



OPEN ACCESS

EDITED BY

Karl Schweizer,
Goethe University Frankfurt, Germany

REVIEWED BY

Romina Antonela Vivaldi,
Staffordshire University, United Kingdom
Nelcida Garcia-Sanchez,
Borough of Manhattan Community College,
United States

*CORRESPONDENCE

W. Miro Ebert
✉ miro.ebert@ur.de

RECEIVED 20 June 2025

ACCEPTED 03 September 2025

PUBLISHED 16 September 2025

CITATION

Ebert WM, Jost L and Jansen P (2025) No
relationship between gender stereotypes and
mental rotation in preschool girls.
Front. Psychol. 16:1650979.
doi: 10.3389/fpsyg.2025.1650979

COPYRIGHT

© 2025 Ebert, Jost and Jansen. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

No relationship between gender stereotypes and mental rotation in preschool girls

W. Miro Ebert*, Leonardo Jost and Petra Jansen

Faculty of Human Sciences, University of Regensburg, Regensburg, Germany

Gender stereotypes about spatial ability have been proposed as a contributing factor to the gender gap in STEM. This goal of this study was to investigate whether implicit gender stereotypes regarding spatial ability are associated with mental rotation (MR) performance in preschool-aged girls, and whether visuospatial working memory (VSWM) plays a moderating role. Fifty-two girls aged 5 to 6.75 years completed a chronometric MR task, a computerized Corsi block-tapping task (assessing VSWM), and a single-target Implicit Association Test measuring associations between gender and toys used in spatial play. Participants did not show significant implicit stereotypes favoring either gender. Contrary to our hypotheses, no evidence for a relationship between implicit stereotypes and MR response times was found. Unexpectedly, stronger implicit associations linking boys with spatial ability were associated with higher MR accuracy. VSWM was positively correlated with both MR accuracy and implicit stereotype scores. Since the hypothesized relationship between implicit stereotypes and MR was not observed, a potential moderation through VSWM was not examined. Our findings diverge from prior research suggesting that stereotype-consistent associations in girls are linked to lower spatial task performance. Taken together, the results cast doubt on the robustness of implicit gender stereotypes about spatial ability in early childhood and highlight the complexity of their potential impact on spatial cognition. Further research is needed to clarify under what conditions, if any, such stereotypes affect performance in young children.

KEYWORDS

spatial ability, gender stereotypes, human sex differences, preschool, children, mental rotation, kindergarten

1 Introduction

Across the lifespan, spatial ability and the skills that comprise it—such as spatial visualization, spatial orientation, perspective taking and mental rotation (MR)—are useful in a wide range of contexts. In childhood, spatial skills are required to engage in activities such as games, puzzles, and sports. As adults, we rely on spatial skills to, for example, navigate our environment, or give directions (Newcombe, 2020). Aside from their relevance to daily life, spatial skills are crucial in specific professional and academic fields. Especially in science, technology, mathematics, and engineering (STEM; Newcombe, 2017), spatial ability has long been considered critical (see, e.g., Super and Bachrach, 1957). Research shows that spatial ability can predict the choice to pursue a career or education in STEM (Wai et al., 2009) and success in these endeavors (Newcombe and Shipley, 2015).

For over two decades, the underrepresentation of women in STEM fields has been a topic of political and academic interest (see, e.g., Hanson, 1996; Wang and Degol, 2013). Although there are around 3 million more women in STEM now than there were in 2011 in the United States, their representation remains low at around 34% of the country's total STEM

workforce (National Science Board, 2024). In Europe, women comprised 41.3% of the employed scientists and engineers in 2019 (European Commission, Directorate-General for Research and Innovation, 2021). The authors suggest that “given the strategic importance of technology (tech) industry to the EU economy, these data indicate that greater effort is needed to increase women’s participation in this field” (European Commission, Directorate-General for Research and Innovation, 2021, p.60). This claim is further underscored by the underrepresentation of women among doctoral graduates in most STEM fields (European Commission, Directorate-General for Research and Innovation, 2021).

Researchers have proposed a variety of explanations for the disparity in STEM (see Christie et al., 2017; Ertl et al., 2017). Gender stereotypes, for instance, have garnered significant attention. Nosek et al. (2009) outline that STEM subjects and professions are often perceived as male domains, which may contribute to sex differences in STEM participation and achievement. Matching this rationale, the authors demonstrated, based on data from 34 countries, that national implicit gender-science stereotyping related to sex differences in eighth-grade math at a national level (Nosek et al., 2009). Moreover, cognitive factors—particularly spatial ability—are often introduced as a possible driver for the gender disparity in STEM. As Levine et al. (2016) discuss, sex differences have been reported on various spatial tasks and are especially pronounced in MR (e.g., Voyer et al., 1995), the cognitive process of imagining an object turning in space to determine its orientation or match it to another object. Mixed findings from infant and child studies (see Levine et al., 2016) leave open questions about when sex differences in spatial ability emerge, how they develop, and what underlying factors contribute to them. As with sex differences in STEM, gender stereotypes have been proposed as a mechanism underlying the sex differences in MR. Research suggests that spatial ability, which is deemed crucial in many STEM contexts, is commonly viewed as a male quality by children (e.g., Vander Heyden et al., 2016; Ebert et al., 2024a) and adults (e.g., van der Ham and Koutzmpi, 2023). Furthermore, numerous studies demonstrate that stereotype related effects can play a role in spatial task performance. Especially concerning MR, research is abundant demonstrating stereotype threat and lift effects in adults and adolescents (e.g., McGlone and Aronson, 2006; Moè and Pazzaglia, 2006). Other studies show an influence of the stereotyped nature of MR stimuli on task performance of adults (Rahe et al., 2021; Rahe and Jansen, 2022). Apart from these effects, a study in university students suggests that unconscious stereotypic beliefs, commonly referred to as implicit stereotypes, correlate with task performance (Sanchis-Segura et al., 2018). Moreover, stereotype endorsement appears to be critical to stereotype threat effects, according to a study in high-school students (Moè and Pazzaglia, 2006). Considering such findings, the relevance of stereotypes in the development of spatial ability, and by consequence, the development of sex differences in STEM, constitutes an interesting research field. A recent study suggests that five- and six-year-old children hold stereotypic beliefs about spatial ability on an implicit and explicit level (Ebert et al., 2024a). No clear relationship between stereotyping and MR performance was found in this study. However, in girls, but not in boys, implicit stereotypes were related to MR accuracy. Girls who implicitly associated boys with spatial ability more strongly solved the task less accurately (Ebert et al., 2024a).

Another factor that has been implicated in spatial ability in general and MR specifically is working memory—the cognitive system

responsible for the temporary maintenance and manipulation of information. MR is believed to involve multiple stages of processing (Shepard and Cooper, 1982; Just and Carpenter, 1985; Heil and Rolke, 2002), some of which entail keeping an object representation in mind and manipulating it (i.e., identification, discrimination, and mental rotation phases). Given these demands, it comes as no surprise that working memory is thought to be centrally involved in MR. Especially the visuospatial sketchpad, which is concerned with information about object features and location (Baddeley and Hitch, 1974), has been the subject of several studies on MR. In fact, studies in children (Lehmann et al., 2014) and adults (e.g., Kaufman, 2007) implicate visuospatial working memory (VSWM) and its subcomponents in MR (Hyun and Luck, 2007; Ebert et al., 2024b).

Crucially, evidence from studies with adults and adolescents also suggests links between stereotype effects and working memory. A prominently proposed mechanism underlying stereotype threat effects concerns disruptions in working memory capacity (Schmader, 2010). The fear of being evaluated negatively and/or confirming a stereotype about one’s group is assumed to occupy part of an individual’s working memory capacity, thereby affecting performance in tasks that rely on this resource. A study by Schmader and Johns (2003) provided evidence that the effect of stereotype threat on women’s math performance was completely mediated by working memory capacity. A study in first-year psychology students by Bonnot and Croizet (2007) suggests that stereotype internalization can affect performance through the same mechanism (i.e., interference in working memory). Adding to these findings, Régner et al. (2010) found working memory capacity to moderate stereotype threat effects in a sample of university students. Only individuals with low working memory capacity were susceptible to stereotype threat induction in their study. Similar results come from a study in adolescents from Hong Kong (Chan and Rosenthal, 2014). These studies used complex span tasks to assess general working memory capacity and mostly focused on mathematical performance as the outcome variable. To the best of our knowledge, this topic has not been researched in a younger population. Regarding studies that include VSWM specifically, a study in adult athletes found moderating effects of VSWM capacity on stereotype threat effects on a simple motor task (Laurin et al., 2022). Thus, while working memory capacity is usually assessed more generally in examining its relation to stereotype effects, the visuo-spatial subsystem has been implicated in such processes empirically. Given the role of VSWM in MR, we believe investigating whether VSWM capacity relates to stereotype effects in MR is worthwhile.

As outlined above, MR shares relationships with stereotypes and working memory. Moreover, evidence suggests that specific stereotype effects may be related with interference in working memory. Therefore, we examine whether implicit stereotyping affects MR in kindergarten girls. If our findings support this notion, we will further investigate whether such an effect would be moderated by working memory capacity. Since, theoretically, VSWM would be expected to affect performance through its relevance to the later perceptual stages (identification and discrimination) and the rotation process, we want to examine both general effects on MR performance and effects emerging in the rotation phase. The corresponding hypotheses are the following.

H1: Stronger implicit beliefs linking spatial ability to males, will predict lower MR performance in girls.

H1.1: Stronger stereotypic beliefs will predict a less efficient rotation process, indicated by steeper increases in response time and/or decreases in accuracy with increasing rotation angle.

H2: The relationship between implicit stereotypes and MR in girls will be moderated by working memory capacity.

H2.1: The relationship between implicit stereotypes and efficiency in the rotation process in girls will be moderated by working memory capacity.

These expectations generally build on the assumption that negative stereotypes about one's group impair performance through disruptions in working memory. Consequently, the same effects could be expected in boys holding negative stereotypes about the spatial ability of their own gender. However, few boys tend to demonstrate such stereotypes, which is why we focus on girls in this study.

2 Method

2.1 Participants

A power analysis conducted with G*Power (Faul et al., 2007) indicated that 53 female participants were needed to achieve a power of 0.80 for detecting small-to-medium effects ($f^2 = 0.12$) at an alpha level of 0.05. Data were gathered as part of an intervention study carried out in the city and district of Regensburg, Germany, with all participants recruited via local kindergartens. Among the children who participated, 54 were female, and their pre-test data were used to examine the stated hypotheses. Two participants had to be excluded, due to missing data. Hence, our final sample consisted of 52 kindergarten girls (mean age: 5.88 years, SD = 0.45, range: 5 to 6.75 years). To acknowledge their participation, children received small gifts, while kindergartens received financial compensation. This study was preregistered at <https://osf.io/32nc9>, and ethical approval was granted by the ethics board of the University Clinic of Regensburg (protocol number: 23–3551-101).

2.2 Materials

The study was run on a Lenovo Thinkpad laptop computer with an external 14-inch touchscreen monitor (1920 x 1080px) placed approximately 45 cm in front of the participants. Children used the touchscreen to solve the working memory task. To implement the MR task and the implicit stereotype measure, we used the opensource software OpenSesame (Mathôt et al., 2012), children used large, colored buttons to respond to stimuli in these tasks.

2.2.1 Visuospatial working memory task

To measure VSWM, we used a computerized Corsi block-tapping test (Hasselhorn et al., 2012). The task involved an array of nine unsystematically arranged grey fields that remained visible throughout. Each trial presented a sequence of at least two positions, indicated by the appearance of a smiley face for 950 milliseconds (ms) in the respective fields. There was a 50 ms interstimulus interval. Immediately after each sequence, participants reconstructed the order by tapping the corresponding fields on touchscreen monitor (see Figure 1A). They had to remember both the positions and the order they appeared in. Task difficulty adapted dynamically: the sequence length increased when participants responded correctly (sometimes requiring several consecutive correct trials) and decreased after multiple consecutive errors. The task included ten critical trials. The longest correctly solved series among these trials, the Corsi span, was used to quantify VSWM. Before the main task, participants completed two or three practice trials, depending on whether the first trial was solved correctly.

2.2.2 Implicit stereotypes measures

We used an adapted single target implicit association test (ST-IAT; Karpinski and Steinman, 2006) to measure implicit associations between gender and toys commonly used in spatial play (Lego, building blocks, puzzle, toy train). Gender was represented by images of faces of boys and girls. Toys used in spatial play were also presented as images. Children completed five blocks (three practice blocks, two critical blocks), for a total of 144 trials. In the first practice block only faces were presented over 16 trials. The second practice block (16

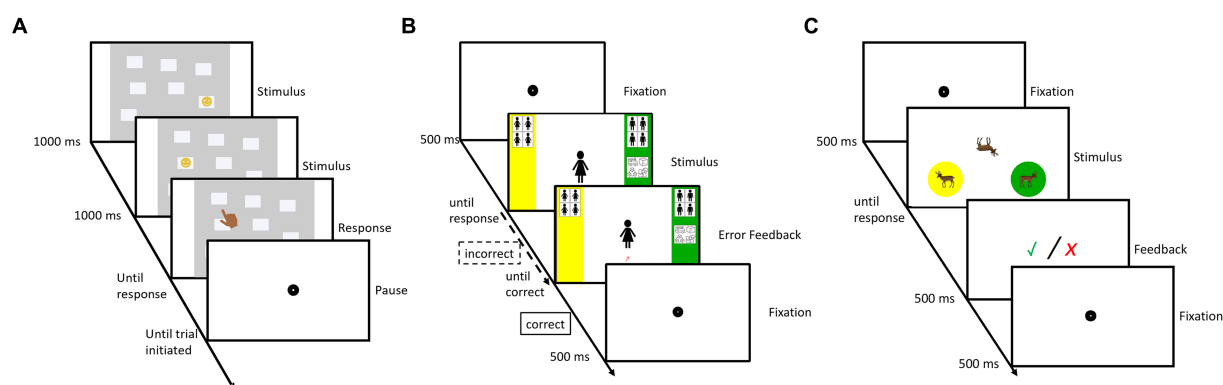


FIGURE 1
Schematic representations of trial sequences. The leftmost (A) schematic displays an exemplary trial sequence of the Corsi block-tapping test. For parsimony, the stimulus presentation (950 ms) and the interstimulus interval of 50 ms are presented together as stimulus. An example of an implicit association test trial is shown in the middle (B). Error feedback was not displayed if the initial answer was correct. The panel on the right (C) provides an example of a trial of the mental rotation task.

trials) included both faces and toys. In this phase, toys shared a response button with the faces of one gender—for example, boys and toys were assigned the left response button, while girls were assigned the right response button. This practice phase was followed by a critical block of 48 trials, using the same response assignment. Next, in a third practice block of 16 trials, the response assignment for the toys was switched. If boys and toys had previously shared a response button, girls and toys would now share a response button instead. This response assignment was retained for the second critical block of 48 trials. When participants gave an incorrect response, a red question mark appeared beneath the stimulus, signaling the need for correction before proceeding with the task. Response times were recorded at the time that the correct answer was given (Cvencek et al., 2011). A fixation dot was presented for 500 ms in between trials (see Figure 1B for schematic representation). Response assignments and block order were counterbalanced. ST-IAT scores, were calculated by subtracting the mean response time in the *boys + toys* block from that in the *girls + toys* block (see, e.g., Bluemke and Frieze, 2008). Positive scores indicated a stereotype favoring boys.

2.2.3 Mental rotation task

Participants completed a chronometric animal MR task. The stimuli were colored drawings of 12 different animals (camel, crocodile, dog, donkey, elephant, grizzly, lion, pig, rhino, sheep, turtle, and zebra) taken from Rossion and Pourtois (2004). In each trial, three drawings of the same animal were presented simultaneously: one in the center of the screen and two reference images at the bottom left and right. The reference images were always upright, facing outward, with the left image in a yellow circle and the right in a green circle. The central image appeared either upright or rotated by 90°, 180°, or 270° and could always be identified as facing left or right. The upright position served as a reference for determining its facing direction. In half the trials it faced left, in the other half, it faced right. Participants had to determine whether the centrally displayed animal matched the left or right reference image in orientation, mentally rotating it if necessary. They pressed the yellow button if it matched the left image and the green button if it matched the right. Before starting the computerized task, children completed five practice trials using printed images and manipulable cut-outs. This was followed by 16 computerized practice trials. The main experiment consisted of 72 critical trials (12 animals $\times$ 2 orientations $\times$ 3 rotation angles), with one rotation angle omitted per animal per orientation to shorten the task compared to previous versions (Ebert et al., 2024a). Each stimulus set remained on screen until a response was given. Feedback was provided after every trial: a green checkmark appeared for 500 ms for correct responses, while a red “x” was shown for incorrect responses. Before the next trial, a fixation dot was displayed for 500 ms (see Figure 1C). Participants could take a break after every eight trials, with a mandatory break after 36 trials. Response time and accuracy were recorded. Similar tasks have been used in previous research with kindergarten children (Hahn et al., 2010; Ebert et al., 2024a).

2.3 Procedure

Three experimenters conducted the study, with at least one of them present during each data collection session. Testing was conducted individually in a quiet room at the participant's

kindergarten. Consent forms, signed by the participants' parents, were handed to the experimenter by kindergarten staff. Before participation, the general procedure was explained to each child, and they were asked if they wanted to participate. All participants completed the tasks in the same order. First, they were introduced to the Corsi Block Task as a game in which they had to memorize routes. They first completed practice trials, and then the main task. Following a short break, they were introduced to the ST-IAT, described as a picture-sorting game. Each block was preceded by specific instructions outlining the response mapping for that block. In between blocks, children were encouraged to take breaks. After another short break, they completed the Animal MR Task. This task was framed as a game requiring them to imagine how animals would look when rotated. To introduce the concept, five practice trials involved manually rotating cut-out figures before moving on to computerized trials.

2.4 Statistical analyses

Per participant, MR response times more than two standard deviations above or below their individual respective mean response time, per rotation angle, were excluded from response time analyses. This applied to 209 observations in our sample. MR response times faster than 300 ms were excluded from response time and accuracy analyses. There were 17 such responses in our data. In addition to the preregistered plan, response time outliers were excluded based on the 1.5*IQR rule since extreme values were present in the data after outlier exclusion. We excluded 201 observations from response time analysis based on this criterion.

To test whether children held implicit gender stereotypes regarding spatial ability, we conducted a two-sided one sample t-test. To examine hypotheses 1 and 1.1 we fit generalized linear mixed models (GLMMs) to our data using the lme4 package (version 1.1.37; Bates et al., 2015) in R (version 4.5.0; R Core Team, 2025). MR accuracy was modeled using a binomial GLMM as recommended by Dixon, 2008. MR response time was modeled in an inverse gaussian GLMM. We included the nested random factors participant and kindergarten in all models before model reduction. Based on research by Barr et al. (2013), Bates et al. (2015), and Matuschek et al. (2017) we reduced the random effects structures of our respective maximal models following a backward selection approach [likelihood ratio tests (LRTs with α -level of 0.2)]. The independent variables were scaled based on their respective standard deviations to facilitate model convergence. Since there were indications of a ceiling effect in MR accuracy (see Supplementary Figure S1), we examined model diagnostics using plots and the DHARMA package (version 0.4.7; Hartig, 2024).

3 Results

Descriptive statistics and correlations among the measured variables are shown in Table 1. There was a significant correlation between MR accuracy and Corsi span ($r = 0.495$, $p < 0.001$). The Corsi span also significantly correlated with IAT scores ($r = 0.314$, $p = 0.023$). No other correlations reached significance.

The girls in our sample did not exhibit implicit stereotypes $t(51) = -1.098$, $p = 0.28$, $d = 0.15$ favoring boys or girls. Table 2 shows

TABLE 1 Descriptive statistics and correlations among study variables.

Variable	Mean	SD	Correlations (df)		
			2	3	4
1. IAT	−36.54	239.90	0.221(50)	0.190 (50)	0.314 (50)*
2. MR Accuracy	0.88	0.13	–	0.168 (50)	0.495 (50)***
3. MR response time	1761.15	343.78		–	0.115 (50)
4. Corsi span	3.90	0.69			–

* $p < 0.05$; *** $p < 0.001$.

model coefficients and test statistics of the inverse gaussian GLMM of MR response times. There was a significant main effect of rotation angle on MR response times ($\chi^2(1) = 58.53$, $p < 0.001$). Greater rotation angles were associated with longer response times. No other effect reached significance in this model.

Table 3 contains model coefficients alongside test statistics of the binomial GLMM of MR accuracy. There were significant main effects of angle ($\chi^2(1) = 19.57$, $p < 0.001$) and IAT score ($\chi^2(1) = 4.46$, $p = 0.035$) in this model. Greater rotation angles were associated with lower accuracy and greater IAT scores were linked to higher accuracy. Since the overall proportion of correct responses was relatively high at 88%, we ran diagnostics to evaluate potential model misspecification. Model diagnostics using the DHARma package (Hartig, 2024) showed no evidence of overdispersion ($p = 0.88$), outliers ($p = 0.85$), or non-uniform residuals ($p = 0.14$). Visual inspection of the Pearson residuals against fitted values revealed a slight fan-shaped pattern and some extreme residuals at high predicted values, consistent with a potential ceiling effect. However, these are typical for binomial models with high probabilities and were not accompanied by signs of poor model fit.

4 Discussion

The current study set out to examine potential relationships between implicit stereotypes regarding spatial ability, VSWM and spatial task performance in preschool-aged kindergarten girls. We assessed VSWM, implicit stereotypes and MR performance of participating children.

4.1 Gender stereotypes

We found no evidence of implicit stereotypes in our sample. In a previous study using the same IAT (Ebert et al., 2024a), implicit stereotypes were observed in a mixed-gender sample, though the effect was small ($d = 0.17$). Notably, when results were broken down by gender in that study, no significant effects emerged (both gender-specific d s = 0.17). In the current study, we observed a similarly small, non-significant effect ($d = 0.15$); interestingly, however, the effect was descriptively in the opposite direction compared to the earlier findings (Ebert et al., 2024a). Hence, the current findings call into question the robustness of the effect observed by Ebert et al. (2024a), and they cast doubt on our earlier conclusion that kindergarten children hold implicit gender stereotypes about spatial ability at preschool age. There are several plausible explanations for the divergent findings across the two studies. One possibility is that

they reflect cohort effects, potentially linked to shifts in public discourse and evolving societal values. For example, Block et al. (2022) demonstrated that even brief exposure to stereotypical or counter-stereotypical narratives can influence implicit gender stereotypes in children aged six to eleven. This suggests that increased awareness and discussion of gender stereotypes among parents, educators, or in media may shape children’s implicit associations—possibly even within short timeframes. Finally, it is worth emphasizing that the effect sizes observed in both studies were very small, raising the possibility that young children may not hold stable or robust implicit stereotypes about spatial ability at this age.

4.2 Visuospatial working memory

We found significant correlations between our measure of visuospatial working memory (VSWM), mental rotation (MR) accuracy, and IAT scores. This aligns with previous research showing a robust relationship between VSWM and MR performance in both children (Lehmann et al., 2014) and adults (Kaufman, 2007; Hofmann et al., 2024). Notably, these studies have specifically linked Corsi span measures to MR outcomes, reinforcing the idea that individual differences in spatial working memory capacity contribute to spatial reasoning abilities.

4.3 The relation of gender stereotypes and mental rotation performance

Our results show no relation of gender stereotypes about spatial ability and MR response times. This aligns with findings from our previous study (Ebert et al., 2024a), in which we also observed no effects of either implicit or explicit gender stereotypes on MR response times in a mixed-gender sample. Also, akin to results from the same study, we did find an effect of implicit gender stereotypes on MR accuracy in the current sample. Interestingly, the direction of the effect was reversed: girls who more strongly associated spatial ability with boys performed more accurately than those who associated it with their own gender. Accordingly, we reject our main hypotheses regarding the relationship between implicit gender stereotypes and MR performance. This unexpected direction of this effect is difficult to reconcile with existing theoretical frameworks (e.g., Schmader, 2010). However, it is plausible that stereotypes are in development in the researched age group, or that they are not applied to the self (Ebert et al., 2024a) and therefore do not impact performance in commonly observed ways.

TABLE 2 Generalized linear mixed model for the dependent variable MR response time.

Predictor	Estimate	95% CI		Test statistic	p-value
Intercept	2412.5	1241.87	3417.89		
IAT	2594.9	−4225.62	6492.42	$\chi^2(1) = 0.58$	0.447
Angle	546.8	497.10	633.15	$\chi^2(1) = 58.53$	<0.001
Age	−913.0	−2015.01	122.51	$\chi^2(1) = 1.89$	0.170
IAT * Angle	4156.2	1498.64	10896.73	$\chi^2(1) = 1.90$	0.168

This model included random intercepts by institution and participant, and random slopes for rotation angle by participant. Confidence intervals are based on bootstrapping of linear mixed model with 1,000 simulations.

TABLE 3 Generalized linear mixed model for the dependent variable MR accuracy.

Predictor	Estimate	95% CI		Test statistic	p-value
Intercept	4.67	−1.03	9.93		
IAT	32.99	1.05	62.60	$\chi^2(1) = 4.46$	0.035
Angle	−1.20	−1.64	−0.74	$\chi^2(1) = 19.57$	<0.001
Age	−0.72	−5.96	4.97	$\chi^2(1) = 0.05$	0.832
IAT * Angle	−13.27	−44.31	18.73	$\chi^2(1) = 0.85$	0.358

This model included random intercepts by institution and participant, and random slopes for rotation angle by participant. Confidence intervals are based on bootstrapping with 1,000 simulations.

4.4 Limitations

It is important to note that the MR tasks used in this study was not identical to the task in Ebert et al. (2024a), where children were asked to decide whether two drawings were identical or mirror images of each other. In contrast, the current study required children to match a target drawing to one of two reference images. This adaptation was intended to increase comprehensibility for young children and to yield more analyzable data, as it eliminated the need to discard mirror-image trials. However, there was some indication of a potential ceiling effect in MR accuracy in the present study, suggesting that the modified task may have been too easy. Nonetheless, model assumptions were carefully examined, and no evidence of substantial model misspecification was found. It should also be noted that the target sample size was not met, limiting our ability to reliably detect effects of interest. The used IAT rests on the assumption that an association between spatial toys (representative of spatial ability) and gender is measured, an assumption that cannot be confirmed within the confines of this study. Combining implicit and explicit measures in future research may in part alleviate this issue.

5 Conclusion

Taken together, our findings call into question, whether implicit gender stereotypes about spatial ability are reliably present in preschool-aged kindergarten girls. Moreover, they do not support the hypothesis that such stereotypes are meaningfully related to spatial task performance. We could not replicate a previously observed effect of implicit stereotypes on MR accuracy. In fact, the direction of the

effect was reversed in the present study, complicating interpretation. Given the small and inconsistent effects across studies, and the methodological differences between them, we urge caution in drawing firm conclusions about the presence or impact of such stereotypes at this early developmental stage. The results do, however, suggest that kindergarten years may hold potential in shaping children’s resilience against stereotypes. If stereotypes are not fully developed at this stage, parents and caretakers may successfully intervene in structured or unstructured ways to foster un-stereotyped thinking. Future research should seek to clarify under what conditions, if any, implicit gender stereotypes influence spatial cognition in young children. This may include larger samples, alternative measures of implicit bias ideally, combined with explicit measures (e.g., structured interview or questionnaire), and longitudinal approaches that capture developmental changes over time. Future studies should also aim for measures of MR with adequate difficulty to avoid floor- and ceiling effects. For instance, the difficulty of the task used in this study could be adapted through the inclusion of more rotation angles. Pilot testing or comparative studies probing different MR tasks to identify sensible difficulty levels could prove helpful. Moreover, investigating the role of parents’ teachers’ or caretakers’ stereotypes could provide further insight into the complex relationships between stereotyping and children’s spatial ability.

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 Ethics Board of the University Clinic of Regensburg. 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.

Author contributions

WE: Writing – original draft, Investigation, Writing – review & editing, Formal analysis, Data curation, Conceptualization. LJ: Methodology, Writing – original draft. PJ: Validation, Project

administration, Methodology, Supervision, Conceptualization, Funding acquisition, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This study was conducted as part of the Marie Skłodowska-Curie Actions “SellSTEM” project (Grant agreement No. 956124) and funded using project resources.

Acknowledgments

The authors would like to thank Antonia Voll and Maria Meisel who significantly contributed to this study through their help with data collection.

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.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

References

- Baddeley, A. D., and Hitch, G. (1974). “Working memory” in Psychology of learning and motivation. ed. G. H. Bower (New York: Academic Press), 47–89.
- Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *J. Mem. Lang.* 68. doi: 10.1016/j.jml.2012.11.001
- Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01
- Block, K., Gonzalez, A. M., Choi, C. J. X., Wong, Z. C., Schmader, T., and Baron, A. S. (2022). Exposure to stereotype-relevant stories shapes children's implicit gender stereotypes. *PLoS One* 17:e0271396. doi: 10.1371/journal.pone.0271396
- Bluemke, M., and Friese, M. (2008). Reliability and validity of the single-target IAT (ST-IAT): assessing automatic affect towards multiple attitude objects. *Eur. J. Soc. Psychol.* 38, 977–997. doi: 10.1002/ejsp.487
- Bonnot, V., and Croizet, J.-C. (2007). Stereotype internalization and women's math performance: the role of interference in working memory. *J. Exp. Soc. Psychol.* 43, 857–866. doi: 10.1016/j.jesp.2006.10.006
- Chan, N. H. W., and Rosenthal, H. E. S. (2014). Working memory moderates stereotype threat effects for adolescents in Hong Kong. *Rev. Int. Psychol. Soc.* 27, 103–118. Available at: <https://shs.cairn.info/revue-internationale-de-psychologie-sociale-2014-3-page-103?lang=fr>
- Christie, M., O'Neill, M., Rutter, K., Young, G., and Medland, A. (2017). Understanding why women are under-represented in science, technology, engineering and mathematics (STEM) within higher education: a regional case study. *Production* 27:e20162205. doi: 10.1590/0103-6513.220516
- Cvencek, D., Greenwald, A. G., and Meltzoff, A. N. (2011). Measuring implicit attitudes of 4-year-olds: the preschool implicit association test. *J. Exp. Child Psychol.* 109, 187–200. doi: 10.1016/j.jecp.2010.11.002
- Dixon, P. (2008). Models of accuracy in repeated-measures designs. *J. Mem. Lang.* 59, 447–456. doi: 10.1016/j.jml.2007.11.004
- Ebert, W. M., Jost, L., and Jansen, P. (2024a). Gender stereotypes in preschoolers' mental rotation. *Front. Psychol.* 15:1284314. doi: 10.3389/fpsyg.2024.1284314
- Ebert, W. M., Jost, L., Jansen, P., Stevanovski, B., and Voyer, D. (2024b). Visual working memory as the substrate for mental rotation: a replication. *Psychon. Bull. Rev.* 32, 1204–1216. doi: 10.3758/s13423-024-02602-4
- Ertl, B., Luttenberger, S., and Paechter, M. (2017). The impact of gender stereotypes on the self-concept of female students in STEM subjects with an under-representation of females. *Front. Psychol.* 8:703. doi: 10.3389/fpsyg.2017.00703
- European Commission, Directorate-General for Research and Innovation. (2021). She figures 2021: Gender in research and innovation: Statistics and indicators. Publications Office of the European Union. Available online at: <https://data.europa.eu/doi/10.2777/06090> (Accessed January 29, 2025).
- 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
- Hahn, N., Jansen, P., and Heil, M. (2010). Preschoolers' mental rotation: sex differences in hemispheric asymmetry. *J. Cogn. Neurosci.* 22, 1244–1250. doi: 10.1162/jocn.2009.21236
- Hanson, S. L. (1996). Lost talent: women in the sciences. Labor and social change series. Philadelphia, PA: Temple University Press.
- Hartig, F. (2024) DHARMA: Residual diagnostics for hierarchical (multi-level / mixed) regression models. Available online at: <https://CRAN.R-project.org/package=DHARMA> (Accessed 18 June, 2025).
- Hasselhorn, M., Schumann-Hengsteler, R., Gronauer, J., Grube, D., Mähler, C., Schmid, I., et al. (2012). AGTB 5–12: Arbeitsgedächtnistestbatterie für Kinder von 5 bis 12 Jahren. Göttingen: Hogrefe.
- Heil, M., and Rolke, B. (2002). Toward a chronopsychophysiology of mental rotation. *Psychophysiology* 39, 414–422. doi: 10.1017/S0048577202001105
- Hofmann, P., Siebertz, M., and Jansen, P. (2024). No role of working memory in the relation between mental rotation and postural stability. *Front. Cogn.* 2:1298371. doi: 10.3389/fcogn.2023.1298371
- Hyun, J.-S., and Luck, S. J. (2007). Visual working memory as the substrate for mental rotation. *Psychon. Bull. Rev.* 14, 154–158. doi: 10.3758/BF03194043
- Just, M. A., and Carpenter, P. A. (1985). Cognitive coordinate systems: accounts of mental rotation and individual differences in spatial ability. *Psychol. Rev.* 92, 137–172. doi: 10.1037/0033-295X.92.2.137
- Karpinski, A., and Steinman, R. B. (2006). The single category implicit association test as a measure of implicit social cognition. *J. Pers. Soc. Psychol.* 91, 16–32. doi: 10.1037/0022-3514.91.1.16

Generative AI statement

The authors declare that Gen AI was used in the creation of this manuscript. The authors used ChatGPT (OpenAI, 2025) to help rephrase and clarify selected paragraphs during the revision of this manuscript. All content was reviewed and verified by the authors.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

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/fpsyg.2025.1650979/full#supplementary-material>

- Kaufman, S. B. (2007). Sex differences in mental rotation and spatial visualization ability: can they be accounted for by differences in working memory capacity? *Intelligence* 35, 211–223. doi: 10.1016/j.intell.2006.07.009
- Laurin, R., Renard-Moulard, M., and Cometti, C. (2022). Stereotype threat effect on a simple motor task: an investigation of the Visuo-spatial working memory. *Res. Q. Exerc. Sport* 93, 423–428. doi: 10.1080/02701367.2020.1826391
- Lehmann, J., Quaiser-Pohl, C., and Jansen, P. (2014). Correlation of motor skill, mental rotation, and working memory in 3- to 6-year-old children. *Eur. J. Dev. Psychol.* 11, 560–573. doi: 10.1080/17405629.2014.888995
- Levine, S. C., Foley, A., Lourenco, S., Ehrlich, S., and Ratliff, K. (2016). Sex differences in spatial cognition: advancing the conversation. *Wiley Interdiscip. Rev. Cogn. Sci.* 7, 127–155. doi: 10.1002/wcs.1380
- Mathôt, S., Schreij, D., and Theeuwes, J. (2012). OpenSesame: an open-source, graphical experiment builder for the social sciences. *Behav. Res. Methods* 44, 314–324. doi: 10.3758/s13428-011-0168-7
- Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., and Bates, D. (2017). Balancing type I error and power in linear mixed models. *J. Mem. Lang.* 94, 305–315. doi: 10.1016/j.jml.2017.01.001
- McGlone, M. S., 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
- Moè, A., and Pazzaglia, F. (2006). Following the instructions! *Learn. Individ. Differ.* 16, 369–377. doi: 10.1016/j.lindif.2007.01.002
- National Science Board. (2024). The STEM labor force: Scientists, engineers, and skilled technical workers. Alexandria, VA: National Science Foundation. Available online at: <https://ncses.nsf.gov/pubs/nsb20245/> (Accessed February 3, 2025).
- Newcombe, N. (2017). Harnessing spatial thinking to support Stem learning. OECD Education Working Papers, Paris: OECD Publishing. doi: 10.1787/7d5dcae6-en
- Newcombe, N. S. (2020). The puzzle of spatial sex differences: current status and prerequisites to solutions. *Child Dev. Perspect.* 14, 251–257. doi: 10.1111/cdep.12389
- Newcombe, N. S., and Shipley, T. F. (2015). “Thinking about spatial thinking: new typology, new assessments” in *Studying visual and spatial reasoning for design creativity*, ed. J. S. Gero (Dordrecht: Springer Netherlands), 179–192.
- Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., et al. (2009). National differences in gender–science stereotypes predict national sex differences in science and math achievement. *Proc. Natl. Acad. Sci.* 106, 10593–10597. doi: 10.1073/pnas.0809921106
- R Core Team (2025) R: A language and environment for statistical computing. Available online at: <https://www.R-project.org/> (Accessed 18 June, 2025).
- Rahe, M., and Jansen, P. (2022). Sex differences in mental rotation: the role of stereotyped material, perceived performance and extrinsic spatial ability. *J. Cogn. Psychol.* 34, 400–409. doi: 10.1080/20445911.2021.2011896
- Rahe, M., Ruthsatz, V., and Quaiser-Pohl, C. (2021). Influence of the stimulus material on gender differences in a mental-rotation test. *Psychol. Res.* 85, 2892–2899. doi: 10.1007/s00426-020-01450-w
- Régner, I., Smeding, A., Gimmig, D., Thinus-Blanc, C., Monteil, J.-M., and Huguet, P. (2010). Individual differences in working memory moderate stereotype-threat effects. *Psychol. Sci.* 21, 1646–1648. doi: 10.1177/0956797610386619
- Rossion, B., and Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart’s object pictorial set: the role of surface detail in basic-level object recognition. *Perception* 33, 217–236. doi: 10.1068/p5117
- Sanchis-Segura, C., Aguirre, N., Cruz-Gómez, Á. J., Solozano, N., and Forn, C. (2018). Do gender-related stereotypes affect spatial performance? Exploring when, how and to whom using a chronometric two-choice mental rotation task. *Front. Psychol.* 9:1261. doi: 10.3389/fpsyg.2018.01261
- Schmader, T. (2010). Stereotype threat deconstructed. *Curr. Dir. Psychol. Sci.* 19, 14–18. doi: 10.1177/0963721409359292
- Schmader, T., and Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *J. Pers. Soc. Psychol.* 85, 440–452. doi: 10.1037/0022-3514.85.3.440
- Shepard, R. N., and Cooper, L. N. (1982). Mental images and their transformations. Cambridge: MIT Press.
- Super, D. E., and Bachrach, P. B. (1957). Scientific careers and vocational development theory: a review, a critique and some recommendations. Oxford, England: Columbia Univer.
- van der Ham, I. J. M., and Koutzmp, V. (2023). Stereotypes and self-reports about spatial cognition: impact of gender and age. *Curr. Psychol.* 42, 26904–26912. doi: 10.1007/s12144-022-03827-z
- Vander Heyden, K. M., van Atteveldt, N. M., Huizinga, M., and Jolles, J. (2016). Implicit and explicit gender beliefs in spatial ability: stronger stereotyping in boys than girls. *Front. Psychol.* 7:1114. doi: 10.3389/fpsyg.2016.01114
- 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
- Wai, J., Lubinski, D., and Benbow, C. P. (2009). Spatial ability for STEM domains: aligning over 50 years of cumulative psychological knowledge solidifies its importance. *J. Educ. Psychol.* 101, 817–835. doi: 10.1037/a0016127
- Wang, M.-T., and Degol, J. (2013). Motivational pathways to STEM career choices: using expectancy-value perspective to understand individual and gender differences in STEM fields. *Dev. Rev.* 33, 304–340. doi: 10.1016/j.dr.2013.08.001