

GEOSPHERE, v. 14, no. 2,

doi:10.1130/GES01494.1

3 figures; 10 tables

CORRESPONDENCE: anne.u.gold@colorado.edu

CITATION: Gold, A.U., Pendergast, P.M., Ormand, C.J., Budd, D.A., Stempien, J.A., Mueller, K.J., and Kravitz, K.A., 2018, Spatial skills in undergraduate students—Influence of gender, motivation, academic training, and childhood play: *Geosphere*, v. 14, no. 2, p. 668–683, doi:10.1130/GES01494.1.

Science Editor: Shanaka de Silva
Guest Associate Editor: Julie Libarkin

Received 10 January 2017
Revision received 14 November 2017
Accepted 20 December 2017
Published online 5 February 2018



This paper is published under the terms of the CC-BY-NC license.

© 2018 The Authors

Spatial skills in undergraduate students—Influence of gender, motivation, academic training, and childhood play

Anne U. Gold¹, Philip M. Pendergast², Carol J. Ormand³, David A. Budd⁴, Jennifer A. Stempien⁴, Karl J. Mueller⁴, and Katherine A. Kravitz⁴

¹Cooperative Institute for Research in Environmental Sciences, University of Colorado at Boulder, UCB 449, Boulder, Colorado 80303, USA

²Institute of Behavioral Science, University of Colorado at Boulder, 483 UCB, Boulder, Colorado 80309, USA

³Science Education Resource Center, Carleton College, One North College St., Northfield, Minnesota 55057, USA

⁴Department of Geological Science, University of Colorado at Boulder, UCB 399, Boulder, Colorado 80309, USA

ABSTRACT

Spatial reasoning is critical for success in science, technology, engineering, and mathematics (STEM) disciplines and particularly for geoscience. To evaluate capabilities among U.S. students, we assessed spatial reasoning skills in 345 introductory geology and upper-level structural geology students. The test measured students' abilities to disembed, visualize, and mentally rotate objects. The results highlight an uneven distribution of spatial skills with a minimum score of 6% and a maximum score of 75% correct responses. Spatial skills are positively correlated with standardized test scores, motivation for learning, STEM major declaration, and number of science courses taken. Our analysis also indicates that the cumulative, informal training of childhood play has the ability to increase spatial reasoning. Spatial skill scores were significantly higher among students who played action, construction, or sports video games in childhood. Male and female students display significant differences in spatial skills, especially for mental rotation, with males outperforming females. However, gender disparities are fully mediated after adjusting for a variety of academic factors and whether students frequently played with construction-based toys. This indicates that gender differences are experiential rather than biological in origin. This study suggests that both formal academic training and extracurricular activities appear to develop spatial skills throughout students' lives and indicates that systematic testing of spatial skills and formal training opportunities for students would facilitate improved spatial reasoning among students. We hypothesize that formal training opportunities for spatial reasoning could increase the potential pool of students who successfully enter STEM careers, including the geosciences, especially among women.

INTRODUCTION

Spatial relations are an integral component of science, technology, engineering, and mathematics (STEM) disciplines. For example, chemists study the spatial structure of molecules; geoscientists, the spatial dimensions and

genesis of landscapes over time; biologists, the three-dimensional DNA structure; and engineers, the design and manufacture of three-dimensional structures. Furthermore, scientific communication utilizes visuospatial representations such as graphs, diagrams, and maps that demand spatial skills for interpretation. Therefore, students' abilities to visualize spatial relations are critical in STEM disciplines, especially in the geosciences.

Spatial reasoning skills are more strongly correlated with entry, performance, and persistence in STEM disciplines than Suite of Assessments (SAT) scores (Shea et al., 2001; Webb et al., 2007; Wai et al., 2009); they even correlate with creative accomplishments as measured by patents and publications (Kell et al., 2013). Despite this convincing evidence that spatial reasoning skills are important for student success in STEM disciplines, spatial reasoning skills are not systematically instructed or tested in K–12 education (National Research Council, 2006; Kastens et al., 2014; Ormand et al., 2014) or assessed among incoming undergraduate students, and they are seldom explicitly included in instruction at the college level. Not surprisingly, spatial skills are unevenly distributed in the population (e.g., National Research Council, 2006; Kastens et al., 2014), which presents a challenge to university instructors teaching spatially demanding courses. This uneven distribution of skills allows some students to perform discipline-specific tasks or interpret scientific communication easily, while others struggle (Kastens et al., 2009, 2014). However, spatial reasoning is not an innate ability. Rather, spatial reasoning has been shown to be trainable with a lasting effect (see meta-analysis in Uttal et al., 2013). For example, spatial reasoning improves through participation in training interventions such as discipline-specific course work (e.g., Sorby, 2001; Kozhevnikov et al., 2007; Titus and Horsman, 2009; Nielsen et al., 2011; Ormand et al., 2014), video gaming (Green and Bavelier, 2003; Feng et al., 2007; Cherney, 2008; Terlecki et al., 2008; Adams and Mayer, 2012), participation in sports (Ozel et al., 2002, 2004; Moreau et al., 2011, 2012; Pietsch and Jansen, 2012), playing music (Pietsch and Jansen, 2012), and playing with construction-based toys (Brosnan, 1998; Wolfgang et al., 2003; Coxon, 2012). However, many prior studies on these ways of training have been short in duration (e.g., interventions of hours to weeks), with limited exploration of the long-term efficacy of these interventions. Furthermore, we don't know how

the combination of formal and informal life experiences influence the development of spatial skills.

Spatial skills have been studied extensively, and we know that spatial reasoning includes a “collection of cognitive skills” (National Research Council, 2006, p. 12). Embodied cognition research finds that “cognitive processes are deeply rooted in the body’s interactions with the world” and that “the mind must be understood in the context of its relationship to a physical body that interacts with the world” (Wilson, 2002, p. 625). Building on the findings that spatial reasoning is trainable and appears to be trained through one’s interactions with the world, but that spatial skills are very unevenly distributed in the population, we sought to measure baseline spatial skills in undergraduate students, including some geoscience majors, at an early stage of their college career. The early stage of students’ college careers is particularly important because a lack in spatial skills may discourage entrance or persistence in STEM majors (Hambrick et al., 2012; Hegarty, 2014). We collected demographic data about students’ lives, high schools, and college experiences to understand the environmental and social factors as well as childhood experiences that may influence baseline spatial reasoning skills. We employed regression analysis to distinguish those formal or informal life experiences that correlate with strong spatial reasoning skills. We hypothesize that social and environmental contexts (such as personal characteristics and academic or college preparation) as well as life experiences (such as childhood play) provide different training experiences for spatial reasoning skills. This study is in contrast to previous efforts, which have focused on the impact of interventions on students’ spatial skills or the predictive nature of early testing. With our approach, we explore the environment that has shaped the cognitive collection of students’ spatial reasoning skills. Our results help identify critical life experiences where spatial skills appear to be trained.

Theoretical Framework

Spatial skills can be defined as the reasoning that “concerns shapes, locations, paths, relations among entities and relations between entities and frames of reference” (Newcombe and Shipley, 2015, p. 179–180). Spatial reasoning skills can be isolated and are not the same as discipline-specific skills (Uttal and Cohen, 2012). Different approaches to categorizing the human understanding of space and spatial relations have been developed based on theoretical findings in cognitive psychology or deduced from statistical analysis of spatial assignments (Hegarty and Waller, 2005; Chatterjee, 2008; Newcombe and Shipley, 2015). Newcombe and Shipley (2015) propose a classification scheme that builds on and expands existing frameworks (e.g., Chatterjee, 2008) using a theoretical approach. They classify a wide variety of spatial skills in geoscience according to Chatterjee’s typology, differentiating between intrinsic (or within objects) and extrinsic (or between objects) as well as static and dynamic spatial skills. In our study, we focus on testing the intrinsic or within objects spatial skills (intrinsic-static and intrinsic-dynamic)

but do not test extrinsic or between objects skills. With this approach, we follow others who have studied spatial skills in geology students (e.g., Titus and Horsman, 2009; Ormand et al., 2014), an example of one spatially demanding STEM discipline. The intrinsic-static skill assessed in this investigation is *disembedding*, the skill to isolate and attend to one aspect of a complex display or scene. Intrinsic-dynamic skills assessed include *spatial visualization or penetrative thinking*, which can also be described as visualizing spatial relations inside an object, and *mental rotation*, visualizing the effect of rotating an object.

Cognition research findings have demonstrated a dynamic connection between cognition and the environment (see reviews by Wilson, 2002; Barsalou, 2008; Lakoff, 2012) and with the social and cultural worlds (Anderson, 2003; Suitner et al., 2015). Spatial reasoning has been described as embodied cognitive processes with a dynamic connection between the brain and the environment (Amorim et al., 2006; Tversky and Hard, 2009; Kessler and Thomson, 2010; Herrera and Riggs, 2013). Following the embodied cognition theory that the brain is dynamically connected to the environment, a review of the existing literature highlights three factors that appear to form or train spatial skills (meta-analysis in Uttal et al., 2013). First, *personal characteristics* such as gender and motivation are one group of factors. Many studies have explored gender differences in spatial skills (e.g., Linn and Petersen, 1985; Baenninger and Newcombe 1989; Voyer et al., 1995; Terlecki et al., 2008). In line with our hypothesis that spatial skills are trained informally throughout life, we include a measure of motivation for learning (Pintrich et al., 1991) as a personal characteristics variable. The second set of factors is broadly grouped under *academic or college preparation*. We explored standardized test scores, selection of major, and impact of course work on spatial skills. The third category that we hypothesized could possibly train spatial reasoning skills is labeled *play experiences*, and includes playing video games, playing with construction-based toys, and playing sports. Many intervention studies exist that indicate that these activities train spatial skills (see summary in Uttal et al., 2013). In this study, we explore which of these variables correlate with spatial reasoning skills in undergraduate students, and also to what extent differences commonly attributed to personal characteristics can be accounted for by academic preparation and play experiences.

Study Design and Student Population

We measured the spatial skills of 345 undergraduates in geology courses from a large U.S. research university with over 30,000 students (Table 1). In the first week of the Fall 2014 semester, all participants completed our spatial skills test instruments during class time. To assess baseline skills of the broad undergraduate population, the majority of participants (80.2%, $n = 277$) were enrolled in introductory geology courses; these are large lecture hall classes that many non-geology major students take to fulfill their science requirement. We also tested a smaller group of undergraduates enrolled in

TABLE 1. DEMOGRAPHIC INFORMATION ABOUT PARTICIPANTS (N = 345)

	Participants (%)
Male	61.2
Female	38.8
White or Caucasian	78.0
Asian or Asian American	5.8
Latino(a) or Hispanic	4.3
Black or African or African American	1.2
Other	5.2
Multi-race	5.5
Non-STEM major	50.9
STEM major	37.5
Undeclared major	11.6

STEM—science, technology, engineering, and mathematics.

an upper-division geology course (12.5%, n = 43), which consists of lectures and weekly labs. Testing these upper-level geology majors allowed us to test whether spatial skills change with increasing course work. This study was part of a larger intervention study, but here we only focus on the analysis of pretest, baseline results. To measure the test-retest effect of taking the test twice under the pre- and post-test design, we also recruited a third group of undergraduate students (7.2%, n = 25) from across campus with the prerequisite of not being geology or engineering majors to serve as an assessment-only comparison group. These students were all lower-level undergraduates and were included in this study as part of the introductory student group.

METHODS

Spatial Skills Instrument Description and Validation

Our spatial skills instrument consisted of three spatial skill tests, each assessing a different key intrinsic spatial skill—mental rotation, disembedding, and penetrative thinking. We administered this three-part instrument following the testing protocol from Ormand et al. (2014) in their study of geology students’ spatial skills. The instrument contains abstract, non-discipline-specific test items so that students would not perceive a lack of content knowledge as a barrier for selecting an answer. Each part of the test instrument included instructions on how to answer the items, including a description of possible strategies to solve the exercises and an example. These instructions were readable at the middle school reading level (grades 6–7, Flesch-Kincaid readability index of 75). Each of the tests was time-limited, with students given three minutes for each test. All questions had a multiple-choice format. The test was completed using pencil and paper; students provided their answers on scantron sheets.

The first part of the test measured students’ *mental rotation skills* (Fig. 1A), consisting of ten items taken from the Purdue Visualization of Rotation Test (Guay, 1976). For each item, students looked at an abstract three-dimensional figure and at a rotated version of the same figure. They were then asked to look at a new three-dimensional figure and choose from five rotated versions of the same figure to identify the one that was rotated in the same way. Items were rotated both around single and multiple axes. The Purdue Visualization of Rotation Test and a revised version have been validated psychometrically by Yoon (2011) and Maeda et al. (2013).

The second test measured students’ *disembedding skills* (Fig. 1B), requiring participants to search a collection of patterns to find a given configuration. Eight items were chosen from the Hidden Figures test (Ekstrom et al., 1976). The instrument was developed and validated by Educational Testing Services (ETS), which rated its difficulty level as “high” (Ekstrom et al., 1976).

The third test measured students’ *penetrative thinking skills* (Fig. 1C) using 15 items from the Planes of Reference test (Titus and Horsman, 2009). Items on this test showed a three-dimensional figure and a cutting plane. Participants had to envision the cutting plane from a view looking straight at the plane surface and identify the correct two-dimensional shapes that outline the intersection of the object and the cutting plane from five answer choices. The test

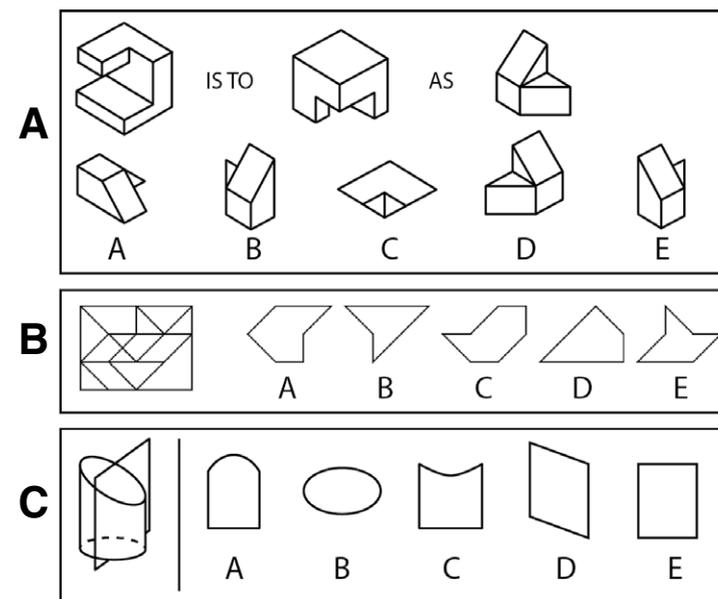


Figure 1. Examples of items from the spatial skills tests used in this study. (A) Mental rotation test item. (B) Disembedding test item. (Item in format of the Educational Testing Services [ETS] test, but actual is copyrighted. Example was drafted by the third author.) (C) Penetrative thinking test item.

was developed based on items from Myers (1953) and Crawford and Burnham (1946) and has been validated by Witkin et al. (1977). It has previously been used with geology students (e.g., Titus and Horsman, 2009; Almquist et al., 2011; Ormand et al., 2014).

All three tests are drawn from the empirical literature and have reasonable surface and content validity. To ensure that the measures were appropriate for our sample, we calculated Cronbach's alpha using imputed values for questions that were left blank. We used full-information maximum likelihood estimation to create ten imputed data sets. Average scores across these data sets are used to calculate reliability estimates. The instrument as a whole demonstrated sufficient reliability ($\alpha = 0.7$), with two of three subscales scoring just below conventional thresholds (mental rotation and penetrative thinking, $\alpha = 0.6$). The final subscale had low internal consistency (disembedding, $\alpha = 0.3$); therefore, while reported, results on this aspect of spatial reasoning are interpreted conservatively.

Data

In addition to the assessment of spatial skills, we collected demographic and educational data for students around students' personal characteristics, academic and college preparation, and play experiences that may have provided spatial skill training. Some academic information (e.g., standardized test scores, college courses completed, and declared major) and demographic data (e.g., gender) were provided by the university's Office of the Registrar. All other student data (e.g., motivation, high school course work, and game playing experiences) were derived from a self-report survey administered to students following the spatial skill assessment.

The widely used and statistically validated Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich et al., 1991) was used to document students' motivations for learning and their learning strategies and cognition. Following Hilpert et al. (2013), we administered a 43-item version of the full MSLQ to calculate values for the six subscales that explore students' affective domains with respect to learning (i.e., self-efficacy, control of learning, intrinsic goals, task value, metacognitive regulation, and effort regulation). Students completed the questionnaire as part of their course homework in the first week of classes.

Academic preparation of the participating students was measured using the index score of the Colorado Commission of Higher Education (CCHHE, 2015); this score combines students' high school grade point averages and performance on standardized tests. The possible index values range from 45 to 146, with study participants having index scores ranging from 75 to 146 (mean $[M] = 114$; standard deviation $[SD] = 12.3$). We centered student indices around the mean, assigning scores higher than 114 positive values (ranging from 1 to 32), indicating achievement above the mean, and scores lower than 114 as negative (ranging from -1 to -39), indicating student achievement below the mean, to aid in interpretation of linear regression models. By doing so, the intercept in a regression analysis indicates the predicted spatial test score for students with average achievement scores.

On the survey, students provided the type and number of pre-college science (physics, chemistry, and biology), engineering, and Earth science (including geography, environmental science, and astronomy) courses they had taken. The registrar's office further provided the completed college-level courses. We counted the number of completed (minimum C grade) science, Earth science, and engineering and/or architecture courses both in college and high school for each participant.

On the survey, students further provided information about their experience with video game play. We asked the students to "list the top three video games they play(ed) the most" and to provide three open-ended answer options. We combined all video games that were listed in the responses. A total of 285 different video games were listed. Following Adams' (2009) video game categories in *Fundamentals of Game Design*, we created a rubric that included common video game types and structures. Combining Adams' definitions with work from Spence and Feng (2010) and Granic et al. (2014), we were able to identify six video game categories (Table 2). Five coders classified the 285 games into these six categories. We compared the five coders' classifications and accepted game assignments with an agreement between coders of 80% or more. In cases where fewer than 80% of the coders agreed, a master coder revisited the code and assigned a category based on an additional review of the code and external sources (e.g., categories that are assigned by the gaming companies or by vending platforms such as Amazon). To create the video gaming variable, we summed for each participant the number of video games in all six categories of those students who played

TABLE 2. VIDEO GAME CATEGORIES AND DESCRIPTIONS USED TO CODE GAMES LISTED BY PARTICIPANTS

Shooter games	Aiming, focus on two places at once, visualization of environment (examples: <i>Call of Duty</i> , <i>Battlefield</i>).
Arcade game and puzzle games	Generally flat, requires some navigation of levels (examples: <i>Asteroids</i> , <i>Galaga</i>).
Sports games	Intended to simulate actual sports, 3D graphics, players interact with simulated athletes, frequent shifts of perspective (examples: <i>FIFA</i> , <i>Madden NFL</i>).
Driving games	Focus on the feeling of movement (examples: <i>Gran Turismo</i> , <i>Mario Kart</i>).
Construction and simulation games	Generally depict the altering of a given world (examples: <i>Tetris</i> , <i>World of Warcraft</i> , <i>Minecraft</i>).
Action games and role-playing games	Navigation of vast worlds, sometimes manipulation of created world; focus on character development (examples: <i>Sky Rim</i> , <i>Mass Effect</i> , <i>Final Fantasy</i>).

video games ($n = 268$). Thus, any student could be assigned a value from 0 to 3 for each of the six game categories.

On the survey, students also indicated how often they had played with construction-based toys, listing Legos®, blocks, connectors, and Magna-Tiles® as examples of construction-based toys. Students selected from five frequency options ranging from “Very Often” to “Never.” We dichotomized the measure to reflect the qualitative difference between playing “very and quite often” versus “sometimes, rarely, and never” into frequent or infrequent play (Table 3).

Lastly, we also asked the students to “list the three sports or physical activities that they participated in the most throughout their life.” Students listed a

TABLE 3. RESPONSES TO THE QUESTION “LOOKING BACK ON YOUR CHILDHOOD, HOW OFTEN DID YOU PLAY WITH CONSTRUCTION-BASED TOYS (I.E., LEGOS, BLOCKS, CONNECTORS, MAGNA-TILES)?”

	Responses (%)		
Very often	26	Frequent play	62
Quite often	36		
Sometimes	26	Infrequent play	38
Rarely	11		
Never	1		

TABLE 4. TOP MOST POPULAR SPORTS LISTED BY THE STUDENTS IN THIS STUDY

Sport or physical activity	Number of times listed by participants
Soccer	114
Running (includes track, cross country, marathon)	74
Basketball	70
Ski and snowboard (includes downhill, backcountry)	56
Football	49
Baseball	42
Swimming	39
Tennis (includes table tennis)	39
Lacrosse/LAX	37
Volleyball	31
Dance (includes ballet)	30

total of 92 unique sports or physical activities. We sorted the sports by the number of times each sport appeared; Table 4 lists the top 11 most popular sports.

Analysis

All descriptive and inferential statistics were calculated using SPSS Statistics Version 23. The spatial score was calculated as a percentage of correctly solved test items. Blank responses were treated as incorrectly solved items. This approach allowed us to compare the number of items a student can correctly solve under a given time restriction. Normality of the distribution of the spatial skill scores (Table 5) was tested to ensure that our analyses are valid. We then examined bivariate relationships between the resultant spatial skill scores and all other potential explanatory variables (i.e, personal characteristics, academic and college preparation, and play experiences). These bivariate associations generally suggested that ordinary least squares (OLS) regression was appropriate for examining and explaining differences in spatial reasoning, with results from residual plots indicating linear relationships and constant error variance. Thus, we proceeded to fit baseline OLS regression models that describe the distribution of spatial skills. We built a series of seven OLS regression models to explore which factors explain the distribution of students’ spatial skills. In each model, we introduced a new explanatory variable. Thus, OLS regression models A through G (Tables 6–9) helped us estimate, on average, how a variety of characteristics influenced observed spatial skill scores in a large undergraduate student sample. In two summary OLS regression models (models H and J; Table 9), we explored the overall factors that appear to influence spatial skills.

RESULTS

Distribution of Participants’ Spatial Skills

The time limit for completion of each spatial test resulted in overall low total scores. Students left, on average, 31% of the questions blank. The mean score for the three-part instrument was 33.9% ($SD = 11.9$) correct responses (low kurtosis and skewness scores, $< \pm 1$, indicate normal distribution and

TABLE 5. DESCRIPTIVE STATISTICS FOR THE SPATIAL SKILLS TEST RESULTS (N = 345)

Measure	Mental rotation (MR) score (%)	Disembedding (DE) score (%)	Penetrative thinking (PT) score (%)	Total score (%)	Blanks on MR test (%)	Blanks on DE test (%)	Blanks on PT test (%)
Mean	39.8	20.3	37.4	34.0	25.0	36.4	32.3
Standard deviation	20.2	15.8	16.0	11.9	19.8	30.0	17.9
Skewness	0.26	0.87	0.24	0.24	--	--	--
Kurtosis	-0.46	0.66	0.19	0.17	--	--	--

TABLE 6. OLS REGRESSION MODELS PREDICTING TOTAL SPATIAL TEST SCORES BY PARTICIPANT CHARACTERISTICS

	Model A1 (total score)	Model A2 (mental rotation)	Model B1 (total score)	Model B2 (mental rotation)	Model C1 (total score)	Model C2 (mental rotation)
<u>Gender</u>						
Female	-3.392** (1.301)	-8.284*** (2.191)	-1.513 (1.385)	-6.598** (2.367)	-3.266* (1.486)	-11.093*** (2.498)
<u>Motivation</u>						
Task value			1.548* (0.736)	2.950* -1.258 (1.516)	1.025 (0.761)	2.758* (1.280)
Intrinsic motivation			-1.229 (0.887)	-1.887 (1.516)	-0.672 (0.918)	-1.197 (1.544)
Control of beliefs			-1.217 (0.958)	-1.241 (1.638)	-1.091 (0.982)	-1.347 (1.651)
Self-efficacy			3.894*** (1.068)	5.223** (1.827)	3.77*** (1.116)	4.645** (1.876)
Metacognition			-1.380 (1.125)	-3.977* (1.923)	-1.427 (1.182)	-4.519** (1.988)
Effort regulation			0.470+ (.871)	1.808 (1.488)	.115 (.942)	1.169 (1.584)
<u>Standardized Test Scores</u>						
Standardized Test Index Score (centered on mean)					0.186** (0.061)	0.383*** (0.102)
Constant	35.301	42.986	35.072	42.686	22.371	24.091

*** p ≤ 0.001; ** p ≤ 0.01; * p ≤ 0.05; + p ≤ 0.10.
 Notes: Unstandardized coefficients. Standard errors in parentheses. n = 184. OLS—ordinary least squares.

TABLE 7. OLS REGRESSION MODELS PREDICTING TOTAL SPATIAL TEST SCORES BY COLLEGE EXPERIENCE

	Model D1 (total score)	Model D2 (mental rotation)	Model E1 (total score)	Model E2 (mental rotation)
<u>Gender</u>				
Female	-4.338** (1.354)	-11.679*** (2.293)	-5.050*** (1.323)	-12.934*** (2.596)
<u>Standardized test scores</u>				
Standardized test index score (centered on mean)	0.216*** (0.054)	0.417*** (0.092)	0.254*** (0.054)	0.432*** (0.110)
<u>Selection of major</u>				
Non-STEM	-4.664*** (1.421)	-6.258** (2.407)		
Undeclared	-4.728* (2.253)	-7.684* (3.814)		
<u>Prior course work</u>				
Number of engineering courses with passing grade			-0.415 (0.375)	0.466 (0.677)
Number of science courses with passing grade			0.572*** (0.177)	1.171*** (0.352)
Constant	38.59	48.522	33.803	47.788

*** p ≤ 0.001; ** p ≤ 0.01; * p ≤ 0.05; + p ≤ 0.10.
 Notes: Unstandardized coefficients. Standard errors in parentheses. n = 211. OLS—ordinary least squares; STEM—science, technology, engineering, and mathematics.

TABLE 8. OLS REGRESSION MODELS PREDICTING TOTAL SPATIAL TEST SCORES BY LIFE EXPERIENCES

	Model F1 (total score)	Model F2 (mental rotation)	Model G1 (total score)	Model G2 (mental rotation)
<u>Gender</u>				
Female	-6.968*** (1.623)	-15.034*** (2.739)	-5.689** (2.011)	-15.590*** (3.307)
<u>Standardized Test Scores</u>				
Standardized Test Index Score (centered on mean)	0.239*** (0.053)	0.443*** (0.090)	0.226*** (0.058)	0.491*** (0.096)
<u>Lifestyle Factors</u>				
<u>Video Gaming</u>				
Shooter games	-1.161 (0.926)	-0.53 (1.563)	-1.086 (1.117)	-1.656 (1.838)
Puzzle games	1.068 (1.040)	1.499 (1.755)	1.258 (1.083)	1.250 (1.781)
Sports games	-1.177 (0.884)	-2.46+ (1.491)	-0.699 (1.060)	-2.797 (1.743)
Driving games	0.891 (1.346)	0.533 (2.271)	1.131 (1.361)	0.579 (2.238)
Construction games	1.043 (1.183)	3.566+ (1.996)	1.241 (1.225)	3.190 (2.015)
Action games	1.938* (0.943)	1.832 (1.591)	2.030* (0.982)	1.290 (1.515)
<u>Toy Play</u>				
Construction-based toys			3.649** (1.516)	5.154* (2.494)
Constant	35.934	44.631	32.397	43.163

*** $p \leq 0.001$; ** $p \leq 0.01$; * $p \leq 0.05$; + $p \leq 0.10$.

Notes: Unstandardized coefficients. Standard errors in parentheses. $n = 211$. OLS—ordinary least squares.

justify use of means rather than medians). The total scores ranged from 6% to 75% correct, showing substantial variation in participants' spatial reasoning skills. Descriptive statistics for the three tests in the instrument are listed in Table 5.

One possible explanation of the distribution of spatial skill scores is that they may be partially explainable by general intelligence. Accordingly, we examined the partial correlations between the three different aspects of spatial thinking and our closest measured proxy for intelligence, standardized test index scores. Participants' mental rotation and penetrative thinking skills showed a significant ($p \leq 0.001$) but weak positive partial correlation of 0.333, indicating that they are somewhat related but different skills associated more closely with each other than with standardized test scores (partial $r = 0.070$). However, disembedding appeared to be a different, unrelated skill (bivariate correlations: mental rotation and disembedding 0.002; penetrative thinking and disembedding 0.076). Low scores on the disembedding test indicated this was the most difficult test, which is corroborated by students' comments after taking the test and is in line with the "high difficulty" rating by ETS.

Personal Characteristics

Demographic and educational data for students are summarized in Table 1. Because the student population is limited in its racial and ethnic diversity, those factors are not considered in the reporting of the results.

Gender

The majority (61.2%, $n = 211$) of the students were males. Our data suggested a statistically significant difference between male and female students in overall spatial skills ($b = -3.4$, $p \leq 0.001$ regression model A1; Table 6 and Fig. 2) and on one of three subscales (mental rotation: $b = -8.3$, $p \leq 0$), with small to medium effect sizes (Cohen's $d = 0.29$ and 0.42) for overall spatial reasoning and mental rotation, respectively. Because of the significant gender difference for mental rotation, we ran all subsequent OLS models with adjustments for gender on both the total spatial skill score and the mental rotation test score.

TABLE 9. COMBINED OLS REGRESSION MODELS PREDICTING TOTAL SPATIAL TEST SCORES

	Model H1 (total score)	Model H2 (mental rotation)	Model J1 (total score)	Model J2 (mental rotation)
<u>Gender</u>				
Female	-3.300* (1.474)	-12.184*** (2.609)	-2.142 (1.61)	-11.159*** (2.617)
<u>Standardized Test Scores</u>				
Standardized Test Index Score (centered on mean)	0.137* (0.06)	0.308** (0.101)	0.146* (0.064)	0.42*** (0.104)
<u>Motivation</u>				
Task orientation	0.025 (0.613)	0.899 (1.032)		
Self-efficacy	2.081** (0.78)	2.284+ (1.317)	2.135** (0.732)	3.595** (1.171)
<u>Selection of Major</u>				
Non-STEM	-4.098** (1.584)	-3.736 (2.670)	-2.555* (1.488)	
Undeclared	-4.124+ (2.362)	-4.885 (3.993)		
<u>Video Games</u>				
Action games	1.701+ (0.94)		1.786+ (0.962)	
Sports games		-1.837 (1.592)		
Construction games		3.511+ (2.072)		3.724* (1.959)
<u>Toy Play</u>				
Construction-based toys			2.712+ (1.598)	3.876 (2.616)
Constant	25.902 (4.279)	30.456 (7.509)	21.974 (4.115)	21.919 (6.430)

*** $p \leq 0.001$; ** $p \leq 0.01$; * $p \leq 0.05$; + $p \leq 0.10$.
Notes: Unstandardized coefficients. Standard errors in parentheses. $n = 225$. OLS—ordinary least squares.

Motivation

Using students' MSLQ scores, we calculated values for the six subscales of that instrument (self-efficacy, control of learning, intrinsic goals, task value, metacognitive regulation, and effort regulation) and entered those subscales as explanatory variables in regression models of spatial skill scores (Table 6). Only two of these six subscales yielded statistically significant results at the $p \leq 0.05$ level after controlling for gender. Students with higher scores on the subscale "task value" (measured on a 7-point Likert scale) scored higher on the spatial skills tests ($b = 1.548$, $p \leq 0.05$, regression model B1, Table 6). "Task value" indicates that students view their learning of a task or content as something important and useful. Similarly, students with higher confidence in

their ability to master a task or learning content ("self-efficacy") demonstrated higher spatial skills ($b = 3.894$, $p \leq 0.0001$). "Effort regulation," a student's ability to continue with their study efforts despite difficulties or distractions (Hilpert et al., 2013), did not quite reach conventional criteria for significance (Table 6). When using mental rotation as the dependent variable instead of the overall score, metacognition showed a negative association with spatial skills ($b = -3.98$, $p \leq 0.05$, regression model B2, Table 6).

In regression model B1, motivation appears to explain the gender gap in overall spatial skills because there is no statistical significance of being female as an explanatory variable. However, when we further control for standardized test scores in regression models C1 and C2 (Table 6), the effect of gender is no longer mediated, and men display higher spatial scores, on average, than

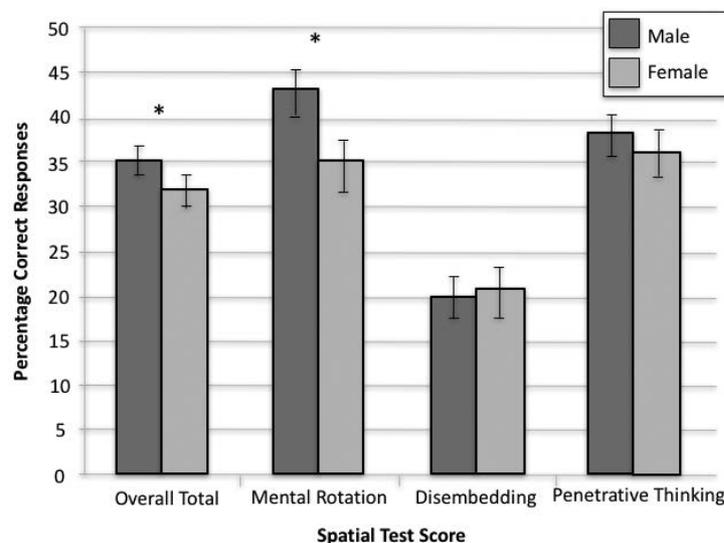


Figure 2. Spatial skill test results by test instrument and gender (males $n = 211$, females $n = 134$; * $p < 0.01$). Error bars represent the 95% confidence interval of the mean.

women with comparable CCHE index scores and levels of motivation. This suggests some suppression of gender differences in spatial test scores when we do not account for previous performance on standardized tests, which may be related to test-taking skills or general intelligence. In fact, women in the sample displayed significantly higher index scores than men ($t = 3.467$, $p < 0.001$), but, as indicated by model C1, this advantage is not enough to account for the gender gap in spatial skills.

Academic and College Preparation

Standardized Test Scores

Regression analysis suggests that students with higher standardized test scores performed significantly better on the spatial skills tests ($b = 0.186$, $p \leq 0.002$, regression model C1, Table 6). However, bivariate correlations between the CCHE academic preparation index and the spatial skills test indicate only a weak positive correlation ($r = 0.242$, $p \leq 0.01$). We interpret these contrasting results to mean that academic preparation test scores cannot fully explain the variation in spatial skills, but may act as a proxy for general intelligence, which is known to correlate with spatial skills (Coyle and Pillow, 2008; Koenig et al., 2008). While selected items on some standardized tests may require spatial reasoning (e.g., Kastens et al., 2014), the standardized test index

used in our study was derived from SAT or ACT scores that have not been shown to accurately measure spatial skills.

Given the theoretical relevance of standardized test scores for spatial reasoning, as well as selection into college (and thus, our sample), we control for academic preparation (i.e., general intelligence) in all subsequent regression models discussed below. However, because this paper is primarily focused on sources of informal spatial skills training rather than innate traits of students, we do not formally discuss the findings for standardized index scores in great detail in subsequent models. It is important to note that any remaining significant associations with spatial skills in the following models are indicative of meaningful differences in spatial skills beyond those that can be explained by general intelligence, alone, as is demonstrated by the persistence of gender differences in models C1 and C2.

Selection of College Major

In this study, 50.7% ($n = 175$) of the participants were non-STEM majors, 37.7% ($n = 130$) were STEM majors (13.3% geology majors; $n = 46$), and 11.6% ($n = 40$) had not yet declared their major. Students pursuing a STEM major displayed ~5% higher spatial skills when compared to the non-STEM majors ($b = -4.7$, $p \leq 0.0001$, regression model D1, Table 7) even after adjusting for standardized test scores (i.e., proxy for general intelligence) and gender. Spatial skill test scores did not differ significantly between the full sample ($N = 345$) and introductory course participants ($n = 276$). The students who had not yet declared a major showed significantly lower spatial scores ($b = -4.7$, $p \leq 0.05$, regression model D1, Table 7) than the STEM majors.

Introductory versus Advanced Geology Course Students

Our data set does not suggest a significant difference between geology majors in introductory geology classes ($n = 17$) and geology majors in advanced geology courses ($n = 29$), the majority of whom are male students (84%). Whether this means spatial skills do not increase with increased training in the discipline at the studied university is unclear. The lack of a significant difference is likely a result of having inadequate statistical power due to the small participant numbers. Differences in spatial reasoning skills between introductory and advanced students would be a fruitful avenue for future research with a larger sample.

Completion of Prior STEM Course Work

Analysis of the number of completed science, Earth science, and engineering and/or architecture courses both in college and in high school for each participant showed, not surprisingly, the number of science courses taken was

moderately correlated with choice of a STEM major ($r = 0.51, p \leq 0.0001$). Thus, we tested the impact of prior STEM course work in regression models E1 (overall skill score) and E2 (mental rotation score) with adjustment for differences in gender and standardized test scores (Table 7). On average, participating students had taken 2.5 science courses ($SD = 1.3$; median [Mdn] = 1.0), 1.4 Earth science courses ($SD = 3.5, Mdn = 1.0$), and 0.5 engineering courses ($SD = 1.8; Mdn = 0.00$). For each additional science course taken, students' spatial scores improved by -0.6% independent of other controls ($b = 0.572, p \leq 0.0001$, model E1, Table 7). No significant correlation was observed for engineering coursework.

Life Experiences and Childhood Play

Experience with Video Gaming

When asked about their video gaming experience and habits, 22.3% of the participating students indicated that they never played video games. Men (79%) were nominally more likely to report playing video games than women (75%), though the difference in play was not significant ($z = 0.821, p = 0.412$). Table 10 shows the categories of games that students most frequently played as well as a weighted average that signifies whether students prefer games from a certain category. Students who indicated that they played action games showed significantly higher spatial skills on the combined test ($b = 1.938, p \leq 0.05$, regression model F1, Table 8) than those who did not play games. Construction games ($b = 3.57, p \leq 0.1$) and sports games ($b = -2.46, p \leq 0.1$) were marginally significant predictors for mental rotation scores. No other video gaming appeared to impact spatial skills in our student sample.

Experience Playing with Construction-Based Toys

Playing with construction-based toys may be an important source of informal spatial training for children that is likely to vary systematically by gender due to the socialization process. Not surprisingly, men (76%) were much more likely to have frequently played with construction-based toys such as Legos as

a child than were women (42%) in the sample, with the difference being highly significant ($z = 5.862, p < 0.001$). Moreover, our data suggest a significant positive association between frequent play with construction-based toys and spatial skills ($b = 3.65, p = 0.017$, regression model G1, Table 8) after controlling for video games, gender, and index test scores, displayed in a 4% increase in spatial skill scores.

Experience with Sports or Physical Activities

Each of the most popular sports or physical activities (Table 4) was entered as a variable into similar regression models (not displayed). However, none of the sports appeared to explain the distribution of spatial skills, because they did not reach statistical significance. It may be that the frequency of play, rather than sport preference, is a more critical source of informal spatial skills training. Future research should consider this dimension of physical activity when investigating its potential association with spatial reasoning.

Combined Experiences

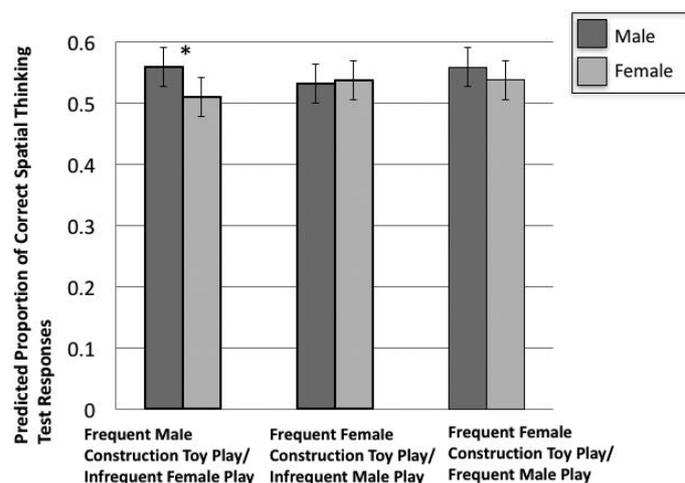
We combine the statistically significant variables identified in models A through G into new multivariate regression models (model H and J, Table 9) to more fully examine how all of these factors, together, explain a wider distribution of students' measured spatial reasoning skills. These final two models (Table 9) include gender, standardized test index scores, two motivation factors (task value and self-efficacy), major, the frequency of playing with construction-based toys, and the frequency of playing certain types of video games (i.e., action, construction, and sports games. Importantly, gender ($b_{\text{female}} = -3.3$, Cohen's $d = 0.28, p < 0.05$), standardized tests ($b = 0.137, p < 0.05$), self-efficacy ($b = 2.1, p < 0.01$), and selection of major ($b_{\text{non-STEM}} = -4.1, p < 0.01$) continue to be important predictors of overall spatial skills (regression model H1). However, the gender effect is fully mediated ($b = -2.142$, Cohen's $d = 0.18, p > 0.10$) once we adjust for the frequency of playing with construction-based toys (model J1; $b = 2.712, p < 0.10$).

TABLE 10. RESPONSES TO THE QUESTION "LIST THE VIDEO GAME THAT YOU PLAYED THE MOST" AS CATEGORIZED INTO SIX CATEGORIES

	Number of participants who played	Weighted value
Shooter games	125	176
Puzzle games	103	138
Sports games	106	159
Driving games	80	87
Construction games	63	86
Action games	133	173

Note: Weighted value includes students listing multiple games of the same type.

Figure 3 depicts what the predicted gender differences in total skill scores (from regression model J1, Table 9) would look like for three different hypothetical scenarios. First is a normative scenario in which men “frequently” played with construction-based toys and women did not. This highlights the gender gap present in the earlier models. The second and third scenarios in Figure 3 demonstrate that there is no gender gap between women who “frequently” played with these toys and men who did not, or between women and men who both played frequently with those toys. Importantly, the other characteristics of the hypothetical male and female students depicted in these scenarios are held constant and reflect the model-predicted spatial skill test scores of a STEM-major student with average CCHSE index scores ($M = 114$), “high” self-efficacy (6 out of 7), and one action game listed among their favorite video games. The height of the bars reflects the resulting predicted values for overall spatial test scores using the regression equation represented by model J1 in Table 9, with only toy and gender manipulated to demonstrate the above-described scenarios; error bars display the 95% confidence interval for gender in the model. The results in Figure 3 imply that, while the variables in models H1 and J1 overall fail to explain the vast majority of the variation in student spatial skill test scores ($R^2 = 0.19$), they do suggest that the widely observed gender disparity in spatial skills may be largely the result of different socialization practices (including different styles of play) during childhood and adolescence.



Note: Predictions based on Model J1 regression results for STEM-major men and women with mean test scores, high self efficiency, and 1 action game listed among their favorite games. Error bars represent the predicted 95% confidence interval for gender. * = Difference is significant, $p < .05$

Figure 3. Model-predicted total spatial skill test score, by gender, for normative and non-normative construction-based toy play scenarios.

DISCUSSION

Prior research demonstrated the importance of spatial reasoning skills for performance, success, and retention in STEM disciplines, particularly in the geosciences. Several meta-analyses have established that spatial skills can be trained through indirect (e.g., course work or video games) and direct interventions (e.g., spatial skills training) (Baenninger and Newcombe, 1989; Uttal et al., 2013). Most studies measured the effect immediately after the intervention; retention of trained skills is much harder to measure and was not part of most prior studies (Uttal et al., 2013). Herein, we focus on the retention of spatial skills through cumulative life experiences as measured in baseline skills early in undergraduate course work. We identified personal characteristics, academic preparations, and life experiences that appear most successful in developing spatial skills by early adulthood. Understanding which factors correlate strongly with spatial skills in undergraduate students enrolled in a geology course provides insight into effective ways that formal and informal training can be shaped.

As summarized in our final regression model (Table 9), our results indicate that among *personal characteristics*, we found gender to be a consistent and significant predictor of spatial skills, among undergraduate students enrolled in geology courses, with women demonstrating lower spatial skills than men even after controlling for several other factors related to spatial reasoning skills. We also found correlations between motivation for learning and spatial skills, suggesting that training and retention of spatial skills throughout life are guided by motivation. Spatial skills are also associated with *academic preparation and training factors* such as standardized test scores, selection of a STEM versus non-STEM major, and number of STEM courses taken. Among other *life experiences*, we find that spatial skills are higher among those who play action-oriented video games and those who played with construction-based toys as children. Importantly, after adjusting for all of these factors, the gender difference in spatial reasoning is no longer significant, implying that gender differences may be a result of greater informal training of boys in childhood and adolescence due to differences in the gender socialization process (Epstein and Ward, 2011).

Many STEM Students Enter College without Adequate Development of Spatial Reasoning Skills

The wide range of spatial reasoning skills displayed among undergraduate students in this study implies that K–12 education does not provide effective formal training of spatial skills for these students. In fact, we likely underestimate the range of spatial skills among graduating high school students because participants in this study have already passed a rigorous research university admissions process. Other studies have suggested significant positive benefits on career outcomes for individuals with higher spatial skills (e.g., Shea et al., 2001; Webb et al., 2007; Wai et al., 2009; Kell et al., 2013). K–12

education provides an obvious opportunity for training spatial skills, especially given the range of simple interventions that are likely effective in improving academic performance when provided at early ages.

The observed weak correlation between standardized test scores and spatial skill performance indicates that a systematic assessment of spatial skills in all students would reveal information about a yet-untested layer of students' academic potential. Similar findings have been reported from longitudinal studies (Shea et al., 2001; Webb et al., 2007; Wai et al., 2009). Thus, quantifying spatial skills could facilitate optimized support for all students (e.g., through remediation or additional training for students with low spatial skill scores) and also provide opportunities for identifying talent that may not be apparent in standardized test scores alone (e.g., Kell et al., 2013). Providing formal training opportunities might also increase the number of students who are cognitively able to succeed in spatial tasks and, thus, increase the potential pool of students who successfully enter a geoscience or STEM career. Also, students with high standardized test scores but low spatial skills could potentially participate in training interventions before entering a geoscience or STEM discipline to increase their success of solving spatially demanding tasks that might otherwise prove difficult.

We find that students with an interest in and motivation for learning (task orientation), especially those with high self-efficacy about their learning success, display significantly higher spatial skills even after controlling for a proxy for general intelligence (standardized test scores). This result supports previous findings that have documented the positive influence of the affective domain (and, thus, motivation) on student learning in general (e.g., Wolters and Pintrich, 1998; Robbins et al., 2004; Dweck, 2006; McConnell and van Der Hoeven Kraft, 2011). Our findings imply that students with an interest in learning and high confidence about their learning success might have developed stronger spatial reasoning skills than their less motivated, less confident peers.

Our results also suggest that STEM college majors have statistically significantly higher spatial skills than students pursuing non-STEM majors. However, our study design does not identify causality. Are students with more developed spatial skills more likely to enter STEM disciplines? Or, do the science courses students take as part of their major train spatial skills in students? A closer look at the students who had not yet declared a major ($n = 40$) provides some insight. A year after the assessment, seven of the students had selected a STEM major, and 14 had selected a non-STEM major (12 students dropped out; eight were still undeclared). An independent samples t -test reveals that while these students' overall spatial skill scores do not exhibit a significant difference by newly declared major ($t = 1.15$, $p = 0.264$), the STEM group's mean scores for mental rotation ($M = 38.5$, $SD = 8.15$, overall; $M = 50.0$, $SD = 17.32$) are higher than those of the non-STEM group ($M = 32.9$, $SD = 11.52$, overall; $M = 37.1$, $SD = 15.90$), implying that some aspect of spatial skills may drive the selection of major. Further work is necessary to confirm this trend with a larger sample size and with a research design more suitable for identifying causality.

Spatial Reasoning Skills Need Not Be Gender Specific

Increasing the diversity of geoscience or STEM graduates (and, subsequently, the geoscience or STEM workforce) is an important goal, both with respect to equity concerns and because the potential for increasing participation in STEM is highest in populations that are currently underrepresented. Females are an underrepresented group frequently targeted by programs designed to increase participation in the STEM workforce (Burke, 2007; Griffith, 2010; Fealing et al., 2015). As predicted by previous research (Linn and Petersen, 1985; Baenninger and Newcombe 1989; Voyer et al., 1995; Parsons et al., 2004; Terlecki et al., 2008), males in our sample significantly outperformed females on spatial skill scores (Tables 6–9). However, the majority of the observed gender difference stems from different mental rotation skills. Depending on model specifications, males achieve 8% to 16% higher performance on mental rotation tests than females with medium to medium-large effect sizes (Cohen's $d = 0.42$ to 0.63). No statistically significant gender differences were apparent for the other spatial skills tested (disembedding and penetrative thinking) or for the overall measure in the most fully developed model (model J1, Table 9).

In thinking about mediating the gender difference, an obvious question is whether observed disparities can be attributed to biological differences between the sexes or to sociocultural experiences inherent in the gender socialization process. Our findings indicate that sociocultural experiences, such as types of childhood toy play, can explain gender differences in spatial skills. In our study, 62% of students (75.6% of males and 41.5% of females) self-reported that they played frequently with construction-based toys. Students who played with construction-based toys as children showed significantly better performance on spatial skills tests, especially the mental rotation test. Once we also accounted for differences in motivation, test scores (i.e., intelligence), and major (model J1), play type fully mediates the gender disparity in total spatial skill scores (Table 9). Further, this mediation is not just an artifact of multicollinearity, since our analysis reveals only a weak correlation between gender and the frequency of playing with construction-based toys ($r = 0.344$, $p \leq 0.01$). The data imply that childhood play with construction-based toys has a lasting training effect on males and females. Our study thus expands on the findings of Richardson (1994) and others (e.g., Moè, 2009) who found that gender differences observed for spatial skills may be attenuated by educational and sociocultural experiences, and that those experiences may be more important than biological predisposition for explaining differences in spatial skills. This is given further support in our sample, which exhibited significantly higher scores among female students ($t = 3.467$, $p < 0.001$) than their male counterparts on a proxy for general intelligence. These results point to the importance (and opportunity!) of sociocultural experiences over biological explanations. Thus, educational interventions, as well as less formal interventions such as encouraging construction-based play in girls, are likely to eliminate gender differences in spatial reasoning that may deter entry and success in geoscience or STEM fields.

In contrast, when we focus on mental rotation skills, the gender difference in our study does not diffuse with frequency of childhood play (model J2; Table 9), likely because the gender difference for mental rotation skills is very strong (model H2; Table 9). A controlled experimental study (Wolfgang et al., 2003) in which nine- to 14-year-olds were engaged in 20 hours of Lego play similarly implied persistence in spatial skill differences between the genders. This may imply that the cumulative effect of childhood play, including play at a young age, may be important for forming spatial reasoning skills for both genders. This result is important because it demonstrates the persistence of the training effects of early childhood activities.

Spatial Thinking Skills Can Be Developed in Many Ways

The cumulative effect of video gaming throughout an individual's life appears to be correlated with an increase in overall spatial skills for those who frequently played action games, with weaker evidence suggesting potential impacts on mental rotation for those playing construction-based and sports-focused video games. This implies that the cumulative training of video game play over many years may have a lasting effect. It also appears to be detectable even outside a controlled setting, though the effect does seem to differ by the types of games being played. Our results corroborate findings from others who suggested that spatial reasoning skills improve with exposure to video gaming, especially with action games (Subrahmanyam and Greenfield, 1994; Feng et al., 2007; Adams and Mayer, 2012), sports games, and construction games (e.g., Okagaki and Frensch, 1994; De Lisi and Wolford, 2002; Sims and Mayer, 2002; Terlecki et al., 2008). However, we did not collect the needed data to test whether these findings are attributable to genre or frequency of play (e.g., students preferring action games may play more frequently than those playing other types of games).

Most of the prior studies on gaming explored the effect of controlled interventions (consisting of video game play) on gains in spatial skills using experimental designs. Such studies suggested that gaming's short-term effect can account for the gender gap in spatial reasoning (Feng et al., 2007; Terlecki et al., 2008). However, the long-term data about video gaming as captured in our study suggest that the mediation effect of gender from video gaming may be more modest than gaming's short-term effect, because we found no evidence for such mediation. While the self-reported data we collected for childhood play are subject to recall bias, there is little reason to suggest that this would systematically influence our findings, which are largely in line with prior work. Video gaming thus may offer one potential avenue for lasting interventions to develop spatial skills. The play-based nature of such training might appeal to many student populations, even students who are difficult to motivate with more traditional learning approaches.

We did not detect a statistically significant correlation between sports play and students' spatial skills. A lack of specificity in the questionnaire may explain this result. Specifically, we did not document the amount of time that

participants spent playing sports. Therefore, the results do not distinguish between athletes who engaged in a sport multiple times a week over years versus students who recreationally played the same sport on an occasional basis. Studies that found improvements in spatial skills with sports play (Ozel et al., 2002, 2004; Moreau et al., 2011, 2012) only included athletes who play sports frequently or students who were enrolled in regular sports programs, which may help account for why we did not observe a significant association of sports with any spatial skill outcomes. The long-term training effects of different sports should be the subject of future research, though it would be important to distinguish between type and frequency of play.

Limitations

The factors explored in the OLS models combine to explain only ~20% of the observed variability in the spatial skill test scores ($R^2 = 19\%$ for total spatial skill test results, $R^2 = 24\%$ for mental rotation test results; models H and J; Table 9). Yet, the statistically significant trends in our data imply that the associations described are meaningful. The remaining variability in spatial skills and/or test performance must be associated with factors not isolated in our study. These could include other training experiences such as prior development of navigation skills, training in music, or other arts experience. It is also possible that our instrument did not accurately measure students' maximum spatial reasoning potential, given the low-stakes character of the test (results were not part of the course grade), and, consequently, students might not have given the tests their best efforts. Alternatively, students subject to test anxiety might have been stressed by the time limit we put on our test.

The three spatial skills tested in this study only address visualization of intrinsic (within objects) spatial relations and not extrinsic (between objects) skills. However, extrinsic skills are also highly relevant for success in some STEM disciplines (Newcombe and Shipley, 2015). This may explain, in part, why two of the three spatial skill tests we conducted had relatively low internal consistency—there may be some omitted factors related to extrinsic skills. Future work should expand testing strategies to include extrinsic spatial skills.

Implications

Our findings suggest an opportunity to train spatial skills and provide all students with the necessary tools to succeed in STEM disciplines, including the geosciences. Newly developed science education standards for K–12 education in the NGSS Lead States, 2013) focus on integrating “scientific practices” and “cross-cutting concepts,” providing an opportunity for regular training of spatial skills. Spatial reasoning, however, is only explicitly described for the Earth science curriculum (standard code: MS-ESS2-2, HS-ESS2-1). Curriculum developers who are creating new classroom materials to support teaching the Next Generation Science Standards (NGSS) may address this shortcoming

by including spatially-demanding components in other STEM curricular materials. For example, mental rotation could be part of engineering-focused instruction, or penetrative thinking could be included in chemistry or life science when discussing molecular or cell structures. Even more simply, informal training through instruction that incorporates construction-based toy play or video gaming has the potential to improve spatial reasoning skills and reduce gender disparities without a need to fundamentally restructure curricula.

CONCLUSIONS

Our sample of 345 undergraduate students at a large U.S. research university documents an uneven distribution in intrinsic spatial skills among students—a set of skills that are shown to be an important predictor for success in STEM disciplines. The wide distribution of spatial skills observed demonstrates the need for both formal and informal spatial skill training in order to provide equal opportunities for students in STEM disciplines, including the geosciences. Our results suggest that the lack of formal spatial skill training is mitigated among certain student groups through informal training during extracurricular activities throughout childhood and early adulthood. We find a statistically significant association between spatial skill scores and gender, motivation for learning, standardized test scores, selection of STEM or non-STEM major, prior STEM course work, playing with construction-based toys, and video gaming.

When several important motivational and play experiences are considered, including childhood play with construction-based toys, female students' average spatial skills scores become statistically indistinguishable from male students' scores. This finding indicates that the informal nature of childhood spatial skill development might systematically disadvantage females, who are often socialized to engage in activities that do not prioritize spatial interaction in the same way as traditionally male-affiliated activities (e.g., video gaming or playing with construction-based toys). Our results also highlight opportunities for reducing the gender gap; for example, early interventions, such as encouraging construction-based play in girls, may prove effective for improving spatial reasoning and facilitating success in STEM fields among the female population. Furthermore, systematic testing of spatial skills and formal training opportunities for students throughout their K–12 and college education might increase the number of students who are cognitively able to succeed in spatial tasks and, thus, increase the potential pool of students who successfully enter a geoscience or STEM career.

ACKNOWLEDGMENTS

The authors would like to thank Thomas F. (Tim) Shipley for valuable input in the study design. We thank Jennifer Stroh, Alyssa Quintanilla, Peter Cirkovic, Alexander Swindell, and Mauricio Munoz for data entry; Samuel Gold for coding video games; and Susan Lynds for copyediting the manuscript. Funding for the study was provided to Anne Gold through a Chancellor's Award of the University of Colorado in Boulder and a fellowship from the German Exchange Council (DAAD). Anke

Friedrich hosted Anne Gold during the academic year 2015–2016 at the Ludwig-Maximilian University in Munich, Germany. Comments and suggestions from anonymous reviewers and Guest Associate Editor Julie Libarkin have significantly improved the quality of the manuscript.

REFERENCES CITED

- Adams, D.M., and Mayer, R.E., 2012, Examining the connection between dynamic and static spatial skills and video game performance, *in* Building Bridges Across Cognitive Sciences Around the World: Proceedings of the 34th Annual Meeting of the Cognitive Science Society (CogSci 2012), no. 1, p. 1254–1259.
- Adams, E., 2009, Fundamentals of Game Design: Berkeley, California, New Riders, 700 p.
- Almquist, H., Stanley, G., Blank, L., Hendrix, M., Rosenblatt, M., Hanfling, S., and Crews, J., 2011, An integrated field-based approach to building teachers' geoscience skills: *Journal of Geoscience Education*, v. 59, p. 31–40, <https://doi.org/10.5408/1.3543926>.
- Amorim, M.A., Isableu, B., and Jarraya, M., 2006, Embodied spatial transformations: "Body analogy" for the mental rotation of objects: *Journal of Experimental Psychology: General*, v. 135, no. 3, p. 327–347, <https://doi.org/10.1037/0096-3445.135.3.327>.
- Anderson, M.L., 2003, Embodied cognition: A field guide: *Artificial Intelligence*, v. 149, no. 1, p. 91–130, [https://doi.org/10.1016/S0004-3702\(03\)00054-7](https://doi.org/10.1016/S0004-3702(03)00054-7).
- Baenninger, M., and Newcombe, N., 1989, The role of experience in spatial test performance: A meta-analysis: *Sex Roles*, v. 20, p. 327–344, <https://doi.org/10.1007/BF00287729>.
- Barsalou, L.W., 2008, Grounded cognition: *Annual Review of Psychology*, v. 59, p. 617–645, <https://doi.org/10.1146/annurev.psych.59.103006.093639>.
- Brosnan, M., 1998, Spatial ability in children's play with Lego blocks: *Perceptual and Motor Skills*, v. 87, p. 19–28, <https://doi.org/10.2466/pms.1998.87.1.19>.
- Burke, R.J., 2007, Women and minorities in STEM: A primer, *in* Burke R.J., and Mattis M.C., eds., Women and Minorities in Science, Technology, Engineering, and Mathematics, Upping the Numbers: Cheltenham, UK, Edward Elgar Publishing Limited, p. 3–27, <https://doi.org/10.4337/9781847206879.00008>.
- CCHE (Colorado Commission of Higher Education), 2015, Admissions Eligibility: <http://higher.ed.colorado.gov/academics/admissions/> (December 2016).
- Chatterjee, A., 2008, The neural organization of spatial thought and language: *Seminars in Speech and Language*, v. 29, p. 226–238, <https://doi.org/10.1055/s-0028-1082886>.
- Cherney, I.D., 2008, Mom, let me play more computer games: They improve my mental rotation skills: *Sex Roles*, v. 59, p. 776–786, <https://doi.org/10.1007/s11199-008-9498-z>.
- Coxon, S.V., 2012, The malleability of spatial ability under treatment of a FIRST LEGO league-based robotics simulation: *Journal for the Education of the Gifted*, v. 35, no. 3, p. 291–316, <https://doi.org/10.1177/0162353212451788>.
- Coyle, T.R., and Pillow, D.R., 2008, SAT and ACT predict college GPA after removing *g*: *Intelligence*, v. 36, no. 6, p. 719–729, <https://doi.org/10.1016/j.intell.2008.05.001>.
- Crawford, A.B., and Burnham, P.S., 1946, Forecasting college achievement: A survey of aptitude tests for higher education: New Haven, Connecticut, Yale University Press, 291 p.
- De Lisi, R., and Wolford, J.L., 2002, Improving children's mental rotation accuracy with computer game playing: *The Journal of Genetic Psychology*, v. 163, p. 272–282, <https://doi.org/10.1080/00221320209598683>.
- Dweck, C., 2006, *Mindset: The New Psychology of Success*: New York, Random House LLC, 320 p.
- Ekstrom, R.B., French, J.W., Harman, H.H., and Derman, D., 1976, *Manual from Kit of Factor-Referenced Cognitive Tests*: Princeton, New Jersey, Educational Testing Service, 216 p.
- Epstein, M., and Ward, L.M., 2011, Exploring parent-adolescent communication about gender: Results from adolescent and emerging adult samples: *Sex Roles*, v. 65, p. 108–118, <https://doi.org/10.1007/s11199-011-9975-7>.
- Fealing, K.H., Lai, Y., and Myers, S.L., Jr., 2015, Pathways vs. pipelines to broadening participation in the stem workforce: *Journal of Women and Minorities in Science and Engineering*, v. 21, p. 271–293, <https://doi.org/10.1615/JWomenMinorScienEng.2015004760>.
- Feng, J., Spence, I., and Pratt, J., 2007, Playing an action video game reduces gender differences in spatial cognition: *Psychological Science*, v. 18, p. 850–855, <https://doi.org/10.1111/j.1467-9280.2007.01990.x>.
- Granic, I., Lobel, A., and Engels, R.C., 2014, The benefits of playing video games: *The American Psychologist*, v. 69, p. 66–78, <https://doi.org/10.1037/a0034857>.

- Green, C.S., and Bavelier, D., 2003, Action video game modifies visual selective attention: *Nature*, v. 423, no. 6939, p. 534–537, <https://doi.org/10.1038/nature01647>.
- Griffith, A. L., 2010, Persistence of women and minorities in STEM field majors: Is it the school that matters?: *Economics of Education Review*, v. 29, p. 911–922.
- Guay, R., 1976, *Purdue Spatial Visualization Test*: West Lafayette, Indiana, Purdue Research Foundation, 18 p.
- Hambrick, D.Z., Libarkin, J.C., Petcovic, H.L., Baker, K.M., Elkins, J., Callahan, C.N., Turner, S.P., Rench, T.A., and LaDue, N.D., 2012, A test of the circumvention-of-limits hypothesis in scientific problem solving: The case of geological bedrock mapping: *Journal of Experimental Psychology: General*, v. 141, p. 397–403, <https://doi.org/10.1037/a0025927>.
- Hegarty, M., 2014, Spatial thinking in undergraduate science education: *Spatial Cognition and Computation*, v. 14, p. 142–167, <https://doi.org/10.1080/13875868.2014.889696>.
- Hegarty, M., and Waller, D., 2005, Individual differences in spatial abilities, in Shah, P., and Miyake, A., eds., *The Cambridge Handbook of Visuospatial Thinking*: Cambridge University Press, p. 121–169.
- Herrera, J.S., and Riggs, E.M., 2013, Relating Gestures and Speech: An analysis of students' conceptions about geological sedimentary processes: *International Journal of Science Education*, v. 35, p. 1979–2003, <https://doi.org/10.1080/09500693.2013.775609>.
- Hilpert, J.C., Stempien, J., van der Hoeven Kraft, K.J., and Husman, J., 2013, Evidence for the latent factor structure of the MSLQ: *SAGE Open*, v. 3, no. 4, doi:10.1177/2158244013510305.
- Kastens, K., Pistoletti, L., and Passow, M.J., 2014, Analysis of spatial concepts, spatial skills and spatial representations in New York State Regents Earth Science examinations: *Journal of Geoscience Education*, v. 62, p. 278–289, <https://doi.org/10.5408/13-104.1>.
- Kastens, K.A., Agrawal, S., and Liben, L.S., 2009, How students and field geologists reason in integrating spatial observations from outcrops to visualize a 3D geological structure: *International Journal of Science Education*, v. 31, p. 365–393, <https://doi.org/10.1080/09500690802595797>.
- Kell, H.J., Lubinski, D., Benbow, C.P., and Steiger, J.H., 2013, Creativity and technical innovation spatial ability's unique role: *Psychological Science*, v. 24, p. 1831–1836, <https://doi.org/10.1177/0956797613478615>.
- Kessler, K., and Thomson, L.A., 2010, The embodied nature of spatial perspective taking: Embodied transformation versus sensorimotor interference: *Cognition*, v. 114, p. 72–88, <https://doi.org/10.1016/j.cognition.2009.08.015>.
- Koenig, K.A., Frey, M.C., and Detterman, D.K., 2008, ACT and general cognitive ability: *Intelligence*, v. 36, p. 153–160, <https://doi.org/10.1016/j.intell.2007.03.005>.
- Kozhevnikov, M., Motes, M.A., and Hegarty, M., 2007, Spatial visualization in physics problem solving: *Cognitive Science*, v. 31, p. 549–579, <https://doi.org/10.1080/15326900701399897>.
- Lakoff, G., 2012, Explaining embodied cognition results: *Topics in Cognitive Science*, v. 4, p. 773–785, <https://doi.org/10.1111/j.1756-8765.2012.01222.x>.
- Linn, M.C., and Petersen, A.C., 1985, Emergence and characterization of sex differences in spatial ability: A meta-analysis: *Child Development*, v. 56, p. 1479–1498, <https://doi.org/10.2307/1130467>.
- Maeda, Y., Yoon, S.Y., Kim-Kang, K., and Imbrie, P.K., 2013, Psychometric properties of the Revised PSVTR for measuring the first year engineering students' spatial ability: *International Journal of Engineering Education*, v. 29, p. 763–776.
- McConnell, D.A., and van Der Hoeven Kraft, K.J., 2011, Affective domain and student learning in the geosciences: *Journal of Geoscience Education*, v. 59, p. 106–110, <https://doi.org/10.5408/1.3604828>.
- Moè, A., 2009, Are males always better than females in mental rotation?: Exploring a gender belief explanation: *Learning and Individual Differences*, v. 19, p. 21–27, <https://doi.org/10.1016/j.lindif.2008.02.002>.
- Moreau, D., Mansy-Dannay, A., Clerc, J., and Guerrien, A., 2011, Spatial ability and motor performance: Assessing mental rotation processes in elite and novice athletes: *International Journal of Sport Psychology*, v. 42, p. 525–547.
- Moreau, D., Clerc, J., Mansy-Dannay, A., and Guerrien, A., 2012, Enhancing spatial ability through sport practice: *Journal of Individual Differences*, v. 33, p. 83–88, <https://doi.org/10.1027/1614-0001/a000075>.
- Myers, C.T., 1953, *Validity study of six spatial relations tests*: Princeton, New Jersey, Educational Testing Service (Research memorandum), 18 p.
- National Research Council, 2006, *Learning to Think Spatially*: Washington, D.C., The National Academies Press, 332 p., <https://doi.org/10.17226/11019>.
- NGSS Lead States, 2013, *Next Generation Science Standards: For States, By States*: Washington, D.C., The National Academies Press, 532 p., <https://doi.org/10.17226/18290>.
- Newcombe, N.S., and Shipley, T.F., 2015, Thinking about spatial thinking: New typology, new assessments, in Gero, J.S., ed., *Studying Visual and Spatial Reasoning for Design Creativity*: Netherlands, Springer, p. 179–192, https://doi.org/10.1007/978-94-017-9297-4_10.
- Nielsen, C.P., Oberle, A., and Sugumaran, R., 2011, Implementing a high school level geospatial technologies and spatial thinking course: *The Journal of Geography*, v. 110, no. 2, p. 60–69, <https://doi.org/10.1080/00221341.2011.534171>.
- Okagaki, L., and Frensch, P.A., 1994, Effects of video game playing on measures of spatial performance: Gender effects in late adolescence: *Journal of Applied Developmental Psychology*, v. 15, p. 33–58, [https://doi.org/10.1016/0193-3973\(94\)90005-1](https://doi.org/10.1016/0193-3973(94)90005-1).
- Ormand, C.J., Manduca, C., Shipley, T.F., Tikoff, B., Harwood, C.L., Atit, K., and Boone, A.P., 2014, Evaluating geoscience students' spatial thinking skills in a multi-institutional classroom study: *Journal of Geoscience Education*, v. 62, p. 146–154, <https://doi.org/10.5408/13-0271>.
- Ozel, S., Larue, J., and Molinaro, C., 2002, Relation between sport activity and mental rotation: Comparison of three groups of subjects: *Perceptual and Motor Skills*, v. 95, p. 1141–1154, <https://doi.org/10.2466/pms.2002.95.3f.1141>.
- Ozel, S., Larue, J., and Molinaro, C., 2004, Relation between sport and spatial imagery: Comparison of three groups of participants: *The Journal of Psychology*, v. 138, no. 1, p. 49–64, <https://doi.org/10.3200/JRPL138.149-64>.
- Parsons, T.D., Larson, P., Kratz, K., Thiebaut, M., Bluestein, B., Buckwalter, J.G., and Rizzo, A.A., 2004, Sex differences in mental rotation and spatial rotation in a virtual environment: *Neuropsychologia*, v. 42, p. 555–562, <https://doi.org/10.1016/j.neuropsychologia.2003.08.014>.
- Pietsch, S., and Jansen, P., 2012, Different mental rotation performance in students of music, sport and education: *Learning and Individual Differences*, v. 22, no. 1, p. 159–163, <https://doi.org/10.1016/j.lindif.2011.11.012>.
- Pintrich, P.R., Smith, D.A.F., Garcia, T., and McKeachie, W.J., 1991, *A Manual for the Use of the Motivated Strategies for Learning Questionnaire (MSLQ)*: Ann Arbor, Michigan, National Center for Research to Improve Postsecondary Teaching and Learning, 76 p.
- Richardson, J.T., 1994, Gender differences in mental rotation: *Perceptual and Motor Skills*, v. 78, p. 435–448, <https://doi.org/10.2466/pms.1994.78.2.435>.
- Robbins, S.B., Lauver, K., Le, H., Davis, D., and Langley, R., 2004, Do psychosocial and study skill factors predict college outcomes?: A meta-analysis: *Psychological Bulletin*, v. 130, p. 261–288, <https://doi.org/10.1037/0033-2909.130.2.261>.
- Shea, D.L., Lubinski, D., and Benbow, C.P., 2001, Importance of assessing spatial ability in intellectually talented young adolescents: A 20-year longitudinal study: *Journal of Educational Psychology*, v. 93, p. 604–614, <https://doi.org/10.1037/0022-0663.93.3.604>.
- Sims, V.K., and Mayer, R.E., 2002, Domain specificity of spatial expertise: The case of video game players: *Applied Cognitive Psychology*, v. 16, p. 97–115, <https://doi.org/10.1002/acp.759>.
- Sorby, S.A., 2001, A course in spatial visualization and its impact on the retention of female engineering students: *Journal of Women and Minorities in Science and Engineering*, v. 7, p. 153–172, <https://doi.org/10.1615/JWomenMinorSciEng.v7i2.50>.
- Spence, I., and Feng, J., 2010, Video games and spatial cognition: *Review of General Psychology*, v. 14, p. 92–104, <https://doi.org/10.1037/a0019491>.
- Subrahmanyam, K., and Greenfield, P.M., 1994, Effect of video game practice on spatial skills in girls and boys: *Journal of Applied Developmental Psychology*, v. 15, no. 1, p. 13–32, [https://doi.org/10.1016/0193-3973\(94\)90004-3](https://doi.org/10.1016/0193-3973(94)90004-3).
- Suitner, C., Giacomantonio, M., and Maass, A., 2015, Embodied Social Cognition, in Wright, J.D., ed., *International Encyclopedia of the Social & Behavioral Sciences (Second Edition)*: Amsterdam, The Netherlands, Elsevier Ltd., p. 409–414.
- Terlecki, M.S., Newcombe, N.S., and Little, M., 2008, Durable and generalized effects of spatial experience on mental rotation: Gender differences in growth patterns: *Applied Cognitive Psychology*, v. 22, p. 996–1013, <https://doi.org/10.1002/acp.1420>.
- Titus, S., and Horsman, E., 2009, Characterizing and improving spatial visualization skills: *Journal of Geoscience Education*, v. 57, p. 242–254, <https://doi.org/10.5408/1.3559671>.
- Tversky, B., and Hard, B.M., 2009, Embodied and disembodied cognition: Spatial perspective-taking: *Cognition*, v. 110, p. 124–129, <https://doi.org/10.1016/j.cognition.2008.10.008>.
- Uttal, D.H., and Cohen, C.A., 2012, Spatial thinking and STEM education: When, why and how?: *Psychology of Learning and Motivation*, v. 57, p. 147–181, <https://doi.org/10.1016/B978-0-12-394293-7.00004-2>.

- Uttal, D.H., Meadow, N.G., Tipton, E., Hand, L.L., Alden, A.R., Warren, C., and Newcombe, N.S., 2013, The malleability of spatial skills: A meta-analysis of training studies: *Psychological Bulletin*, v. 139, p. 352–402, <https://doi.org/10.1037/a0028446>.
- Voyer, D., Voyer, S., and Bryden, M.P., 1995, Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables: *Psychological Bulletin*, v. 117, p. 250–270, <https://doi.org/10.1037/0033-2909.117.2.250>.
- Wai, J., Lubinski, D., and Benbow, C.P., 2009, Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance: *Journal of Educational Psychology*, v. 101, p. 817–835, <https://doi.org/10.1037/a0016127>.
- Webb, R.M., Lubinski, D., and Benbow, C.P., 2007, Spatial ability: A neglected dimension in talent searches for intellectually precocious youth: *Journal of Educational Psychology*, v. 99, p. 397–420, <https://doi.org/10.1037/0022-0663.99.2.397>.
- Wilson, M., 2002, Six views of embodied cognition: *Psychonomic Bulletin & Review*, v. 9, p. 625–636, <https://doi.org/10.3758/BF03196322>.
- Witkin, H.A., Moore, C.A., Goodenough, D.R., and Cox, P.W., 1977, Field-dependent and field-independent cognitive styles and their educational implications: *Review of Educational Research*, v. 47, p. 1–64, <https://doi.org/10.3102/00346543047001001>.
- Wolfgang, C., Stannard, L., and Jones, I., 2003, Advanced constructional play with LEGOs among preschoolers as a predictor of later school achievement in mathematics: *Early Child Development and Care*, v. 173, p. 467–475, <https://doi.org/10.1080/0300443032000088212>.
- Wolters, C.A., and Pintrich, P.R., 1998, Contextual differences in student motivation and self-regulated learning in mathematics, English, and social studies classrooms: *Instructional Science*, v. 26, p. 27–47, <https://doi.org/10.1023/A:1003035929216>.
- Yoon, S.Y., 2011, Psychometric properties of the Revised Purdue Spatial Visualization Tests: Visualization of Rotations (The Revised PSVT:R) [Ph.D. thesis]: Lafayette, Indiana, Purdue University: Retrieved from ProQuest Dissertations and Theses (Order Number: 3480934).