise in this sample. There were participants with a higher SCL than others. Therefore, the study endeavored to examine if the rotational axis (whether rotated in picture-plane or in-depth), the type of stimuli (whether concrete or abstract) and the gender-stereotyped nature of the stimuli (whether male-stereotyped, female-stereotyped or gender-neutral objects) would lead to differences in accuracy and reaction time across groups of "high" and "low" SCL participants. It was found that rotational axis, stimulus type and the gender-stereotyped nature of stimuli did lead to significant differences in accuracy as a main effect, but they did not significantly interact with SCL. Therefore, as SCL differs individually, each participant's emotional reactivity measured by SCL appears to have adjusted accordingly during the cognitive task so that it did not result in differences in performance. Therefore, being a "high" or "low" SCL individual did not impact accuracy and did not interact with performance by stimulus type, rotational axis or gender-stereotyped stimuli.

It was, however, found that there was a tendency toward significant differences in reaction time (RT) on gender-stereotyped objects. Moreover, this difference was significant between SCL groups. Specifically, RT in the "high" SCL group was significantly higher on the gender-neutral objects than on the female-stereotyped objects.

There are two possible explanations for this result. Emotional arousal increases when an individual is presented with stimulus which is relevant and salient to them. Through this increase, the observer assigns a higher priority to that stimulus. It may therefore be that objects represented in the gender-neutral stimuli were more relevant to participants in this study. These stimuli contained everyday objects, which also have recreational or educational significance, such as the scissors, the bicycle and the pot. Participants were likely very familiar with these objects, hence, they were relevant and important to them, resulting in more time spent at attempting to solve the MR test items containing these objects. The female-stereotyped objects on the other hand, such as the bow, the pram and the handbag, may have been assigned a lower relevance and priority by participants because they do not regularly use these objects or they are devalued due to negative associations arising from stereotype threat. This may have resulted in a tendency for participants to dismiss the items containing these objects more quickly.

Another explanation for longer RT on the gender-neutral objects, but one which was not investigated in this study, is the role of occlusion. Two-dimensional representations of three-dimensional objects do not allow the viewer to scan the object from multiple angles. This results in parts of it being hidden. Hidden or occluded parts of objects represented using two-dimensional media such as paper-and-pencil tests or tests undertaken on a computer screen or mobile device make it difficult for the participant to know the shape of the object without seeing it from another angle. The resulting occlusion is not only known to contribute to item difficulty and longer reaction times, but also to gender differences in performance on MR tasks (Felix et al. 2011; Nolte et al. 2022).

4.4. Gender-Stereotyped Objects and Perceived Difficulty of Stimuli

As expected, participants identified the 'Car' and the 'Cube', which are often classified as male-stereotyped and associated with construction and transportation activities (often referred to as "boys' toys") as more masculine (Ruthsatz et al. 2019). On the perceived difficulty questionnaire, the 'Hammer', the 'Crocodile' and the 'Letter_F' were rated as easier. Items with animals and letters are found on the picture-rotation test, which was designed to test mental rotation skills in pre-school children (Quaiser-Pohl 2003).

Upon an analysis of accuracy and reaction time during the MR task on objects identified in the stereotyped nature of stimuli and the perceived difficulty questionnaires, there were significant differences found within the sample. Participants achieved higher scores and spent more time on the male-stereotyped 'Hammer'. Similarly, the highest reaction time was found on the male-stereotyped 'Car'. These objects appear to have been highly salient and participants may therefore have prioritized them by spending more time attempting to solve them. As expected, easy items such as the 'Crocodile' were solved accurately and quickly in this sample.

4.5. Relevance for the Development of a CAT-MR

In this study, items with abstract and gendered stimuli and objects rotated in-depth appeared to be more difficult and required a longer amount of time to solve than others types of items, a result also found in previous research (Neuburger et al. 2011, 2012b; Ruthsatz et al. 2017).

Salience and relevance of stimuli are also important factors when choosing items containing these objects for a CAT, as they may be given higher priority by participants, resulting in longer reaction times and, hence, in many cases, a higher accuracy. Similarly, items with stimuli considered easier or less relevant to participants may result in shorter reaction times, which may then impact accuracy or the ability to correctly solve items containing these objects.

Although a purely anecdotal observation in this study, the visibility of features of some stimuli, such as the hidden edges of cube figures or gender-neutral objects, has been found to be related to the level of difficulty in solving them. Therefore, presenting the stimuli in another format such as real-life objects or three-dimensional models on the computer screen may improve accuracy. Such items could also be incorporated into a CAT.

Item banks in CAT are files of various suitable test items that are coded by subject area, instructional level and other pertinent item characteristics such as item difficulty and discriminating power (Gronlund 1998). In creating an item bank for the CAT-MR, the results of this and any consequent replicated study regarding those items to be used in a CAT-MR need to be considered.

4.6. Limitations and Outlook

This was a pilot study, so the sample size was a significant limitation. Therefore, the study is now being replicated with a bigger sample size.

As occlusion may be limiting participants' performance on the MR task, it may be useful to develop an MR task using real-life objects or three-dimensional models of stimuli to be used in future experiments.

No data regarding test or spatial anxiety were collected, but this could be beneficial in providing more insight into participants' experience of and their feelings about the task. Furthermore, academic self-concept could provide useful data on baseline perceived academic ability in primary school children. Self-concept questions regarding mathematical ability are included in this questionnaire and could yield valuable additional information. Mathematical ability is highly correlated with spatial ability in adults and children (Rahe and Quaiser-Pohl 2021)

There was unfortunately a loss of data due to technical issues. This has now been resolved so that it should not occur during the larger data collection. An additional physiological method of measuring emotional arousal during an MR task has been added for the main study.

Due to the above-mentioned limitations, a number of queries arose, which the pilot study was unable to answer, but which are important for the development of the CAT-MR: (a) whether the MR task was too difficult and too long for some participants; (b) the issue of how the test was constructed, i.e., gender-stereotyped items and increasing item difficulty in the second part of the task; (c) whether some participants may have performed better on a test in which item selection is adapted to their ability, e.g., a CAT-MR; and (d) whether taking a CAT-MR might impact positively on emotional reactivity, accuracy and speed. These queries will be addressed in more detail in the larger study.

5. Conclusions

Reducing the disadvantages in cognitive testing is paramount to preserving the reliability and validity of psychometric instruments designed to measure spatial abilities such as mental rotation. Moreover, providing individuals exposed to stereotype threat with a means to cope effectively with negative emotions such as anxiety can restore executive resources, improve cognition and thus, test performance (Johns et al. 2008). This can also lead to diminishing the gender difference in spatial test performance (Voyer et al. 1995; Wraga et al. 2006). Therefore, the identification of the effects of gender-stereotyped stimuli and emotional arousal arising during testing, as well as item difficulty and the chronological order of items on spatial ability performance warrant ongoing investigation. This research is key to ensuring that these factors do not compromise measurement accuracy nor contribute to increasing gender differences, but rather serve to measure spatial ability accurately, regardless of gender. Test construction as a science benefits from such research as does the field of gender and STEM. In the pursuit of the development of psychological assessment approaches such as a CAT-MR, the authors seek to identify and decrease bias in tests on spatial ability, thus contributing positively to the reduction in the gender gap in STEM.

Supplementary Materials: PsychoPy® is a free, cross-platform, open-source package allowing researchers to run a wide range of experiments in the behavioral sciences. It can be downloaded at www.psychopy.org/download.html (accessed on 6 April 2023).

Author Contributions: Conceptualization, M.L.-M. and C.M.Q.-P.; methodology, M.L.-M.; software, M.S.; validation, M.L.-M. and C.M.Q.-P.; formal analysis, M.L.-M.; investigation, M.L.-M.; resources, M.S. and V.R.; data curation, M.L.-M.; writing—original draft preparation, M.L.-M.; writing—review and editing, M.L.-M. and C.M.Q.-P.; visualization, M.L.-M.; supervision, C.M.Q.-P.; project administration, M.L.-M.; funding acquisition, C.M.Q.-P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Marie Sklodowska-Curie Actions International Training Network (MSCA ITN) "SellSTEM", grant number 956124.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of the University of Koblenz, Germany (Ethikkommission/Antrag Lennon-Maslin/Beschluss vom 9 November 2022). Furthermore, according to section 67, paragraph 6 of the Education Act of the State of Rhineland-Palatinate, Germany, approval was sought and provided by the Aufsichts- und Dienstleistungsdirektion Rheinland Pfalz (ADD RLP), Willy-Brandt-Platz 3, 54290 Trier, Germany, the authorities who oversee schools in this state.

Informed Consent Statement: Informed consent was obtained from all participants involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to ethical and privacy reasons stated in the parental consent information form.

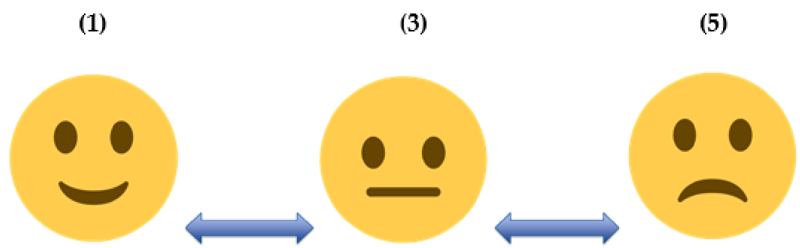
Acknowledgments: Many thanks to Martina Rahe for support and advice regarding the statistical analyses. A special thanks to Emil Ruthsatz, who kindly tested a trial version of our experiment before we undertook the study in a school. Thanks also to Heike Ruf-Urmersbach, who provided advice and support regarding the schools, and to our research assistant Annika Volkmann, who has been a consistent support for our project throughout.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Stereotype Questionnaire

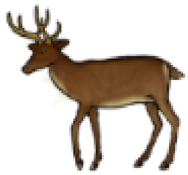
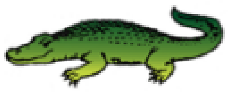
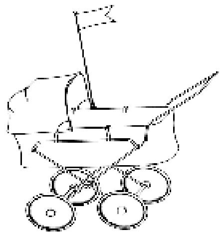
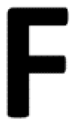
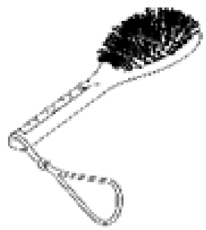
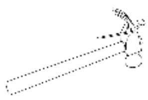
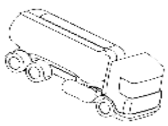
On a sliding 5-point scale, participants were asked to rate how gender-stereotyped they considered the stimuli used in the MR task.

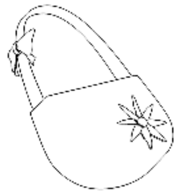
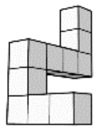
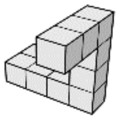
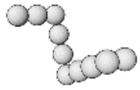
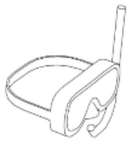
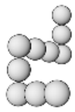
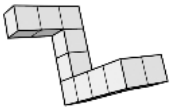
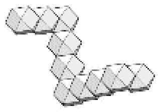
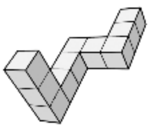
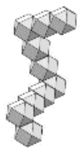
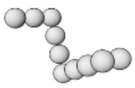
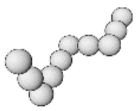
Appendix B. Difficulty Questionnaire

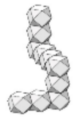


On a sliding 5-point scale, participants were asked to rate how difficult they found various items on the MR task.

Appendix C. Items Used in the MR Task

MRT 1	MRT 2
Stag 	Car_3D 
Crocodile 	Screwdriver_3D 
Letter_g 	Pram_3D 
Letter_F 	Brush_3D 
Hammer_2D 	Scissors_3D 
Lorry_2D 	Pot_3D 

MRT 1	**MRT 2**
Handbag_2D 	Cubes_3D_#5 
Bow_2D 	Cubes_3D_#3 
Bike_2D 	Pellets_3D_#1 
Goggles_2D 	Pellets_3D_#5 
Cubes_2D_#1 	Polyhedron_3D_#1 
Cubes_2D_#4 	Polyhedron_3D_#11 
Pellets_2D_#1 	
Pellets_2D_#5 	

MRT 1	MRT 2
Polyhedron_2D_#11 	
Polyhedron_2D_#1 	

References

Arrighi, Linda, and Markus Hausmann. 2022. Spatial anxiety and self-confidence mediate sex/gender differences in mental rotation. *Learning & Memory* 29: 312–20. [CrossRef]

Braithwaite, Jason, Derrick G. Watson, Robert Jones, and Mickey Rowe. 2013. Guide for Analysing Electrodermal Activity & Skin Conductance Responses for Psychological Experiments. CTIT Technical Reports Series. Available online: https://www.semanticscholar.org/paper/Guide-for-Analysing-Electrodermal-Activity-%26-Skin-Braithwaite-Watson/b99d1f004e4194ac6ef86a86bb0918a11152a01e (accessed on 6 April 2023).

Caissie, Andre F., François Vigneau, and Douglas A. Bors. 2009. What does the Mental Rotation Test Measure? An Analysis of Item Difficulty and Item Characteristics. *The Open Psychology Journal* 2: 94–102. [CrossRef]

Ceci, Stephen J., Donna K. Ginther, Shulamit Kahn, and Wendy M. Williams. 2014. Women in Academic Science. *Psychological Science in the Public Interest* 15: 75–141. [CrossRef] [PubMed]

Chesler, Naomi C., Gilda Barabino, Sangeeta N. Bhatia, and Rebecca Richards-Kortum. 2010. The Pipeline Still Leaks and More Than You Think: A Status Report on Gender Diversity in Biomedical Engineering. *Annals of Biomedical Engineering* 38: 1928–35. [CrossRef]

Christopoulos, George I., Marilyn A. Uy, and Wei Jie Yap. 2019. The Body and the Brain: Measuring Skin Conductance Responses to Understand the Emotional Experience. *Organizational Research Methods* 22: 394–420. [CrossRef]

Chuesathuchon, Chaowprapha. 2008. Computerized Adaptive Testing in Mathematics for Primary Schools in Thailand. Ph.D. and Masters's theses. Available online: https://ro.ecu.edu.au/theses/1591 (accessed on 20 October 2022).

Costanzi, Marco, Beatrice Cianfanelli, Daniele Saraulli, Stefano Lasaponara, Fabrizio Doricchi, Vincenzo Cestari, and Clelia Rossi-Arnaud. 2019. The Effect of Emotional Valence and Arousal on Visuo-Spatial Working Memory: Incidental Emotional Learning and Memory for Object-Location. *Frontiers in Psychology* 10: 2587. [CrossRef]

Damasio, Antonio R. 2001. Reflections on the Neurobiology of Emotion and Feeling. In *The Foundations of Cognitive Science*. Edited by João Branquinho. Oxford: Clarendon Press, pp. 99–108.

Deng, Yaling, Lei Chang, Meng Yang, Meng Huo, and Renlai Zhou. 2016. Gender Differences in Emotional Response: Inconsistency between Experience and Expressivity. *PLoS ONE* 11: e0158666. [CrossRef]

Eggen, Theo J. H. M., and Angela J. Verschoor. 2006. Optimal Testing With Easy or Difficult Items in Computerized Adaptive Testing. *Applied Psychological Measurement* 30: 379–93. [CrossRef]

Felix, Michael C., Joshua D. Parker, Charles Lee, and Kara I. Gabriel. 2011. Real Three-Dimensional Objects: Effects on Mental Rotation. *Perceptual and Motor Skills* 113: 38–50. [CrossRef]

Fladung, Anne-Katharina, and Markus Kiefer. 2016. Keep calm! Gender differences in mental rotation performance are modulated by habitual expressive suppression. *Psychological Research* 80: 985–96. [CrossRef]

Fritts, Barbara E., and Jacob M. Marszalek. 2010. Computerized adaptive testing, anxiety levels, and gender differences. *Social Psychology of Education* 13: 441–58. [CrossRef]

Geršak, Vesna, Helena Smrtnik Vitulić, Simona Prosen, Gregor Starc, Iztok Humar, and Gregor Geršak. 2019. Use of wearable devices to study activity of children in classroom; Case study—Learning geometry using movement. *Computer Communications* 150: 581–88. [CrossRef]

Gronlund, Norman E. 1998. *Assessment of Student Achievement*, 6th ed. Needham Heights: Allyn & Bacon Publishing.

James, William. 1890. *The Principles of Psychology*. New York: Henry Holt and Company, vol. 1. [CrossRef]

Jansen, Petra, Claudia Quaiser-Pohl, Sarah Neuburger, and Vera Ruthsatz. 2014. Factors Influencing Mental-Rotation with Action-based Gender-Stereotyped Objects—The Role of Fine Motor Skills. *Current Psychology* 34: 466–76. [CrossRef]

Johns, Michael, Michael Inzlicht, and Toni Schmader. 2008. Stereotype threat and executive resource depletion: Examining the influence of emotion regulation. *Journal of Experimental Psychology: General* 137: 691–705. [CrossRef]

Kersh, Joanne, Beth M. Casey, and Jessica Mercer Young. 2008. Research on spatial skills and block building in girls and boys. In *Contemporary Perspectives on Mathematics in Early Childhood Education*. Charlotte: Information Age Publishing, pp. 233–51.

Liesefeld, Heinrich René, Xiaolan Fu, and Hubert D. Zimmer. 2015. Fast and careless or careful and slow? Apparent holistic processing in mental rotation is explained by speed-accuracy trade-offs. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 41: 1140–51. [CrossRef]

Linden, Wim J., Wim J. van der Linden, and Cees A. W. Glas. 2000. *Computer Adaptive Testing: Theory and Practice*. New York: Kluwer Academic Publishers. Available online: https://research.utwente.nl/en/publications/computer-adaptive-testing-theory-and-practice (accessed on 20 October 2022).

Linn, Marcia C., and Anne C. Petersen. 1985. Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development* 56: 1479–98. [CrossRef]

Makarova, Elena, Belinda Aeschlimann, and Walter Herzog. 2019. The Gender Gap in STEM Fields: The Impact of the Gender Stereotype of Math and Science on Secondary Students' Career Aspirations. *Frontiers in Education* 4: 60. [CrossRef]

Martin, Andrew J., and Goran Lazendic. 2018. Computer-adaptive testing: Implications for students' achievement, motivation, engagement, and subjective test experience. *Journal of Educational Psychology* 110: 27–45. [CrossRef]

Martin, Carol Lynn, Diane N. Ruble, and Joel Szkrybalo. 2002. Cognitive theories of early gender development. *Psychological Bulletin* 128: 903–33. [CrossRef]

Maselli, Ilaria, and Miroslav Beblavý. 2014. Why Too Few Students Do Maths and Science (SSRN Scholarly Paper No. 2381924). Available online: https://papers.ssrn.com/abstract=2381924 (accessed on 20 October 2022).

Monahan, John S., Maureen A. Harke, and Jonathon R. Shelley. 2008. Computerizing the Mental Rotations Test: Are gender differences maintained? *Behavior Research Methods* 40: 422–27. [CrossRef]

Najafpour, Ebrahim, Naser Asl Aminabadi, Sara Nuroloyuni, Zahra Jamali, and Sajjad Shirazi. 2017. Can galvanic skin conductance be used as an objective indicator of children's anxiety in the dental setting? *Journal of Clinical and Experimental Dentistry* 9: e377–e383. [CrossRef]

Neuburger, Sarah, Petra Jansen, Martin Heil, and Claudia Quaiser-Pohl. 2011. Gender differences in pre-adolescents' mental-rotation performance: Do they depend on grade and stimulus type? *Personality and Individual Differences* 50: 1238–42. [CrossRef]

Neuburger, Sarah, Petra Jansen, Martin Heil, and Claudia Quaiser-Pohl. 2012a. A Threat in the Classroom. *Zeitschrift Für Psychologie* 220: 61–69. [CrossRef]

Neuburger, Sarah, Vera Heuser, Petra Jansen, and Claudia Quaiser-Pohl. 2012b. Influence of Rotational Axis and Gender-Stereotypical Nature of Rotation Stimuli on the Mental-Rotation Performance of Male and Female Fifth Graders. In *Spatial Cognition VIII: Paper Presented at International Conference, Spatial Cognition 2012, Kloster Seeon, Germany, August 31–September 3*. Berlin: Springer, vol. 7463, pp. 220–29. [CrossRef]

Neuburger, Sarah, Vera Ruthsatz, Petra Jansen, and Claudia Quaiser-Pohl. 2015. Can girls think spatially? Influence of implicit gender stereotype activation and rotational axis on fourth graders' mental-rotation performance. *Learning and Individual Differences* 37: 169–75. [CrossRef]

Newcombe, Nora. 2017. *Harnessing Spatial Thinking to Support Stem Learning*. Chiyoda: OECD. [CrossRef]

Newcombe, Nora S., and Andrea Frick. 2010. Early Education for Spatial Intelligence: Why, What, and How. *Mind, Brain, and Education* 4: 102–11. [CrossRef]

Nolte, Nils, Florian Schmitz, Jens Fleischer, Maximilian Bungart, and Detlev Leutner. 2022. Rotational complexity in mental rotation tests: Cognitive processes in tasks requiring mental rotation around cardinal and skewed rotation axes. *Intelligence* 91: 101626. [CrossRef]

Ortner, Tuulia M., and Juliane Caspers. 2011. Consequences of Test Anxiety on Adaptive Versus Fixed Item Testing. *European Journal of Psychological Assessment* 27: 157–63. [CrossRef]

Quaiser-Pohl, Claudia. 2003. The Mental Cutting Test "Schnitte" and the Picture Rotation Test-Two New Measures to Assess Spatial Ability. *International Journal of Testing* 3: 219–31. [CrossRef]

Rahe, Martina, and Claudia Quaiser-Pohl. 2019. Cubes or Pellets in Mental-Rotation Tests: Effects on Gender Differences and on the Performance in a Subsequent Math Test. *Behavioral Sciences* 10: 12. [CrossRef]

Rahe, Martina, and Claudia Quaiser-Pohl. 2021. Can (perceived) mental-rotation performance mediate gender differences in math anxiety in adolescents and young adults? *Mathematics Education Research Journal* 35: 255–79. [CrossRef]

Ramirez, Gerardo, Elizabeth A. Gunderson, Susan C. Levine, and Sian L. Beilock. 2012. Spatial Anxiety Relates to Spatial Abilities as a Function of Working Memory in Children. *The Quarterly Journal of Experimental Psychology* 65: 474–87. [CrossRef] [PubMed]

Ruthsatz, Vera, Martina Rahe, Linda Schürmann, and Claudia Quaiser-Pohl. 2019. Girls' Stuff, boys' stuff and mental rotation: Fourth graders rotate faster with gender-congruent stimuli. *Journal of Cognitive Psychology* 31: 225–39. [CrossRef]

Ruthsatz, Vera, Mirko Saunders, Michelle Lennon-Maslin, and Claudia Quaiser-Pohl. forthcoming. *Male? Female? Neutral! Using Novel Polyhedral Figures as Gender-Neutral Stimuli in a Mental Rotation Test*. Koblenz: University of Koblenz.

Ruthsatz, Vera, Sarah Neuburger, Martina Rahe, Petra Jansen, and Claudia Quaiser-Pohl. 2017. The gender effect in 3D-Mental-rotation performance with familiar and gender-stereotyped objects—A study with elementary school children. *Journal of Cognitive Psychology* 29: 717–30. [CrossRef]

Ruthsatz, Vera, Sarah Neuburger, Petra Jansen, and Claudia Quaiser-Pohl. 2014. 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, Paper Presented at International Conference, Spatial Cognition 2014, Bremen, Germany, September 15–19*. Berlin: Springer, pp. 370–82. [CrossRef]

Ruthsatz, Vera, Sarah Neuburger, Petra Jansen, and Claudia Quaiser-Pohl. 2015. Cars or dolls? Influence of the stereotyped nature of the items on children's mental-rotation performance. *Learning and Individual Differences* 43: 75–82. [CrossRef]

Sanchis-Segura, Carla, Naiara Aguirre, Álvaro Javier Cruz Gómez, Noemí Solozano, and Cristina Forn. 2018. Do Gender-Related Stereotypes Affect Spatial Performance? Exploring When, How and to Whom Using a Chronometric Two-Choice Mental Rotation Task. *Frontiers in Psychology* 9: 1261. [CrossRef] [PubMed]

Saunders, Mirko, and Claudia Michaela Quaiser-Pohl. 2020. Identifying solution strategies in a mental-rotation test with gender-stereotyped objects by analyzing gaze patterns. *Journal of Eye Movement Research 13*. [CrossRef] [PubMed]

Schmader, Toni. 2010. Stereotype Threat Deconstructed. *Current Directions in Psychological Science* 19: 14–18. [CrossRef]

Schmader, Toni, Michael Johns, and Chad Forbes. 2008. An integrated process model of stereotype threat effects on performance. *Psychological Review* 115: 336–56. [CrossRef]

Schmand, Ben. 2019. Why are neuropsychologists so reluctant to embrace modern assessment techniques? *The Clinical Neuropsychologist* 33: 209–19. [CrossRef]

SellSTEM MSCA ITN. 2021. Available online: https://sellstem.eu/ (accessed on 1 September 2021).

Shepard, Roger N., and Christine Feng. 1972. A chronometric study of mental paper folding. *Cognitive Psychology* 3: 228–43. [CrossRef]

Sohn, Jin-Hun, Estate Sokhadze, and Shigeki Watanuki. 2001. Electrodermal and Cardiovascular Manifestations of Emotions in Children. *Journal of Physiological Anthropology and Applied Human Science* 20: 55–64. [CrossRef]

Steele, Claude M., and Joshua Aronson. 1995. Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology* 69: 797–811. [CrossRef] [PubMed]

Storbeck, Justin, and Gerald L. Clore. 2008. Affective Arousal as Information: How Affective Arousal Influences Judgments, Learning, and Memory. *Social and Personality Psychology Compass* 2: 1824–43. [CrossRef]

Tabachnick, Barbara G., and Linda S. Fidell. 2013. *Using Multivariate Statistics: Pearson New International Edition*, 6th ed. Unterschleissheim: Microsoft Press.

Turkileri, Nilgun, David T. Field, Judi A. Ellis, and Michiko Sakaki. 2021. Emotional arousal enhances the impact of long-term memory in attention. *Journal of Cognitive Psychology* 33: 119–32. [CrossRef]

Tyng, Chai M., Hafeez Ullah Amin, Mohamad N. M. Saad, and Aamir S. Malik. 2017. The Influences of Emotion on Learning and Memory. *Frontiers in Psychology* 8: 1454. [CrossRef] [PubMed]

Vandenberg, Steven G., and Allan R. Kuse. 1978. Mental Rotations, a Group Test of Three-Dimensional Spatial Visualization. *Perceptual and Motor Skills* 47: 599–604. [CrossRef] [PubMed]

Veldkamp, Bernard P., and Angela J. Verschoor. 2019. Robust Computerized Adaptive Testing. In *Theoretical and Practical Advances in Computer-Based Educational Measurement. Methodology of Educational Measurement and Assessment*. Berlin: Springer, pp. 291–305. [CrossRef]

Vispoel, Walter P. 1993. Computerized Adaptive and Fixed-Item Versions of the Ited Vocabulary Subtest. *Educational and Psychological Measurement* 53: 779–88. [CrossRef]

Voyer, Daniel, Susan Voyer, and M. Philip Bryden. 1995. Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin* 117: 250–70. [CrossRef]

Walton, Gregory M., and Geoffrey L. Cohen. 2003. Stereotype Lift. *Journal of Experimental Social Psychology* 39: 456–67. [CrossRef]

World Economic Forum. 2019. The Global Gender Gap Report 2020. Available online: https://gdc.unicef.org/resource/global-gender-gap-report-2020 (accessed on 20 October 2022).

Wraga, MaryJane, Lauren Duncan, Emily C. Jacobs, Molly Helt, and Jessica Church. 2006. Stereotype susceptibility narrows the gender gap in imagined self-rotation performance. *Psychonomic Bulletin & Review* 13: 813–19. [CrossRef]

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 23 May 2024
DOI 10.3389/feduc.2024.1300598

OPEN ACCESS

EDITED BY
Sheryl Sorby,
University of Cincinnati, United States

REVIEWED BY
Danielle Harris,
University of Canberra, Australia
Lauren C. Pagano,
Northwestern University, United States

*CORRESPONDENCE
Michelle Lennon-Maslin
✉ mlennonm@uni-koblenz.de

RECEIVED 23 September 2023
ACCEPTED 07 May 2024
PUBLISHED 23 May 2024

CITATION
Lennon-Maslin M, Quaiser-Pohl C and Wickord L-C (2024) Beyond numbers: the role of mathematics self-concept and spatial anxiety in shaping mental rotation performance and STEM preferences in primary education.
Front. Educ. 9:1300598.
doi: 10.3389/feduc.2024.1300598

Beyond numbers: the role of mathematics self-concept and spatial anxiety in shaping mental rotation performance and STEM preferences in primary education

Michelle Lennon-Maslin[1]*, Claudia Quaiser-Pohl[2] and Lea-Christin Wickord[2]

[1]Department of Developmental Psychology and Psychological Assessment, Institute of Psychology, University of Koblenz, Koblenz, Germany, [2]Department of Educational Sciences, Institute of Psychology, University of Koblenz, Koblenz, Germany

Introduction: Factors such as low self-concept and anxiety have been shown to negatively impact mathematical achievement and spatial skills, as well as enjoyment of math-related subjects. Understanding these factors is crucial for promoting STEM interest and performance, particularly among primary school students.

Methods: This cross-sectional study examines the influence of gender, childhood development stage, maths self-concept, spatial anxiety, perceived difficulty, mental rotation performance, and STEM preferences in a sample of 144 primary school students (mean age $M = 8.47$), comprising 70 girls and 74 boys. Data were collected through four questionnaires and a computerized Mental Rotation Task (MRT).

Results: Girls and tweens (9-to-11-year-olds) exhibit lower maths self-concept, impacting their preference for maths as a STEM subject. Girls also demonstrate higher spatial anxiety and perceived difficulty of the MRT compared to boys. Maths self-concept is significantly associated with spatial anxiety and perceived difficulty, while gender is not. Maths self-concept shows marginal effects on students' accuracy on the MRT, with evidence of a mediating effect of spatial anxiety.

Discussion: These findings underscore the importance of maths self-concept in shaping STEM preferences, particularly among girls and tweens. Additionally, maths self-concept serves as a mitigating factor for spatial anxiety and perceived difficulty in spatial tasks among primary school children. The study also suggests that spatial anxiety may contribute to gender disparities in mathematics and STEM-related domains. Further research is needed to explore interventions targeting maths self-concept and spatial anxiety to promote equitable STEM engagement amongst primary school students.

KEYWORDS

mathematics self-concept, spatial anxiety, STEM subject preferences, stage of childhood development, mental rotation performance, primary education

1 Introduction

> "I must tell you what my opinion of my own mind and powers is exactly … I believe myself to possess a most singular combination of qualities exactly fitted to make me pre-eminently a discoverer of the hidden realities of nature."
>
> Ada Noel King, *The Bride of Science* (Woolley, 2000).

Augusta Ada King Noel, Countess of Lovelace, was a British pioneer in the fields of mathematics and science. She was born in London in 1815, into an era in which fundamental rights for women had not yet been established. Ada, who was also the daughter of the renowned poet Lord Byron, often found herself overshadowed by her father's reputation for literary greatness and notoriety (Chiaverini, 2017). However, with her mother's encouragement, Ada pursued a keen interest in mathematics and computation carving her niche in the field and achieving greatness in her own right. Throughout her short life, Ada, who referred to herself as the "Enchantress of Numbers," demonstrated a strong and undeterred self-belief in her intellectual ability as the opening quote illustrates.

The focus of this cross-sectional study is therefore to examine the role of self-concept, specifically maths self-concept, and its association with preference for STEM (Science, Technology, Engineering and Mathematics) subjects and spatial ability in a sample of primary school children at two stages of childhood development. Furthermore, the association between maths self-concept, spatial anxiety, the perceived difficulty of the spatial task and their role in spatial ability are investigated.

1.1 Mathematical skills: nature or nurture?

While it might come last in the acronym "STEM," maths is the foundation of science, technology, and engineering. Educational research has frequently focused on STEM in the context of science with less focus on maths in classroom activities (Maass et al., 2019; Larsen et al., 2022). However, maths alongside its models serves as the language of science and the means through which scientific concepts are interpreted and communicated (Just and Siller, 2022). Traditionally, maths skills were thought to be fixed and innate: Recent evidence reveals distinct individual differences in the brains and genes of individuals with a propensity towards maths (Chen et al., 2017; Skeide et al., 2020). From a biopsychosocial perspective, however, the role of the environment, attitudes and even emotions experienced around engagement with the subject also play an important role in the development of mathematical ability (Petrill et al., 2009). Contrary to the notion of fixed abilities, research supports the idea that maths skills, even in low-performing school children, are malleable (Lopez-Pedersen et al., 2023). Moreover, findings demonstrate that experts in the field are characterized not by giftedness alone, but by extended and intense training which leads to enhancement of numerical cognition (Sella and Cohen Kadosh, 2018). That said, it is also essential to understand and appreciate how attitudes and emotions impact maths learning and achievement. They must also be taken into account in order improve outcomes in this field for all children (Dowker et al., 2016).

1.2 Maths self-concept and its impact on maths achievement

There is evidence that humans are born sensitive to numeracy (Berger et al., 2006) and that in fact, attitudes and perceptions as well as classroom experience shape children's beliefs about their mathematical ability (Attard, 2013). For many, an aversion towards maths is attributed to negative experiences during their school years leading to decreased engagement with the subject (Boaler, 2015; Hawes et al., 2022). Maths frequently serves as that academic domain in which children encounter significant challenges, characterized by profound setbacks, a circumstance further compounded by the responses of educators and attitudes adopted by parents (Rossnan, 2006; Boaler, 2015; Davadas and Lay, 2017). These experiences shape what is referred to as maths self-concept, or how children think about themselves in relation to maths, which in turn impacts their engagement with and achievement in the subject (Cvencek et al., 2020). Maths self-concept is an important factor accounting for differences in school children's mathematical achievement with the socialization process having a significant effect on this (Manger and Eikeland, 1998; Eccles, 2009; Else-Quest et al., 2010).

In a longitudinal study conducted over 6 years in Germany, Marsh et al. (2018) found that maths self-concept was predictive of secondary school students' maths scores and grades (Marsh et al., 2018). A study conducted over 9 years in Australia found maths self-concept to be a significant predictor of post-school preference for STEM studies (Parker et al., 2014). The Programme for International Student Assessment (PISA) conducted a large-scale, longitudinal study across 35 member countries of the Organisation for Economic Co-operation and Development (OECD) and 31 partner countries and economies whose finding show that students who have low maths self-concept perform worse than students who are more confident in their maths ability (OECD, 2013).

The Reciprocal Effects Model (REM) posits a mutually reinforcing relationship between self-concept and academic achievement over time (Marsh, 1990; Marsh and Craven, 2006). According to this theoretical framework, proficiency in mathematics fosters a positive self-concept in students, which, in turn, contributes to enhanced academic performance (Sewasew et al., 2018). Encouraging active involvement and participation in mathematics is therefore anticipated to yield positive effects on students' self-concept. For example, studies conducted in the United States and China with pre- and primary school children found a reciprocal relationship between maths self-concept, interest, and achievement (Fisher et al., 2012; Cai et al., 2018).

All of these findings emphasize the critical role of maths self-concept in shaping educational trajectories.

1.3 Gender differences in maths self-concept and the impact on STEM as a career choice

Unlike the abundance of confidence demonstrated by Countess Lovelace with regards to her intellectual and scientific prowess, many of todays' girls and women do not appear so self-assured when it comes to their mathematical ability. Although women now excel in the STEM field, poor maths self-concept may explain why they continue to be under-represented in university programs which are the most mathematically intensive, such as engineering, computer science and the physical sciences (Ceci et al., 2014). Frequently, students with low maths self-concept lack faith in their maths ability which in turn can have a major impact on their preference for STEM careers (Goldman and Penner, 2016).

There is evidence from research conducted with adults that despite similar levels of maths achievement, women demonstrate lower maths self-concept than men (Sax et al., 2015). Initially, girls display similar or even superior maths achievement compared to boys, but a shift occurs as they advance through the school system, especially in the upper levels of the skills distribution (Fryer and Levitt, 2010; Cimpian et al., 2016). In a longitudinal study conducted in the USA, spanning children from kindergarten through primary school, initially both girls and boys demonstrated comparable proficiency in mathematics. However, by third grade, significant gender differences were found in both maths confidence and achievement particularly at the top of the attainment distribution (Fryer and Levitt, 2010). The gender gap observed at the upper tiers of maths achievement is especially disconcerting, given that this subgroup often generates future STEM professionals (Cimpian et al., 2016).

Moreover, international assessments reveal persistent gender gaps in maths self-concept, with girls expressing less confidence in mastering challenging mathematical concepts compared to boys (OECD, 2013). Notably, countries such as Switzerland, Denmark, and Germany revealed particularly pronounced gender disparities in maths self-concept (OECD, 2013).

1.4 STEM stereotypes influence tweens' and girls' preference for maths and related careers

Stereotype threat (ST) describes situations in which individuals perceive themselves to be at risk of conforming to negative stereotypes about their ingroup (Aronson et al., 2014). ST is a factor which has been found to negatively influence attractivity of STEM subjects such as maths, physics, computer science and associated professions for young people (Garriott et al., 2017). A stereotype frequently associated with individuals who excel in such fields is that of the "nerd" or "geek," the characteristics of whom include disproportionate intelligence, awkwardness in social circumstances, unattractive appearance and romantically unsuccessful (Starr and Leaper, 2019). Findings confirm that the "STEM-nerd" or enthusiast stereotype was endorsed among secondary school students in the USA (Garriott et al., 2017), as well as middle-school students in Germany (Kessels, 2005). This attitude is problematic because it may undermine some young people's interest in the field due to the threat of confirming a negative stereotype (Starr, 2018).

A period in which young people begin to explore their identities and in which self-concept and stereotypes become more salient is the pre-adolescent or tween stage of childhood development (Starr, 2018). Tweens are young people around the age of 9 to 12 years, a period directly preceding adolescence and following middle childhood (McArthur et al., 2021). Tweens are on the cusp of a significant educational transition between primary and secondary school. In many countries, they are required to make important choices regarding subjects they will focus on in secondary education. Moreover, tweens are beginning to develop an awareness of gender identity as well as societal norms, roles, and hierarchies (McArthur et al., 2021). They are more cognizant of negative cultural stereotypes such as those surrounding STEM professions and this can affect their preference for and engagement with associated subjects such as maths (McGuire et al., 2022). At this stage, however, awareness of stereotypes is new and not yet established making it ideal for introducing interventions such as peer engagement programs to foster positive socialization towards STEM (Space Science Institute, 2023).

The STEM-enthusiast stereotype has also been found to negatively influence girl's identification with related career fields and associated professions (Starr, 2018). Although male STEM professionals such as computer scientists, are often unfairly stereotyped as socially awkward outsiders who neglect their appearance, there remains a societal expectation for women to be attractive, sociable, and skilled in romantic pursuits (Cheryan et al., 2013). Mentoring programs for young women in the tween and adolescent stage, such as the Ada Lovelace Project based in Germany (Ada Lovelace-Projekt, 2023), employ female role models to exercise a positive influence on STEM's appeal for these groups (Quaiser-Pohl, 2012; Quaiser-Pohl and Endepohls-Ulpe, 2012; Quaiser-Pohl et al., 2012, 2014a).

1.5 Mental rotation and its relationship with mathematics skills

Mental Rotation (MR), the ability to rotate mental representations of two- and three-dimensional objects, is a spatial skill extensively studied in psychology and education (Quaiser-Pohl et al., 2014b; Neuburger et al., 2015; Buckley et al., 2018). Studies have shown that boys perform significantly better on mental rotation tasks than girls, making it an important topic to research (Linn and Petersen, 1985; Voyer et al., 1995).

The relationship between spatial and mathematical ability has been widely investigated and is well-established (Hawes et al., 2015; Mix, 2019; Lowrie et al., 2020). The findings from a recent meta-analysis conducted by Hawes et al. (2022) align with previous research and theoretical assertions, indicating that engaging in spatial training serves as an effective method for improving both comprehension and performance in mathematics (Hawes et al., 2022).

As with mathematical and other spatial skills, mental rotation (MR) ability can be improved through practice and training (Uttal et al., 2013). This training can have a spillover effect on maths performance, also enhancing these skills. Furthermore, mental rotation has been linked to higher-level mathematical skills and is a significant predictor of mathematical ability in pre-and primary school children (Moe, 2018; Rahe and Quaiser-Pohl, 2019).

Mental rotation performance has also been identified as a significant predictor of academic success and career choice in STEM fields (Moe, 2018; Moe et al., 2018). However, gender disparities in mental rotation performance can be exacerbated by factors such as stereotype threat and spatial anxiety, particularly among girls. Addressing these barriers is crucial for promoting equitable participation in STEM fields and fostering a diverse talent pool.

1.6 Stereotype threat, child spatial anxiety, and mental-rotation performance

Gender-stereotyped beliefs influence boys' and girls' preferences and behaviors early in life, including their play and choice of toys (Ruthsatz et al., 2019). Boys' preference for construction toys and other games that are related to object manipulation in space may lead to gender differences as a result of more practice and confidence acquired during play (Moe et al., 2018). An increased familiarity with spatially related objects and activities appears to be advantageous to male participants on a mental rotation task consisting of items with frequently used cubical figures (Ruthsatz et al., 2014). Female performance on spatial tasks may also be compromised by the fear of confirming existing gender negative stereotypes about spatial abilities (McGlone and Aronson, 2006; Campbell and Collaer, 2009). Research on spatial ability consistently finds the most significant and dependable cognitive gender gaps on mental rotation tests, where men typically outperform women by around one standard deviation (Wraga et al., 2006). Noteworthy, however, is the effect of emotions such as anxiety on mental rotation performance and cognitive processes in general. For instance, pronounced gender differences in spatial anxiety and spatial self-efficacy have been found to mediate gender differences in mental rotation performance, particularly on more demanding tasks (Arrighi and Hausmann, 2022).

Child Spatial Anxiety (CSA) or feelings of nervousness at the prospect of engaging in spatial tasks, is known to affect mental rotation performance in young children (Ramirez et al., 2012). Worry associated with maths and complex cognitive tasks such as mental manipulation can limit executive resources such as working memory (WM) and can lead to performance deficits in these domains (Engle, 2002). WM is a short-term memory store which actively holds information in the mind needed to complete complex tasks such as mental rotation or numerical operations (Alloway and Passolunghi, 2011). An interaction between working memory and spatial anxiety appears to contribute to gender differences in mental rotation performance. Girls are further disadvantaged by the added effects of stereotype threat which usurps working memory capacity needed to perform well on spatial tasks (Schmader et al., 2008; Schmader, 2010; Ramirez et al., 2012). Consequently, girls' performance may be impacted by the fluctuating spatial anxiety during related tasks.

In the tween stage of development, spatial ability undergoes significant growth and development, surpassing the levels observed in middle childhood (Hodgkiss et al., 2021). The development of mental rotation in pre-adolescence is associated with changes in the brain, particularly in regions related to spatial processing and cognitive control. The parietal cortex, which plays a crucial role in spatial cognition, undergoes maturation during this period. Structural changes, such as increased grey matter density and synaptic pruning, contribute to improved spatial abilities (Modroño et al., 2018). This developmental period therefore presents a prime opportunity for the expansion and enhancement of cognitive abilities, which are essential for excelling in STEM subjects during secondary education. Subjects such as advanced mathematics, computer science, physics, and chemistry benefit greatly from development in spatial cognition. This developmental juncture sets the stage for not only cognitive enhancement but also potential challenges such as the emergence of negative attitudes and emotions related to spatial tasks. Hence, it is important that in girls and pre-adolescents, spatial skills are actively nurtured and interest in mathematics sustained.

1.7 Perceived difficulty of a mental-rotation task and its relationship to affect, self-concept and performance

Transitioning from the intricate dynamics of spatial anxiety, stereotype and gender influences on mental rotation performance, the focus shifts to the perceived difficulty of a mental rotation task. This can also have consequences for students' affect, self-concept, and cognitive performance (Nuutila et al., 2021). High perceived difficulty of a task can lead to increased frustration and negative emotional reaction which in turn impairs performance (Brunstein and Schmitt, 2010). However, self-efficacy and interest in the subject matter can buffer these effects leading to increased time and effort spent attempting to solve associated tasks (van Steensel et al., 2019).

The difficulty of a mental rotation task depends on its characteristics, for example, on the rotational axis, stimulus complexity, dimensionality and familiarity of stimuli (Nolte et al., 2022). These factors can influence both the number of items answered correctly but also the time spent attempting to solve them. Response time as well as accuracy is often measured in studies of mental rotation as longer processing time can lead to increased accuracy on the task (Liesefeld et al., 2014). Indeed, the removal of time limitations imposed on mental rotation tasks has been found to influence gender differences in performance due to diversity in processing strategies (Voyer, 2011).

Women are known to approach mental rotation tasks more cautiously and analytically, a processing style known to be more time-consuming (Peters, 2005). A study of the eye-movement patterns of primary school students found that girls tend to analyse parts of the stimulus whereas boys approach these more holistically (Taragin et al., 2019). A holistic strategy involves a comprehensive approach, where all parts of a test stimulus are viewed as a whole. An analytic strategy involves breaking down the problem into individual components and finding solutions for each component separately (Li and O'Boyle, 2013). Therefore, if children perceive a mental rotation task as more difficult, this may impair their performance. Raised negative affect such as anxiety and higher frustration levels interfere with response time. Test takers may take too long to solve items, which can compromise their performance on time-limited tests. Alternatively, they might not spend enough time on and put sufficient effort into solving items which may lead to guessing the solutions (Liesefeld et al., 2014; Neuburger et al., 2015).

1.8 Rationale and aim of the current study

In summary, this study investigates the effect of children's maths self-concept, their spatial task-related anxiety, and their perception of task difficulty on mental rotation performance across two age groups. Moreover, the role of gender, childhood developmental stage, maths self-concept, spatial anxiety, and perceived difficulty in shaping children's preferences for STEM or maths and their performance on spatial tasks is examined. Our study aims to provide insights for improving STEM engagement and reducing gender gaps in STEM fields.

1.8.1 Research questions and hypotheses

Based on our review of the literature, in this study we will examine the following research questions:

Q.1. Are there gender and age differences in children's maths self-concept, spatial anxiety, perceived difficulty of spatial tasks, mental rotation performance and preference for STEM subjects?

Hypothesis 1. *We expect there will be significant gender differences in maths self-concept, spatial anxiety, perceived difficulty, accuracy and response times on the mental rotation task and preference for the STEM subject maths in primary school students. Specifically, girls will have significantly lower maths self-concept, higher spatial anxiety, and higher perceived difficulty than boys. Girls will have lower scores and longer response times on the mental rotation task than boys. Girls are less likely to choose the STEM subject maths than boys.*

Hypothesis 2. *We expect there will be significant age differences in maths self-concept, spatial anxiety, perceived difficulty, accuracy and response times on the mental rotation task and preference for the STEM subject maths in primary school students. Specifically, tweens (9-to-11-year-olds) will have significantly lower maths self-concept, lower spatial anxiety and lower perceived difficulty than students in middle childhood (6-to-8-year-olds). Tweens will have significantly higher scores and shorter response times on the mental rotation task than students in middle childhood. Tweens are less likely to choose the STEM subject maths than students in middle childhood.*

Q.2. Does students' preference for STEM or non-STEM subjects predict maths self-concept?

Hypothesis 3. *We expect students who prefer the STEM subject maths at school will have higher maths self-concept than students who prefer non-STEM subjects.*

Q.3. How does maths self-concept relate to spatial anxiety and perceived difficulty?

Hypothesis 4. *We expect maths self-concept is significantly associated with spatial anxiety and perceived difficulty of the mental rotation task in primary school students. Specifically, as maths self-concept increases. Spatial anxiety decreases. Additionally, as maths self-concept increases, perceived difficulty decreases.*

Q.4. Does maths self-concept predict mental rotation performance? Is this association mediated by spatial anxiety and perceived task difficulty?

Hypothesis 5. *We expect maths self-concept is a significant predictor of accuracy and response time on the mental rotation task and that this association is mediated by spatial anxiety and perceived difficulty.*

2 Materials and methods

2.1 Participants

A total of one-hundred-and-forty-eight students were recruited from primary schools in the state of Rhineland Palatinate in Germany ($N = 148$). All of the students belonged to first, second, third, and fourth-grade classes, respectively, and were enrolled in five primary schools in Rhineland Palatinate, Germany. One-hundred-and-forty-seven students ($N = 147$) took the mental rotation task and completed the questionnaires. Due to a technical error however, one student could not complete the questionnaires during the data collection. As three children were no longer enrolled in primary education, their data were removed from the sample leaving a total of one-hundred-and-forty-four ($N = 144$) whose data were analysed for this study. Seven children included in the sample were tested in the home environment as opposed to the school setting. Upon analysis, however, their responses and performance on the various measures did not significantly differ from that of the rest of the participants.

Included in the sample were 74 students who identified as boys and 70 as girls. The average age of the students was $M = 8.47$ ($SD = 1.12$) years old. The gender variable was dichotomous and contained boys (M) coded 0 and girls (F) coded 1. A variable for the stage of childhood development students was also dichotomous and contained two groups: Group 1 consisting of 72 children aged 6 to 8 years old ($N = 72$) in Middle Childhood, coded 0; Group 2 consisting of 72 children aged 9 to 11 years old ($N = 72$) in the Tween or pre-adolescent stage, coded 1. The age distribution by gender is shown in Table 1 and across two stages of childhood development in the bar chart in Figure 1.

Every student entered their preferred subject at school into the online questionnaire. Preferred subjects were then divided into three categories: A STEM group was created based on preference for Maths as preferred subject and coded 0; A Non-STEM group was based on preference for Art, Music, and languages as preferred subjects, coded 8; An 'Other' group represented students' preference for Physical Education (P.E.) and was coded 9.

Schools from which the data were collected, were situated in areas with diverse socio-economic and cultural populations, in both urban and rural areas.

An *a-priori* G*Power analysis was used to calculate the minimum number of participants needed for the planned statistical analyses with a medium effect size and power of 0.8. The sample size was found to be adequate and the number of participants in each of the groups was balanced, with the exception of the one-way analysis of variance (ANOVA) (Faul et al., 2007). The G*Power protocol is appended to this report (Appendix 1).

2.2 Material and instrumentation

2.2.1 Demographic data

An online questionnaire, created in PsychoPy®, was presented to each participant at the beginning of the experiment before

TABLE 1 Gender (sex) * stage of childhood development crosstabulation count.

		Stage of childhood development		Total
		Middle childhood	Tweens	
Sex	Boys	35	39	74
	Girls	37	33	70
Total		72	72	144

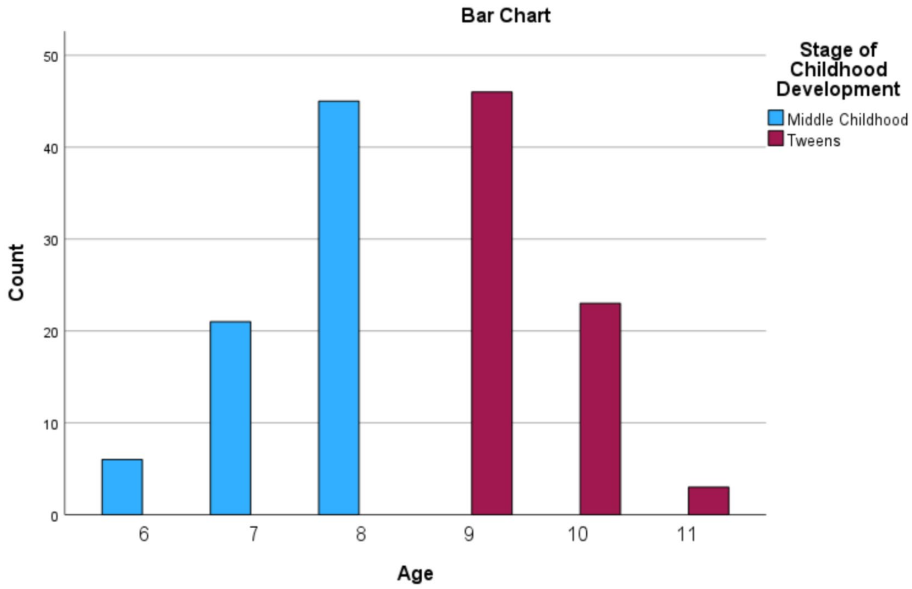


FIGURE 1
Age distribution of students across two stages of childhood development.

undertaking the mental rotation task (MRT) in order to collect data relating to participants' age and gender.

2.2.2 Self-report questionnaires

In order to avoid priming or activation of stereotypes, four self-report questionnaires were presented following the MRT. The purpose of these questionnaires was to assess the students' levels of maths self-concept, spatial anxiety and perceived difficulty of stimuli, all of which were presented to the students in the German language.

2.2.2.1 Preferred subject at school

Students were asked to name their favourite subject and a blank space was provided for them to type this. It was explained that students could pick a preferred subject from all of the subjects at school.

2.2.2.2 Academic self-concept questionnaire in primary school children

Students were asked to complete the Academic Self-Concept in Primary School Children (ASKG) questionnaire (Ehm, 2014). This is a German language self-report measure based on the revised hierarchical self-concept model from Marsh et al. (1988). It allows primary school students to self-assess their mathematical, reading and writing skills on three subscales on a 7-point sliding scale. Each subscale consists of 6 items. The mathematics self-concept subscale was analysed for this study. An example of an item from this subscale is as follows: "I like Maths…" Participants are then asked to choose from a scale of 1 labelled "not at all" to 7 labelled "very much," how much they like or dislike mathematics at school. Researchers explained to students that the higher up the sliding scale they choose, the more they demonstrate enjoyment of mathematics for example. A reliability analysis of the full ASKG conducted on the data from our sample yielded a Cronbach's Alpha of $\alpha = 0.90$. Cronbach's Alpha for the mathematics self-concept subscale was also $\alpha = 0.90$.

2.2.2.3 Child spatial anxiety

The Child Spatial Anxiety Questionnaire (CSAQ) (Ramirez et al., 2012) consists of 8 items in which students are asked to rate how anxiety-provoking they find various spatial tasks. 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 a 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 anxious 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 $\alpha = 0.65$. Despite reliability of the CSAQ being marginally below the acceptable range of $\alpha = 0.70$, it was nevertheless chosen for its practicality and suitability in the school setting.

2.2.2.4 Perceived difficulty of stimuli

The Perceived Difficulty of Stimuli (PDQ) scale, adapted from the Stereotyped Nature of Stimuli questionnaire (Neuburger et al., 2015), uses emojis on a sliding 5-point scale to assess perceived difficulty of items in the MRT. Participants are presented with 6 stimuli included in items from the MRT and are asked to rate the difficulty of these items on the emoji scale. The first point on the scale (1) represents easy or happy face emoji, and the fifth (5) difficult or sad face emoji. The third point (3) represents neither easy nor difficult or neutral face emoji. A reliability analysis of the PDQ conducted on the data from our sample yielded a Cronbach's Alpha of $\alpha = 0.72$.

2.2.3 Mental rotation task

A computerized mental rotation task (MRT; Vandenberg and Kuse 1978) was programmed in PsychoPy® software and installed on

Microsoft Pro 8 Surface tablets, each with a keyboard and a mouse. The task was programmed to record both accuracy and response time in seconds. Items included stimuli for younger children, i.e., animals, letters and cubes rotated in picture-plane (Quaiser-Pohl, 2003) as well as abstract and concrete stimuli rotated in picture-plane and in-depth. Abstract items consisted of stimuli such as cubes, pellets (Ruthsatz et al., 2014) and polyhedral figures (Ruthsatz et al., forthcoming). Concrete items consisted of male and female-stereotyped stimuli (Ruthsatz et al., 2015) and gender-neutral stimuli (Ruthsatz et al., forthcoming). Examples of some of the items used in the task are shown in Figure 2.

The task 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 (see sample items in Appendix 5). MRT 2 was considered more difficult due to the rotational axes of the stimuli (Neuburger Ruthsatz et al., 2012). 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 the full MRT conducted on the data from our sample yielded a Cronbach's Alpha of $\alpha = 0.86$. A Cronbach's Alpha analysis on MRT 1 yielded $\alpha = 0.71$ and on MRT 2 $\alpha = 0.86$.

2.3 Procedure

Approval for the pilot study was provided by the Ethics Committee of the University of Koblenz and the state authorities in Rhineland Palatine overseeing schools. Informed consent was sought and provided by parents and guardians of all students involved in the study. The class teacher and the principal also gave permission for the 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 at any stage of the experiment with no consequences. The students were tested by two female researchers in a separate classroom with access to their teacher, if required. The room had adequate lighting and individual seating arrangements. The researchers explained the MRT to the students by rotating objects such as a pair of scissors, a toy and a wooden object. They demonstrated that rotating the object 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 the study and the significance of mental rotation in everyday life and for school work 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 (Jansen et al., 2013).

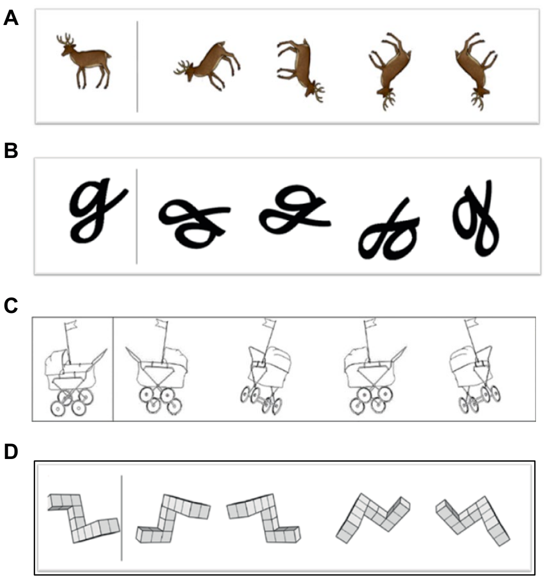


FIGURE 2
Examples of MRT items used with concrete objects: **(A)** animal stimulus and **(B)** letter stimulus (Quaiser-Pohl, 2003); **(C)** female-stereotyped stimulus "Pram" (Ruthsatz et al., 2015; Rahe and Quaiser-Pohl, 2019) and with an abstract object **(D)** "Cube" figure (Ruthsatz et al., 2014).

2.4 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.

Independent samples *t*-tests were conducted to investigate potential age and gender differences in the following variables: Maths self-concept (MSC); Child spatial anxiety (CSA); Perceived difficulty of stimuli (PD) and Accuracy and Response Times on the mental rotation task (MRT). Data for all dependent variables were examined for normality, outliers were identified and removed, and missing data were imputed with the median for each variable. The independent variables were found to be normally distributed.

A Chi-Squared test of association was used to check which groups were less likely to choose the preferred STEM subject, girls or boys and tweens or children in middle childhood. The assumptions for 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 had less than 5.

One-way ANOVA was used to assess differences in maths self-concept based on preferred school subjects was conducted. Assumptions of normal distribution of data were met but homogeneity of variance was not due to unequal sizes of STEM, Non-STEM, and Other (physical education) subject groups. Therefore, the ANOVA was run on a subsample of $N = 97$ with equal numbers of participants in each group.

A two-way multivariate analysis of variance (MANOVA) was used to assess the interaction effects of gender and stage of childhood development on child spatial anxiety, perceived difficulty, and mental rotation performance (accuracy and response time in the MRT). Assumptions for the MANOVA, such as homogeneity of variances (verified with Levene's test) and the absence of multivariate outliers (confirmed with Box's test), were satisfied (Tabachnick and Fidell, 2013).

Multiple Regression was performed to test the relationship between gender, age, maths self-concept, spatial anxiety, and perceived difficulty of the MRT and the relationship of maths-self-concept to accuracy and response time on the task. A mediation analysis was performed utilizing the PROCESS procedure in SPSS Version 4.2 (Hayes, 2022) to test whether the effects of maths self-concept on accuracy and response time were mediated by spatial anxiety and perceived difficulty of the task. Assumptions for regression and mediation analyses were validated, including the continuous nature of the criterion variable, absence of substantial outliers, normal distribution of residual scores, and the absence of multicollinearity (tolerance values below 0.2). Furthermore, for mediation analysis, independence of observations and absence of measurement error were insured. For moderator analyses, homoscedasticity was checked in PROCESS and was found not to be violated (MacKinnon, 2008).

3 Results

3.1 Gender differences in maths self-concept, spatial anxiety, perceived difficulty, mental rotation performance and preference for the STEM subject maths in primary school children

> *H.1. Girls will have significantly lower maths self-concept, higher spatial anxiety, and higher perceived difficulty than boys. Girls will have lower scores and longer response times on the mental rotation task than boys. Girls are less likely to choose the STEM subject maths than boys.*

An independent samples *t*-test with the dependent variables 'Maths Self-Concept, Child Spatial Anxiety and Perceived Difficulty' and independent variable 'Sex' found that there was a statistically significant difference in maths self-concept of boys ($M = 28.85$, $SD = 7.06$) and girls ($M = 25.57$, $SD = 7.67$) ($t(141) = 2.67$, $p = 0.009$, *CI* (95%) 0.850 → 5.72) with a small to medium effect size ($d = 0.45$). Therefore, girls showed significantly lower maths self-concept than boys. There was a statistically significant difference in spatial anxiety of boys ($M = 16.07$, $SD = 4.94$) and girls ($M = 18.20$, $SD = 5.05$) ($t(141) = -2.52$, $p = 0.013$, *CI* (95%) −3.80 → −0.46) with a small to medium effect size ($d = -0.42$). There was a statistically significant difference in perceived difficulty of boys ($M = 11.71$, $SD = 3.41$) and girls ($M = 13.10$, $SD = 3.10$) ($t(141) = -2.35$, $p = 0.014$, *CI* (95%) −2.44 → −0.28) with a small to medium effect size ($d = -0.42$). Therefore, girls showed significantly higher spatial anxiety and perceived difficulty than boys. There was no statistically significant difference between girls and boys in accuracy and response times on the mental rotation task (Table 2).

A Chi-Square test for association with the criterion variable 'Subject_Preference' and predictor variable 'Sex' revealed a statistically significant relationship between students' preference for the STEM subject and gender ($X2$ (1, $N = 141$) = 7.58, $p = 0.023$). The association between preference for the STEM subject and gender was moderately positive. Specifically, girls were less likely than boys to choose STEM subject maths (Figure 3).

3.2 Age differences in maths self-concept, spatial anxiety, perceived difficulty, mental rotation performance and preference for the STEM subject maths in primary school children

> *H.2. Tweens (9-to-11-year-olds) will have significantly lower maths self-concept, lower spatial anxiety and lower perceived difficulty than students in middle childhood (6-to-8-year-olds). Tweens will have significantly higher scores and shorter response times on the mental rotation task than students in middle childhood. Tweens are less likely to choose the STEM subject maths than students in middle childhood.*

An independent samples *t*-test with the dependent variable 'Maths Self-Concept, Child Spatial Anxiety and Perceived Difficulty' and independent variable 'Stage of Childhood Development' found that there was a statistically significant difference in maths self-concept of the tween group ($M = 25.79$, $SD = 7.57$) and the middle childhood group ($M = 28.67$, $SD = 7.24$) ($t(141) = 2.32$, $p = 0.022$, *CI* (95%) 0.43 → 5.33) with a small to medium effect size ($d = 0.39$). Therefore, students in the tween group demonstrated significantly lower maths self-concept than students in the middle childhood group. There was no difference in spatial anxiety and perceived difficulty between the tween group and the middle childhood group. The was a statistically significant difference between the tween group ($M = 0.35$, $SD = 0.20$) and the middle childhood group ($M = 0.27$, $SD = 1.67$) ($t(141) = -2.66$, $p = 0.009$, *CI* (95%) −0.141 → −0.021) with a small to medium effect size ($d = -0.45$) in accuracy on the mental

TABLE 2 Independent samples test: gender differences in maths self-concept (MSC), spatial anxiety (CSA) and perceived difficulty of the mental rotation task (PD).

	Levene's test for equality of variances		*t*-test for equality of means				
	F	Sig.	*t*	df	Significance	95% Confidence interval of the difference	
					Two-sided *p*	Lower	Upper
MSC	1.997	0.162	2.667	141	0.009	0.85045	5.72108
CSA	0.002	0.969	−2.522	141	0.013	−3.78837	−0.45922
PD	0.223	0.638	−2.504	141	0.014	−2.43947	−0.28575

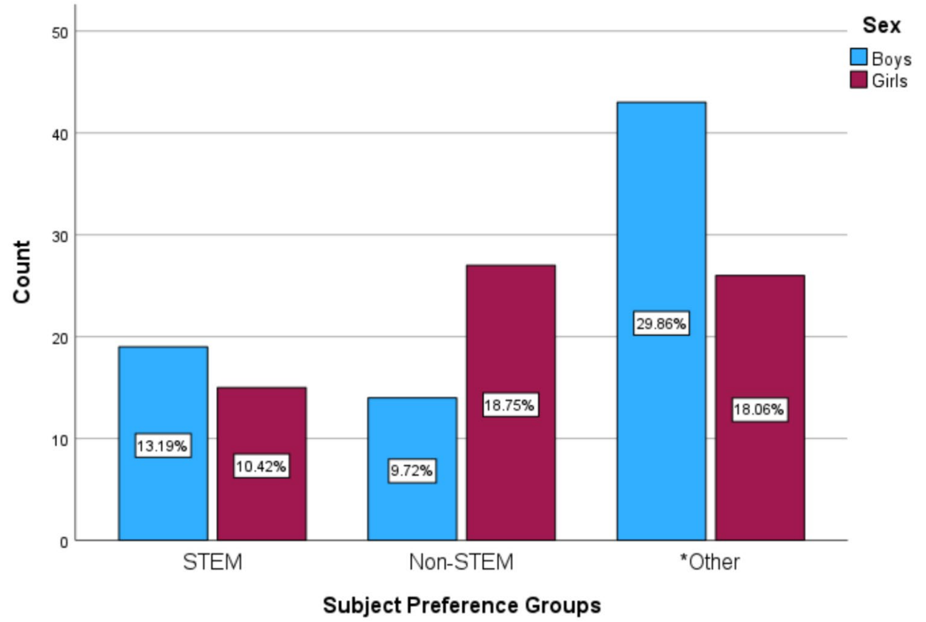


FIGURE 3
Percentage of STEM or non-STEM as the preferred subject at school by gender (sex). *'Other' represents physical education (P.E.) as the preferred subject.

TABLE 3 Independent samples test: age differences in maths self-concept (MSC), and accuracy on the mental rotation task.

	Levene's test for equality of variances		*t*-test for equality of means				
	F	Sig.	*t*	df	Significance	95% confidence interval of the difference	
					Two-sided *p*	Lower	Upper
MSC	0.862	0.355	2.234	141	0.022	0.42953	5.32764
Accuracy	5.115	0.025	−2.657	136.725	0.009	−0.14144	−0.02073

rotation test but there was no difference in response times. Therefore, tweens scored higher than students in middle childhood on the mental rotation task (Table 3).

A Chi-Square test for association with the criterion variable 'Subject_Preference' and predictor variable 'Stage of Childhood Development' revealed a statistically significant relationship between students' preference for the STEM subject and developmental stage ($X2$ (1, $N = 142) = 15.66$, $p < 0.001$). The association between preference for the STEM subject and developmental stage was moderately positive. Specifically, tweens were less likely than students in middle childhood to choose the STEM subject maths (Figure 4).

A two-way MANOVA was used to test the interaction effects of Gender and Stage of Childhood Development on Spatial Anxiety, Perceived Difficulty, Accuracy and Response Time on the Mental Rotation Task (MRT). Following a Bonferroni adjustment at 0.025, there were no significant interaction effects of gender and age on perceived difficulty or on accuracy and response time in the MRT. There were, however, significant interaction effects of gender and age on spatial anxiety (F (1, 143) = 5.81, $p = 0.017$) with a small to medium effect size ($\eta p^2 = 0.04$). Pairwise comparisons highlight there were no differences in spatial anxiety between girls and boys in the tween group but spatial anxiety in girls in the middle childhood group was significantly higher than in boys in the same group (*Mean Difference* = 3.33) (Figure 5).

3.3 Students' preference for STEM or non-STEM subjects predicts maths self-concept

> *H.3. Students who prefer the STEM subject maths at school will have higher maths self-concept than students who prefer non-STEM subjects.*

A one-way analysis of variance with dependent variable 'Maths Self-Concept' and Between Subjects Factor 'STEM_Subject' was conducted on a subsample of the data ($N = 97$) in order to ensure group homogeneity in the STEM, Non-STEM (Art, Music, Languages) and Other (Physical Education or P.E.) groups. All groups had thirty-two participants ($N = 32$) with forty-nine boys

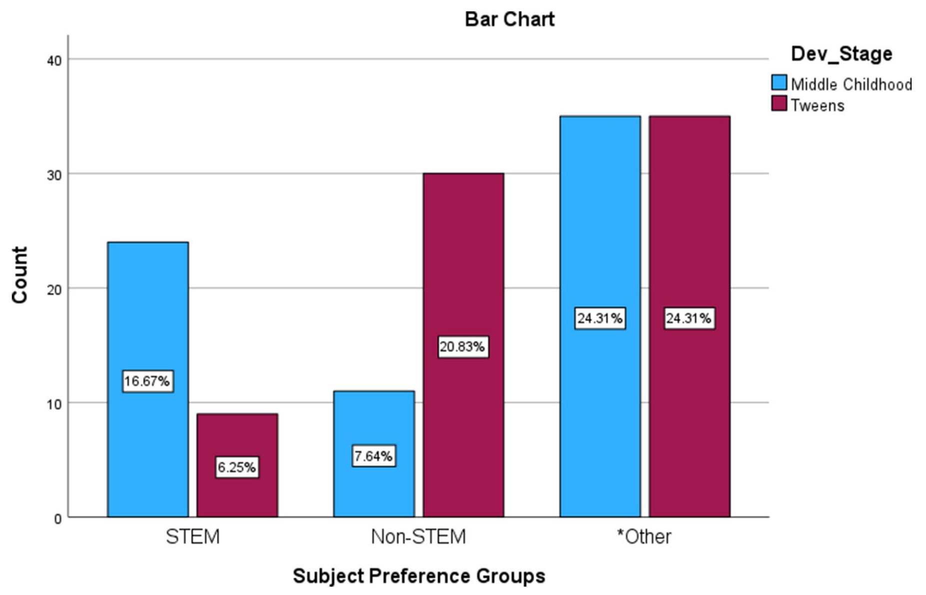


FIGURE 4
Percentage of STEM or non-STEM as the preferred subject at school by stage of childhood development (Dev_Stage). *'Other' represents physical education (P.E.) as the preferred subject.

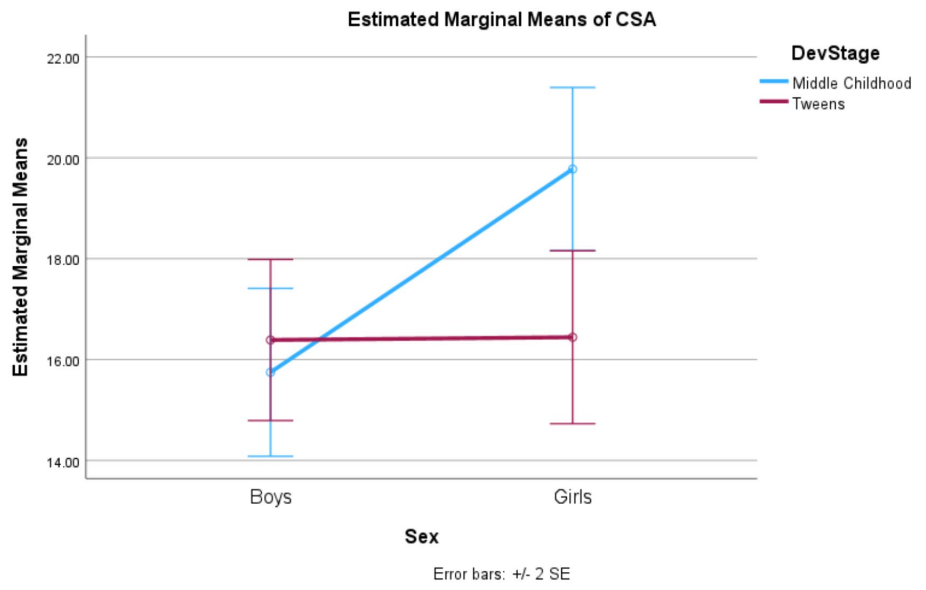


FIGURE 5
Interaction effects of gender and developmental stage on spatial anxiety in primary school students.

($N = 49$) and fifty-one girls ($N = 51$). Fifty-two students were in the Middle Childhood stage of development ($N = 52$) and there were forty-eight tweens ($N = 48$). The ANOVA was run to examine the relationship between students' preferred subject at school and their levels of maths self-concept. The results indicated a statistically significant difference in maths self-concept on this subsample among the three preferred subject groups (F (2, 95) = 10.70, $p < 0.001$) with a large effect size ($\eta^2 = 0.19$). This large effect size indicates that 19%, so a substantial amount of the variance, in maths self-concept can be explained by students' subject preferences. Subsequent Tukey HSD *post hoc* analyses were conducted to further explore these differences and this revealed the following significant differences in maths self-concept:

In the STEM subject preference group, maths self-concept was significantly higher than in the non-STEM group (*Mean difference* = 8.22, $p < 0.001$, 95% *CI* [3.92, 12.53]).

In the STEM subject preference group, maths self-concept was not significantly higher than in the Other (P.E.) group (*Mean difference* = 2.78, $p = 0.277$, 95% *CI* [−1.52, 7.10]).

In the Other (P.E.) preference group, maths self-concept was significantly higher than in the non-STEM group (*Mean difference* = 5.44, $p = 0.009$, 95% *CI* [1.13, 9.75]).

10.3389/feduc.2024.1300598

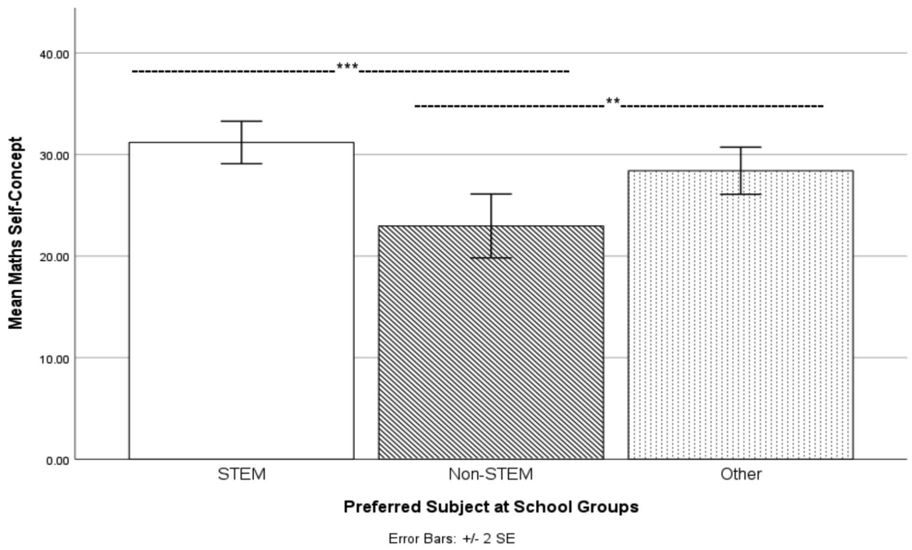


FIGURE 6
Differences in maths self-concept by preference for STEM, non-STEM and other subjects at school. STEM represents maths, non-STEM represents art, music & languages and other represents physical education (P.E.). (*** demonstrates statistical significance at $p < 0.001$; ** demonstrates statistical significance at $p < 0.05$).

In summary, students who indicated a preference for the STEM subject Maths or for P.E. demonstrated significantly higher levels of maths self-concept compared to those who preferred a non-STEM subject (Figure 6).

3.4 Maths self-concept is significantly associated with spatial anxiety and perceived difficulty of the mental rotation task

> *H.4. Maths self-concept is significantly associated with spatial anxiety and perceived difficulty of the mental rotation task in primary school students. Specifically, as maths self-concept increases, spatial anxiety decreases. Additionally, as maths self-concept increases, perceived difficulty decreases.*

Multiple regression with criterion variable 'Child Spatial Anxiety' and predictor variables 'Sex', 'Age' and 'Maths Self-Concept' was used to analyse the association of maths self-concept with spatial anxiety in primary school children. The results of the regression indicated that three predictors explained 18% of the variance in Child Spatial Anxiety ($R^2adj = 0.183$, F (3, 142) = 11.60, $p < 0.001$). Age ($\beta = -0.25$, $p = 0.003$, 95% $CI = -1.83 \rightarrow -0.45$) and Maths Self-Concept ($\beta = -0.38$, $p = <$ 0.001, 95% $CI = -0.36 \rightarrow -0.15$) are significantly associated with spatial anxiety in primary school children but gender is not ($\beta = 0.11$, $p = 0.174$, 95% $CI = -0.49 \rightarrow 2.68$) The negative beta coefficient for age and maths self-concept suggests that, as students' age and maths self-concept increase, spatial anxiety tends to decrease (Table 4).

Multiple regression with criterion variable 'Perceived Difficulty' and predictor variables 'Sex', 'Age' and 'Maths Self-Concept' was used to analyse the association of maths self-concept with perceived difficulty of the mental rotation task in primary school children. The results of the regression indicated that three predictors explained 7% of the variance in Perceived Difficulty ($R^2adj = 0.073$, F (3, 142) = 4.72, $p = 0.004$). Maths Self-Concept ($\beta = -0.23$, $p = 0.007$, 95% $CI = -0.18 \rightarrow -0.03$) is significantly associated with Perceived Difficulty but Gender ($\beta = 0.15$, $p = 0.076$, 95% $CI = -0.10 \rightarrow 2.10$) and Age ($\beta = -0.08$, $p = 0.341$, 95% $CI = -0.72 \rightarrow -0.25$) and are not. The negative beta coefficient for maths self-concept suggests that, as students' maths self-concept increases, their perceived difficulty of the task tends to decrease (Table 5).

3.5 Maths self-concept predicts accuracy and response time on the mental rotation task and is mediated by spatial anxiety and perceived difficulty

> *H.4. Maths self-concept is a significant predictor of accuracy and response time on the mental rotation task and this association is mediated by spatial anxiety and perceived difficulty.*

3.5.1 Accuracy

A simple mediation was performed utilizing the PROCESS procedure in SPSS Version 4.2 (Hayes, 2022) was conducted to examine the indirect effect of Maths Self-Concept (MSC) on Accuracy on the MRT (Acc_MRT) through two proposed mediators: Spatial Anxiety (CSA) and Perceived Difficulty of the MRT (PD).

The total effect of MSC on Acc_MRT was found to be marginally significant, $\beta = 0.0038$, $SE = 0.0019$, $t(141) = 1.96$, $p = 0.052$, 95% CI [0.0000, 0.0077], completely standardized indirect effect (CSI) = 0.1542. The direct effect of MSC on Acc_MRT was not significant, $\beta = 0.0029$, SE = 0.0020, $t(141) = 1.42$, $p = 0.159$, 95% CI [−0.0011, 0.0068], completely standardized direct effect (CSD) = 0.1153.

TABLE 4 Coefficients: age and maths self-concept (MSC) predict spatial anxiety (CSA) in primary school children.

Model		Unstandardized coefficients		Standardized coefficients	t	Sig.	95.0% confidence interval for B	
		B	Std. error	Beta			Lower bound	Upper bound
1	(Constant)	33.358	3.769		8.851	<0.001	25.906	40.810
	Sex	1.094	0.801	0.107	1.366	0.174	−0.489	2.677
	Age	−1.156	0.357	−0.252	−3.234	0.002	−1.863	−0.449
	MSC	−0.256	0.054	−0.376	−4.717	<0.001	−0.364	−0.149

a. Dependent Variable: CSA.

TABLE 5 Coefficients: maths self-concept (MSC) predicts perceived difficulty (PD) of the MRT in primary school children.

Model		Unstandardized coefficients		Standardized coefficients	t	Sig.	95.0% confidence interval for B	
		B	Std. error	Beta			Lower bound	Upper bound
1	(Constant)	16.683	2.597		6.425	<0.001	11.549	21.817
	Sex	0.987	0.552	0.149	1.790	0.076	−0.103	2.077
	Age	−0.235	0.246	−0.079	−0.956	0.341	−0.722	0.252
	MSC	−0.103	0.037	−0.233	−2.743	0.007	−0.177	−0.029

a. Dependent variable: PD.

Indirect effects analysis revealed that the total indirect effect of MSC on Acc_MRT through all mediators combined was significant, $\beta = 0.0010$, $BootSE = 0.0009$, 95% CI [−0.0006, 0.0031], $CSI = 0.0389$. Specifically, CSA mediated this relationship with a significant path a of the indirect effect, $\beta = 0.0003$, $BootSE = 0.0009$, 95% CI [−0.0014, 0.0022], $CSI = 0.0133$. However, path a of the indirect effect through PD was not significant, $\beta = 0.0006$, $BootSE = 0.0006$, 95% CI [−0.0003, 0.0021], $CSI = 0.0256$. Moreover, the specific indirect effect contrast (*C1*) comparing the mediation effects of CSA and PD showed a non-significant indirect effect, $\beta = -0.0003$, $BootSE = 0.0012$, 95% CI [−0.0029, 0.0020], $CSI = -0.0123$.

These findings suggest that the relationship between Maths Self-Concept and Accuracy on the MRT is partially mediated by Spatial Anxiety, but not by Perceived Difficulty (Figure 7).

3.5.2 Response time

A mediation analysis was conducted to explore the indirect effect of Maths Self-Concept (MSC) on Response Times on the Mental Rotation Task (RT_MRT) through Spatial Anxiety (CSA) and Perceived Difficulty of the MRT (PD).

The total effect of MSC on RT_MRT was not statistically significant, $\beta = 0.1016$, $SE = 0.0768$, $t(141) = 1.32$, $p = 0.188$, 95% CI [−0.0504, 0.2535], completely standardized indirect effect (CSI) = 0.1086. The direct effect of MSC on RT_MRT was also not significant, $\beta = 0.0811$, $SE = 0.0824$, $t(141) = 0.98$, $p = 0.327$, 95% CI [−0.0818, 0.2441], $CSI = 0.0867$.

Further examination of the indirect effects revealed that the total indirect effect of MSC on RT_MRT through all mediators combined was not significant, $\beta = 0.0204$, $BootSE = 0.0337$, 95% CI [−0.0428, 0.0924], $CSI = 0.0218$. Specifically, while path a of the indirect effect through CSA was significant, $\beta = 0.0441$, $BootSE = 0.0344$, 95% CI [−0.0189, 0.1158], CSI = 0.0471, path a of the indirect effect through PD was not significant, $\beta = -0.0236$, $BootSE = 0.0236$, 95% CI [−0.0759, 0.0158], $CSI = -0.0253$. Moreover, the specific indirect effect contrast (*C1*) comparing the mediation effects of CSA and PD showed a non-significant indirect effect, $\beta = 0.0677$, $BootSE = 0.0484$, 95% CI [−0.0213, 0.1700], $CSI = 0.0724$.

These results suggest that neither Spatial Anxiety nor Perceived Difficulty significantly mediate the relationship between Maths Self-Concept and Response Time on the MRT.

4 Discussion

4.1 Gender and age differences in children's maths self-concept, spatial anxiety, and perceived difficulty

In this study, notable gender disparities emerged in maths self-concept, spatial anxiety, and perceived difficulty of mental rotation tasks. Specifically, primary school girls exhibited markedly lower maths self-concept compared to boys, alongside reporting higher levels of spatial anxiety and perceived difficulty in spatial tasks. These findings underscore potential discrepancies in confidence and comfort levels among male and female students when navigating mathematical and spatial concepts during primary education.

Variances in emotional expression between genders are believed to stem from a combination of biologically influenced temperamental predispositions and socialization processes (Sanchis-Sanchis et al., 2020). Boys and girls tend to adhere to gender-specific norms regarding the expression of emotions; for instance, girls often exhibit more positive emotions but also display higher rates of negative internalizing emotions such as sadness and anxiety (Brody and Hall, 2008). Moreover, gender disparities in emotional expression can be influenced by the nature of the task children are engaged in (Chaplin and Aldao, 2013). Boys' increased exposure to

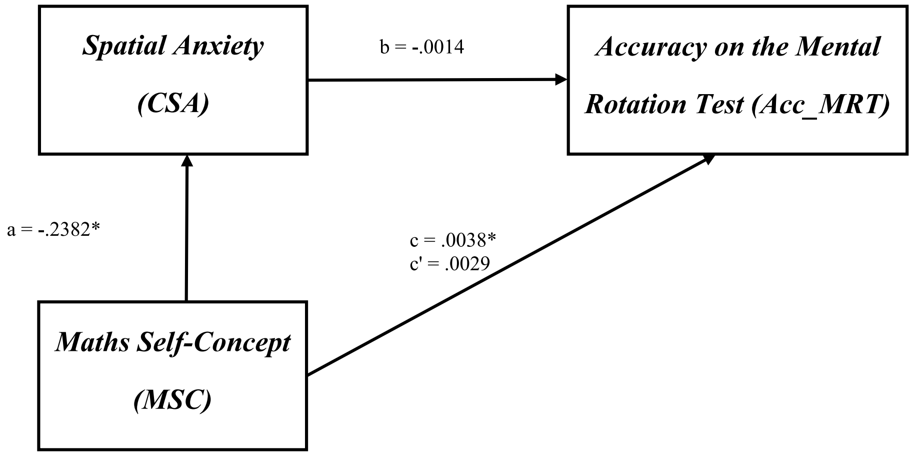


FIGURE 7
Simple mediation diagram: association between maths self-concept (MSC) and accuracy on the mental rotation task (Acc_MRT) as mediated by spatial anxiety (CSA). Paths a and b represent the indirect (mediated) effects of MSC on Acc_MRT through CSA. Path c represents the total (mediated and direct) effect of maths self-concept on accuracy in the MRT. Path c′ represents the direct effect of MSC on Acc_MRT and is calculated controlling for the indirect, mediated effect. Paths a and c are significant (*) indicating the association between MSC and Acc_MRT is partially mediated by CSA.

spatially-related objects and activities may contribute to their reduced spatial anxiety and perceived difficulty on mental rotation tasks (Ruthsatz et al., 2014). Furthermore, parental and teacher stereotypes, as well as their own anxieties regarding spatial tasks and related subjects, can potentially impede the development of girls' spatial skills and their inclination towards STEM disciplines (Gunderson et al., 2012; Garcia et al., 2021; Rocha et al., 2022). Hence, it is imperative to address these issues through targeted teacher training initiatives and support programs for parents of primary school children.

In the tween group, noticeably lower maths self-concept was observed compared to the middle childhood group, although no significant disparities emerged in spatial anxiety and perceived difficulty between these developmental phases. A decline in maths self-concept in the tween group may be a result of an emerging awareness of negative stereotypes associated with individuals who excel in maths and STEM (Starr, 2018). Fear of confirming negative stereotypes about maths and STEM can demotivate young people from engaging in these subjects preceding adolescence when social norms, hierarchies and a desire to fit in among peers is important (Cheryan et al., 2013).

Several factors, including developmental growth, exposure, and experience, may elucidate outcomes for spatial anxiety and perceived task difficulty. The tween stage signifies a period marked by substantial cognitive and socio-emotional development. Notably, neurobiological maturation during this stage enhances emotion regulatory processes, enabling older children to better manage task-related anxiety (Sanchis-Sanchis et al., 2020). As children progress through this stage, they typically experience improvements in cognitive abilities, including spatial reasoning skills, potentially contributing to a diminished perceived difficulty in spatial tasks (McGlone and Aronson, 2006; McArthur et al., 2021; McGuire et al., 2022). Furthermore, tweens are likely to have greater exposure to spatial tasks and activities, both within and outside of educational settings, compared to younger children. This increased exposure fosters familiarity with spatial tasks, thereby potentially reducing levels of spatial anxiety and perceived difficulty (Neuburger et al., 2011; Peterson et al., 2020).

These insights underscore the dynamic interplay between cognitive development, emotional regulation, and experiential learning for girls and during the tween stage, shaping perceptions and abilities related to mathematical and spatial domains.

4.2 Gender and age differences in children's' preference for STEM subjects, and mental rotation performance

Our analysis also yielded significant results with regards to students' preference for maths as a subject in primary school. Our findings show girls were less likely to choose maths as were children in the tween group. We observed that girls in primary education demonstrate a persistent belief that maths is not a subject with which they identify. Socialization and environmental factors may contribute significantly to children's relationship with maths and this may account for differences in interest and achievement in the STEM field beyond the school years. Moreover, low maths self-concept in the tween group suggests that as children progress through development, their preferences for STEM subjects change, warranting further investigation into the underlying factors. These shifts may be evidence that pre-adolescents as well as girls stereotypically believe maths and other STEM subjects to be male-dominated domains which are only for the most mathematically-gifted. The large effect size yielded by the ANOVA indicates that students' preferences for school subjects had a substantial impact on their maths self-concept. Specifically, students who expressed a preference for maths demonstrated higher levels of maths self-concept compared to those who preferred non-STEM subjects such as art, music, or languages.

Contrary to the bulk of research and despite girls' lower maths self-concept, higher spatial anxiety and perceived difficulty, we found no significant gender differences in mental rotation performance, neither accuracy nor response time, among students. Gender-specific differences in spatial experiences, role models, ability-related self-concept, and socio-economic status are well-established factors that exert a significant influence on mental rotation performance (Ruthsatz and Quaiser-Pohl, 2013). This study was conducted in Germany, known for its strong emphasis on education and social welfare. In this country, significant investments are made to provide equal access to resources and opportunities for all students, regardless of gender or socio-economic background (OECD, 2020).

Other factors may also be influential, such as the nature of mental rotation task used in the study: This included stimuli that appeal to both males and females (Ruthsatz et al., 2015) and stimuli such as polyhedral figures and other objects known to be gender-neutral (Ruthsatz et al., 2014, forthcoming). Additionally, time limits on the task, particularly on the second subscale consisting of stimuli rotated in-depth were longer than on a standard MRT to allow for gender-diverse processing strategies (Taragin et al., 2019; Saunders and Quaiser-Pohl, 2020). These factors may have diminished the power of negative emotions such as anxiety as well as stereotype threat, resulting in more equity across the task.

Despite dwindling interest in maths, tweens demonstrated higher accuracy in the mental rotation task compared to students in middle childhood, indicating potential advancements in spatial cognition and emotion suppression strategies as children progress through development and education (Neuburger et al., 2011; Sanchis-Sanchis et al., 2020).

Regarding girls' and tweens' lower maths self-concept, our findings suggest these groups may be at risk of a decline in interest and hence in achievement in STEM subjects. A trend towards disengagement from maths and related subjects at primary level has implications for childrens' academic and professional options for the future. If students' subject preferences are based on poor self-concept rather than actual ability, the STEM field will lose many potential candidates before they enter secondary school. Interventions such as peer mentoring to counteract negative attitudes and stereotypes for girls and tweens are helpful in maintaining their engagement with maths and STEM going forward (Quaiser-Pohl et al., 2014a; Ada Lovelace-Projekt, 2023; Space Science Institute, 2023). Moreover, as reciprocal relations between active participation in mathematics practice and maths self-concept have been observed, it is important to sustain and encourage students' engagement at primary level and beyond (Sewasew et al., 2018).

An unexpected outcome of this study was children's preference for sport and physical education (P.E. or "Sport" in German). Indeed, it was the most popular choice in our sample overall with more boys choosing P.E. as their preferred subject than girls. Moreover, maths self-concept in students who indicated a preference for physical education (P.E.) was significantly higher than those who preferred the non-STEM subjects but not substantially higher than in those who preferred maths. A close link between physical activity, mathematical skills and the development of spatial ability has already been identified in research (Jansen and Pietsch, 2022). Findings demonstrate that sport and regular exercise lead to enhancement of children's visuo-spatial skills (Morawietz and Muehlbauer, 2021). Moreover, specific physical activities such as kinetics and creative dance can improve mental rotation in primary school children (Jansen et al., 2013; Pietsch et al., 2017). Therefore, promoting sports and offering attractive physical education options for girls may be beneficial in improving mental rotation during the primary school years.

The findings of the two-way MANOVA examining the interplay between gender and age on aspects of the mental rotation task, spatial anxiety and the perceived difficulty of the task revealed interesting insights. Despite employing a stringent Bonferroni adjustment, no significant interaction effects emerged for perceived difficulty, accuracy, or response time in the MRT. However, a notable interaction effect was observed for spatial anxiety, underscoring the nuanced influence of gender and age on this domain. Specifically, girls in the middle childhood stage exhibited significantly higher levels of spatial anxiety compared to boys in the same developmental phase, contrasting with the absence of such disparity in the tween group. This outcome suggests that gender differences in spatial anxiety may manifest more prominently during certain developmental stages, possibly reflecting social or cognitive factors that vary across age groups. The effect size, although modest, implies a meaningful contribution of this interaction to understanding individual differences in spatial cognition during childhood. These findings underscore the importance of considering developmental nuances when exploring gender-related disparities in cognitive and emotional processes, highlighting avenues for future research to delve deeper into the underlying mechanisms driving such variations.

We conclude that while gender and age play a role in maths self-concept and spatial anxiety, other factors such as prior experience in tasks or activities involving mental rotation, such as playing certain video games or engaging in activities such as puzzles or construction play may be influencing performance on the spatial task and further research is needed to explore these (Quaiser-Pohl et al., 2006; Milani et al., 2019; Ruthsatz et al., 2019; Liu et al., 2020). Factors such as stress or fatigue, during test administration can also affect cognitive performance and might be explored using physiological measures such as skin conductance measurement (Memar and Mokaribolhassan, 2021). Higher motivation and interest in completing the task may also lead to improved performance (Moè, 2016). Variations in learning styles, cognitive strategies, and problem-solving approaches may also affect how individuals approach and perform on mental rotation tasks (Janssen and Geiser, 2010; Khooshabeh et al., 2013; Moè, 2013). Moreover, sample characteristics or methodological constraints may have affected the results warranting further exploration of individual differences or cognitive factors that might influence childrens' performance on spatial tasks.

4.3 Predictive effects of gender, age, maths self-concept on spatial anxiety and perceived difficulty of the spatial task

Due to its long-established and robust link to mathematics, the focus of this study was ultimately the impact of motivational factors such as maths self-concept on spatial ability. By considering gender, age, maths self-concept, we aimed to explore potentially predictive influences on spatial anxiety and perceived difficulty of the task in primary school children. Our findings reveal that the three predictors collectively account for a significant proportion of the variance in spatial anxiety and perceived difficulty. While age, maths self-concept emerged as significant for spatial anxiety in the regression model, gender was not. Specifically, older children and those with higher levels of maths self-concept tended to exhibit lower levels of spatial anxiety. Furthermore, maths self-concept accounted for a noteworthy proportion of the variance in students' perceived difficulty of the mental rotation task but gender and age did not.

Contrary to expectations, gender was not found to be a significantly predictor of spatial anxiety nor perceived difficulty in our sample. While gender differences in spatial abilities and anxiety have been documented in previous research (Ramirez et al., 2012; Arrighi and Hausmann, 2022), our findings suggest that other factors, such as age, maths self-concept may play a more prominent role in influencing spatial anxiety among this particular sample. The effect of gender on spatial anxiety and indeed perceived difficulty of the task may be moderated by, for example, environmental circumstances, individual learning experiences, or socio-cultural influences, factors not included in our analysis. Greater access to educational materials, extracurricular activities, and supportive family environments that foster the development of spatial skills can mitigate gender disparities and stereotype threat by providing girls with opportunities to develop and practice spatial reasoning abilities (Quaiser-Pohl and Endepohls-Ulpe, 2012; Ruthsatz and Quaiser-Pohl, 2013). Additionally, as outlined in the previous section, the nature of the mental rotation task itself may have reduced the risk of provoking stereotype threat. A pilot study in which primary school children were asked to rate the stereotyped nature of stimuli and the perceived difficulty of the same task found students identified objects as gendered, neutral, or difficult and this in turn impacted accuracy and response time on these items during the task (Lennon-Maslin et al., 2023).

Additionally, both researchers collecting the data were female, a feature of the study which may have contributed to the non-significant gender effects. The researchers may potentially have been perceived as positive science role models mitigating stereotype threat in girls. Additionally, the testing took place in small mixed gender groups of students from the same class, who were therefore familiar with one another. Moreover, securing participation from schools for this large research project posed a significant challenge. However, those schools who agreed to participate typically employed principals and teaching staff who were dedicated to fostering students' interest in STEM subjects. Moreover, parents, by the very nature of providing consent for their children to participate in the study, are actively nurturing and supporting their interest in science and research.

To conclude, interventions aimed at addressing socio-economic disparities, such as providing equal access to resources and support, can help level the playing field and improve girls' performance on mental rotation tasks. However, interventions to overcome anxiety and improve spatial skills should focus on all children regardless of gender whose maths self-concept has been identified as low. This may in turn reduce their perceived difficulty of spatial tasks and increase the likelihood of engagement with related activities and STEM subjects.

4.4 Mediating effects of spatial anxiety and perceived difficulty on the relationship between maths self-concept and accuracy on the mental rotation task

Concerning on the mental rotation task, the total effect of maths self-concept approached statistical significance. Moreover, path a of the indirect effect through spatial anxiety was significant, indicating that higher levels of maths self-concept are associated with higher scores on the mental rotation task via reduced spatial anxiety. This finding highlights the importance of fostering positive self-concept in maths from an early age. Children who feel confident in their mathematical abilities may experience less spatial anxiety, which can positively impact their performance on spatial tasks such as a mental rotation test. Therefore, interventions aimed at reducing spatial anxiety could potentially improve children's performance in both spatial and mathematical domains (Lauer et al., 2018; Garcia et al., 2021).

4.5 Mediating effects of spatial anxiety and perceived difficulty on the relationship between maths self-concept and response times on the mental rotation task

The results of the mediation analysis regarding the relationship between maths self-concept and response time on the mental rotation task revealed that neither the direct nor the total effect of maths self-concept on accuracy was statistically significant. Additionally, the indirect effects through spatial anxiety and perceived difficulty were not significant, suggesting that these variables do not mediate the relationship between maths self-concept and response time on the mental rotation task. Other psychological constructs or variables not included in the analysis might be more relevant in mediating this relationship. For example, motivation or task-specific strategies can influence both maths self-concept and performance on spatial tasks in primary school children (Moè, 2016). Furthermore, the relationship between maths self-concept, anxiety, and response time on the mental rotation task may be bidirectional or reciprocal: A recent study found that performance on mental rotation tasks mediates gender differences in maths anxiety (Rahe and Quaiser-Pohl, 2019), creating complex feedback loops that are not adequately captured in a simple mediation model. Therefore, further research investigating the interplay of other influential factors alongside maths self-concept and demographic variables such as gender and age could provide a more comprehensive understanding of response times on mental rotation tasks.

4.6 Implications of mediation findings for spatial ability development

The results of the mediation analyses underscore the interconnectedness of various cognitive skills. Positive self-concept in mathematics is known to influence mathematical performance but also indirectly affects spatial performance through reduced spatial anxiety. This suggests that interventions targeting one skill area may have spillover effects on other related skills. Specifically, exposure to mathematical and spatial activities in a supportive environment may raise comfort levels in children who might otherwise avoid these pursuits. Therefore, addressing spatial anxiety may have implications for improving performance on spatial tasks among primary school children, emphasizing the importance of considering affective factors in educational interventions aimed at enhancing mathematical abilities. However, this warrants further investigation in future research. These results highlight the need for a nuanced understanding of the factors influencing performance on cognitive tasks such as mental rotation and suggests avenues for future research and intervention development.

4.7 Limitations and future research

One limitation of this study is that maths performance was not measured. This will be collected in a future study. Moreover, the measurement of other emotional mediators such as maths anxiety may also contribute significantly and provide additional evidence to compliment maths self-concept. Future research may also benefit from the exploration of physiological measures of emotional arousal during a mental rotation task to counteract reliance on self-report.

Upon analysis of the data from our sample, the Child Spatial Anxiety Questionnaire, yielded internal reliability in the lower range. Removal of items with a reliability rating below the acceptable threshold of $\alpha = 0.70$ did not lead to an increased Cronbach's Alpha indicating that 8 items were not sufficient to assess the construct reliably. Ramirez et al. (2012) also refer to a low reliability rating for the CSAQ based on the number of items. These were purposely kept small in order to reduce fatigue in child participants (Ramirez et al., 2012). We also felt it was important to minimize participant fatigue, especially considering the age group and the number of other measures involved. Furthermore, we are not aware of any other measure of child spatial anxiety suitable for use in the busy school setting.

Although reliability of the Perceived Difficulty of items on the MRT questionnaire was within the acceptable range, there were more male-stereotyped objects (the hammer and the cubical figure) than female and neutral objects (the pellet and the polyhedral figure) included in this instrument. Upon analysis, gender differences in response times in favour of boys were found on the male-stereotyped "hammer" in the MRT. It may be, that this item triggered stereotype threat in girls and stereotype lift in boys. The questionnaire will be amended to include another female-stereotyped object to avoid the risk of triggering stereotype threat or lift in future studies.

The results of the regression analysis should be interpreted with caution as the study was cross-sectional. A future research project might adopt a longitudinal approach which follows a cohort of students across their primary school years, in order to study the effects of the predictor variables over time.

With regards to the ANOVA results, the sample size fell short of the G*Power *A priori* calculation of optimal sample size. Therefore, the ANOVA was run on a subsample in order to avoid violating assumptions of equal variances in groups. The effect size was however noteworthy with a large proportion of the variance in students' maths self-concept being explained by their preference for subjects at school. Important to note however, that students did not choose any other preferred STEM subject, so that the results can effectively only be applied to maths with potential implications for STEM.

To minimize the risk of introducing bias or stereotype threat and improve the validity of the study, collecting demographic data such as gender and age might best be done in conclusion rather than preceding testing.

5 Conclusion

This study provides insights into critical factors influencing children's attitudes and performance in mathematics, spatial tasks, and STEM subjects. One key finding is the persistent gender disparity in maths self-concept, spatial anxiety, perceived difficulty of spatial tasks, and preference for maths as a STEM subject, highlighting enduring negative stereotypes in these fields. Similarly, low maths self-concept and loss of interest and enthusiasm for maths in late primary school children warrants attention in order to prevent disengagement from STEM-related subjects in further education. Counteracting stereotypes early is essential to ensure childrens' academic choices are based on ability rather than attitude. The findings emphasize the importance of addressing spatial anxiety and promoting positive maths self-concept to enhance performance on spatial tasks among primary school children. Interventions aimed at improving maths self-concept may indirectly benefit spatial task performance by alleviating spatial anxiety. Further investigation is warranted to explore the complex interplay between cognitive, affective, and demographic factors influencing performance on spatial tasks in primary school children. Additionally, longitudinal studies could provide insights into the developmental trajectories of maths self-concept and spatial skills. In conclusion, while maths self-concept plays a role in performance on spatial tasks, its influence is mediated by factors such as spatial anxiety, with gender showing limited moderating effects. Understanding these dynamics is crucial for developing targeted interventions to effectively support children's mathematical self-concept and spatial skills development in educational settings.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Ethikkommission, Universität Koblenz. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin. Written informed consent was obtained from the minor(s)' legal guardian/next of kin for the publication of any potentially identifiable images or data included in this article.

Author contributions

ML-M: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. CQ-P: Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Writing – review & editing. L-CW: Formal analysis, Supervision, Writing - review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was funded by the Marie Sklodowska-Curie Actions International Training Network (MSCA ITN) "SellSTEM" (Spatially Enhanced Learning linked to Science, Technology, Engineering and Maths), grant number 956124.

Acknowledgments

I would like to express my sincere thanks to the principals, teachers, parents and the students of the five schools who took part in our study. Without your participation, this research would not have been possible. I would also like to thank the students of the empirical psychological training course at the Institute of Psychology for their assistance with the acquisition of schools and data. Also, thanks to Mirko Saunders for programming the experiment and providing technical support. Thanks also to research assistants Annika Volkmann and Amelie Tietz for their hard work throughout the project. Finally, this article is dedicated to the Ada Lovelace Project who have been inspiring girls and young women in STEM for two and a half decades. I wish you, the mentors and their mentees many more successful years ahead.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2024.1300598/full#supplementary-material

References

Ada Lovelace-Projekt. (2023). *MINT-mentoring—Ada-Lovelace-Projekt*. Available at: https://ada-lovelace.de/mint-mentoring/

Alloway, T., and Passolunghi, M. (2011). The relationship between working memory, IQ and mathematical skills in children. *Learn. Individ. Differ.* 21, 133–137. doi: 10.1016/j.lindif.2010.09.013

Aronson, E., Wilson, T. D., and Akert, R. M. (2014). *Social psychology*. *8th* Edn. Boston: Pearson Education.

Arrighi, L., and Hausmann, M. (2022). Spatial anxiety and self-confidence mediate sex/gender differences in mental rotation. *Learn. Mem.* 29, 312–320. doi: 10.1101/lm.053596.122

Attard, C. (2013). "If I had to pick any subject, it wouldn't be maths": foundations for engagement with mathematics during the middle years. *Math. Educ. Res. J.* 25, 569–587. doi: 10.1007/s13394-013-0081-8

Berger, A., Tzur, G., and Posner, M. I. (2006). Infant brains detect arithmetic errors. *Proc. Natl. Acad. Sci.* 103, 12649–12653. doi: 10.1073/pnas.0605350103

Boaler, J. (2015). *The elephant in the classroom: Helping children learn and love Maths*. London: Souvenir Press.

Brody, L. R., and Hall, J. A. (2008). "Gender and emotion in context" in *Handbook of emotions*. eds. L. F. Barrett, M. Lewis and J. M. Haviland-Jones. *3rd* ed (New York: The Guilford Press), 395–408.

Brunstein, J. C., and Schmitt, C. H. (2010). "Chapter 6 – Assessing individual differences in achievement motivation with the implicit association test: predictive validity of a chronometric measure of the self-concept "me = successful"" in *Implicit motives*. eds. O. Schultheiss and J. Brunstein (Oxford: Oxford University Press).

Buckley, J., Seery, N., and Canty, D. (2018). A heuristic framework of spatial ability: a review and synthesis of spatial factor literature to support its translation into STEM education. *Educ. Psychol. Rev.* 30, 947–972. doi: 10.1007/s10648-018-9432-z

Cai, D., Viljaranta, J., and Georgiou, G. K. (2018). Direct and indirect effects of self-concept of ability on math skills. *Learn. Individ. Differ.* 61, 51–58. doi: 10.1016/j.lindif.2017.11.009

Campbell, S., and Collaer, M. (2009). Stereotype threat and gender differences in performance on a novel visuospatial task. *Psychol. Women Q.* 33, 437–444. doi: 10.1111/j.1471-6402.2009.01521.x

Ceci, S. J., Ginther, D. K., Kahn, S., and Williams, W. M. (2014). Women in academic science: a changing landscape. *Psychol. Sci. Public Interest* 15, 75–141. doi: 10.1177/1529100614541236

Chaplin, T. M., and Aldao, A. (2013). Gender differences in emotion expression in children: a meta-analytic review. *Psychol. Bull.* 139, 735–765. doi: 10.1037/a0030737

Chen, H., Gu, X., Zhou, Y., Ge, Z., Wang, B., Siok, W. T., et al. (2017). A genome-wide association study identifies genetic variants associated with mathematics ability. *Sci. Rep.* 7:40365. doi: 10.1038/srep40365

Cheryan, S., Plaut, V. C., Handron, C., and Hudson, L. (2013). The stereotypical computer scientist: gendered media representations as a barrier to inclusion for women. *Sex Roles* 69, 58–71. doi: 10.1007/s11199-013-0296-x

Chiaverini, J. (2017). *Enchantress of numbers: A novel of Ada Lovelace*. New York: Penguin.

Cimpian, J. R., Lubienski, S. T., Timmer, J. D., Makowski, M. B., and Miller, E. K. (2016). Have gender gaps in math closed? Achievement, teacher perceptions, and learning behaviors across two ECLS-K cohorts. *AERA Open* 2:233285841667361. doi: 10.1177/2332858416673617

Cvencek, D., Paz-Albo, J., Master, A., Herranz Llácer, C. V., Hervás-Escobar, A., and Meltzoff, A. N. (2020). Math is for me: a field intervention to strengthen math self-concepts in Spanish-speaking 3rd grade children. *Front. Psychol.* 11:593995. doi: 10.3389/fpsyg.2020.593995

Davadas, S. D., and Lay, Y. F. (2017). Factors affecting students' attitude toward mathematics: a structural equation modeling approach. *EURASIA J. Math. Sci Tech. Educ.* 14, 517–529. doi: 10.12973/ejmste/80356

Dowker, A., Sarkar, A., and Looi, C. Y. (2016). Mathematics anxiety: what have we learned in 60 years? *Front. Psychol.* 7:508. doi: 10.3389/fpsyg.2016.00508

Eccles, J. S. (2009). Who am I and what am I going to do with my life? Personal and collective identities as motivators of action. *Educ. Psychol.* 44, 78–89. doi: 10.1080/00461520902832368

Ehm, J.-H. (2014). *Akademisches Selbstkonzept im Grundschulalter. Entwicklungsanalyse dimensionaler Vergleiche und Exploration differenzieller Unterschiede*. pedocs. Available at: http://nbn-resolving.de/urn:nbn:de:0111-opus-95657

Else-Quest, N. M., Hyde, J. S., and Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychol. Bull.* 136, 103–127. doi: 10.1037/a0018053

Engle, R. W. (2002). Working memory capacity as executive attention. *Curr. Dir. Psychol. Sci.* 11, 19–23. doi: 10.1111/1467-8721.00160

Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. (2007). G*power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi: 10.3758/BF03193146

Fisher, P. H., Dobbs-Oates, J., Doctoroff, G. L., and Arnold, D. H. (2012). Early math interest and the development of math skills. *J. Educ. Psychol.* 104, 673–681. doi: 10.1037/a0027756

Fryer, R. G., and Levitt, S. D. (2010). An empirical analysis of the gender gap in mathematics. *Am. Econ. J. Appl. Econ.* 2, 210–240. doi: 10.1257/app.2.2.210

Garcia, N. L., Hall, L. V., and Pruden, S. M. (2021). Individual differences in young children's spatial ability: a systematic review. *PsyArXiv.* [Preprint]. doi: 10.31234/osf.io/5mc2y

Garriott, P. O., Hultgren, K. M., and Frazier, J. (2017). STEM stereotypes and high school students' math/science career goals. *J. Career Assess.* 25, 585–600. doi: 10.1177/1069072716665825

Goldman, A. D., and Penner, A. M. (2016). Exploring international gender differences in mathematics self-concept. *Int. J. Adolesc. Youth* 21, 403–418. doi: 10.1080/02673843.2013.847850

Gunderson, E. A., Ramirez, G., Levine, S. C., and Beilock, S. L. (2012). The role of parents and teachers in the development of gender-related math attitudes. *Sex Roles* 66, 153–166. doi: 10.1007/s11199-011-9996-2

Hawes, Z. C. K., Gilligan-Lee, K. A., and Mix, K. S. (2022). Effects of spatial training on mathematics performance: a meta-analysis. *Dev. Psychol.* 58, 112–137. doi: 10.1037/dev0001281

Hawes, Z., Moss, J., Caswell, B., and Poliszczuk, D. (2015). Effects of mental rotation training on children's spatial and mathematics performance: a randomized controlled study. *Trends Neurosci. Educ.* 4, 60–68. doi: 10.1016/j.tine.2015.05.001

Hayes, A. F. (2022). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. *3rd* Edn. New York, London: Taylor & Francis.

Hodgkiss, A., Gilligan-Lee, K., Thomas, M., Tolmie, A., and Farran, E. (2021). The developmental trajectories of spatial skills in middle childhood. *Br. J. Dev. Psychol.* 39, 566–583. doi: 10.1111/bjdp.12380

Jansen, P., Kellner, J., and Rieder, C. (2013). The improvement of mental rotation performance in second graders after creative dance training. *Creat. Educ.* 4, 418–422. doi: 10.4236/ce.2013.46060

Jansen, P., and Pietsch, S. (2022). Sports and mathematical abilities in primary school-aged children: how important are spatial abilities? An explorative study. *Curr. Psychol.* 41, 7132–7141. doi: 10.1007/s12144-020-01190-5

Jansen, P., Schmelter, A., Quaiser-Pohl, C., Neuburger, S., and Heil, M. (2013). Mental rotation performance in primary school age children: are there gender differences in chronometric tests? *Cogn. Dev.* 28, 51–62. doi: 10.1016/j.cogdev.2012.08.005

Janssen, A., and Geiser, C. (2010). On the relationship between solution strategies in two mental rotation tasks. *Learn. Individ. Differ.* 20, 473–478. doi: 10.1016/j.lindif.2010.03.002

Just, J., and Siller, H.-S. (2022). The role of mathematics in STEM secondary classrooms: a systematic literature review. *Educ. Sci.* 12:629. doi: 10.3390/educsci12090629

Kessels, U. (2005). Fitting into the stereotype: how gender-stereotyped perceptions of prototypic peers relate to liking for school subjects. *Eur. J. Psychol. Educ.* 20, 309–323. doi: 10.1007/BF03173559

Khooshabeh, P., Hegarty, M., and Shipley, T. F. (2013). Individual differences in mental rotation: piecemeal versus holistic processing. *Exp. Psychol.* 60, 164–171. doi: 10.1027/1618-3169/a000184

Larsen, D. M., Kristensen, M. L., Seidelin, L., and Svabo, C. (2022). Mathematics as the focal point of STEM teaching: Twelfth congress of the European Society for Research in mathematics education. *Twelfth Congress of the European Society for Research in Mathematics Education*. Available at: https://hal.science/hal-03745468

Lauer, J. E., Esposito, A. G., and Bauer, P. J. (2018). Domain-specific anxiety relates to children's math and spatial performance. *Dev. Psychol.* 54, 2126–2138. doi: 10.1037/dev0000605

Lennon-Maslin, M., Quaiser-Pohl, C. M., Ruthsatz, V., and Saunders, M. (2023). Under my skin: reducing Bias in STEM through new approaches to assessment of spatial abilities considering the role of emotional regulation. *Soc. Sci.* 12:356. doi: 10.3390/socsci12060356

Li, Y., and O'Boyle, M. (2013). How sex and college major relate to mental rotation accuracy and preferred strategy: an electroencephalographic (EEG) investigation. *Psychol. Rec.* 63, 27–42. doi: 10.11133/j.tpr.2013.63.1.003

Liesefeld, H., Fu, X., and Zimmer, H. (2014). Fast and careless or careful and slow? Apparent holistic processing in mental rotation is explained by speed-accuracy trade-offs. *J. Exp. Psychol. Learn. Mem. Cogn.* 41, 1140–1151. doi: 10.1037/xlm0000081

Linn, M. C., and Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: a meta-analysis. *Child Dev.* 56, 1479–1498. doi: 10.2307/1130467

Liu, X., Heqing, H., Yu, K., and Dou, D. (2020). Can video game training improve the two-dimensional mental rotation ability of young children? A randomized controlled trial. In HCI in Games: Second International Conference, HCI-Games 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings 22. Springer International Publishing (pp. 305–317).

Lopez-Pedersen, A., Mononen, R., Aunio, P., Scherer, R., and Melby-Lervåg, M. (2023). Improving numeracy skills in first graders with low performance in early numeracy: a randomized controlled trial. *Remedial Spec. Educ.* 44, 126–136. doi: 10.1177/07419325221102537

Lowrie, T., Resnick, I., Harris, D., and Logan, T. (2020). In search of the mechanisms that enable transfer from spatial reasoning to mathematics understanding. *Math. Educ. Res. J.* 32, 175–188. doi: 10.1007/s13394-020-00336-9

Maass, K., Geiger, V., Ariza, M. R., and Goos, M. (2019). The role of mathematics in interdisciplinary STEM education. *ZDM* 51, 869–884. doi: 10.1007/s11858-019-01100-5

MacKinnon, D. (2008). *Introduction to statistical mediation analysis*. New York: Routledge.

Manger, T., and Eikeland, O.-J. (1998). The effect of mathematics self-concept on girls' and boys' mathematical achievement. *Sch. Psychol. Int.* 19, 5–18. doi: 10.1177/0143034398191001

Marsh, H., Byrne, B. M., and Shavelson, R. (1988). A multifaceted academic self-concept: its hierarchical structure and its relation to academic achievement. *J. Educ. Psychol.* 80, 366–380. doi: 10.1037/0022-0663.80.3.366

Marsh, H. W., Pekrun, R., Murayama, K., Arens, A. K., Parker, P. D., Guo, J., et al. (2018). An integrated model of academic self-concept development: academic self-concept, grades, test scores, and tracking over 6 years. *Dev. Psychol.* 54, 263–280. doi: 10.1037/dev0000393

Marsh, H. W. (1990). Causal ordering of academic self-concept and academic achievement: a multiwave, longitudinal panel analysis. *J. Educ. Psychol.* 82, 646–656. doi: 10.1037/0022-0663.82.4.646

Marsh, H. W., and Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: beyond seductive pleasure and unidimensional perspectives. *Perspect. Psychol. Sci.* 1, 133–163. doi: 10.1111/j.1745-6916.2006.00010

McArthur, B. A., Madigan, S., and Korczak, D. J. (2021). Tweens are not teens: the problem of amalgamating broad age groups when making pandemic recommendations. *Can. J. Public Health* 112, 984–987. doi: 10.17269/s41997-021-00585-6

McGlone, M., and Aronson, J. (2006). Stereotype threat, identity salience, and spatial reasoning. *J. Appl. Dev. Psychol.* 27, 486–493. doi: 10.1016/j.appdev.2006.06.003

McGuire, L., Hoffman, A. J., Mulvey, K. L., Hartstone-Rose, A., Winterbottom, M., Joy, A., et al. (2022). Gender stereotypes and peer selection in STEM domains among children and adolescents. *Sex Roles* 87, 455–470. doi: 10.1007/s11199-022-01327-9

Memar, M., and Mokaribolhassan, A. (2021). Stress level classification using statistical analysis of skin conductance signal while driving. *SN Appl. Sci.* 3:64. doi: 10.1007/s42452-020-04134-7

Milani, L., Grumi, S., and Di Blasio, P. (2019). Positive effects of videogame use on visuospatial competencies: the impact of visualization style in preadolescents and adolescents. *Front. Psychol.* 10:1226. doi: 10.3389/fpsyg.2019.01226

Mix, K. (2019). Why are spatial skill and mathematics related? *Child Dev. Perspect.* 13, 121–126. doi: 10.1111/cdep.12323

Modroño, C., Navarrete, G., Nicolle, A., González-Mora, J. L., Smith, K. W., Marling, M., et al. (2018). Developmental grey matter changes in superior parietal cortex accompany improved transitive reasoning. *Think. Reason.* 25, 151–170. doi: 10.1080/13546783.2018.1481144

Moè, A. (2013). Cognitive styles and mental rotation ability in map learning. *Cogn. Process.* Available at: https://www.academia.edu/15149262/Cognitive_styles_and_mental_rotation_ability_in_map_learning

Moè, A. (2016). Teaching motivation and strategies to improve mental rotation abilities. *Intelligence* 59, 16–23. doi: 10.1016/j.intell.2016.10.004

Moe, A. (2018). Mental rotation and mathematics: gender-stereotyped beliefs and relationships in primary school children. *Learn. Individ. Differ.* 61, 172–180. doi: 10.1016/j.lindif.2017.12.002

Moe, A., Jansen, P., and Pietsch, S. (2018). Childhood preference for spatial toys. Gender differences and relationships with mental rotation in STEM and non-STEM students. *Learn. Individ. Differ.* 68, 108–115. doi: 10.1016/j.lindif.2018.10.003

Morawietz, C., and Muehlbauer, T. (2021). Effects of physical exercise interventions on spatial orientation in children and adolescents: a systematic scoping review. *Front. Sports Act. Living* 3:664640. doi: 10.3389/fspor.2021.664640

Neuburger, S., Jansen, P., Heil, M., and Quaiser-Pohl, C. (2011). Gender differences in pre-adolescents' mental-rotation performance: do they depend on grade and stimulus type? *Personal. Individ. Differ.* 50, 1238–1242. doi: 10.1016/j.paid.2011.02.017

Neuburger Ruthsatz, V., Jansen, P., and Quaiser-Pohl, C. (2012). Influence of rotational Axis and gender-stereotypical nature of rotation stimuli on the mental-rotation performance of male and female fifth graders. In *Spatial cognition: Vol. LNAI 7463* (p. 229).

Neuburger, S., Ruthsatz, V., Jansen, P., and Quaiser-Pohl, C. (2015). Can girls think spatially? Influence of implicit gender stereotype activation and rotational axis on fourth graders' mental-rotation performance. *Learn. Individ. Differ.* 37, 169–175. doi: 10.1016/j.lindif.2014.09.003

Nolte, N., Schmitz, F., Fleischer, J., Bungart, M., and Leutner, D. (2022). Rotational complexity in mental rotation tests: cognitive processes in tasks requiring mental rotation around cardinal and skewed rotation axes. *Intelligence* 91:101626. doi: 10.1016/j.intell.2022.101626

Nuutila, K., Tapola, A., Tuominen, H., Molnár, G., and Niemivirta, M. (2021). Mutual relationships between the levels of and changes in interest, self-efficacy, and perceived difficulty during task engagement. *Learn. Individ. Differ.* 92:102090. doi: 10.1016/j.lindif.2021.102090

OECD (2013). *Mathematics self-beliefs and participation in mathematics-related activities*. Paris: OECD, 87–112.

OECD (2020). *Education policy outlook in Germany*. Paris: OECD.

Parker, P. D., Marsh, H. W., Ciarrochi, J., Marshall, S., and Abduljabbar, A. S. (2014). Juxtaposing math self-efficacy and self-concept as predictors of long-term achievement outcomes. *Educ. Psychol.* 34, 29–48. doi: 10.1080/01443410.2013.797339

Peters, M. (2005). Sex differences and the factor of time in solving Vandenberg and Kuse mental rotation problems. *Brain Cogn.* 57, 176–184. doi: 10.1016/j.bandc.2004.08.052

Peterson, E., Weinberger, A., Uttal, D., Kolvoord, B., and Green, A. (2020). Spatial activity participation in childhood and adolescence: consistency and relations to spatial thinking in adolescence. *Cogn. Res. Princ. Implic.* 5:43. doi: 10.1186/s41235-020-00239-0

Petrill, S. A., Kovas, Y., Hart, S. A., Thompson, L. A., and Plomin, R. (2009). The genetic and environmental etiology of high math performance in 10-year-old twins. *Behav. Genet.* 39, 371–379. doi: 10.1007/s10519-009-9258-z

Pietsch, S., Böttcher, C., and Jansen, P. (2017). Cognitive motor coordination training improves mental rotation performance in primary school-aged children. *Mind Brain Educ.* 11, 176–180. doi: 10.1111/mbe.12154

Quaiser-Pohl, C. (2003). The mental cutting test 'Schnitte' and the picture rotation test-two new measures to assess spatial ability. *Int. J. Test.* 3, 219–231. doi: 10.1207/S15327574IJT0303_2

Quaiser-Pohl, C. (2012). "Women's choices in STEM – statistical data and theoretical approaches explaining the gender gap" in *Women's choices in Europe. Influence of gender and education, occupational career and family development*. eds. C. Quaiser-Pohl and M. Endepohls-Ulpe (Münster, New York: Waxmann), 53–61.

Quaiser-Pohl, C., Endenpohls-Ulpe, M., Rasic, R., Gnosa, T., and Sander, E. (2012). Mentoring beim Übergang in die berufliche Ausbildung am Beispiel des Ada-Lovelace-Projekts. *Diskurs Kindheits- und Jugendforschung / Discourse J. Childhood Adolesc. Res.* 7:Article 2, Available at: https://www.budrich-journals.de/index.php/diskurs/article/view/7113

Quaiser-Pohl, C., and Endepohls-Ulpe, M. (Eds.) (2012). *Women's choices in Europe. Influence of gender on education, occupational career and family development*. Münster, New York: Waxmann.

Quaiser-Pohl, C., Endepohls-Ulpe, M., and Meyer, C. (2014a). *How does mentoring with female pupils work? Results of multivariate analyses of the effects of a German mentoring program in the STEM field.* 2nd gender and STEM conference, Berlin, Germany.

Quaiser-Pohl, C., Geiser, C., and Lehmann, W. (2006). The relationship between computer-game preference, gender, and mental-rotation ability. *Personal. Individ. Differ.* 40, 609–619. doi: 10.1016/j.paid.2005.07.015

Quaiser-Pohl, C., Neuburger, S., Heil, M., Jansen, P., and Schmelter, A. (2014b). Is the male advantage in mental-rotation performance task independent? On the usability of chronometric tests and paper-and-pencil tests in children. *Int. J. Test.* 14, 122–142. doi: 10.1080/15305058.2013.860148

Rahe, M., and Quaiser-Pohl, C. (2019). Mental-rotation performance in middle and high-school age: influence of stimulus material, gender stereotype beliefs, and perceived ability of gendered activities. *J. Cogn. Psychol.* 31, 594–604. doi: 10.1080/20445911.2019.1649265

Ramirez, G., Gunderson, E. A., Levine, S. C., and Beilock, S. L. (2012). Spatial anxiety relates to spatial abilities as a function of working memory in children. *Q. J. Exp. Psychol.* 65, 474–487. doi: 10.1080/17470218.2011.616214

Rocha, K., Lussier, C. M., and Atit, K. (2022). What makes online teaching spatial? Examining the connections between K-12 teachers' spatial skills, affect, and their use of spatial pedagogy during remote instruction. *Cogn. Res. Princ. Implic.* 7:25. doi: 10.1186/s41235-022-00377-7

Rossnan, S. (2006). Overcoming math anxiety. *Mathitudes* 1:4.

Ruthsatz, V., Neuburger, S., Jansen, P., and Quaiser-Pohl, C. (2014). "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, Vol. 8684*. eds. C. Freksa, B. Nebel, M. Hegarty and T. Barkowsky (Heidelberg, New York. Dordrecht, London: Springer International Publishing), 370–382.

Ruthsatz, V., Neuburger, S., Jansen, P., and Quaiser-Pohl, C. (2015). Cars or dolls? Influence of the stereotyped nature of the items on children's mental-rotation performance. *Learn. Individ. Differ.* 43, 75–82. doi: 10.1016/j.lindif.2015.08.016

Ruthsatz, V. S. N., and Quaiser-Pohl, C. (2013). The social relevance and the socio-cultural origins of gender differences in spatial abilities. *Folia Sociol.* 43, 17–32.

Ruthsatz, V., Quaiser-Pohl, C. M., Saunders, M., and Lennon-Maslin, M. (forthcoming). *Male? Female? Neutral! Using novel polyhedral figures as gender-neutral stimuli in a mental rotation test.*

Ruthsatz, V., Rahe, M., Schürmann, L., and Quaiser-Pohl, C. (2019). Girls' stuff, boys' stuff and mental rotation: fourth graders rotate faster with gender-congruent stimuli. *J. Cogn. Psychol.* 31, 225–239. doi: 10.1080/20445911.2019.1567518

Sanchis-Sanchis, A., Grau, M. D., Moliner, A.-R., and Morales-Murillo, C. P. (2020). Effects of age and gender in emotion regulation of children and adolescents. *Front. Psychol.* 11:946. doi: 10.3389/fpsyg.2020.00946

Saunders, M., and Quaiser-Pohl, C. M. (2020). Identifying solution strategies in a mental-rotation test with gender-stereotyped objects by analyzing gaze patterns. *J. Eye Mov. Res.* 13:5. doi: 10.16910/jemr.13.6.5

Sax, L. J., Kanny, M. A., Riggers-Piehl, T. A., Whang, H., and Paulson, L. N. (2015). "But I'm not good at math": the changing salience of mathematical self-concept in shaping women's and men's STEM aspirations. *Res. High. Educ.* 56, 813–842. doi: 10.1007/s11162-015-9375-x

Schmader, T. (2010). Stereotype threat deconstructed. *Curr. Directions Psychol. Sci.* 19, 14–18. doi: 10.1177/0963721409359292

Schmader, T., Johns, M., and Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychol. Rev.* 115, 336–356. doi: 10.1037/0033-295X.115.2.336

Sella, F., and Cohen Kadosh, R. (2018). What expertise can tell about mathematical learning and cognition. *Mind Brain Educ.* 12, 186–192. doi: 10.1111/mbe.12179

Sewasew, D., Schroeders, U., Schiefer, I. M., Weirich, S., and Artelt, C. (2018). Development of sex differences in math achievement, self-concept, and interest from grade 5 to 7. *Contemp. Educ. Psychol.* 54, 55–65. doi: 10.1016/j.cedpsych.2018.05.003

Skeide, M. A., Wehrmann, K., Emami, Z., Kirsten, H., Hartmann, A. M., Rujescu, D., et al. (2020). Neurobiological origins of individual differences in mathematical ability. *PLoS Biol.* 18:e3000871. doi: 10.1371/journal.pbio.3000871

Space Science Institute. (2023). Tweens and STEM. *STAR Library Network*. Available at: https://www.starnetlibraries.org/resources/tweens-and-stem/

Starr, C. R. (2018). "I'm not a science nerd!": STEM stereotypes, identity, and motivation among undergraduate women. *Psychol. Women Q.* 42, 489–503. doi: 10.1177/0361684318793848

Starr, C. R., and Leaper, C. (2019). Do adolescents' self-concepts moderate the relationship between STEM stereotypes and motivation? *Soc. Psychol. Educ.* 22, 1109–1129. doi: 10.1007/s11218-019-09515-4

Tabachnick, B. G., and Fidell, L. S. (2013). *Using multivariate statistics: Pearson new international. Edition (6th Edition)* Edn. Harlow; Pearson: Microsoft Press.

Taragin, D., Tzuriel, D., and Vakil, E. (2019). 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.* 12:10.16910/jemr.12.8.2. doi: 10.16910/jemr.12.8.2

Uttal, D., Meadow, N., Tipton, E., Hand, L., Alden, A., Warren, C., et al (2013). The malleability of spatial skills: a meta-analysis of training studies. *Psychol. Bull.* 139, 352–402. doi: 10.1037/a0028446

Vandenberg, S. G., and Kuse, A. R. (1978). Mental rotations a group test of three-dimensional spatial visualization. *Percept. Mot. Skills.* 47, 599–604. doi: 10.2466/pms.1978.47.2.599

van Steensel, R., Oostdam, R., and van Gelderen, A. (2019). Affirming and undermining motivations for reading and associations with reading comprehension, age and gender. *J. Res. Read.* 42, 504–522. doi: 10.1111/1467-9817.12281

Voyer, D. (2011). Time limits and gender differences on paper-and-pencil tests of mental rotation. *Psychon. Bull. Rev.* 18, 267–277. doi: 10.3758/s13423-010-0042-0

Voyer, D., Voyer, S., and Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. *Psychol. Bull.* 117, 250–270. doi: 10.1037/0033-2909.117.2.250

Woolley, B. (2000). *The bride of science: romance, reason, and Byron's daughter (first edition)*. New York: McGraw-Hill.

Wraga, M., Duncan, L., Jacobs, E. C., Helt, M., and Church, J. (2006). Stereotype susceptibility narrows the gender gap in imagined self-rotation performance. *Psychon. Bull. Rev.* 13, 813–819. doi: 10.3758/BF03194002

behavioral sciences

MDPI

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

check for updates

Citation: Lennon-Maslin, M.; Quaiser-Pohl, C.M. "It's Different for Girls!" The Role of Anxiety, Physiological Arousal, and Subject Preferences in Primary School Children's Math and Mental Rotation Performance. *Behav. Sci.* **2024**, *14*, 809. https://doi.org/10.3390/bs14090809

Academic Editor: Jennifer LaCosse

Received: 29 July 2024
Revised: 6 September 2024
Accepted: 11 September 2024
Published: 12 September 2024

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**	
Class	2nd Grade	16	21	37
	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 α = 0.86.

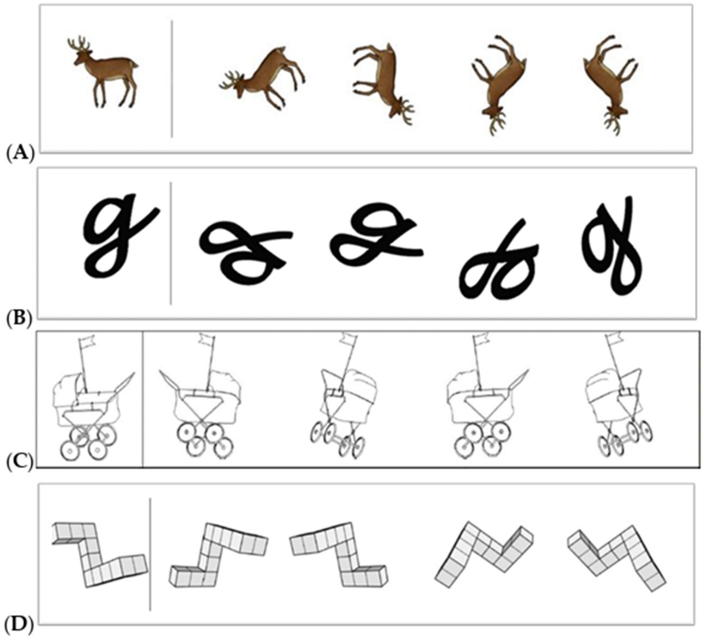


Figure 1. Examples of MRT items used with concrete objects: (**A**) animal stimulus and (**B**) letter stimulus [85]; (**C**) female-stereotyped stimulus "Pram" [21,23]; and with an abstract object (**D**) "Cube" figure [22].

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 α = 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 (α = 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
Sex	Math_PT	1643.461	1	1643.461	4.630	0.033	0.035
	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.

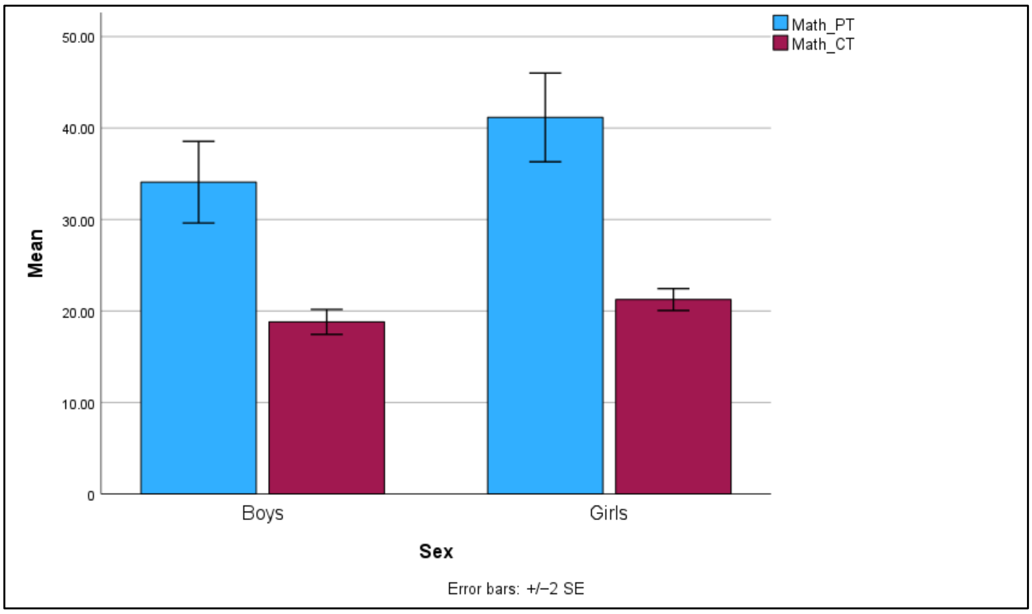


Figure 2. Gender (sex) differences in percentage scored (blue bar = Math_PT) and completion times (red bar = Math_CT) on the math task. Error bars represent the standard error of the mean (SEM). The error bars indicate similar variability in scores and completion times between boys and girls.

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.

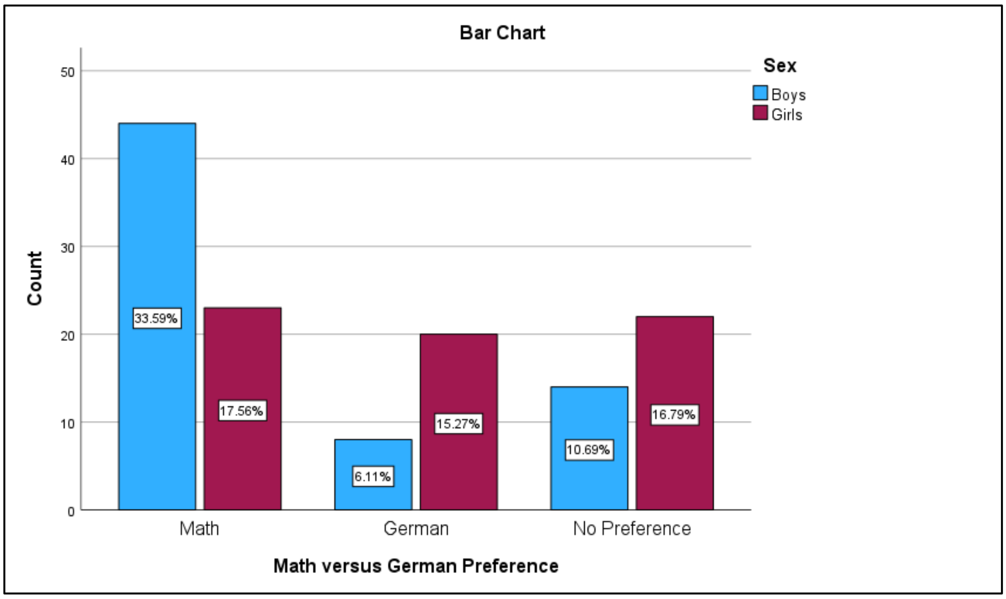


Figure 3. Percentage of math vs. German as the preferred school subject by gender (sex).

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'* Λ = 0.10, $F(2, 59) = 0.09$, $p = 0.915$, $\eta^2 = 0.00$), neither was math anxiety (*Wilks'* Λ = 0.95, $F(2, 59) = 1.55$, $p = 0.220$, $\eta^2 = 0.05$), nor spatial anxiety (*Wilks'* Λ = 0.98, $F(2, 59) = 0.72$, $p = 0.489$, $\eta^2 = 0.02$), nor subject preference (*Wilks'* Λ = 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'* Λ = 0.84, $F(2, 58) = 5.53$, $p = 0.006$, $\eta^2 = 0.16$) on girls' performance on the math task. Math anxiety (*Wilks'* Λ = 0.83, $F(2, 58) = 5.93$, $p = 0.005$, $\eta^2 = 0.17$) and subject preference (*Wilks'* Λ = 80, $F(4, 116) = 3.45$, $p = 0.011$, $\eta^2 = 0.11$) also had significant effects, but spatial anxiety did not (*Wilks'* Λ = 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
Girls	ZGSR_Math	Wilks' Lambda	0.840	5.532	2.000	58.000	0.006	0.160
	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
Girls	ZGSR_Math	Math_CT	158.365	1	158.365	8.480	0.005	0.126
	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).

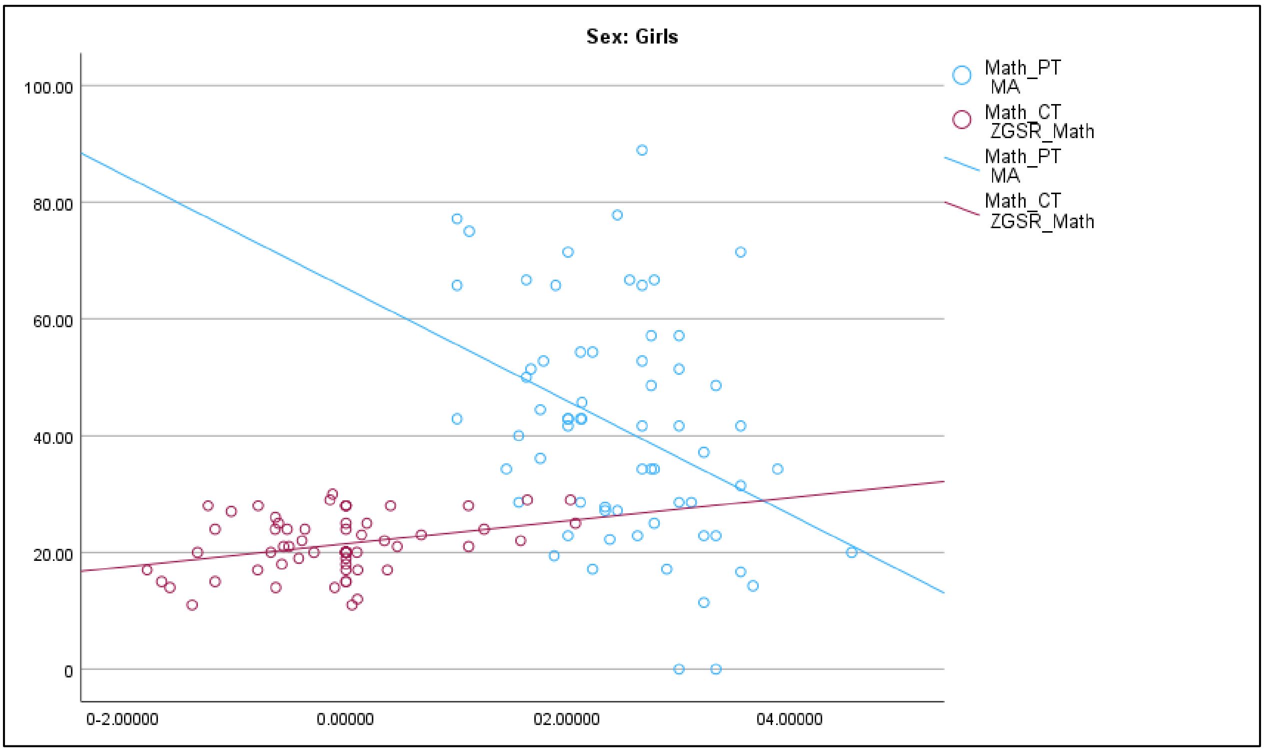


Figure 4. Effects of math anxiety on percentage scores and emotional reactivity (GSR) on completion times for girls. X-Axis = standardized scores; Y-Axis = percentage scores and completion times; blue circles: percentage scores (Math_PT) vs. math anxiety (MA); blue line: trend line for percentage scores (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'* $\Lambda = 0.96$, $F(2, 121) = 2.70$, $p = 0.071$, $\eta^2 = 0.04$. Neither math anxiety, *Wilks'* $\Lambda = 0.96$, $F(2, 121) = 2.64$, $p = 0.075$, $\eta^2 = 0.04$, nor spatial anxiety, *Wilks'* $\Lambda = 0.10$, $F(2, 121) = 0.23$, $p = 0.798$, $\eta^2 = 0.004$, had a significant effect. A significant effect of gender, *Wilks'* $\Lambda = 0.94$, $F(2, 121) = 3.73$, $p = 0.027$, $\eta^2 = 0.06$, but no significant effect of subject preference, *Wilks'* $\Lambda = 0.96$, $F(4, 242) = 1.11$, $p = 0.350$, $\eta^2 = 0.02$, was found. There was no significant interaction effect for gender and subject preference, *Wilks'* $\Lambda = 0.98$, $F(4, 242) = 0.71$, $p = 0.582$, $\eta^2 = 0.01$ (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, $F(1, 122) = 0.27$, $p = 0.603$, $\eta^2 = 0.00$, but the effect of math anxiety was significant, $F(1, 122) = 5.26$, $p = 0.023$, $\eta^2 = 0.04$. The effects of spatial anxiety, $F(1, 122) = 0.310$, $p = 0.579$, $\eta^2 = 0.003$, gender, $F(1, 122) = 0.02$, $p = 0.885$, $\eta^2 = 0.00$, and subject preference, $F(2, 122) = 0.138$, $p = 0.871$, $\eta^2 = 0.00$, 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, $F(1, 122) = 5.43$, $p = 0.021$, $\eta^2 = 0.04$. The effects of math anxiety, $F(1, 122) = 0.80$, $p = 0.373$, $\eta^2 = 0.01$, and spatial anxiety, $F(1, 122) = 0.274$, $p = 0.602$, $\eta^2 = 0.00$, were not significant. The effect of gender was significant, $F(1, 122) = 7.12$, $p = 0.009$, $\eta^2 = 0.06$, but subject preference was not, $F(2, 122) = 2.06$, $p = 0.131$, $\eta^2 = 0.03$ (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, $F(2, 59) = 1.57$, $p = 0.217$, $\eta^2 = 0.050$). Neither math anxiety (*Wilks'* Λ = 0.94, $F(2, 59) = 1.74$, $p = 0.185$, $\eta^2 = 0.06$), spatial anxiety (*Wilks'* Λ = 0.99, $F(2, 59) = 0.36$, $p = 0.699$, $\eta^2 = 0.01$), nor subject preference (*Wilks'* Λ = 0.912, $F(4, 118) = 1.40$, $p = 0.239$, $\eta^2 = 0.04$) show significant effects.

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

For boys' response time, neither emotional reactivity measured by GSR ($F(1, 60) = 3.17$, $p = 0.080$, $\eta^2 = 0.05$), nor math anxiety ($F(1, 60) = 0.08$, $p = 0.783$, $\eta^2 = 0.00$), nor spatial anxiety, ($F(1, 60) = 0.60$, $p = 0.442$, $\eta^2 = 0.01$) nor subject preference ($F(2, 60) = 1.91$, $p = 0.156$, $\eta^2 = 0.06$) 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, $F(2, 58) = 1.70$, $p = 0.191$, $\eta^2 = 0.05$). Neither math anxiety (*Wilks'* Λ = 0.97, $F(2, 58) = 0.974$, $p = 0.384$, $\eta^2 = 0.03$), nor spatial anxiety (*Wilks'* Λ = 0.98, $F(2, 58) = 0.72$, $p = 0.490$, $\eta^2 = 0.02$), nor subject preference (*Wilks'* Λ = 0.98, $F(4, 116) = 0.28$, $p = 0.893$, $\eta^2 = 0.01$) showed significant effects.

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

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

Author Contributions: Conceptualization, M.L.-M. and C.M.Q.-P.; methodology, M.L.-M.; software, validation, M.L.-M. and C.M.Q.-P.; formal analysis, M.L.-M.; investigation, M.L.-M.; resources, data curation, M.L.-M.; writing—original draft preparation, M.L.-M.; writing—review and editing, M.L.-M. and C.M.Q.-P.; visualization, M.L.-M.; supervision, C.M.Q.-P.; project administration, M.L.-M.; funding acquisition, C.M.Q.-P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Marie Sklodowska-Curie Actions International Training Network (MSCA ITN) "SellSTEM", grant number 956124.

Institutional Review Board Statement: This study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board (or Ethics Committee) of the University of Koblenz, Germany (Ethik-kommission/Antrag Lennon-Maslin/Beschluss vom 9 November 2022). Furthermore, according to section 67, paragraph 6 of the Education Act of the State of Rhineland-Palatinate, Germany, approval was sought and provided by the Aufsichts- und Dienstleistungsdirektion Rheinland Pfalz (ADD RLP), Willy-Brandt-Platz 3, 54290 Trier, Germany, the authorities who oversee schools in this state.

Informed Consent Statement: Written informed consent has been obtained from parents and guardians of all child participants to publish this paper.

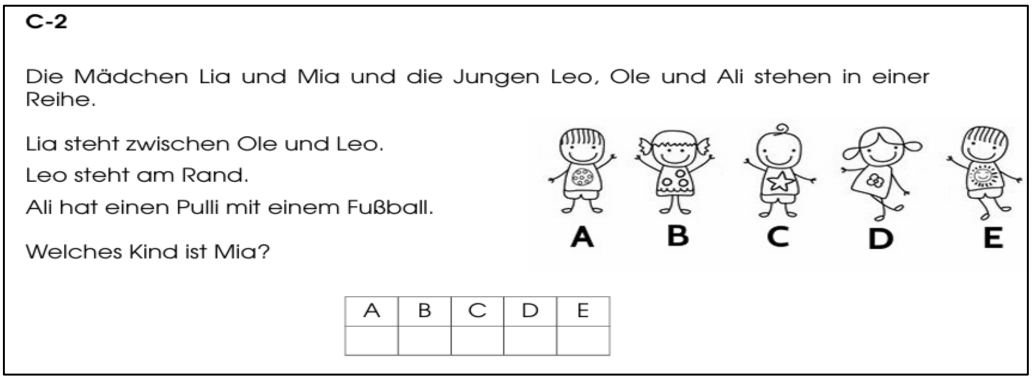
Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to ethical and privacy reasons stated in the parental consent information form.

Acknowledgments: We would like to express our sincere thanks to the principals, teachers, parents, and the students of the five schools who took part in our study. Without your participation, this research would not have been possible. Also, thanks to Vera Ruthsatz for permission to use the gendered and neutral stimuli in the mental rotation task, and to Mirko Saunders for programming the experiment and providing technical support. Also, thanks to our student research assistants for their support throughout the project. We used PsychoPy®, a free, open-source package allowing researchers to run a wide range of experiments in the behavioral sciences. It can be downloaded at www.psychopy.org/download.html (accessed on 6 April 2023).

Conflicts of Interest: The authors declare no conflicts of interest.

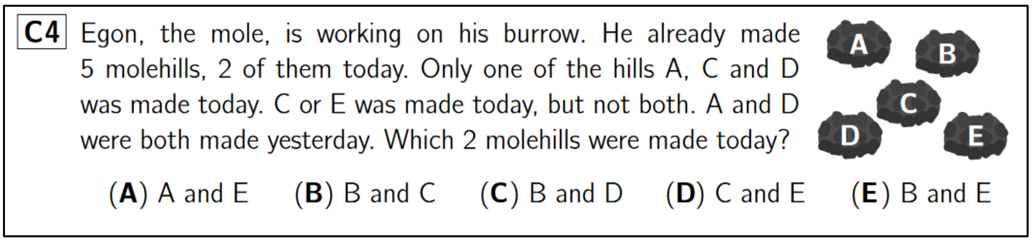
Appendix A. Sample Problems from Math Task

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

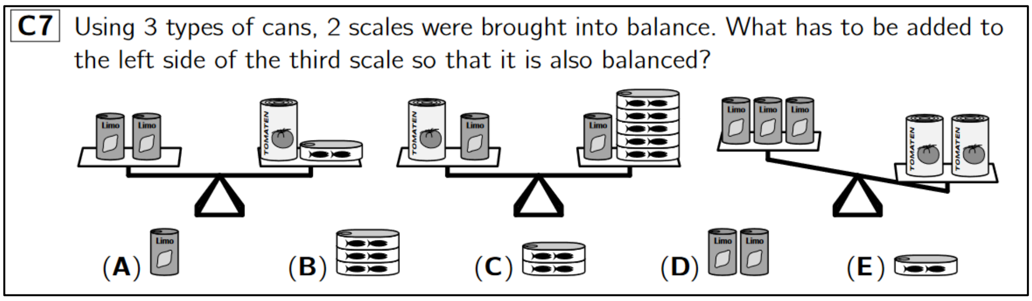


Translation from German: "The girls Lia and Mia and the boys Leo, Ole and Ali are standing in a row. Lia is standing between Ole and Leo. Leo is standing on the edge. Ali has a sweater with a soccer ball on it. Which child is Mia?"

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



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



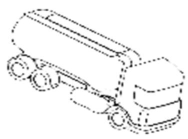
All problems are available in several languages from the German Kangaroo and Mini-Kangaroo competition website, but were presented to participants in German.

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

MRT 1

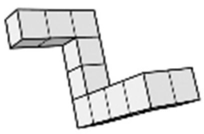
Letter_g

Lorry_2D



Goggles_2D

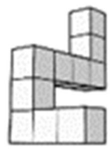
Cubes_2D_#1



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.