Published in final edited form as:
Dev Sci. 2023 March ; 26(2): e13302. doi:10.1111/desc.13302.

Tracing the origins of the STEM gender gap: The contribution of childhood spatial skills

Jing Tian[1], **Kexin Ren**[1], **Nora S. Newcombe**[1], **Marsha Weinraub**[1], **Deborah Lowe Vandell**[2], **Elizabeth A. Gunderson**[1]

[1]Department of Psychology and Neuroscience, Temple University, Philadelphia, Pennsylvania, USA

[2]School of Education, University of California, Irvine, California, USA

Abstract

Despite some gains, women continue to be underrepresented in many science, technology, engineering, and math (STEM) fields. Using a national longitudinal dataset of 690 participants born in 1991, we tested whether spatial skills, measured in middle childhood, would help explain this gender gap. We modeled the relation between 4th-grade spatial skills and STEM majors while simultaneously accounting for competing cognitive and motivational mechanisms. Strong spatial skills in 4th grade directly increased the likelihood of choosing STEM college majors, above and beyond math achievement and motivation, verbal achievement and motivation, and family background. Additionally, 4th-grade spatial skills indirectly predicted STEM major choice via math achievement and motivation in the intervening years. Further, our findings suggest that gender differences in 4th-grade spatial skills contribute to women's underrepresentation in STEM majors.

Keywords

motivation; achievement; gender differences; spatial skills; STEM major

1 | INTRODUCTION

Despite the tremendous progress women have made in education and the workforce over the past 50 years, men continue to dominate science, technology, engineering, and mathematics (STEM) domains (Martinez & Christnacht, 2021; U.S. Department of Education, National Center for Education Statistics, 2020). From 1970 to 2019, the proportion of women in the U.S. workforce increased from 38% to 48%, yet women's representation in STEM fields increased from 8% to only 27% in the same time period (Martinez & Christnacht, 2021).

Correspondence: Jing Tian and Elizabeth A. Gunderson, Department of Psychology and Neuroscience, Temple University; Philadelphia, PA, 19122, USA. jing.tian@temple.edu and liz.gunderson@temple.edu.

CONFLICT OF INTEREST
The authors declare no conflict of interest.

SUPPORTING INFORMATION
Additional supporting information can be found online in the Supporting Information section at the end of this article.

Women are similarly underrepresented at earlier points in the STEM pipeline: In the class of 2018–2019, women earned only 36% of STEM bachelor's degrees (U.S. Department of Education, National Center for Education Statistics, 2020). Why do fewer women than men pursue STEM fields in college and beyond?

Considerable efforts have been made to identify cognitive, motivational, sociocultural, and environmental factors contributing to women's lower participation in STEM college majors (Wang & Degol, 2017). Research on cognitive factors has traditionally centered around mathematics ability in adolescence. Strong mathematical competence is fundamental for STEM achievement, and high-school students with higher math achievement are more likely to choose STEM college majors (Guo et al., 2015). However, though boys have long been believed to have stronger math ability than girls, an analysis of U.S. state assessments showed no gender difference in math performance among boys and girls in grades 2 through 11 (Hyde et al., 2008). Further, although some have hypothesized that the STEM major gender gap could be explained by an overrepresentation of males in the upper tail of math ability, meta-analyses revealed minimal differences in the proportions of males and females in the top tail (Hyde et al., 2008; Lindberg et al., 2010). Taken together, recent research suggests that females have largely reached parity with males in math ability, and math ability cannot explain the STEM gender gap.

One cognitive factor that is often overlooked when explaining the STEM gender gap is spatial skills. Spatial skills refer to abilities involved in mentally manipulating objects in space and reasoning about spatial relationships, which are critical for success in many STEM disciplines. For example, in chemistry, determining if a molecule is chiral requires spatial skills to examine whether the mirror image of the molecule can be superimposed on itself. In astronomy, understanding the moon's phases requires spatial skills to visualize the movements of the earth, the moon, and the sun. Spatial skills correlate with achievement in various STEM disciplines, such as math (Atit et al., 2021), chemistry (Wu & Shah, 2004), and engineering (Sorby, 2009). Longitudinally, adolescents with stronger spatial skills are more likely to subsequently earn STEM degrees and work in STEM fields, even after considering the influence of math ability (Shea et al., 2001; Wai et al., 2009; Webb et al., 2007).

Critically, males outperform females on many spatial tasks, with the most profound difference in mental rotation (Lauer et al., 2019; Voyer et al., 1995). A recent meta-analysis showed that small male advantage in mental rotation emerged by 6 years of age, and the advantage increased with age through at least early adulthood (Lauer et al., 2019). Further, in the few specific math areas where a male advantage exists, including number line estimation and the Mathematics Scholastic Aptitude Test (SAT-M), spatial skills appear to contribute to such gender differences (Casey et al., 1997; Geary et al., 2020; Tian et al., 2022). Spatial skills also mediate the gender difference in science achievement among middle school students (Ganley et al., 2014). Therefore, unlike math achievement, spatial skills can plausibly explain the gender gap in STEM majors.

Here, we aim to trace the origins of the gender gap in STEM college majors to spatial skills in middle childhood, a much earlier age than previous research. In addition to spatial skills, we simultaneously examined multiple cognitive and motivational factors that have been shown to influence STEM participation. Doing so allowed us to better isolate the relation between spatial skills and STEM major choice.

One predictor we included was math achievement. Math achievement robustly predicts entering STEM majors within gender (Guo et al., 2015). Further, math achievement is closely related to spatial skills (Atit et al., 2021), making it a potential mediator between spatial skills and STEM major choices.

We also included a motivational predictor, math ability self-concept. Math ability self-concept is the perception of one's capability to succeed in math and predicts STEM major choice above and beyond math achievement (Guo et al., 2015). Males often report higher math ability self-concepts than females while controlling for math achievement (Skaalvik & Skaalvik, 2004). Therefore, math ability self-concepts may help explain the STEM gender gap.

Further, we included verbal achievement and verbal ability self-concept. This choice was motivated by a recent hypothesis for the STEM gender gap: Females' higher verbal ability, compared to their own math ability and to males' verbal ability, enhances females' verbal ability self-concept and steers them towards non-STEM fields (Parker et al., 2012). Consistent with this view, adolescents' verbal achievement and ability self-concept negatively predicted STEM major choices after accounting for math ability and ability self-concept (Guo et al., 2015).

We used data from the National Institute of Child Health and Human Development Study of Early Child Care and Youth Development (SECCYD)—a national longitudinal study of 1364 children born in 1991 across ten sites in the U.S. This dataset is well-suited for our purpose: It is a contemporary dataset that measured spatial skills in middle childhood and included cognitive and motivational measures from birth to age 26 years.

In the SECCYD, spatial skills were measured by the Block Design subtest of the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999). This test asks children to construct designs matching sample models or pictures using cubes of red, white, and half red and half white faces. Block design is often viewed as a relatively comprehensive measure of spatial skills because it requires various components of spatial ability, including disembedding and mental rotation. Moreover, the block design test is a reasonably good predictor of everyday spatial abilities (Groth-Marnat & Teal, 2000). Although an early meta-analysis of 15 studies reported no gender difference in block design performance (Voyer et al., 1995), a male advantage was present in standardization samples of the test covering age ranges both younger and older than the children in our study (Irwing, 2012; Jirout & Newcombe, 2015; Piffer, 2016). Specifically, a male advantage was found on the Block Design subtest in the 4 to 7-year-old standardization sample of the Wechsler Preschool and Primary Scale of Intelligence (WPPSI-IV; Wechsler, 2012) after controlling for cognitive abilities measured by other subtests (e.g., verbal comprehension, processing speed, fluid reasoning, and working memory; Jirout & Newcombe, 2015). In the U.S. standardization sample of Wechsler Adult Intelligence Scale-III and –IV (WAIS-III and WAIS-IV), males had significantly higher scores than females on the Block Design subtest (Irwing, 2012; Piffer, 2016). We, therefore, expected the 4th-grade boys in our sample to outperform girls on the Block Design test.

We used structural equation modeling (SEM) to capture the interconnected pathways among childhood spatial skills, math achievement, math ability self-concept, verbal achievement, verbal ability self-concept, and STEM college major choice while controlling for demographic characteristics. Access to the SECCYD data used in the present study is limited to qualified researchers. We pre-registered the analysis plan (https://osf.io/7n5du) and made the analysis code and model output publicly available (https://osf.io/n3y46/).

2 | METHOD

2.1 | Participants

The SECCYD began in 1991 with 1364 children and their families. Participants were assessed at several time points from 1991 through 2009 during SECCYD Phases I through IV and during two follow-up studies (i.e., Phases V and VI). The main outcome variable of interest in the current study is the major field of study in college, which was collected in Phase VI when the participants were 26 years old.

Phase VI included 814 participants from the initial sample. They were asked about their highest degree earned. Those who indicated that they had completed at least some college were asked to report their major field of study. Our final analytic sample included the 690 participants (383 females) who completed some college and reported their major field of study. Based on caregivers' report, among participants in our analytic sample, 83% were White, 8% were Black or African American, 3% were Hispanic, 1% were Asian or Pacific Islander, and 4% were of other race or ethnicity.

The analytic sample included fewer boys and more White children than the non-college sample (children not in our analytic sample but who participated in Phase VI; N = 124) and the drop-out sample (children in the initial sample but who dropped out before Phase VI; $N = 550$). Moreover, children in the analytic sample were from families of higher SES than the children in the non-college and drop-out samples as indicated by the mother's years of education and family income-to-needs ratio (Supplemental Materials [SM] Table S1).

2.2 | Measures

2.2.1 | Outcome

Major field of study.: When participants were 26 years old, they reported their major field of study by choosing among 35 choices or providing written responses on Qualtrics or paper questionnaires. We first categorized each of the 35 provided major fields of study and participants' written response as a non-STEM field (e.g., philosophy and religious studies, history, etc.), a non-math-intensive STEM filed (e.g., psychology, biological and biomedical sciences, etc.), or a math-intensive STEM field (e.g., physical sciences, engineering, etc.). This categorization was done based on a coding scheme (SM, Section A) developed using the STEM Designated Degree Program List compiled by the Department of Homeland

Security of the US (U.S. Department of Homeland Security, 2016), the American Freshman Survey: National Norms 2016 (Eagan et al., 2017), and prior research (Goldman & Hewitt, 1976). Participants were then categorized as a non-STEM major, a non-math-intensive STEM major, or a math-intensive STEM major based on their major field(s) of study. Participants reporting multiple fields of study were categorized based on their most math-intensive major (math intensiveness: math-intensive STEM > non-math-intensive STEM > non-STEM). This variable was treated as an ordinal categorical variable in the analyses.

2.2.2 | Predictors

Spatial skills.: Participants' spatial skills were assessed in 4th grade using the Block Design subtest of WASI (Wechsler, 1999) during a lab visit. Children were scored based on the correctness of the design completed within a limited time. The raw scores were then converted to T-scores (mean = 50, SD = 10), which we used in our analyses. Published internal consistency for children on WASI subtests ranges from 0.81 to 0.97 (Stano, 2004).

Math and verbal achievement.: Math and verbal achievement were assessed when participants were in 5th grade and at age 15 using the Applied Problems subtest and the Passage Comprehension subtest of the Woodcock-Johnson Psycho-Educational Battery-Revised (WJ-R; Woodcock & Johnson, 1989), respectively. Published internal consistency reliabilities range from 0.94 to 0.98 for each WJ-R subtest (Woodcock & Johnson, 1989). We used the W scores, which have equal interval units, of these two subtests in our main analyses.

Math and verbal ability self-concept.: Ability self-concept in math and verbal domains was measured in 6th grade and at age 15 using three questions per subject area. The questions were adapted from the Self and Task Perception Questionnaire (Jacobs et al., 2002). Participants were asked to respond on a seven-point Likert scale to questions such as, "How good at [math/reading] are you?" We used responses on these questions to construct a latent variable of ability self-concept within each domain at each time point (see Section 3.2 for details).

2.2.3 | Covariates—We included the following variables as covariates: gender of the participant, family income-to-needs ratio, and maternal education. Participants' gender was reported by caregivers when the participant was 1-month old. Female was coded as 0, and male was coded as 1.

The family income-to-needs ratio was calculated based on total family income, total number of household members, and the number of children living in the home measured in 1st and 3rd grades. This information was obtained via phone interviews with the participants' mothers. We used the mean value of the family income-to-needs ratios at the two-time points as a continuous variable in the analyses. In the case where family income-to-needs ratio was missing at either time point, data from the other timepoint was used in the analyses.

Maternal education was measured by mother's years of education (e.g., 12 = high school graduate or GED; 16 = undergraduate degree) when the participants were 1-month old during interviews at the participants' home. This variable was entered as a continuous variable in our analyses.

2.3 | Analytic approach

We used SEM to examine the pathways from spatial skills to STEM college major choice. Prior to conducting all SEM analyses, we scaled all continuous variables to make all variables on similar scales. Ability self-concept items were scaled by first subtracting 4 (mean score on the 1–7 scales) and then dividing by 3 (maximum score). Other continuous variables were scaled by subtracting the variable's mean and then dividing by the maximum value in our sample.

In the first step of model fitting, we fit the measurement model of math and verbal ability self-concept and tested levels of measurement invariance across the two-time points (i.e., 6th grade and age 15). The maximum likelihood robust estimator was used. The most-constrained measurement model that yielded model fit similar to that of the configural model was used in fitting subsequent structural models.

In the second step, we fit a model only including paths from achievement and ability self-concept in math and verbal domains at age 15 to STEM major (Model 1). In the third step, we fit a model by adding math and verbal achievement in 5th grade and math and verbal ability self-concept in 6th grade to Model 1 (Model 2). Finally, we included 4th-grade spatial skills (Model 3, the Full Model).[1] In all structural models, we set variables measured earlier to predict variables measured later and entered covariates as predictors of each variable. If the coefficient of a particular covariate in a particular path was not significant, that covariate was not included in the corresponding path in later steps.

Additionally, we ran multi-group SEM to test whether boys and girls showed similar pathways from spatial skills to STEM majors in the Full Model. As a preliminary step, we tested measurement invariance of the ability self-concept measurement model across the two gender groups. We then used the most constrained model that yielded a similar fit to the configural model in the subsequent multi-group SEM. To test for gender differences in the strength of specific pathways, we defined the parameters of interest in the multigroup SEM and tested whether those parameters differed from zero.

We used R (R Core Team, 2018) and MPlus (Muthén & Muthén, 1998–2017) for our analyses. In SEM analyses, we used the weighted least square mean and variance adjusted (WLMSV) estimator because our outcome variable, STEM major, was ordinal categorical. We used full information maximum likelihood method to handle missing data and a bias-corrected bootstrap of 500 iterations to obtain the 95% confidence intervals of model estimates. Confidence intervals excluding zeros were evaluated as indicating significant effects. We used root mean square error of approximation (RMSEA), comparative fit index (CFI), and Tucker-Lewis index (TLI) as indices of the goodness of model fit. These indices are commonly used in combination for evaluating model fit as they are sensitive and biased

[1]We ran additional exploratory analysis including math and science course-taking in 9th grade and high school. Details of these variables and results of this analysis are reported in the SM (see Figure S3 and Table S10 for model results).

to different aspects of model specification (Kenny & McCoach, 2003). For RMSEA, a value below 0.10 was considered "good" and below 0.05 was "very good" (Steiger & Lind, 1980). We considered a model with RMSEA value below 0.08 as of interest. For both TLI and CFI, we considered a value of 0.95 or above as indicating good fit (Hu & Bentler, 1999).

3 | RESULTS

3.1 | Descriptive analyses

Among the 690 participants in our analytic sample, 471 were categorized as non-STEM majors (286 females), 112 were categorized as non-math-intensive STEM majors (68 females), and 107 were categorized as math-intensive STEM majors (29 females). Consistent with prior research, a greater proportion of males than females chose STEM majors, specifically math-intensive STEM majors (Figure 1).

Among the 662 children who completed the spatial skills measure, boys, on average, outperformed girls (Cohen's $d = 0.28$, $p < 0.001$) (Figure 2). This effect size is comparable to the effect size of gender differences in mental rotation among children of similar ages reported in a meta-analysis (Lauer, Ilksoy et al., 2018) and in the US standardization samples of the WAIS-III and WAIS–IV on the Block Design subtest (Irwing, 2012; Piffer, 2016). Descriptively, in the current study, among children in the top quartile of spatial skills in 4th grade, 43% later entered a STEM major, compared to only 22% of those in the bottom quartile of childhood spatial skills (Figure 3).

Boys and girls did not differ in verbal achievement in 5th grade or at age 15 (ps ≥ 0.610). In contrast, boys had higher math achievement than girls in both 5th grade and at age 15 (ps < 0.001). See Tables S2 and S3 for descriptive statistics for all variables.

3.2 | Measurement model fitting and measurement invariance testing

The configural model of math and verbal ability self-concept in 6th grade and at age 15 had a RMSEA of 0.062, a CFI of 0.973, and a TLI of 0.950 (see Table S4 for the model results), all indicating good model fit. We then fit the metric model by constraining the factor loadings of corresponding items across the two-time points to be equal, which also had good model fit (RMSEA = 0.061, CFI = 0.971, and TLI = 0.952; see Table S5 for the model results). Chi-square difference test indicated that the metric model was significantly worse than the configural model ($p = .018$). The modification indices in the metric model suggested that releasing the factor loading constraint on the verbal item, "How good at reading are you?" across the time points would yield the largest decrease in model fit chi-square. We thus released this constraint and fit a partial metric model. This partial metric model had good model fit (RMSEA = 0.060, CFI = 0.972, and TLI = 0.953; see Table S6 for the model results) and was not significantly different from the configural model ($p = 0.104$). We used this partial metric model in subsequent SEM analyses.

3.3 | Full model fitting

Results of Models 1 and 2 are reported in the SM (Figures S1 and S2; Tables S7 and S8). The Full Model (Figure 4), which tested the pathways from 4th-grade spatial skills to STEM college majors while accounting for math and verbal achievement and ability self-concepts, had reasonably good fit (RMSEA = 0.038, CFI = 0.956, and TLI = 0.931; see Table S9 for the model results). In this model, spatial skills directly predicted STEM major choice (standardized path coefficient = 0.13, with 95% bootstrapped confidence interval (CI) = [0.02, 0.24]). Critically, the gender difference in 4th-grade spatial skills partially accounted for the gender gap in STEM major choice (standardized indirect effect from gender to spatial skills to STEM major choice = 0.02, with 95% bootstrapped CI [0.004, 0.047]; Table S9).

To understand the mechanisms relating childhood spatial skills to STEM major choice, we next examined the specific indirect pathways. Three specific indirect pathways from spatial skills to STEM major were significant (Table S9): via age-15 math ability self-concept (standardized effect = 0.023, 95% bootstrapped CI = [0.006, 0.055]); via 6th-grade math ability self-concept to age-15 math ability self-concept (standardized effect = 0.008, 95% bootstrapped CI = [0.001, 0.023]); and via 5th-grade math achievement, to 6th-grade math ability self-concept, to age-15 math ability self-concept (standardized effect = 0.009, 95% bootstrapped CI = [0.002, 0.021]). Notably, all three significant indirect paths involved age-15 math ability self-concepts, suggesting that strong early spatial skills elevate the likelihood of choosing STEM majors not only directly but also indirectly by enhancing math ability self-concept in the intervening years.

3.4 | Multi-group analyses: Gender

We used multi-group SEM to examine whether the pathways from spatial skills to STEM major differed among boys and girls. First, we tested measurement invariance of the measurement model used in the main analyses (the partial metric model) across the two gender groups. We constrained factor structure (configural), factor loadings (metric), factor intercepts (scalar), and residual variances (strict) one by one across the two groups (parameters constrained in earlier models were also constrained in later models). Chi-square difference tests suggested that the scalar model was not significantly different from the configural model, but the strict model was. Therefore, we used the scalar model in the multi-group SEM.

We then ran multi-group SEM on the Full Model. We defined three parameters of interest: the direct effect of 4th-grade spatial skills on STEM major choice, the indirect effect of 4th-grade spatial skills on STEM major choice via age-15 math ability self-concept, and the indirect effect of 4th-grade spatial skills on STEM major choice via 6th-grade and age-15 math ability self-concept. We obtained the bootstrapped confidence interval of the difference between the parameter estimate for females and the parameter estimate for males. None of these parameters were significantly different from zero (see Table S11 for parameter estimates). Therefore, whereas there was an initial gender difference in spatial skills, spatial skills predicted STEM major choice to a similar extent and via similar mechanisms among boys and girls.

4 | DISCUSSION

In the current study, we examined the long-term relation of childhood spatial skills to STEM college major choice and asked whether gender differences in childhood spatial skills help explain the gender gap in STEM majors. We simultaneously accounted for multiple cognitive and motivational mechanisms of STEM major choice in a comprehensive model. We found that spatial skills in 4th grade both directly and indirectly predicted STEM major choice. Importantly, girls' less strong spatial skills in 4th grade partially accounted for their underrepresentation in STEM majors.

The present findings demonstrated notable long-term relations between spatial skills and later STEM major choice. A few prior studies found that spatial skills predicted later STEM participation beyond math ability (Shea et al., 2001; Wai et al., 2009; Webb et al., 2007). However, children from these previous analyses were either intellectually precocious (top 0.5% for their age group in Shea et al., 2001 and top 3% for their age group in Webb et al., 2007) or much older when spatial skills were assessed (9th–12th graders in Wai et al., 2009) than children in the current study. Further, none of these prior studies accounted for motivational factors that have also been shown to predict STEM participation (Guo et al., 2015). With a more generalizable sample, we tested the influence of childhood spatial skills on STEM major choice while accounting for multiple cognitive and motivational factors. In this comprehensive model, spatial skills in 4th grade directly predicted STEM major choice, with a standardized path coefficient of 0.13. Assuming average levels on all predictors in our model except for spatial skills, this effect size means that the probability of entering a math-intensive STEM major was 50% greater for a student with high (one standard deviation above the mean) than low (one standard deviation below the mean) spatial skills in 4th grade. Similarly, the probability of entering a non-math-intensive STEM major was 24% greater for a student with high than low spatial skills (see SM, Section B for calculations). This effect size is notable, both due to the wide time span between 4th grade and college years and because it occurs over and above the effects of math and verbal achievement and ability self-concepts.

In addition to this direct pathway, indirect pathways in our model suggest that strong early spatial skills also elevate the likelihood of choosing STEM majors by enhancing math achievement and ability self-concept in the intervening years. In contrast to the mounting evidence on the close relation between spatial skills and math achievement (see Atit et al., 2021 for a meta-analysis), less is known about how spatial skills relate to affective and motivational factors in math. A few studies showed that children and adults with stronger spatial skills reported lower levels of math anxiety and that men's advantage in spatial skills helped explain their lower levels of math anxiety compared to women (Lauer, Esposito et al., 2018; Maloney et al., 2012; Sokolowski et al., 2019). These findings have led some researchers to stress the importance of the interplay between spatial skills and affective factors in math for understanding students' STEM participation (Sokolowski et al., 2019). Our findings contribute to this line of work by identifying significant indirect pathways from spatial skills to STEM majors via math ability self-concepts. More research is needed to examine *how* and *why* spatial skills predict affective and motivational factors in math, which can inform mechanisms of STEM participation.

Further, our findings suggest that the gender gap in STEM college majors starts to take shape as early as middle childhood, in the form of a male advantage in spatial skills. In addition to those considered in the current study, many factors potentially contribute to women's low participation in STEM, such as goal orientation, career preferences, and gender stereotypes (see Wang & Degol, 2017 for a review). Although gender differences in some of these factors may emerge in childhood (e.g., gender-science stereotypes, Miller et al., 2018), most investigations of the relations between these factors and STEM participation have focused on adolescence—a period proximate to choosing college majors. Our findings constitute, to our knowledge, the first evidence that the gender gap in STEM college majors starts to develop as early as middle childhood.

Several limitations of the current study suggest possible future directions. First, our findings are correlational in nature. Although the SEM results provided evidence that childhood spatial skills predict later STEM major choices, no causal inference should be made. Whether enhancing childhood spatial skills promotes pursuing STEM majors needs to be explored in experimental studies. A recent meta-analysis on spatial interventions provides hopeful findings: 29 spatial skill interventions on average produced significant transfer effects (Hedge's $g = 0.28$) on math skills (Hawes et al., 2022). Another limitation of the current study is that our sample may not be nationally representative. The initial SECCYD sample included participants from a range of family backgrounds and geographic areas, but was not designed to be nationally representative, and our analytical sample was less diverse and from higher SES background than that initial sample. This is consistent with prior studies indicating that people who drop out from longitudinal studies and people who do not go to college are usually from lower SES backgrounds (Guo et al., 2015; Gustavson et al., 2012). However, future research is needed to examine whether strong childhood spatial skills increase the likelihood of pursuing STEM degrees or working in STEM fields in more general populations.

Taken together, these findings shed new light on women's underrepresentation in STEM fields and provide new hope for addressing it. Research has shown that spatial play, such as building with blocks and solving jigsaw puzzles, can facilitate children's spatial development (Jirout & Newcombe, 2015; Levine et al., 2012). Children's spatial skills can also be improved through intentional training. For example, 6- to 8-year-old children improved their spatial skills after playing computerized spatial training games over 6 weeks as part of their regular classroom activities (Hawes et al., 2015). A meta-analysis of 206 spatial training studies reported a 0.47 standard deviation improvement in spatial skills and showed that males and females both benefited from spatial training (Uttal et al., 2013). Therefore, enhancing spatial skills in childhood—such as through spatial play at home and spatial activities in schools—shows great promise for setting more children, especially girls, on a pathway toward STEM achievement in adulthood.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank Gregory R. Hancock and Elizabeth Tipton for their advice on our statistical analyses. National Science Foundation grant 1760144 to N. S. Newcombe, M. Weinraub, and E. A. Gunderson supported the current study. A cooperative agreement (5 U10 HD027040) between the study investigators that included M. Weinraub and D. L.

Vandell and the Eunice Kennedy Shriver National Institute of Child Health and Human Development supported the design and data collection of the Study of Early Child Care and Youth Development (SECCYD) from birth through age 15 years. Charles Stewart Mott Foundation grant G-2017-00786 to D. L. Vandell supported the SECCYD age 26 study design, data collection, and analyses.

DATA AVAILABILITY STATEMENT

The SECCYD was administered in six phases. Data from Phases I – IV are publicly available on https://www.icpsr.umich.edu/web/ICPSR/series/00233. For privacy concerns, some data files are only available to researchers who agree to the terms and conditions of a Restricted Data Use Agreement. Data from Phases V and VI (i.e., age-26 assessment) may be obtained by contacting one of the principal investigators with the SECCYD, Deborah Lowe Vandell. We have pre-registered the selection and scoring of variables and analytic decisions (https://osf.io/7n5du). We have also made the analysis code and model output publicly available (https://osf.io/n3y46/).

REFERENCES

Atit K, Power JR, Pigott T, Lee J, Geer EA, Uttal DH, & Sorby SA (2021). Examining the relations between spatial skills and mathematical performance: A meta-analysis. Psychonomic Bulletin & Review, 29(3), 699–720. 10.3758/s13423-021-02012-w [PubMed: 34799844]

Casey MB, Nuttall RL, & Pezaris E (1997). Mediators of gender differences in mathematics college entrance test scores: A comparison of spatial skills with internalized beliefs and anxieties. Developmental Psychology, 33(4), 669–680. 10.1037/0012-1649.33.4.669 [PubMed: 9232382]

Eagan MK, Stolzenberg EB, Zimmerman HB, Aragon MC, Sayson HW, & Rios-Aguilar C (2017). The American freshman: National norms fall 2016. Los Angeles: Higher Education Research Institute, UCLA. Retrived from https://www.heri.ucla.edu/monographs/TheAmericanFreshman2016.pdf

Ganley CM, Vasilyeva M, & Dulaney A (2014). Spatial ability mediates the gender difference in middle school students' science performance. Child Development, 85(4), 1419–1432. 10.1111/cdev.12230 [PubMed: 24673201]

Geary DC, Scofield JE, Hoard MK, & Nugent L (2020). Boys' advantage on the fractions number line is mediated by visuospatial attention: Evidence for a parietal-spatial contribution to number line learning. Developmental Science, 24(3), e13063. 10.1111/desc.13063 [PubMed: 33185311]

Goldman RD, & Hewitt BN (1976). The Scholastic Aptitude Test "explains" why college men major in science more often than college women. Journal of Counseling Psychology, 23(1), 50–54. 10.1037/0022-0167.23.1.50

Groth-Marnat G, & Teal M (2000). Block design as a measure of everyday spatial ability: A study of ecological validity. Perceptual and Motor Skills, 90, 522–526. [PubMed: 10833749]

Guo J, Parker PD, Marsh HW, & Morin AJS (2015). Achievement, motivation, and educational choices: A longitudinal study of expectancy and value using a multiplicative perspective. Developmental Psychology, 51(8), 1163–1176. 10.1037/a0039440 [PubMed: 26053150]

Gustavson K, von Soest T, Karevold E, & Røysamb E (2012). Attrition and generalizability in longitudinal studies: Findings from a 15-year population-based study and a Monte Carlo simulation study. BMC Public Health, 12(1), 918. 10.1186/1471-2458-12-918 [PubMed: 23107281]

Hawes Z, Gilligan-Lee KA, & Mix KS (2022). Effects of spatial training on mathematics performance: A meta-analysis. Developmental Psychology, 58(1), 112–137. 10.1037/dev0001281 [PubMed: 35073120]

Hawes Z, Moss J, Caswell B, & Poliszczuk D (2015). Effects of mental rotation training on children's spatial and mathematics performance: A randomized controlled study. Trends in Neuroscience and Education, 4(3), 60–68. 10.1016/j.tine.2015.05.001

Hu L, & Bentler PM (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal, 6(1), 1–55. 10.1080/10705519909540118

Hyde JS, Lindberg SM, Linn MC, Ellis AB, & Williams CC (2008). Gender similarities characterize math performance. Science, 321(5888), 494–495. 10.1126/science.1160364 [PubMed: 18653867]

Irwing P (2012). Sex differences in g: An analysis of the US standardization sample of the WAIS-III. Personality and Individual Differences, 53(2), 126–131. 10.1016/j.paid.2011.05.001

Jacobs JE, Lanza S, Osgood DW, Eccles JS, & Wigfield A (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. Child Development, 73(2), 509–527. 10.1111/1467-8624.00421 [PubMed: 11949906]

Jirout JJ, & Newcombe NS (2015). Building blocks for developing spatial skills: Evidence from a large, representative U.S. sample. Psychological Science, 26(3), 302–310. 10.1177/0956797614563338 [PubMed: 25626442]

Kenny DA, & McCoach DB (2003). Effect of the number of variables on measures of fit in structural equation modeling. Structural Equation Modeling: A Multidisciplinary Journal, 10(3), 333–351. 10.1207/S15328007SEM1003_1

Lauer JE, Esposito AG, & Bauer PJ (2018). Domain-specific anxiety relates to children's math and spatial performance. Developmental Psychology, 54(11), 2126–2138. 10.1037/dev0000605 [PubMed: 30265030]

Lauer JE, Ilksoy SD, & Lourenco SF (2018). Developmental stability in gender-typed preferences between infancy and preschool age. Developmental Psychology, 54(4), 613–620. 10.1037/dev0000468 [PubMed: 29154646]

Lauer JE, Yhang E, & Lourenco SF (2019). The development of gender differences in spatial reasoning: A meta-analytic review. Psychological Bulletin, 145(6), 537–565. 10.1037/bul0000191 [PubMed: 30973235]

Levine SC, Ratliff KR, Huttenlocher J, & Cannon J (2012). Early puzzle play: A predictor of preschoolers' spatial transformation skill. Developmental Psychology, 48(2), 530–542. 10.1037/a0025913 [PubMed: 22040312]

Lindberg SM, Hyde JS, Petersen JL, & Linn MC (2010). New trends in gender and mathematics performance: A meta-analysis. Psychological Bulletin, 136(6), 1123–1135. 10.1037/a0021276 [PubMed: 21038941]

Maloney EA, Waechter S, Risko EF, & Fugelsang JA (2012). Reducing the sex difference in math anxiety: The role of spatial processing ability. Learning and Individual Differences, 22(3), 380–384. 10.1016/j.lindif.2012.01.001

Martinez A, & Christnacht C (2021). Women making gains in STEM occupations but still underrepresented. Retrieved from https://www.census.gov/library/stories/2021/01/women-making-gains-in-stem-occupations-but-still-underrepresented.html

Miller DI, Nolla KM, Eagly AH, & Uttal DH (2018). The development of children's gender-science stereotypes: A meta-analysis of 5 decades of U.S. draw-a-scientist studies. Child Development, 89(6), 1943–1955. 10.1111/cdev.13039 [PubMed: 29557555]

Muthén LK, & Muthén BO (1998–2017). Mplus user's guide. (Eighth). Muthén & Muthén.

Parker PD, Schoon I, Tsai Y-M, Nagy G, Trautwein U, & Eccles JS (2012). Achievement, agency, gender, and socioeconomic background as predictors of postschool choices: A multicontext study. Developmental Psychology, 48(6), 1629–1642. 10.1037/a0029167 [PubMed: 22799584]

Piffer D (2016). Sex differences in intelligence on the American WAIS-IV. Mankind Quarterly, 57(1), 25–33. 10.46469/mq.2016.57.1.3

R Core Team. (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Retrieved from https://www.rproject.org/

Shea DL, Lubinski D, & Benbow CP (2001). Importance of assessing spatial ability in intellectually talented young adolescents: A 20-year longitudinal study. Journal of Educational Psychology, 93(3), 604–614. 10.1037/0022-0663.93.3.604

Skaalvik S, & Skaalvik EM (2004). Gender differences in math and verbal self-concept, performance expectations, and motivation. Sex Roles, 50(3/4), 241–252. 10.1023/B:SERS.0000015555.40976.e6

Sokolowski HM, Hawes Z, & Lyons IM (2019). What explains sex differences in math anxiety? A closer look at the role of spatial processing. Cognition, 182(October 2018), 193–212. 10.1016/j.cognition.2018.10.005 [PubMed: 30343180]

Sorby SA (2009). Educational research in developing 3-D spatial skills for engineering students. International Journal of Science Education, 31(3), 459–480. 10.1080/09500690802595839

Stano JF (2004). Wechsler abbreviated scale of intelligence. Rehabilitation Counseling Bulletin, 48(1), 56–57.

Steiger JH, & Lind JC (1980). Statistically based tests for the number of common factors. Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA.

Tian J, Dam S, & Gunderson EA (2022). Spatial skills, but not spatial anxiety, mediate the gender difference in number line estimation. Developmental Psychology, 58(1), 138–151. 10.1037/dev0001265 [PubMed: 35073121]

U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (2020). Fall 2010 through Fall 2019, Completions component. Retrieved from https://nces.ed.gov/programs/digest/d20/tables/dt20_318.45.asp

U.S. Department of Homeland Security. (2016). DHS STEM designated degree program list. Retrieved from https://www.ice.gov/sites/default/files/documents/Document/2016/stem-list.pdf

Uttal DH, Meadow NG, Tipton E, Hand LL, Alden AR, Warren C, &Newcombe NS (2013). The malleability of spatial skills: A meta-analysis of training studies. Psychological Bulletin, 139(2), 352–402. 10.1037/a0028446 [PubMed: 22663761]

Voyer D, Voyer S, & Bryden MP (1995). Magnitude of sex differences in spatial abilities. Psychological Bulletin, 117(March), 250–270. [PubMed: 7724690]

Wai J, Lubinski D, & Benbow CP (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. Journal of Educational Psychology, 101(4), 817–835. 10.1037/a0016127

Wang M-T, & Degol JL (2017). Gender gap in Science, Technology, Engineering, and Mathematics (STEM): Current knowledge, implications for practice, policy, and future directions. Educational Psychology Review, 29(1), 119–140. 10.1007/s10648-015-9355-x [PubMed: 28458499]

Webb RM, Lubinski D, & Benbow CP (2007). Spatial ability: A neglected dimension in talent searches for intellectually precocious youth. Journal of Educational Psychology, 99(2), 397–420. 10.1037/0022-0663.99.2.397

Wechsler D (1999). Wechsler abbreviated scale of intelligence.

Wechsler D (2012). Wechsler preschool and primary scale of intelligence— fourth edition. 10.1007/978-3-319-91280-6_866

Woodcock RW, & Johnson MB (1989). Woodcock-Johnson Psycho-Educational Battery- Revised.

Wu H-K, & Shah P (2004). Exploring visuospatial thinking in chemistry learning. Science Education, 88(3), 465–492. 10.1002/sce.10126

Research Highlights

- Using a national longitudinal dataset, we found 4th-grade spatial skills directly predicted STEM college major choice after accounting for multiple cognitive and motivational mechanisms.
- Strong spatial skills in 4th grade also elevated STEM major choice via enhanced math achievement and motivation in the intervening years.
- Gender differences in 4th-grade spatial skills contributed to women's underrepresentation in STEM college majors.

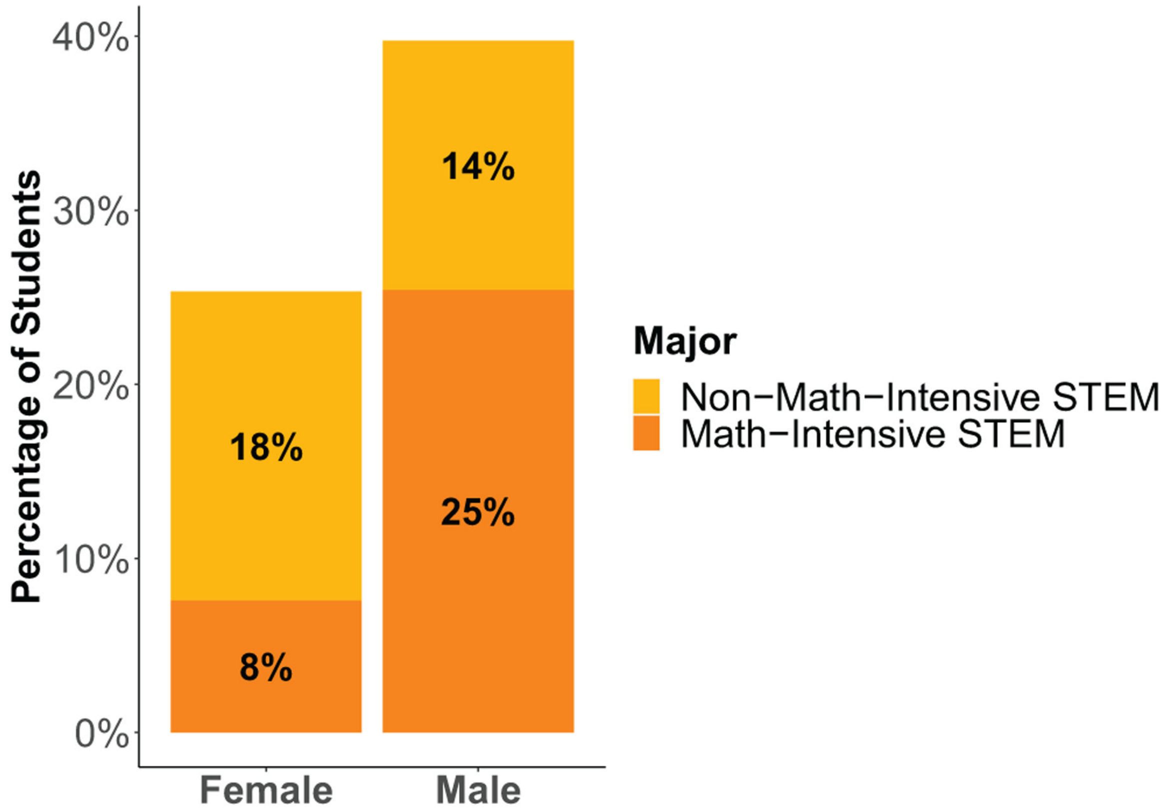

FIGURE 1.

Percentage of females ($N = 383$) and males ($N = 307$) entering STEM majors

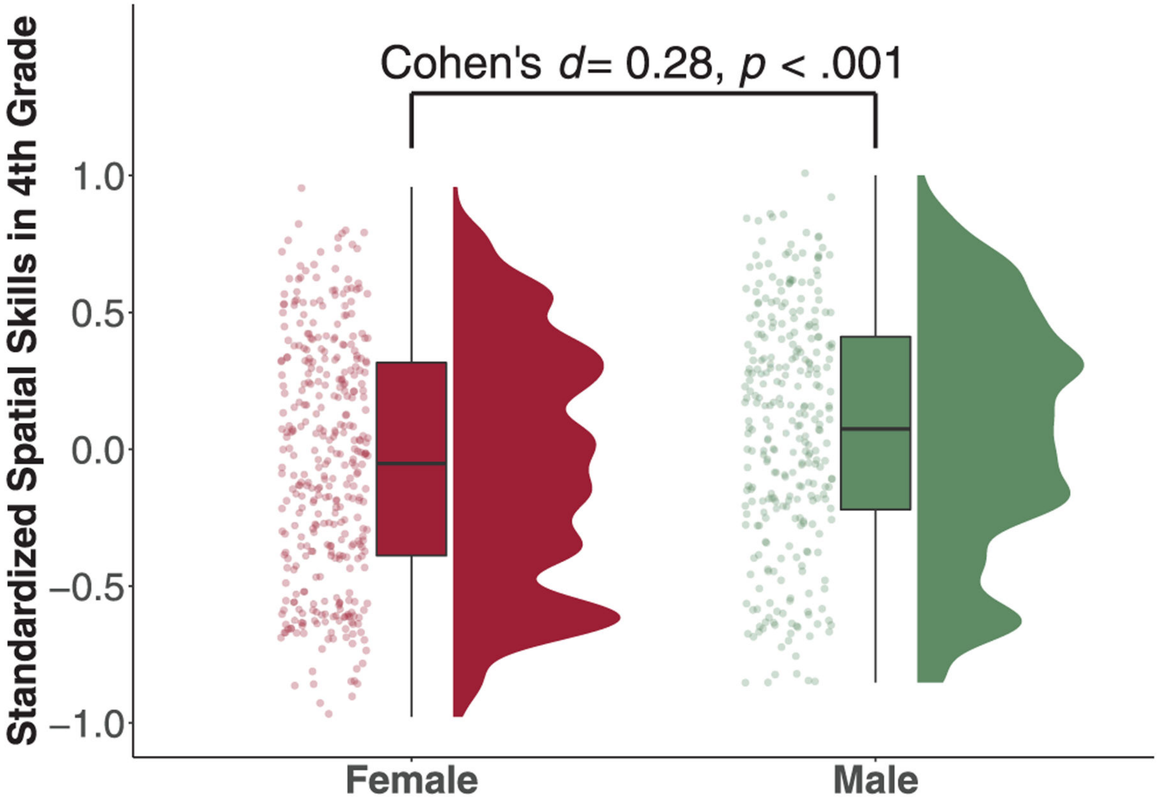

FIGURE 2.

Spatial skills of female and male students in middle childhood. Boxplots show median, lower and upper quartiles, and minimum and maximum values in each gender group

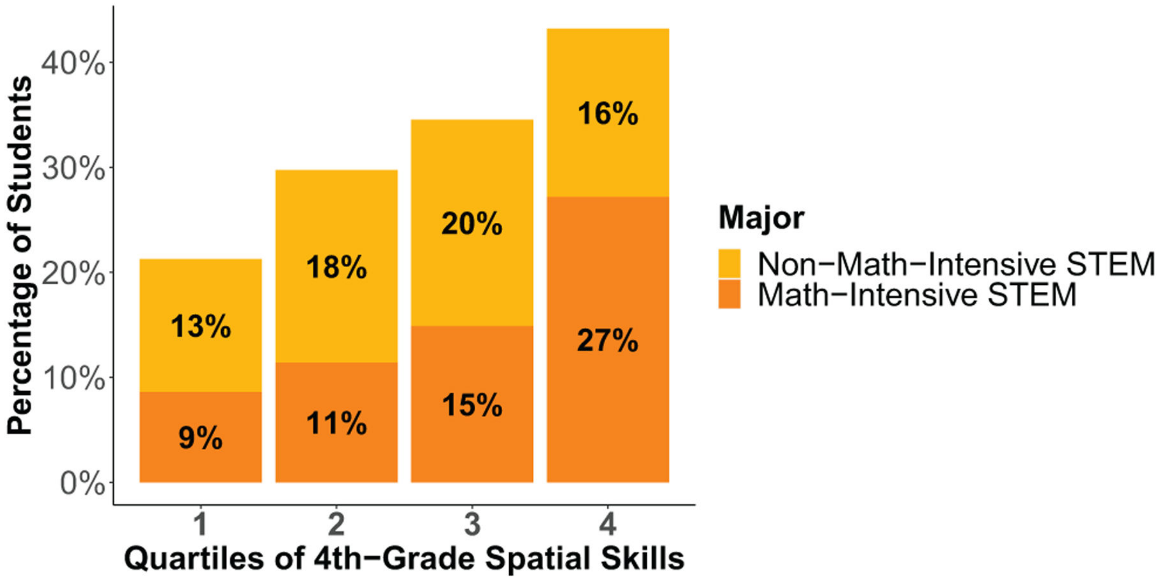

FIGURE 3.

Percentage of students with different levels of spatial skills entering STEM majors. For descriptive purposes, the graph categorizes students into quartiles based on 4th-grade spatial skill. Note that our inferential analyses considered spatial skills as a continuous variable

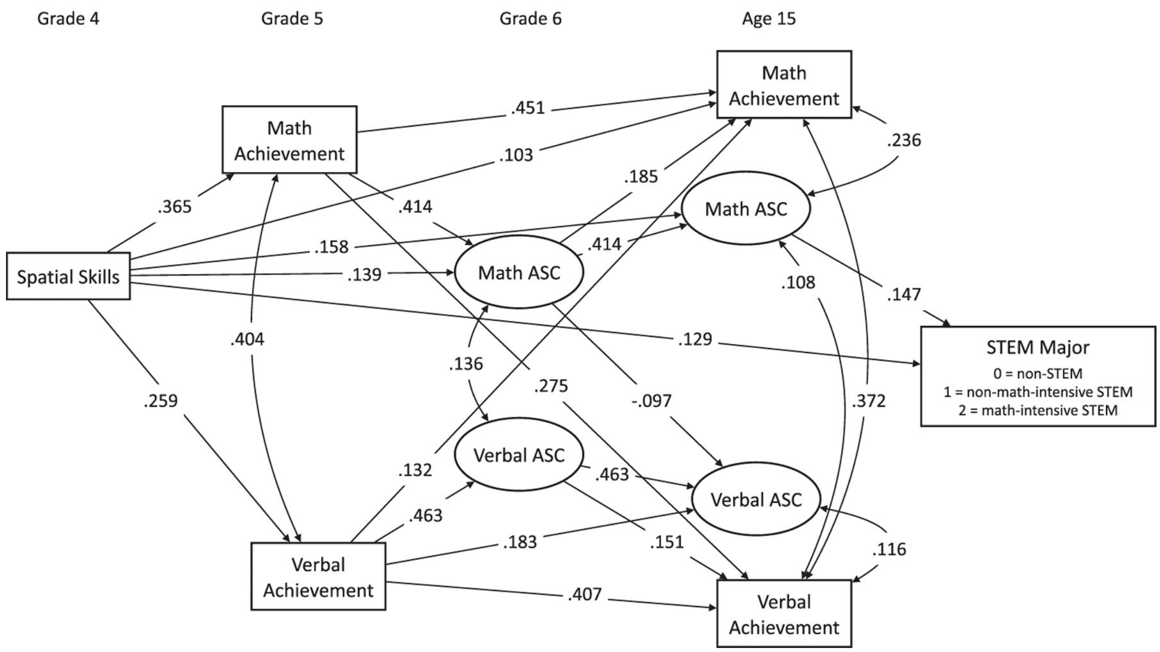

FIGURE 4.

The Full Model, predicting STEM major with 4th-grade spatial skills, 5th-grade and age-15 math and verbal achievement, and 6th-grade and age-15 math and verbal ability self-concept. "ASC" stands for ability self-concept. Path coefficients shown are standardized estimates of significant effects. For simplicity, non-significant paths, paths including covariates (i.e., gender, family income-to-needs ratio, and maternal education), and the measurement part of the model are omitted from the figure