Heliyon 9 (2023) e15247

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Persistent gender differences in spatial ability, even in STEM experts

Elina S. Tsigeman [a], Maxim V. Likhanov [b], Anna V. Budakova [a], Aydar Akmalov [c], Ildar Sabitov [d], Evgeniia Alenina [a], Ksenia Bartseva [e], Yulia Kovas [f,*]

[a] Department of Psychology, Tomsk State University, Tomsk, Russia
[b] State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China
[c] Kazan Open University of Talents 2.0, Kazan, Republic of Tatarstan, Russia
[d] Ulyanovsk State University, Ulyanovsk, Russia
[e] Laboratory for Social and Cognitive Informatics, Higher School of Economics, Saint Petersburg, Russia
[f] Department of Psychology, Goldsmiths, University of London, London, UK

A R T I C L E I N F O

Keywords:
Spatial ability
Spatial cognition
STEM education
Gender differences
Art and sport

A B S T R A C T

Background: Spatial ability (SA) shows wide variability. One proposed explanation for the observed individual difference in SA is variability in interest and engagement in activities that promote spatial ability. Research also robustly shown that males on average outperform females in most aspects of SA. Previous studies have identified a number of activities that can potentially contribute to both individual and gender differences in SA, including tinkering with electronics, particular sports activities, and designing. However, the findings regarding these links are inconsistent. One way to investigate these links is to compare the groups that are intensively engaged with these activities.

Aim: The present study aims to evaluate the robustness of these links by comparing SA in adolescents with expertise in STEM, arts, and sports, with their unselected peers. We also aimed to assess whether gender differences in SA are still present in expert groups.

Methods: The data on ten small-scale SA tests was collected in an unselected sample of adolescents (N = 864, Mean age = 15.4, SD = 1.1); as well as in 3 samples of adolescents with expertise in STEM (N = 667, Mean age = 15, SD = 1.2); in Arts (N = 280, Mean age = 15, SD = 1.2) and in Sports (N = 444, Mean age = 14.3, SD = 0.7).

Results: Out of the three expert groups, only STEM experts on average outperformed the unselected group on all SA tasks. The STEM experts also outperformed Arts and Sports experts. Gender differences persisted in all expert groups, with moderate effect sizes.

Discussion: Findings support previously established links between spatial ability and STEM-related expertise. In contrast, such links were not found for expertise in arts and sports. Consistent with previous research, we found gender differences in SA for all samples, which persisted in STEM experts.

* Corresponding author.
E-mail address: y.kovas@gold.ac.uk (Y. Kovas).

https://doi.org/10.1016/j.heliyon.2023.e15247
Received 26 July 2022; Received in revised form 29 March 2023; Accepted 30 March 2023
Available online 5 April 2023

Abbreviations

SA Spatial ability
STEM Science, technology, engineering, and mathematics.

1. Introduction

Spatial ability (SA) is an ability to produce, recall, store, and modify spatial relations among objects [1]. SA has been linked to performance in different domains, such as everyday problem-solving [1,2] and maths achievement at school [3]. Research has also found links between SA (measured at age of 13) and long-term outcomes in STEM-related careers (measured as a number of patents and publications [4,5]). SA's links with sports and arts were identified [6–8], but they are less robust.

Individual differences in different aspects of SA are explained by both genetic and environmental factors [1,9]. Research has also shown that both genetic and environmental factors contribute to links between SA and educational and other outcomes [10,11]. Different mechanisms may underlie links between SA and achievement in different areas, such as STEM, arts, and sports. For example, both genetic and environmental factors may contribute directly to some aspects of cognitive processing associated with SA (e.g., visualisation or manipulation of spatial information). In turn, people with greater SA may show greater performance and engagement in STEM domain, arts, or sports. This phenomenon has been discussed in the framework of *the relevant skills hypothesis* [12]. People with lower SA might avoid specific SA-demanding activities in accordance with *the default hypothesis* [12]. In addition, genetic and environmental factors may contribute to motivation and interest in particular domains [13], which may lead to greater engagement in STEM or arts or sports, and consequently to greater SA.

Some experimental research has shown that engaging in STEM, arts, sports, and games with spatial content can enhance SA [6,8,14,15]. For example, one study showed enhanced mental rotation after 10 months of wrestling training in students [16] and the other showed improved mental rotation after 5 weeks of dance training in young children [17]. However, other studies did not find such effects [18,19]. The positive effect of training on SA is consistent with a number of causal explanations. SA enhancement might occur as a result of deliberate practice [20] of a domain activity (e.g. targeted training of aiming in sports or training of perspective view in drawing) or as a result of near transfer – when practice of a multidimensional activity (such as music performance) might indirectly enhance SA [6].

Quasi-experimental studies also produced mixed findings. For example, one study compared scores on self-report measures of SA of students majoring in arts (fine art, painting, sculpture, design, and illustration), STEM, and humanities. Arts students scored higher than the other groups in visual imagery, object-spatial imagery, and object recognition ability task, with modest effect sizes [21]. Comparison of athletes and non-athletes on the ability to perform mental rotation and spatial orientation tasks yielded athletes' advantage in these tasks with large effect sizes in several studies [7,22,23]. However, there are inconsistent results on effects of sports training. Some studies show that the link between SA and sports depends on intensity and time of engagement. One study showed that longer period of sports experience is associated with better spatial orientation ($r = 0.26$ [24]). Similarly, longer PE (>10 h per week) was associated with better mental rotation in comparison to shorter PE (90 min per week) in schoolchildren [25]. In contrast, one study found no difference in mental rotation between elite and recreational athletes [26], suggesting no "dose effect".

Beyond expertise, another key factor contributing to individual differences in SA is gender. A wealth of studies demonstrated males of all ages on average outperform females in most aspects of small-scale SA, including mental rotation, paper folding, and cross-sections [8,12,18,25,27–32], with medium to large effect sizes (Hedge's g from 0.62 to 1.34 in meta-analysis [32]). A recent meta-analysis has also found robust gender differences in large-scale SA, namely navigation ability [33]. Less consistent gender differences were found for spatial visualisation tasks [34]. Further, gender differences in SA might be moderated by age. An early meta-analysis concluded that gender differences in SA appear in childhood and stay persistent over the lifespan [35]. However, a recent meta-analysis on navigation skills showed smaller gender differences in participants younger than 13 years in comparison to older ones [33].

Spatial ability level, expertise in different areas (e.g., STEM), engagement in SA-related activity (e.g., videogames), gender, and age may interact. For example, one study with undergraduate students examined the effects of videogames training on mental rotation ability [36]. Males outperformed females before the video games training ($d = 0.54$), but this difference disappeared after the training. Further analyses showed that the spatial ability gain was greater in females. In contrast, another study with 7–8 year old children showed somewhat greater gains for males in mental rotation, following a robotics-based training (programming a robot to orient according to a map [37]. This result may be explained by age differences between the samples, as greater female gains were found in undergraduate students and greater male gains in 7–8 year olds. An additional moderator may be neuroticism and anxiety, which are on average higher in females [38] and have been shown to affect self-efficacy and performance [39]. These affective characteristics may be involved in complex reciprocal processes. For example, if one struggles with the task, it might lead to an increased anxiety, which in turn might negatively affect performance. Therefore, lower average spatial ability in females may further be impacted by anxiety triggers. If this is correct, then the gender differences in spatial ability should be smaller in high-performing samples. Moreover, neuroticism was previously shown to be negatively correlated with performance in timed tasks [40]. As spatial ability tasks are usually timed, greater neuroticism in females may contribute to the observed gender differences in performance on spatial tasks. Age may be an additional factor in this complex system. For example, one study has shown that anxiety increases in females (but not in males) during the transition from childhood to adolescence [41].

The differences in effect may also be related to gender differences in self-perceived ability, interest, and enjoyment of SA-related activity, which in turn may affect engagement. For example, one study found that males report greater self-efficacy, enjoyment, and engagement in spatial activities [42]. Another study [18] showed that females reported more visual-arts experience and less involvement in spatial-orientation activities than males, with weak effects ($\eta p2 = 0.05$). In both studies, the link between activities and SA was negligible. This is consistent with another study, which found little evidence that increased female participation in sports increase spatial ability and engagement in spatially related occupations [43].

Only a handful of studies explored interaction between spatial ability, gender, and expertise in SA-related areas. For example, one study found male advantage in SA in an unselected sample that was not observed in a selected sample of athletes [22]. The authors suggested that sports activity reduced gender differences, presumably via hormonal modulation mechanisms (i.e., via increased levels of androgens in females). The role of hormones (prenatal and postnatal) in SA has been widely researched, with mixed findings. Some research supports hormonal explanations of SA advantage in males. For example, some studies of individuals with congenital adrenal hyperplasia (a disease that affects production of sex steroids) indirectly support the hormonal explanation. In these studies (e.g. Refs. [44,45]) females with this disorder outperformed unaffected sisters; and affected males performed lower than unaffected brothers on SA tasks. These results may stem from both direct effect of androgens on SA and indirect effect from androgens on spatial activity interests. However, several studies, including one meta-analysis, did not support the role of androgens in SA and suggest other causal processes, including socialisation [30,42,46]. Moreover, it is difficult to apply hormonal explanation to potential positive effects of arts and music on SA.

Overall, the mechanisms underlying individual and group differences in SA remain poorly understood, as previous research has multiple limitations: many studies were underpowered (i.e. used insufficient sample sizes); focused only on links between SA and STEM; investigated a limited number of SA facets (mainly mental rotation and paper folding tasks); and did not explore "the dosage effect" of engagement.

The present paper aims to explore expertise by gender interactions, addressing limitations of the previous literature. The study used a battery of 10 SA tests to collect data from a large sample of male and female adolescents, with expertise in STEM, arts, or sports; and their unselected peers.

Several predictions were made based on previous research:

1. Students with expertise in STEM, arts, or sports will outperform the unselected sample on all SA tasks;
2. Males will outperform females in all SA tasks;
3. There will be reduced or absent gender differences in expert groups of adolescents.

2. Methods

2.1. Participants

2.1.1. Expert samples

One thousand, three hundred and sixty-nine adolescents (*Mean age* = 14.84; *SD* = 1.16) with expertise in different areas were recruited at the educational centres in Russia. Expertise in this study was operationalised as participation in intensive extracurricular activities in a specific domain. For STEM experts this included clubs, summer schools, boot camps and additional tutoring, participation in domain-specific competitions, e.g. subject Olympiads (see Ref. [40]). For Sports experts, criteria included participation in sport teams, clubs, sports leagues, sport camps and competitions. For Arts experts, criteria included training in performing arts, fine arts, ballet, or literature; participation in clubs, competitions, exhibitions, and concerts. All experts were actively involved in their domain activities for at least a year and had a track record of high performance in a respective area.

2.1.2. Unselected sample

Eight hundred and sixty-four schoolchildren (413 females; *Mean age* = 15.40; *SD* = 1.13) were recruited from public schools in two cities of Russia with no selection criteria. Table 1 presents the samples' composition.

Table 1
Composition of all samples.

	STEM (N = 656)		Arts (N = 280)		Sports (N = 443)		Unselected (N = 864)		Arts and Sports combined * (N = 718)	
	N Males	N Females	N Males	N Females	N Males	N Females	N Males	N Females	N Males	N Females
	390 (59.4%)	250 (38.1%)	56 (20%)	223 (79.6%)	368 (83%)	71 (16%)	451 (52%)	413 (47.8%)	424 (59%)	294 (40.9%)
Age range	13–18	13–18	13–18	13–18	13–16	13–17	13–18	13–18	13–18	13–18
Mean age (SD)	15.0 (1.21)	15.14 (1.26)	14.49 (1.30)	15.08 (1.15)	14.27 (0.58)	14.47 (1.07)	15.41 (1.15)	15.39 (1.12)	14.31 (0.77)	14.96 1.16)

Note: 16 participants from STEM group, 1 participant from Arts group and 4 participants from Sports group did not indicate their gender; *see section 3.1 for explanation on combined sample. Significant age differences were found in Arts sample and Combined non-STEM sample.

2.2. Procedure

The Ethics committee of the Interdisciplinary Research at Tomsk State University approved the study (code of ethical approval: 16012018-5). The participants and their parents or legal guardians received information regarding goals and procedures of the study and the voluntary basis of their participation. Only the students whose parents or legal guardians provided written consent forms participated in the study. Additionally, assent was obtained from the adolescent participants before the testing session. Participants did not receive any compensation for their participation.

All testing took place at the educational centres, where adolescents from different schools were taking part in some extra curriculum activity. All participants completed the same battery on individual laptops in a similar laboratory condition under supervision of a researcher. The testing took approximately 90 min.

2.3. Materials

Before the start of the cognitive tests, the participants filled in a computerised socio-demographic inventory providing their age, gender, and area of expertise.

2.4. King's Challenge battery

Ten small-scale SA timed tasks from a gamified online battery "King's challenge" were used to assess different facets of SA [1]. The sum of correct answers for each task is used to assess each facet of SA. See tests description in Table 2. For a more detailed battery description, see Rimfeld and colleagues [1]. The battery was adapted to Russian and validated in student [27,47] and adolescent samples [48].

2.5. Statistical analysis

All statistical analyses were conducted using IBM SPSS (Statistical Package for the Social Sciences), version 23 and RStudio [49]. All variables were standardised and screened for univariate outliers within groups. The threshold of $Z = 3.29$ was used as recommended in Field (2007) to exclude outliers. Overall, there were fewer than 5% of univariate outliers; and fewer than 5% of multivariate outliers according to Mahalanobis distance within each group. Skewness and kurtosis of all variables varied within an acceptable range (i.e. below the cut-off of 2 as recommended by George and Mallery [50]). However, distributions deviated from normality according to the Kolmogorov-Smirnov test; therefore, Pillai's Trace statistic was used in MANOVAs as the most robust to violation of the assumption.

3. Results

Means and standard deviations for each sample are presented in Fig. 1. Additional details, including post-hoc analyses, are presented in Appendix (Tables A.1-A.2).

3.1. Sample differences

Before testing the first hypothesis, we created residuals scores after controlling for gender and age. The gender composition of the

Table 2
Test battery description.

Task name	N of items	Time limit per item (sec)	Description	Reliability (Cronbach's alpha)				
				Full sample	STEM	Art	Sport	Unselected
3D to 2D drawing	5	45	sketching a 2D layout of a 3D object from a specified viewpoint	.94	.90	.91	.91	.93
2D to 3D drawing	7	70	sketching a 3D object from a 2D layout	.87	.79	.81	.84	.84
Cross-sections	15	20	visualising cross-sections of objects	.85	.79	.83	.78	.81
Elithorn mazes	10	7	joining as many dots as possible from an array	.87	.82	.82	.85	.89
Mazes	10	25	searching for a way through a 2D maze	.68	.60	.56	.66	.68
Mechanical reasoning	16	25	multiple-choice naive physics questions	.63	.60	.45	.48	.58
Paper folding	15	20	visualising holes in a piece of paper which it is folded, pierced, and unfolded	.89	.83	.86	.83	.87
Pattern assembly	15	20	mentally combining pieces of objects together to make a complete figure	.78	.68	.75	.74	.76
Perspective-taking	15	20	visualising objects from a different perspective	.89	.89	.85	.87	.88
Shape rotation	15	20	mentally rotating objects	.86	.82	.80	.81	.84

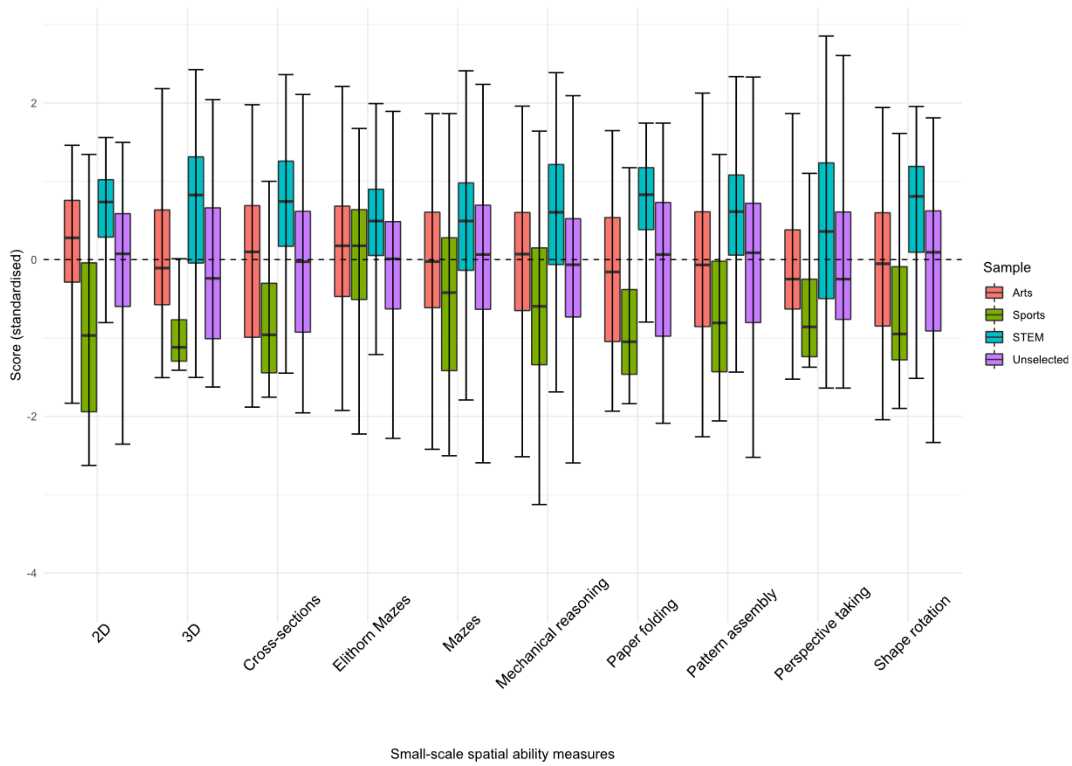

Fig. 1. Standardised residuals for correct answers on SA tasks in all samples (gender and age differences regressed out).

Table 3
Sample x Gender univariate effects.

Subtest (range of scores)	Samples mean (SD): STEM Males	STEM Females	non-STEM *** Males	non-STEM *** Females	Unselected Males	Unselected Females	Sample × Gender interaction F	η_p^2	Sample F	η_p^2	Gender F	η_p^2
3D to 2D drawing (0–5)	4.37 (.65)	4.11 (.76)	2.24 (1.44)	3.27 (1.26)	3.63 (1.19)	3.10 (1.31)	83.83**	0.07	242.04**	0.19	1494.98**	0.42
2D to 3D drawing (0–7)	4.08 (1.77)	3.43 (1.61)	.83 (1.11)	2.18 (1.60)	2.69 (1.93)	2.00 (1.68)	80.67**	0.07	281.17**	0.21	417.35**	0.17
Cross-sections (0–15)	9.10 (3.34)	8.00 (3.53)	3.16 (2.93)	5.00 (3.73)	6.67 (3.73)	5.33 (3.45)	42.54**	0.04	250.24**	0.20	638.68**	0.24
Elithorn mazes (0–10)	8.40 (.91)	7.84 (.97)	7.13 (1.61)	7.13 (1.46)	7.63 (1.59)	6.86 (1.76)	12.74**	0.01	81.87**	0.07	5819**	0.74
Mazes (0–10)	6.51 (1.78)	5.95 (1.94)	4.18 (2.16)	4.91 (1.96)	5.56 (2.31)	5.21 (2.09)	17.20**	0.01	98.46**	0.09	1306.41**	0.39
Mechanical reasoning (0–16)	12.26 (2.24)	10.28 (2.35)	8.20 (2.52)	8.56 (2.34)	10.58 (2.52)	8.58 (2.40)	50.58**	0.04	213.55**	0.17	4216.76**	0.67
Paper folding (0–15)	11.48 (3.33)	10.89 (3.18)	3.99 (3.39)	7.17 (4.07)	8.32 (4.36)	7.64 (4.25)	53.28**	0.05	320.17**	0.24	643.57**	0.24
Pattern assembly (0–15)	8.18 (2.69)	7.13 (2.72)	3.96 (2.91)	5.32 (3.00)	6.81 (3.23)	5.40 (2.99)	40.73**	0.03	164.22**	0.14	903.42**	0.31
Perspective-taking (0–15)	7.39 (4.63)	3.84 (3.61)	2.52 (3.16)	3.81 (3.27)	5.77 (4.43)	2.79 (2.99)	46.70**	0.04	91.12**	0.08	693.72**	0.25
Shape rotation (0–15)	10.47 (3.57)	8.95 (3.63)	4.42 (3.55)	5.98 (3.81)	8.03 (4.12)	6.24 (3.94)	38.80**	0.03	210.47**	0.17	862.87**	0.30

Note: *p < .006, **p < .001, *** Art and Sport combined; Means and SD represent scores before age was regressed out. Higher SA for females in comparison with males in the non-STEM-selected sample for some of the measures might be explained by effects of age: females were slightly older than males in this sample (see Table 1) and age was correlated with SA. For the main analysis, presented in Section 3.3., age was regressed out, resulting male advantage for all measures except for 2D to 3D drawing for which no significant gender differences were found.

samples was uneven (See Table 1). Therefore, we regressed out variance in SA tests scores, associated with gender. We also regressed out age to control for differences between the samples (See Table 1); and between age-SA correlation ($r = [0.12–0.20]$, $ps < 0.01$).

Further, we explored the hypothesis regarding differences in SA between individuals with various expertise. MANOVA on standardised residuals was significant (*Pillai's Trace* = 0.484, F (30, 6045) = 29.90, $p < .001$, $\eta_p^2 = 0.16$). Post-hoc comparisons with Bonferroni correction showed that STEM experts significantly differed from Arts and Sports; and the unselected sample (See Table A.1 and A.2 in Appendix).

We then conducted ten univariate ANOVAs to see patterns of results by expertise groups for specific SA facets. All ten models were significant at 0.006 (corrected for multiple comparisons; see Table SA1). We also conducted a non-parametric analysis to confirm the results using the Kruskal-Wallis H test, as Levene's test was significant for all facets (F varied from 2.79 to 62.91, all $ps < .05$) and the data for all facets violated normality assumption. The differences were significant, with χ^2 varying from 138.35 to 452.83, $p < .001$, for all comparisons.

Post hoc Bonferroni comparisons showed significant differences between the samples, with η_p^2 varying from 0.08 to 0.26. For detailed information, see Table SA1 and Fig. 1. Overall, STEM experts outperformed all other samples on all SA tasks. The Arts experts performed similar to the unselected sample, while Sports experts showed lower scores on almost all tasks than other samples. Therefore, the hypothesis that expertise is associated with SA level was supported only for STEM expertise. For further analyses, we combined two expert groups (Arts and Sports) into one expert sample – non-STEM-experts. This also made samples more balanced in terms of gender distribution (See Table 1).

3.2. Gender differences

Further, we aimed to investigate gender differences across samples. We explored gender differences in non-STEM-experts (combined Arts-and-Sports sample), STEM-experts, and unselected samples in performance on ten SA tasks in three (Sample) by two (Gender) MANOVA, using standardised residuals with age regressed out.

Main effect of the sample was again significant (*Pillai's Trace* = 0.344, F (20, 3990) = 41.48, $p < .001$, $\eta_p^2 = 0.172$). Univariate analyses with Post Hoc Bonferroni correction showed that STEM-experts outperformed two other samples on all tasks, irrespective of gender. The combined Arts-and-Sports sample scored lower than STEM and unselected samples on all tasks (see Table 3; See Table A.3 for raw means and SDs for males and females in Arts and Sports samples).

Main effect of gender was significant (Pillai's trace = 0.807, F (10.00, 1991.00) = 832.16, $p < .001$, $\eta_p^2 = 0.80$). Univariate analyses revealed significant gender differences in all tasks with moderate-to-strong male advantage (η_p^2 vary from 0.17 to 0.67; see Table 3 for

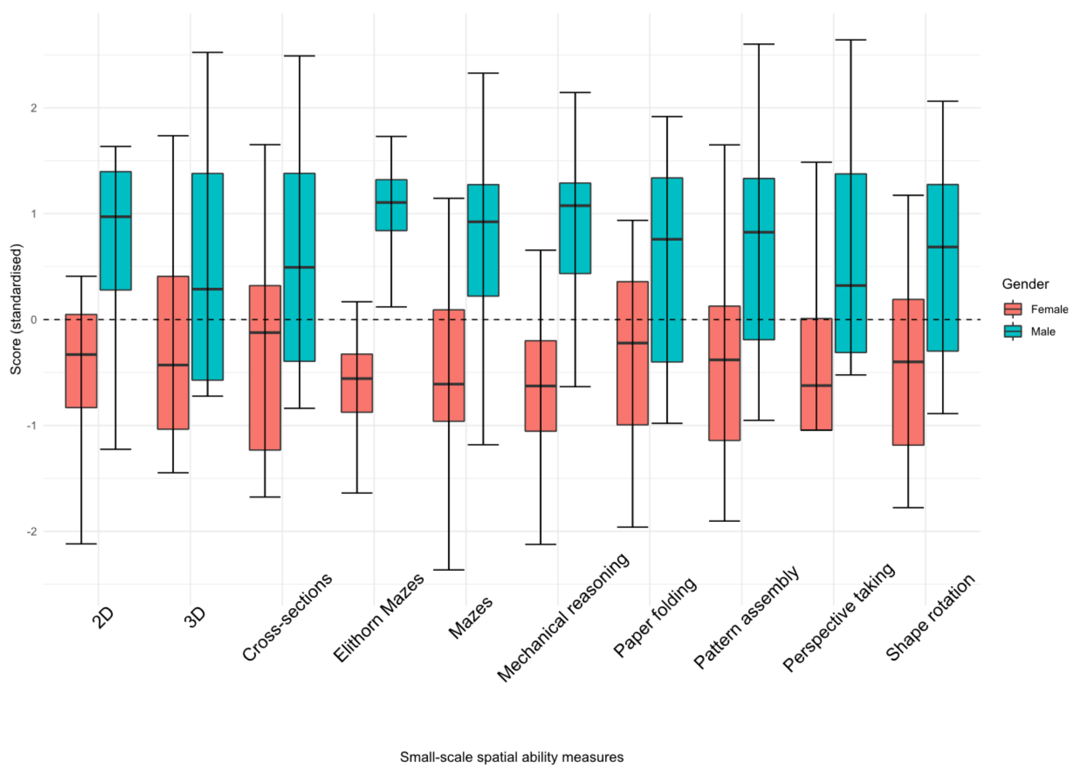

Fig. 2. Standardised residuals of correct answers on SA tasks in Genders (age differences are regressed out).

raw means and Fig. 2 for standardised residuals).

3.3. Sample by gender interaction

Interaction between sample and gender factors was also significant (*Pillai's trace* = 0.117, F (20, 3984) = 12.39, $p < .001$, η_p^2 = 0.059). Univariate analyses revealed significant interaction in all tasks, with η_p^2 varying from 0.01 to 0.07.

We conducted three additional MANOVAs to explore gender differences in each sample. We found significant differences between males and females in the STEM-selected sample (*Pillai's trace* = 0.908, F (10, 584) = 579.42, $p < .001$, $\eta_p^2 = 0.90$). Univariate analyses revealed significant differences in all tasks, with η_p^2 varying from 0.27 to 0.87, suggesting strong male advantage.

A similar pattern was found in the Unselected sample (*Pillai's trace* = 0.819, F (10, 737) = 332.96, $p < .001$, $\eta_p^2 = 0.81$). Males outperformed females on all tasks, with η_p^2 varying from 0.27 to 0.72.

Similar differences of a bit lower magnitude were found in the non-STEM expert sample consisting of artists and athletes (Pillai's trace = 0.747, F (10, 652) = 192.27, $p < .001$, $\eta_p^2 = 0.74$). Despite female advantage in raw means (that can be attributed to age differences), strong male advantage was found in almost all tests, with η_p^2 varying from 0.06 to 0.67 after controlling for age. The only exception was the 3D drawing task, in which the difference was nonsignificant. See Table 3 for raw means and Fig. 3 for standardised residuals.

4. Discussion

The present paper aimed to explore links between spatial ability, gender, and expertise (in STEM, arts, or sports). First, we examined whether adolescents who are experts in these areas scored higher in SA than their unselected peers. The effect of expertise was only observed for STEM, replicating a wealth of previous findings of positive associations between SA and STEM [4,5,12,27,51]. The advantage in SA demonstrated by STEM experts further supports adding SA assessment to maths and verbal tests and achievement measures used for talent identification [4,5,51–53].

The absence of effect expertise on SA in Arts and Sports might be explained by the use of only small-scale SA tasks in the present study [31]. These tasks might not capture the potentially superior ability of artists and athletes to process visuospatial information in a natural environment. For example, several studies have shown that visual attention is higher in athletes, including hockey players, compared with non-athletes [22,54,55]. Moreover, some research suggests that athletes can demonstrate superior performance in

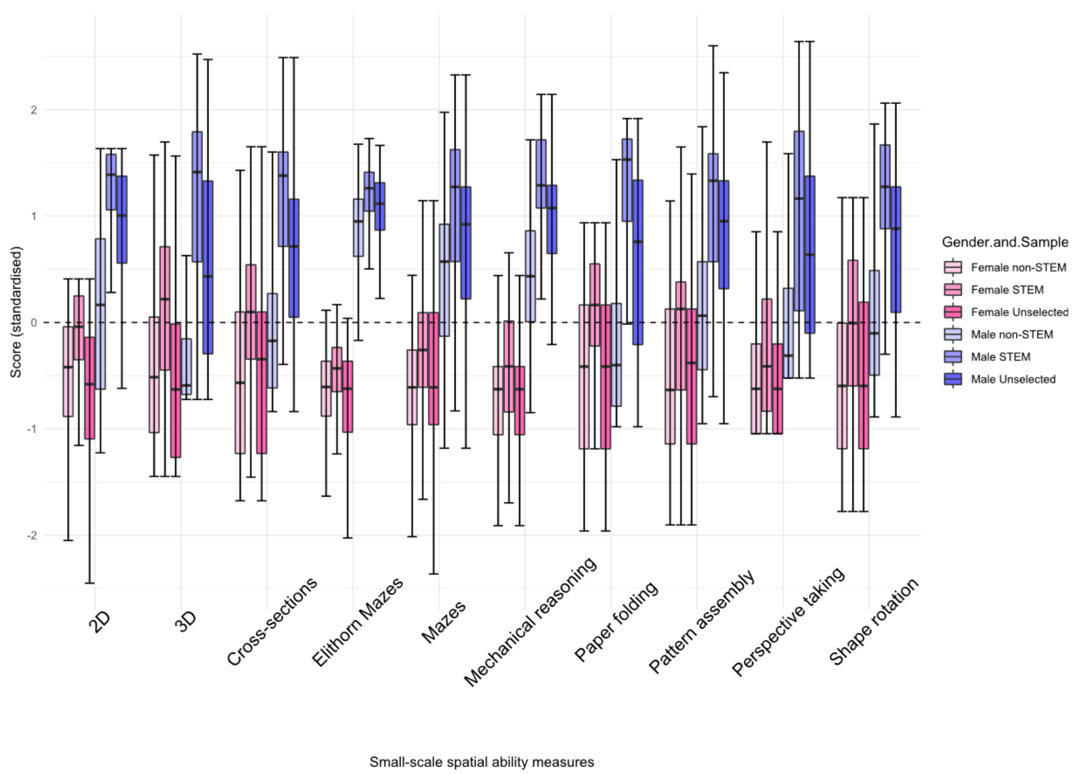

Fig. 3. Standardised residuals of correct answers on SA tasks in samples and genders (age differences are regressed out).

small-scale SA, but only in a task that is directly relevant to their sport. For example, a recent study [56] found that basketball experts outperformed novice players (ηp2 = 0.08) on a mental rotation task that has been modified to be directly relevant to basketball. Similarly, the highest visuo-spatial working memory performance was evident in chess experts when stimuli were chess-specific [57]. These findings are consistent with *the investment of effort theory* [58] which assumes that direct investment in training of relevant abilities might inhibit abilities in other domains [59].

It is also possible that specific abilities are related to different types of sports and arts activities. For example, large-scale SA (e.g. spatial orientation, distance estimation [31]) might be more important for team sports in which athletes often need to process spatial information regarding positions of partners, opponents, or important field landmarks. In addition, sports performance may enhance visuospatial working memory. For example, one study [60] shown that dancers and athletes with high expertise (>10 years of experience) have larger visuospatial working memory capacity than non-experts (<10 years of experience), with moderate effect size (η_p^2 = 0.08–0.15). The effect of expertise on working memory was also found for artistic domains. For example, it was shown that musicians have larger working memory capacity, including for visual information, compared to nonmusicians [61]. Another study found higher visuo-spatial working memory performance in adolescents with expertise in arts compared with athletes [62]. In the current study, the Sports and Arts groups were heterogeneous: figure skaters, hockey, and chess players for Sports; and ballet dancers, musicians, and fine arts performers for the Arts sample. Further fine-grained research is needed to identify specific aspects of SA associated with specific domains of expertise.

Beyond specificity of SA with regard to expertise, other factors might have contributed to the absence of main effect of Sport and Arts expertise on SA in our study. For example, motivation to complete computerised cognitive tests could be lower in Arts and Sports experts compared to STEM experts, who also have more practice with such tasks. Further, reliability of the measures was suggested to contribute to differences in performance between samples [63]. Thus, computerised tests might have higher reliability in STEM experts and lower reliability in Sports and Arts experts. However, we found no evidence of drop in reliability for the Art and Sports samples (see Table 2). In addition, inconsistencies in the literature regarding reported effects of expertise might be linked with age effects and/or "dosage" of the training. For example, the effect of some activity, e.g. sport training or learning a second language, may require longer time to grow in strength and/or may manifest in adulthood [24,64].

In our study, males outperformed females in all SA tasks with moderate effects after controlling for age. The strongest male advantage was found in Mechanical reasoning and Elithorn mazes, which replicates a previous study with the same battery where these two tests also demonstrated the largest gender differences [27]. Elithorn mazes had the shortest presentation time, but was found to be the easiest task in the battery [48]; and showed on average lower correlations with other spatial ability tests [65]. Moreover, Elithorn mazes and Mechanical reasoning demonstrated somewhat lower reliability in comparison with other tests [47]. Further research is needed to explore whether these or other factors may contribute to greater gender differences found for these tasks.

Gender differences were present for all expert samples and the unselected sample. Our findings are in line with previous studies that consistently replicate male advantage in SA (see e.g. Refs. [29,34]). Multiple hypotheses regarding the origins of observed gender differences have been explored, including evolutionary hunter-gatherer theory [66]; differences in brain structure and function [32]; hormonal differences [29,30,67]; differences in cognitive strategies used to solve spatial tasks [32]; differences in experience with spatial games and toys [29,67]; cultural settings and social stereotypes [67–69]; and differences in neuroticism that may be particularly influential for timed tests [29,40]. In addition, lack of interest in STEM [70] and computer science [71] in females might contribute to the obtained gender gap in results, as the tests were completed on a computer. However, this does not explain observed gender differences in the expert STEM group, who are proficient computer users. In our study, we expected that gender differences in SA would be moderated by expertise. Specifically, we expected smaller gender differences in STEM experts in comparison to the unselected sample, as at least one previous study suggested a reduced gender gap in SA after training [36]. However, our study demonstrated that gender differences in SA persist at the expert level. Although, STEM experts demonstrated on average higher SA, intensive and prolonged practice in STEM activities did not narrow the gap between males and females in SA. In fact, even greater male advantage was found in the sample of STEM ($\eta_p^2 = 0.90$) and non-STEM experts ($\eta_p^2 = 0.81$), compared with the unselected sample ($\eta_p^2 = 0.74$). Moreover, females with expertise in STEM performed worse on average compared with unselected males (*Cohen's d* = [0.55–1.43], all $ps < .001$).

The results are consistent with the view that high SA is a "relevant skill" for STEM achievement and expertise [12], but is not necessary or sufficient for such achievement. Our results are also consistent with the *default hypothesis* [12], according to which people avoid specific areas because they have weaknesses or perceived weaknesses in a specific ability [72]. This hypothesis has been put forward as one of the explanations for under-representation of females in STEM fields [73]. These processes may also partly explain lower proportion of females in STEM group in our study (38%) and lower SA in Arts and Sports groups.

Interestingly, the differences between males and females were found for all tests in non-STEM experts except 3D drawing task. It is possible that this SA facet is particularly trained in fine arts experts. In our study, most experts in fine arts (almost 80%) were females.

4.1. Limitations

The current research addressed several limitations of previous studies. Specifically, we explored large samples of adolescents with different expertise and a large control sample. However, our expert samples were heterogeneous, which precluded more fine-grained analysis. While we used a comprehensive battery of 10 S A tests, this battery did not include any measures of large-scale SA and visual attention. In addition, the gender disproportion of Arts and Sports samples did not allow us to investigate gender differences in these two samples separately. It is also worth noting that our groups differed in age range and average age, and therefore the results on raw means (showing female advantage in some tests) could not be meaningfully interpreted. Once the age was controlled for, this advantage disappeared. Moreover, the samples of the current study consisted of adolescents only. As age was shown to moderate the relationship between gender and spatial ability [32], further research is needed to replicate current results in younger children and adults.

4.2. Conclusions

Overall, findings support previously established links between spatial ability and STEM-related expertise. In contrast, such links were not found for expertise in arts and sports. Consistent with previous research, we found gender differences in SA for all samples, including STEM experts, after controlling for age differences. Future research needs to reconcile the studies that show effects of brief interventions (videogames, sports activities, etc.) on SA and studies that suggest negligible impact of prolonged intensive training.

Author contribution statement

Elina S. Tsigeman: Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper. Maxim V. Likhanov: Conceived and designed the experiments; Analyzed and interpreted the data; Wrote the paper. Anna V. Budakova; Aydar Akmalov; Ildar Sabitov: Performed the experiments; Contributed reagents, materials, analysis tools or data; Wrote the paper. Evgeniia Alenina; Ksenia Bartseva: Performed the experiments; Wrote the paper. Yulia Kovas: Conceived and designed the experiments; Analyzed and interpreted the data; Wrote the paper.

Data availability statement

The authors do not have permission to share data.

Additional information

Supplementary content related to this article has been published online at [URL].

Declaration of competing interest

All authors declare no conflict of interest.

Funding information

This study was supported by the Tomsk State University Development Programme (Priority 2030).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e15247.

References

[1] K. Rimfeld, N.G. Shakeshaft, M. Malanchini, M. Rodic, S. Selzam, K. Schofield, et al., Phenotypic and genetic evidence for a unifactorial structure of spatial abilities, Proc. Natl. Acad. Sci. USA 114 (2017) 2777–2782.
[2] M. Hegarty, D.R. Montello, A.E. Richardson, T. Ishikawa, K. Lovelace, Spatial abilities at different scales: individual differences in aptitude-test performance and spatial-layout learning, Intelligence 34 (2006) 151–176.
[3] K. Atit, J.R. Power, T. Pigott, J. Lee, E.A. Geer, D.H. Uttal, et al., Examining the relations between spatial skills and mathematical performance: a meta-analysis, Psychon. Bull. Rev. 29 (2021) 699–720.
[4] H.J. Kell, D. Lubinski, C.P. Benbow, J.H. Steiger, Creativity and technical innovation: spatial ability's unique role, Psychol. Sci. 24 (2013) 1831–1836.
[5] D.L. Shea, D. Lubinski, C.P. Benbow, Importance of assessing spatial ability in intellectually talented young adolescents: a 20-year longitudinal study, J. Educ. Psychol. 93 (2001) 604–614.
[6] L. Hetland, Learning to make music enhances spatial reasoning, J. Aesthetic. Educ. 34 (2000) 179.
[7] S. Ozel, J. Larue, C. Molinaro, Relation between sport and spatial imagery: comparison of three groups of participants, J. Psychol. 138 (2004) 49–64.
[8] S. Pietsch, P. Jansen, Different mental rotation performance in students of music, sport and education, Learn. Indiv Differ 22 (2012) 159–163.
[9] M.J. King, D.P. Katz, L.A. Thompson, B.N. Macnamara, Genetic and environmental influences on spatial reasoning: a meta-analysis of twin studies, Intelligence 73 (2019) 65–77.
[10] M. Malanchini, K. Rimfeld, N.G. Shakeshaft, A. McMillan, K.L. Schofield, M. Rodic, et al., Evidence for a unitary structure of spatial cognition beyond general intelligence, Npj Sci Learn 5 (2020) 9.
[11] M.G. Tosto, K.B. Hanscombe, C.M.A. Haworth, O.S.P. Davis, S.A. Petrill, P.S. Dale, et al., Why do spatial abilities predict mathematical performance? Dev. Sci. 17 (2014) 462–470.
[12] E. Winner, M.B. Casey, D. Dasilva, R. Hayes, Spatial abilities and reading deficits in visual art students, Empir. Stud. Arts 9 (1991) 51–63.
[13] D.G. Betsworth, J. Thomas Bouchard Jr., C.R. Cooper, H.D. Grotevant, J.-I.C. Hansen, S. Scarr, et al., Genetic and environmental influences on vocational interests assessed using adoptive and biological families and twins reared apart and together, J. Vocat. Behav. 44 (1994) 263–278.
[14] T. Lowrie, T. Logan, D. Harris, M. Hegarty, The impact of an intervention program on students' spatial reasoning: student engagement through mathematics-enhanced learning activities, Cogn. Res. Princ. Implic. 3 (2018) 50.

[15] D.H. Uttal, N.G. Meadow, E. Tipton, L.L. Hand, A.R. Alden, C. Warren, et al., The malleability of spatial skills: a meta-analysis of training studies, Psychol. Bull. 139 (2013) 352–402.
[16] D. Moreau, J. Clerc, A. Mansy-Dannay, A. Guerrien, Enhancing spatial ability through sport practice: evidence for an effect of motor training on mental rotation performance, J. Indiv. Differ. 33 (2012) 83–88.
[17] P. Jansen, J. Kellner, C. Rieder, The improvement of mental rotation performance in second graders after creative dance training, Creativ. Educ. 4 (2013) 418–422.
[18] D.W. Chan, Gender differences in spatial ability: Relationship to spatial experience among Chinese gifted students in Hong Kong, Roeper Rev. 29 (2007) 277–282.
[19] G. Sala, F. Gobet, Does far transfer exist? Negative evidence from chess, music, and working memory training, Curr. Dir. Psychol. Sci. 26 (2017) 515–520.
[20] K.A. Ericsson, R.T. Krampe, C. Tesch-Römer, The role of deliberate practice in the acquisition of expert performance, Psychol. Rev. 100 (1993) 363–406.
[21] O. Blazhenkova, M. Kozhevnikov, Visual-object ability: a new dimension of non-verbal intelligence, Cognition 117 (2010) 276–301.
[22] A. Notarnicola, G. Maccagnano, V. Pesce, S. Tafuri, G. Novielli, B. Moretti, Visual- spatial capacity: gender and sport differences in young volleyball and tennis athletes and non-athletes, BMC Res. Notes 7 (2014) 57.
[23] S. Ozel, J. Larue, C. Molinaro, Relation between sport activity and mental rotation: comparison of three groups of subjects, Percept. Mot. Skills 95 (2002) 1141–1154.
[24] S. Stoyanova, N. Ivantchev, K. Petrova, Spatial orientation in sportsmen, Eur Sci J ESJ 12 (2016) 88.
[25] P. Jansen, J. Ellinger, J. Lehmann, Increased physical education at school improves the visual-spatial cognition during adolescence, Educ. Psychol. 38 (2018) 964–976.
[26] H. Heppe, A. Kohler, M.-T. Fleddermann, K. Zentgraf, The relationship between expertise in sports, visuospatial, and basic cognitive skills, Front. Psychol. 7 (2016).
[27] E.A. Esipenko, E.P. Maslennikova, A.V. Budakova, K.R. Sharafieva, V.I. Ismatullina, I.V. Feklicheva, et al., Comparing spatial ability of male and female students completing humanities vs. technical degrees, Psychol. Russ. State Art 11 (2018) 37–49.
[28] C.M. Ganley, M. Vasilyeva, Sex differences in the relation between math performance, spatial skills, and attitudes, J. Appl. Dev. Psychol. 32 (2011) 235–242.
[29] Y. Maeda, S.Y. Yoon, A meta-analysis on gender differences in mental rotation ability measured by the purdue spatial visualization tests: visualization of rotations (PSVT:R), Educ. Psychol. Rev. 25 (2013) 69–94.
[30] T. Toivainen, G. Pannini, K.A. Papageorgiou, M. Malanchini, K. Rimfeld, N. Shakeshaft, et al., Prenatal testosterone does not explain sex differences in spatial ability, Sci. Rep. 8 (2018), 13653.
[31] L. Wang, A.S. Cohen, M. Carr, Spatial ability at two scales of representation: a meta-analysis, Learn. Indiv Differ 36 (2014) 140–144.
[32] L. Yuan, F. Kong, Y. Luo, S. Zeng, J. Lan, X. You, Gender differences in large-scale and small-scale spatial ability: a systematic review based on behavioral and neuroimaging research, Front. Behav. Neurosci. 13 (2019) 128.
[33] A. Nazareth, X. Huang, D. Voyer, N. Newcombe, A meta-analysis of sex differences in human navigation skills, Psychon. Bull. Rev. 26 (2019) 1503–1528.
[34] D. Voyer, S. Voyer, M.P. Bryden, Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables, Psychol. Bull. 117 (1995) 250–270.
[35] M.C. Linn, A.C. Petersen, Emergence and characterization of sex differences in spatial ability: a meta-analysis, Child Dev. 56 (1985) 1479–1498.
[36] I.D. Cherney, K. Bersted, J. Smetter, Training spatial skills in men and women, Percept. Mot. Skills 119 (2014) 82–99.
[37] J.A. González-Calero, R. Cózar, R. Villena, J.M. Merino, The development of mental rotation abilities through robotics-based instruction: an experience mediated by gender, Br. J. Educ. Technol. 50 (2019) 3198–3213.
[38] M.V. Likhanov, E.S. Tsigeman, K.A. Papageorgiou, A.F. Akmalov, I.A. Sabitov, Y.V. Kovas, Ordinary extraordinary: elusive group differences in personality and psychological difficulties between STEM-gifted adolescents and their peers, Br. J. Educ. Psychol. 91 (2021) 78–100.
[39] N. Schmitt, The interaction of neuroticism and gender and its impact on self-efficacy and performance, Hum. Perform. 21 (2007) 49–61.
[40] K.A. Byrne, C.D. Silasi-Mansat, D.A. Worthy, Who chokes under pressure? The Big Five personality traits and decision-making under pressure, Pers. Indiv. Differ. 74 (2015) 22–28.
[41] C.M. Ohannessian, S. Milan, A. Vannucci, Gender differences in anxiety trajectories from middle to late adolescence, J. Youth Adolesc. 46 (2017) 826–839.
[42] H. Frenken, K.A. Papageorgiou, T. Tikhomirova, S. Malykh, M.G. Tosto, Y. Kovas, Siblings' sex is linked to mental rotation performance in males but not females, Intelligence 55 (2016) 38–43.
[43] M. Baker, K. Cornelson, Title IX and the spatial content of female employment—out of the lab and into the labor market, Lab. Econ. 58 (2016) 128–144.
[44] S.A. Berenbaum, K.L.K. Bryk, A.M. Beltz, Early androgen effects on spatial and mechanical abilities: evidence from congenital adrenal hyperplasia, Behav. Neurosci. 126 (2012) 86–96.
[45] E. Hampson, J.F. Rovet, Spatial function in adolescents and young adults with congenital adrenal hyperplasia: clinical phenotype and implications for the androgen hypothesis, Psychoneuroendocrinology 54 (2015) 60–70.
[46] M.L. Collaer, M. Hines, No evidence for enhancement of spatial ability with elevated prenatal androgen exposure in congenital adrenal hyperplasia: a meta-analysis, Arch. Sex. Behav. 49 (2020) 395–411.
[47] M.V. Likhanov, V.I. Ismatullina, A.Y. Fenin, W. Wei, K. Rimfeld, E.P. Maslennikova, et al., The factorial structure of spatial abilities in Russian and Chinese students, Psychol. Russ. State Art 11 (2018) 96–114.
[48] A.V. Budakova, M.V. Likhanov, A.V. Zhurbitskiy, E.O. Sitnikova, E.M. Bezrukovaa, Y.V. Kovas, Measuring spatial ability for talent identification, educational assessment, and support: evidence from adolescents with high achievement in science, arts, and sports, Psychol. Russ. State Art 14 (2021) 59–85.
[49] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2017. Available: https://www.R-project.org/.
[50] Darren George, Mallery Paul, SPSS for Windows Step by Step : a Simple Guide and Reference, 17.0 Update, tenth ed., Allyn & Bacon, Boston, 2010. Available: http://lib.ugent.be/catalog/rug01:001424067.
[51] J. Wai, D. Lubinski, C.P. Benbow, Spatial ability for STEM domains: aligning over 50 years of cumulative psychological knowledge solidifies its importance, J. Educ. Psychol. 101 (2009) 817–835.
[52] L. Andersen, Visual–spatial ability: important in STEM, ignored in gifted education, Roeper Rev. 36 (2014) 114–121.
[53] J.M. Lakin, J. Wai, Spatially gifted, academically inconvenienced: spatially talented students experience less academic engagement and more behavioural issues than other talented students, Br. J. Educ. Psychol. 90 (2020) 1015–1038.
[54] K. Clark, M. Maddocks, Enhanced visual attention in university hockey players, J. Vis. 18 (2018) 1120.
[55] J.T. Enns, J.C. Richards, Visual attentional orienting in developing hockey players, J. Exp. Child Psychol. 64 (1997) 255–275.
[56] M. Weigelt, D. Memmert, The mental rotation ability of expert basketball players: identifying on-court plays, Res. Q. Exerc. Sport 92 (2021) 137–145.
[57] E.T. Smith, J.C. Bartlett, D.C. Krawczyk, C. Basak, Are the advantages of chess expertise on visuo-spatial working-memory capacity domain specific or domain general? Mem. Cognit. 49 (2021) 1600–1616.
[58] R.B. Cattell, Intelligence: its Structure, Growth and Action, Elsevier, 1987.
[59] T.R. Coyle, Sex differences in tech tilt: support for investment theories, Intelligence 80 (2020), 101437.
[60] E.M. Barhorst-Cates, Spatial working memory is enhanced for movement experts in traditional and embodied tasks, Spatial Cognit. Comput. 19 (2019) 69–91.
[61] F. Talamini, B. Carretti, M. Grassi, The working memory of musicians and nonmusicians, Music Percept. 34 (2016) 183–191.
[62] E. Tsigeman, S. Silas, K. Frieler, M. Likhanov, R. Gelding, Y. Kovas, et al., The Jack and Jill adaptive working memory task: construction, calibration and validation, PLoS One 17 (2022), e0262200.
[63] N.S. Newcombe, M. Hegarty, D. Uttal, Building a cognitive science of human variation: individual differences in spatial navigation, Top. Cogn. Sci. 15 (2023) 6–14.

[64] M. Likhanov, O. Bogdanova, E. Alenina, T. Kolienko, Y. Kovas, No evidence of a positive effect of learning Chinese language as an L2 on spatial ability, Sci. Rep. 13 (2023) 1262.
[65] M. Likhanov, E. Maslennikova, G. Costantini, A. Budakova, E. Esipenko, V. Ismatullina, et al., This is the way: network perspective on targets for spatial ability development programmes, Br. J. Educ. Psychol. 92 (2022) 1597–1620.
[66] I. Silverman, J. Choi, M. Peters, The hunter-gatherer theory of sex differences in spatial abilities: data from 40 countries, Arch. Sex. Behav. 36 (2007) 261–268.
[67] H.B. Yılmaz, On the development and measurement of spatial ability, Electronic Journal of Elementary Education 1 (2017) 83–96.
[68] M. Hoffman, U. Gneezy, J.A. List, Nurture affects gender differences in spatial abilities, Proc. Natl. Acad. Sci. USA 108 (2011) 14786–14788.
[69] M.R. Tarampi, N. Heydari, M. Hegarty, A tale of two types of perspective taking: sex differences in spatial ability, Psychol. Sci. 27 (2016) 1507–1516.
[70] P.M. Sadler, G. Sonnert, Z. Hazari, R. Tai, Stability and volatility of STEM career interest in high school: a gender study, Sci. Educ. 96 (2012) 411–427.
[71] S. Beyer, Why are women underrepresented in Computer Science? Gender differences in stereotypes, self-efficacy, values, and interests and predictors of future CS course-taking and grades, Comput. Sci. Educ. 24 (2014) 153–192.
[72] G. Stoet, D.C. Geary, The gender-equality paradox in science, technology, engineering, and mathematics education, Psychol. Sci. 29 (2018) 581–593.
[73] H. Nimmesgern, Why are women underrepresented in STEM fields? Chem. Eur J. 22 (2016) 3529–3530.