




# Spatial Training and Calculus Ability: Investigating Impacts on Student Performance and Cognitive Style

McCunn, L. J., PhD

Vancouver Island University, Nanaimo, British Columbia, Canada

 <https://orcid.org/0000-0002-2246-6229>

Cilli-Turner, E., PhD

University of La Verne, La Verne, California, United States

Contact: [lindsay.mccunn@viu.ca](mailto:lindsay.mccunn@viu.ca)

## Abstract

Undergraduate calculus is a foundational mathematics sequence that previews the sophistication students will need to succeed in higher-level courses. However, students often struggle with concepts in calculus because they are more abstract and visual than those in other foundational mathematics courses. Additionally, women continue to be underrepresented in the STEM fields. This study builds on previous work indicating a malleability in spatial ability by testing whether improvement occurs in students' spatial and mathematics ability after implementing spatial training in calculus courses. The researchers also measured associations between spatial training and self-reported cognitive style. While spatial training did not significantly improve calculus and spatial skills, the researchers measured impacts on the psychological correlate of cognitive learning style. This is important as non-spatial cognitive preferences can result in students not recognizing the dynamic information encapsulated in graphical images, which is a large component of calculus. These results may have practical implications for teaching calculus at the undergraduate level and may, with further research, help to narrow the gender gap in the STEM fields.

Keywords: *Calculus; spatial training; cognitive style; gender*

Date Submitted: January 24, 2020 | Date Published: October 4, 2020

### Recommended Citation

McCunn, L. J., & Cilli-Turner, E. (2020). Spatial training and calculus ability: Investigating impacts on student performance and cognitive style. *Journal of Educational Research and Practice*, 10, 317–337.  
<https://doi.org/10.5590/JERAP.2020.10.1.20>

## Introduction

A concentrated effort is being made to achieve greater diversity in students graduating with bachelor's degrees in Science, Technology, Engineering, and Mathematics (STEM) (Redmond & Gutke, 2019). Unfortunately, the demographics of STEM graduates have been stagnant in recent years (Stieff & Uttal, 2015). In particular, women are often underrepresented in STEM fields in North America and Europe (Nimmegern, 2016; Schlenker, 2015), and those who do undertake STEM studies seem to have a higher probability of not working

---

Note: We acknowledge and thank the School of Interdisciplinary Arts and Sciences at the University of Washington Tacoma for partially funding this research. Corresponding author: Lindsay J. McCunn, PhD, Department of Psychology, Vancouver Island University, 900 Fifth Street, Nanaimo, BC V9R 5S5, [lindsay.mccunn@viu.ca](mailto:lindsay.mccunn@viu.ca). Emily Cilli-Turner, PhD, University of La Verne, Founders Hall 108C, La Verne, CA, 91750, [ecilli-turner@laverne.edu](mailto:ecilli-turner@laverne.edu).

---

in STEM occupations after graduation when compared to men (Organization for Economic Cooperation and Development [OECD], 2012).

Although several psychosocial reasons for the underrepresentation of women choosing to study the STEM fields exist (e.g., Nimmegern, 2016; Saucerman & Vasquez, 2014; Skolnik, 2015), one important reason may be that females can perform spatial tasks differently than males—starting as young as four years of age (Dawson, 2019; Levine et al., 1999; Voyer, et al., 1995; Voyer et al., 2017). Men often (but not always) have been found to show higher levels of accuracy and organization of spatial information compared to women (e.g., Acredolo, 1988; Brown et al., 1998; Cutmore et al., 2000; Gärling et al., 1981; Gifford, 2007; Hawes & Ansari, 2020; Kirasic et al., 1984; Lehnung et al., 2003; Ward et al., 1986; Webley & Whalley, 1987). A number of studies have found insignificant differences between the performances of men and women on various spatial tasks, but results often depend on the type of task and the age of the study (see Casey, 2013). Indeed, research from the 1980s suggests that men seem to perform consistently better than women on spatial perception and mental rotation tasks (Linn & Petersen, 1985), but recent studies have shown that this gender difference may be diminishing (Dawson, 2019).

More recent research concerning differences between men and women with respect to their participation in the STEM fields indicates that women are 1.5 times more likely to abandon the STEM pipeline after calculus compared to men—and that a lack of mathematical confidence to persist in mathematics is a key reason why (Ellis et al., 2016). Gender disparity in STEM fields may relate less to ability (Lachance & Mazzocco, 2006; Spelke, 2005) and more to confidence and limited support structures (Hargreaves et al., 2008; Redmond & Gutke, 2019). These results mirror those outlined in a report completed by the American Association of University Women in 2010 highlighting the role confidence plays when women choose (and subsequently succeed in) STEM fields (e.g., Cooper & Robinson, 1991; Dweck, 2006; Lent et al., 1986; Pajares, 1996, 2005; Saucerman & Vasquez, 2014; Zimmerman & Martinez-Pons, 1990).

Although the body of literature on spatial ability indicates that this skill is an important facet of intelligence (Johnson & Bouchard, 2005), and that it can be used as a key performance indicator for success in calculus and the STEM fields in general (Dawson, 2019; National Science Board, 2010; Sorby et al., 2013; Wang, 2019), more must be done to understand these relationships. Burte et al. (2017) note that spatial skills may need to develop prior to entering university and that “developing spatial skills well before high school may have a more pronounced impact on STEM outcomes” (p. 2). Recently, an association between spatial ability and success in the STEM fields has been outlined in a landmark longitudinal study (i.e., “Project Talent”) after following the lives of high school students from the 1950s to the present day (Wai et al., 2009). The study indicates that those with high scores on spatial tests have been more likely to enjoy the STEM fields and to gravitate toward STEM careers, over and above the effects of mathematical and verbal ability (Wai et al.). A previous study by the same authors found that mathematically-talented participants who chose careers in math and science had excelled in object-based skills earlier in life (Lubinsky & Benbow, 2006).

Fortunately, a lack of natural spatial skills does not dictate a lack of success in the STEM fields. Encouraging evidence exists suggesting that one can train to improve spatial ability (Stieff & Uttal, 2015; Uttal, 2009), and that such training may help close the gender gap in spatial cognition (Hill et al., 2010; Newcombe, 2010; Newcombe, 2017). Spatial training refers to the explicit teaching of spatial skills, often through the use of mental rotation tasks. At this time, the body of literature is not conclusive concerning whether spatial training associates with an increase in spatial ability or whether it has direct effects on performance in the STEM fields. For example, Hawes et al. (2017) explored geometric spatial learning in young learners’ (e.g., Kindergarten to grade 2) performance in spatial language, visual-spatial reasoning, 2D mental rotation, and symbolic number comparison. Similarly, Taylor and Hutton (2013) implemented a program for elementary-aged children involving origami and pop-up paper engineering to encourage visuospatial thinking (see also Burte et al., 2017). And, Stieff et al. (2016) examined the impacts of teaching students to use gestures to

support spatial thinking in chemistry. While these and other findings support the use of gestures and visuospatial training programs in facilitating learning in STEM disciplines, they do not use the spatial training format utilized in the present research with undergraduate calculus students.

In fact, although studies done with students enrolled in engineering courses have found an association between spatial training and academic performance, as well as a closing of the gender gap (Sorby et al., 2013), few studies have been done to correlate spatial training and performance in mathematics. Visuospatial training has been shown to reduce the gender gap on spatial tasks (Baenninger & Newcombe, 1989; Newcombe, 2017; Rodán et al., 2019; Vasta et al., 1996) and math performance (Dawson, 2019; Linn & Hyde, 1989). This suggests that, in some format, spatial training may affect confidence and the retention of females in STEM fields (Casey et al., 2001; Linn & Hyde).

Another example is a thirty-year old study by Ferrini-Mundy (1987) that asked undergraduate students to complete spatial training exercises during a calculus course. Although a significant increase in calculus performance was not found, female students were better able to visualize solids of revolution after spatial training when compared to male students (Ferrini-Mundy). Other research has found that while spatial training can reduce the gender gap in performance on spatial tasks, it fails to eliminate it (Uttal, 2009). Thus, a call has been made in the spatial cognition literature to extend the line of enquiry into the potential for spatial training to close the gender gap in the STEM fields by investigating new variables and extending periods of spatial training (Casey, 2013).

A correlate to spatial and mathematical ability may be the psychological construct of cognitive style. A cognitive style represents consistency in an individual's manner of cognitive functioning (i.e., information acquisition and processing) (Harvard Mental Imagery and Human-Computer Interaction Lab, 2013). Researchers at Harvard University have investigated cognitive styles that describe individuals' preferences to, or self-assessments of, the use of object, spatial, or verbal modes of information processing. Because the human visual system distinctly processes properties about objects (color, shape) and space (location and spatial relations), Kozhevnikov et al. (2005) have used neuropsychological evidence to propose the Object-Spatial-Verbal theoretical model of cognitive style.

The theoretical model outlines three independent dimensions (object imagery, spatial imagery, and verbalization) to explain that object visualizers prefer to construct vivid, concrete, and detailed images of individual objects while, in contrast, spatial imagers schematically represent spatial relations of objects and spatial transformations. Verbalizers prefer to process and represent information verbally and rely on non-visual strategies (Kozhevnikov et al., 2005). In their examination of the three types of cognitive styles, Kozhevnikov et al. found that verbalizers performed at an intermediate level on imagery tasks (i.e., not at the low level one might expect) and that object visualizers "encode and process images holistically as a single perceptual unit, while spatial visualizers generate and process images analytically, part-by-part" (p. 710).

It seems that individuals often prefer to use one style over another (Kozhevnikov et al., 2005), and that one's preference for a style has been shown to directly relate to performance on either mathematical, object imagery ability, or spatial ability tests (MM Virtual Design, 2016). Thus, cognitive styles may assist math educators with tailoring material, assignments, and visualization media to students' individual differences in cognitive style and decision-making based on visual stimuli. Information presented in ways that satisfy numerous cognitive styles could augment student engagement with information presented in the classroom and, perhaps, encourage a willingness to contribute and collaborate with others.

Presumably, learning and performance based on visual information presented in a manner congruent to one's cognitive style would be more consistent and effective and, perhaps, help close the gender gap in STEM. Women tend to report higher object imagery ratings (Blajenkova et al. 2006) and have a negligible advantage in verbal ability (Hyde & Linn, 1988). Casey (2013) points out that one reason why large gender differences

are found for mental rotation tasks is because verbal strategies are often less effective than holistic mental rotation approaches used more often by men. Thus, measuring cognitive style in association with individuals' spatial ability and an understanding of calculus may afford information about whether those with predominant verbal cognitive styles are women, as well as whether their performance in calculus improves with spatial training. Since success in calculus often predicts success in STEM fields (and calculus involves spatial reasoning), it may be that congruence between cognitive style and material presented or assigned in calculus courses can increase learning and retention (as suggested by Blajenkova & Kozhevnikov, 2008) and, perhaps, increase confidence—a psychological factor that has the potential to significantly narrow the gender gap in the STEM fields.

## Study Aims

This research involves two studies that aim to augment the body of literature regarding the usefulness of spatial training in undergraduate mathematics and explore the following three research questions: (1) What are the impacts of spatial training on undergraduate students' performance in a calculus course? (2) Does spatial training affect male and female students differently? (3) What are the associations between spatial training, calculus ability, and students' cognitive style?

Because academic success in at least one term of calculus is often required in undergraduate STEM programs, calculus courses can serve as “gatekeepers” for STEM fields (Bressoud et al., 2009). If these studies reveal that spatial training is beneficial to performance in a calculus course, math educators may have reason to recommend a spatial training module be added to courses that require students to think abstractly. If results reveal this training to be particularly effective for female students, an argument may exist that spatial training be studied further with respect to gender differences in calculus performance. Finally, if cognitive styles change over the course of a term to become more spatial as students receive spatial training, or if one style associates with strong performance over another style, it may be reasonable to include an assessment of cognitive style at the start of STEM courses in order for instructors to tailor the delivery of information to students depending on the predominant style of the group (or to assist individual students with styles that are more or less likely to respond to spatial requirements of a course). Overall, connecting research on cognitive styles with spatial training methodology may help educators diversify the pool of students studying STEM at the undergraduate level while bolstering the success of those who enter into the STEM fields with a non-spatial cognitive style.

## Methods

### Context

The first of two studies (Study 1) took place in a summer quarter (June to August, 2016) with student participants enrolled in a third-quarter calculus course at a mid-sized state-funded university in the United States. This course covers calculus of sequences and series, vectors and parametric equations, properties of three-dimensional surfaces, as well as integration techniques and approximation, applications of integration, and differential equations.

The second study (Study 2) took place in the following fall quarter (September to December, 2016) in a second-quarter calculus course at the same institution. This course focuses on the calculus of functions of a single variable and emphasizes integral calculus, applications, and problem solving using the tools of calculus. During the same quarter, a control group of students was formed from a second-quarter calculus course taught by a different instructor at the same institution. Students in the control group completed all of the assessments involved in the study but did not receive spatial training.

Students in both of the treatment groups were taught by the same instructor using an inquiry-oriented pedagogy with spatial training incorporated into the courses. Because we assessed introductory-level calculus knowledge, students in either course could be expected to understand items in the knowledge assessment used in both studies.

## Participants

Participants in Study 1 were undergraduate students who had already completed two quarters of calculus and, thus, had some prior knowledge of the subject. Seventeen students (8 males, 9 females) attended class for both rounds of data collection but all but one took part in the study ( $n = 16$ ; 8 males, 8 females, mean age = 21 years,  $SD = 2.02$ ).

Five students (31% of the sample) reported to be concurrently taking a differential equations course during the summer term, while 11 participants (69% of the sample) reported that they were not receiving other forms of math training at the time of the study). These students also reported that they expected to work an average of 12 hours a week ( $SD = 5.00$ ) on studying the course material.

Participants in the treatment group in Study 2 were 13 undergraduate students (8 males, 5 females, mean age = 23 years,  $SD = 4.07$ ) who had completed one less quarter of calculus than those in Study 1 (but had at least one quarter of calculus instruction and so had prior knowledge of some of the course's concepts). Two students (6% of the treatment sample) reported to be concurrently taking another math course during the term, while 11 participants (33% of the treatment sample) reported that they were not receiving other forms of math training at the time of the study. These students also reported that they expected to work an average of 11 hours a week ( $SD = 4.70$ ) on studying the course material.

Participants in the control group were seven undergraduate students (4 males, 3 females, mean age = 26 years,  $SD = 6.31$ ) who, like in Study 2, had only one quarter of previous calculus instruction. Students reported that they expected to work an average of 8.83 hours a week ( $SD = 3.25$ ) on studying the course material (one student in the sample did not respond to this question). Only one student in the control sample reported to be concurrently taking another math course during the fall term.

In addition, participants were in either the treatment or control group based on their course enrollment and not for any other reason. Therefore, our samples are not random, and no matched pairing of participants occurred. A limitation of this study is that the small sample size limits its power and generalizability. However, the study is constrained by the number of students enrolled in the quarters the authors had access to for this research. Our intention is to explore initial evidence for (or against) associations between spatial training and undergraduate students' cognitive style.

## Materials

### The Calculus Concept Inventory

Epstein (2013) designed the Calculus Concept Inventory (CCI) to evaluate how students think about the fundamental concepts in calculus and we used it in the present study to gather a baseline of students' knowledge of calculus as well as to determine the effects of spatial training on conceptual knowledge of calculus at the end of the term. The CCI contains 22 questions about limits and differential calculus only, many of which are visual and require an interpretation of a graph. The use of the CCI was deliberate: we did not want to measure mastery of concepts learned in the current course but to measure instead whether spatial training could impact understanding of previously learned visual topics.

After this study was undertaken, Gleason et al. (2016) published a psychometric analysis outlining weaknesses in the validity of the CCI to measure conceptual knowledge in calculus. Their findings explain that usage of

terminology (e.g., “derivative”) and notation (e.g.,  $f'(x)$ ) in the CCI items may afford an advantage to students who have already seen calculus. However, we believe that the use of this instrument does not impact the results in the present study because the CCI was used differently than it was originally intended by its developers. In the context of the present study, the CCI is used as a measure of ability in introductory-level calculus, rather than as a measure of understanding for particular concepts. Additionally, all students in the present study had previously taken calculus, minimizing the possibility of an advantage to some students over others.

The researchers tested each CCI variable in each condition in both studies for normality based on recommendations by Kline (1997). All items met the criteria for acceptable skewness (values between +3 and -3) and acceptable kurtosis (values between +8 and -8) in each case and, thus, parametric tests were used in the analyses.<sup>1</sup>

#### The Purdue Spatial Visualization Test Rotations

To test students’ spatial ability, the researchers also administered a shortened version (15 items) of the Purdue Spatial Visualization Test: Rotations (PSVT:R) (Guay, 1977). The PSVT:R is a multiple choice test that, per shape, asks students to choose from 5 possibilities a shape that is equivalently rotated as a given shape. This test established each student’s baseline spatial ability to assess improvement throughout the quarter and to determine whether spatial training had any effect on students’ cognitive style.

Each PSVT:R variable in each condition in both studies was tested for normality based on recommendations by Kline (1997). All items met the criteria for acceptable skewness (values between +3 and -3) and acceptable kurtosis (values between +8 and -8) in each case and, thus, parametric tests were used in the analyses.

#### Object-Spatial Imagery and Verbal Questionnaire

The researchers administered the Object-Spatial Imagery and Verbal Questionnaire (OSIVQ) developed by Blazhenkova and Kozhevnikov (2009) (but copyrighted jointly by MM Virtual Design, LLC and Rutgers University) to determine students’ predominant cognitive style. The OSIVQ is a “self-report questionnaire designed to distinguish between three different types of people: 1) object imagers who prefer to construct vivid, concrete and detailed images of individual objects (e.g., visual artists), 2) spatial imagers who prefer to use imagery to schematically represent spatial relations among objects and to perform complex spatial transformations (e.g., scientists), and 3) verbalizers who prefer to use verbal-analytical tools to solve cognitive tasks (e.g., philosophers and linguists)” (MM Virtual Design, 2016, paragraph 1). The OSIVQ consists of 45 questions (an equal number concerning object imagery, spatial imagery ability, and verbal ability) to assess object imagers, spatial imagers, and verbalizers and takes approximately 10 minutes. Each item asked for responses using a 5-point Likert scale ranging from “totally disagree” (1) to “totally agree” (5). Four items (one spatial, three verbal) from the OSIVQ required reverse coding.

Each OSIVQ variable in each condition in both studies was tested for normality based on recommendations by Kline (1997). All items met the criteria for acceptable skewness (values between +3 and -3) and acceptable kurtosis (values between +8 and -8) in each case and, thus, parametric tests were used in the analyses.

#### Spatial training

All participants (except those in the control group) received spatial training during the term consisting of the administration and discussion of several tests of spatial ability, such as those described in Wai et al. (2009). In

---

<sup>1</sup> Aside from there being acceptable normality shown in the data, Mann-Whitney tests were not used for comparisons between two groups of five-point Likert scale data because both the Mann-Whitney test and the two-sample t-test often provide the same protection against false negatives, as well as against false positives. These patterns appear to hold true for sample sizes of 10, 30, and 200 per group (de Winter & Dodou, 2010).

addition, exercises from a spatial training workbook developed by Sorby et al. (2013) were used with permission.

In Study 1, students completed an average of 15 minutes of spatial training during each day of class. Exercises in the workbook ranged from assessments of what a given shape would look like when rotated around a given axis, to asking students to draw an object from different angles using different cross-sections, to showing a 2-D expression of an object when asking students to draw an analogous 3-D object. During spatial training, students were asked to discuss the exercises in small groups and come to a consensus on the correct answer before answers were discussed among the class at large.

In Study 2, students completed portions of the spatial training workbook collaboratively in small groups during class time, as was done in Study 1. However, to allow for longer, more in-depth training than in Study 1, students were assigned additional portions of the workbook as homework.

## Procedure

In Study 1, participants completed all tasks in the course's assigned classroom for both rounds of data collection. The four tasks were offered to participants in the following order: consent form; OSIVQ; Visualization and Rotation Purdue Spatial Visualization Test; Calculus Concept Inventory. Thus, the tasks were not completed simultaneously: only when a task was completed was the next task offered to a participant by one of the researchers. Each participant was asked to create a unique codename for him or herself to include on the front page of each task in order to afford direct measurement of change in task performance over time. No calculators or other electronic devices were used during task completion. Each participant was dismissed from the classroom after he or she had completed all tasks.

The amount of time the researchers anticipated that participants would need to read and sign an informed consent form and comfortably complete all tasks precluded regular teaching on the days that data collection took place. Thus, while participants completed each task without significant time pressure (between 45 and 90 minutes) in Study 1, an argument was made for changing this procedure in Study 2. Therefore, in Study 2, for both data collection rounds, participants were given time in the designated classroom to complete the consent form (approximately 5 minutes) and the CCI (approximately 30 minutes). The PSVT:R and the OSIVQ were then offered to participants on their own time (without a time constraint) via "Canvas," the online learning platform used at the institution.

An additional alteration applied to the procedure in Study 2 was an increase in our study's shortened version of the PSVT:R's difficulty level. In Study 1, we chose 15 relatively easy items from the larger PSVT:R to assess students' spatial rotation ability. However, after Study 1 was complete, we felt that these items were insufficiently challenging for students because many received a perfect score, making it difficult to measure any knowledge gain. Therefore, in Study 2, some items were replaced to create an assessment of rotation ability that increased in difficulty level (i.e., while the first 5 items remained relatively easy, the middle 5 items had a moderate level of difficulty and the 5 final items were quite difficult). We chose all items, however, from Guay's (1977) PSVT:R.

Finally, the control group used the same procedure and assessments as Study 2 (except for the lack of spatial training). A small incentive (i.e., no more than three extra points toward a student's course grade) was offered to encourage completion of the online tasks in the control group and in Study 2's treatment group.

## Results

### Responses on the CCI

#### Study 1

Students' mean score on the CCI at the start of the term was 8.86 ( $SD = 3.98$ ) and 9.25 at the end of the term ( $SD = 4.96$ ). Despite an increase in average score, paired-samples  $t$ -tests revealed no significant improvement over the term,  $t(15) = -0.96, p > .05$ .<sup>1</sup>

#### Study 2

Students' mean score on the CCI at the start of the term was 6.69 ( $SD = 4.39$ ) and 8.77 at the end of term ( $SD = 4.66$ ) for the treatment group. Similar to Study 1, despite an increase in mean score, paired-samples  $t$ -tests revealed no significant improvement,  $t(12) = -1.87, p > .05$ .

Control group participants' mean score on the CCI at the start of the term was 8.71 ( $SD = 4.72$ ) and 8.29 at the end of term ( $SD = 4.46$ ). Paired-samples  $t$ -tests revealed no significant improvement over the term,  $t(6) = 0.57, p > .05$ . Moreover, independent samples  $t$ -tests done with average scores on the CCI in the treatment and control groups did not reveal significant differences at the start or end of term (all  $ps > .05$ ).

### Gender Differences on the CCI

#### Study 1

Although male participants' average scores were higher on the CCI than women's at the start of term ( $M = 9.63, SD = 4.47$  and  $M = 7.75, SD = 7.75$ , respectively), an independent-samples  $t$ -test revealed that they were not significantly higher,  $t(14) = 0.94, p > .05$ . This result was also borne out at the end of term whereby male students' average scores on the CCI were insignificantly higher than female students' scores ( $M = 10.25, SD = 2.09$  and  $M = 8.25, SD = 1.40$ , respectively;  $t(14) = 0.80, p < .05$ ).

#### Study 2

As in Study 1, male participants' average scores were higher on the CCI than women' at the start of term ( $M = 7.38, SD = 5.26$  and  $M = 5.60, SD = 2.61$ , respectively), and an independent-samples  $t$ -test revealed that they were not significantly higher,  $t(11) = 0.70, p > .05$ . Again, similar to Study 1, male students' average scores on the CCI were insignificantly higher than female students' scores after the term was complete ( $M = 8.87, SD = 4.76$  and  $M = 8.60, SD = 5.03$ , respectively;  $t(11) = 0.09, p < .05$ ).

Unlike results in Study 1, or in the treatment group in Study 2, male participants' average scores on the CCI in the control group were lower than women's at the start of term ( $M = 7.75, SD = 6.29$  and  $M = 13.33, SD = 4.51$ , respectively), but not significantly so,  $t(5) = -1.30, p > .05$ . However, at the end of term, males' average scores on the CCI were significantly higher than women's scores,  $M = 11.25, SD = 2.87$  and  $M = 4.33, SD = 2.52$ , respectively;  $t(5) = 3.31, p < .05, d = 2.56$ .

### Responses on the PSVT:R

#### Study 1

Similar to results pertaining to the CCI, students' mean score on the PSVT:R at the start of the term was 9.81 ( $SD = 3.25$ ) and 9.88 at the end of the term ( $SD = 3.74$ ). Paired-samples  $t$ -tests revealed no significant improvement over the term,  $t(15) = -0.12, p > .05$ .

---

<sup>1</sup> One-way ANOVA tests were not used because not all samples sizes were greater than 15. A Bonferroni-adjusted alpha level was used where appropriate.



## Study 2

Mirroring results of Study 1, students' mean score on the rotation task in Study 2, at the start of the term, was 10.39 ( $SD = 3.20$ ) and 10.54 at the end of term ( $SD = 3.15$ ). Paired-samples  $t$ -tests revealed no significant improvement over the term,  $t(12) = -0.20, p > .05$ .

Control group participants' mean score on the PSVT:R at the start of the term was 8.29 ( $SD = 3.86$ ) and 8.00 at the end of term ( $SD = 4.65$ ). Again, paired-samples  $t$ -tests revealed no significant improvement in rotation ability over the term  $t(6) = -.22, p > .05$ , and no statistically significant differences were found between average scores on the PSVT:R at the start or end of term (all  $ps > .05$ )

## Gender Differences on the PSVT:R

### Study 1

Similar to scores on the CCI, male participants' average scores on the PSVT:R were higher than women's at the start of term ( $M = 9.87, SD = 4.29$  and  $M = 9.75, SD = 2.05$ , respectively), and at the end of the term ( $M = 10.5, SD = 1.45$  and  $M = 9.25, SD = 1.24$ , respectively). However, independent samples  $t$ -tests showed that these differences were not significant at the start of term,  $t(14) = 0.07, p > .05$  or at the end of term,  $t(14) = 0.66, p < .05$ .

### Study 2

Just as in Study 1, male participants' average scores were higher on the PSVT:R than women's at the start of the term ( $M = 11.00, SD = 2.88$  and  $M = 9.40, SD = 3.78$ , respectively) and an independent-samples  $t$ -test revealed that they were not significantly higher,  $t(11) = 0.40, p > .05$ . At the end of the term, male students' average scores on the PSVT:R were insignificantly higher than female students' scores ( $M = 10.75, SD = 3.81$  and  $M = 10.20, SD = 2.05$ , respectively;  $t(11) = 0.29, p < .05$ ).

In the control group, just as in Study 1, as well as in the treatment group in Study 2, male participants' average scores were higher on the PSVT:R than women's at the start of term ( $M = 12.25, SD = 1.89$  and  $M = 10.67, SD = 2.52$ , respectively). However, an independent-samples  $t$ -test revealed that they were not significantly higher,  $t(5) = 0.96, p > .05$ . At the end of the term, male students' average scores on the PSVT:R continued to be higher than female students' scores ( $M = 11.25, SD = 0.96$  and  $M = 3.67, SD = 3.79$ , respectively). However, while these means appear to be significantly different, an independent samples  $t$ -test revealed that they are not because Levene's test for equality of variances between the two means was statistically significant,  $F = 9.50, p < .05, t(2.19) = 3.39, p < .05$ .

## Responses on the OSIVQ

### Study 1

On average, more students at the start of the term self-identified as object learners ( $M = 48.56, SD = 7.00$ ) than they did as spatial learners ( $M = 47.94, SD = 8.69$ ) or verbal learners ( $M = 41.13, SD = 4.41$ ). Paired sample  $t$ -tests conducted using Bonferroni-adjusted alpha levels of .02 per test (.05/3) revealed that scores on the object subscale were significantly higher than on the verbal subscale,  $t(15) = 3.97, p = .001, d = 0.99$ . In addition, scores on the spatial subscale were significantly higher than those on the verbal subscale,  $t(15) = -2.98, p < .01, d = 0.75$ . No significant differences were found between scores on the spatial and object subscales,  $t(15) = 0.22, p > .05$ .

After the summer term, the object style remained predominant for the class as a whole ( $M = 50.81, SD = 7.87$ ) and, just as at the start of the term, the second-most common cognitive style among the class was spatial ( $M = 50.75, SD = 8.96$ ), followed by verbal ( $M = 42.75, SD = 4.16$ ). Also similar to the start of the term, these differences were statistically significant: students self-scored significantly higher on the object style subscale compared to the verbal style subscale,  $t(15) = 3.57, p < .01, d = 0.89$ . They also self-scored significantly higher

on the spatial subscale compared to the verbal subscale,  $t(15) = 3.55, p < .01, d = 0.89$ . Again, no significant differences were found between scores on the spatial and object subscales,  $t(15) = 0.02, p > .05$ .

Finally, students did not self-score significantly better or worse on the object or verbal subscales over the duration of the term,  $t(15) = -1.42, p > .05$ , and  $t(15) = -1.45, p > .05$ , respectively. While students did not self-score significantly higher on the spatial subscale at the end of the term after receiving spatial training when analyzed using the Bonferroni-adjusted alpha level of .02, the  $t$  statistic did surpass a conventional alpha level of .05,  $t(15) = -2.59, p < .05, d = 0.65$ .

## Study 2

On average, more students at the start of the autumn term self-identified as spatial learners ( $M = 46.23, SD = 10.44$ ) than they did as verbal learners ( $M = 44.92, SD = 6.22$ ) or object learners ( $M = 44.08, SD = 11.39$ ). Paired sample  $t$ -tests conducted using Bonferroni-adjusted alpha levels of .02 per test (.05/3) revealed that scores on the spatial subscale were not significantly higher than on the verbal subscale,  $t(12) = 0.66, p > .05$ . Scores on the spatial subscale were also not significantly higher than those on the object subscale,  $t(12) = -5.56, p > .05$ , and no significant differences were found between scores on the verbal and object subscales,  $t(12) = 0.27, p > .05$ .

At the end of the term, the spatial cognitive style remained predominant for the class as a whole ( $M = 45.30, SD = 11.68$ ) but, in contrast to the start of term, the verbal style had the same mean value as the spatial style ( $M = 45.30, SD = 6.48$ ), with the object style revealed as the least dominant ( $M = 44.54, SD = 11.68$ ).

Similar to the start of the term, paired sample  $t$ -tests revealed no significant differences between self-reported object and spatial cognitive styles,  $t(12) = -0.24, p > .05$ , object and verbal styles,  $t(12) = -0.40, p > .05$ , or spatial and verbal styles,  $t(12) = 0.00, p > .05$ .

In the control group, on average, more students at the start of the autumn term self-identified as verbal learners ( $M = 47.57, SD = 8.66$ ) than they did as object learners ( $M = 44.14, SD = 2.55$ ) or spatial learners ( $M = 41.86, SD = 13.63$ ). However, paired sample  $t$ -tests conducted using Bonferroni-adjusted alpha levels of .02 per test (.05/3) revealed that none of these average scores on the three subscales OSIVQ were significantly different from each other (all  $ps > .05$ ).

After the term, the verbal cognitive style remained predominant for the class as a whole ( $M = 46.71, SD = 6.73$ ), with the object style being the next-most dominant, followed by the spatial style ( $M = 42.43, SD = 12.15$  and  $M = 41.14, SD = 10.33$ , respectively). Paired-samples  $t$ -tests using Bonferroni-adjusted alpha levels of .02 per test (.05/3) revealed no statistical differences among cognitive styles (all  $ps > .05$ ).

Students did not self-score significantly higher or lower on any of the cognitive style subscales over the course of the term (all  $ps > .05$ ). Moreover, independent samples  $t$ -tests revealed no statistically significant differences between the treatment and control groups with respect to average scores on the three subscales of the OSIVQ before or after the term, all  $ps > .05$ .

## Gender Differences on the OSIVQ

### Study 1

At the start of the summer term, men identified mostly, on average, as spatial learners ( $M = 49.75, SD = 2.74$ ) and least as verbal learners ( $M = 42.75, SD = 1.40$ ). This was also the case at the end of the term (see Table 1). In contrast, the highest average score among the three cognitive styles for women was on the object subscale ( $M = 50.13, SD = 2.26$ ) while the lowest was on the verbal subscale ( $M = 39.50, SD = 1.58$ ) and remained so at the end of term.

Although participants' general perceptions of dominant cognitive styles remained stable during the term, scores increased on each subscale except that women's scores on the object style subscale decreased slightly (but insignificantly,  $t(7) = 0.43$ ,  $p > .05$ ) over time.

Table 1: *Descriptive Statistics for Test Variables Per Gender Type in Study 1*

Variable	Gender	Mean	Standard Deviation
		Start of Term	Start of Term
CCI (scored out of 22)	Male	9.63	4.47
	Female	7.75	3.45
PSVT:R (scored out of 15)	Male	9.88	4.29
	Female	9.75	2.05
OSIVQ: Spatial (scored out of 75)	Male	49.75	2.74
	Female	46.13	3.44
OSIVQ: Object (scored out of 75)	Male	47.00	2.67
	Female	50.13	2.26
OSIVQ: Verbal (scored out of 75)	Male	42.75	1.40
	Female	39.50	1.58
		End of Term	End of Term
CCI (scored out of 22)	Male	10.25	2.09
	Female	8.25	1.40
PSVT:R (scored out of 15)	Male	10.50	1.45
	Female	9.25	1.24
OSIVQ: Spatial (scored out of 75)	Male	54.00	2.19
	Female	47.50	3.70
OSIVQ: Object (scored out of 75)	Male	52.38	3.42
	Female	49.25	2.04
OSIVQ: Verbal (scored out of 75)	Male	45.38	0.80
	Female	40.13	1.42

After the first round of data collection, independent samples  $t$ -tests revealed no significant differences between male and female participants' scores on the three OSIVQ subscales (all  $ps > .05$ ). However, at the end of the term, a significant difference between men and women's scores on the verbal subscale of the OSIVQ was revealed,  $t(14) = 3.22$ ,  $p < .01$ ,  $d = 4.56$ . No other significant differences were revealed between genders on the OSIVQ subscales after the term was complete (all  $ps > .05$ ).

Finally, scores on the spatial subscale of the OSIVQ did not correlate significantly with high scores on the PSVT:R at the start of the term ( $r = .27$ ,  $p > .05$ ) but did so at the end of term ( $r = .62$ ,  $p = .01$ , Fisher's  $Z_r = 0.73$ ). No other significant correlations were revealed between scores on the PSVT:R and other subscales of the OSIVQ, nor were there any significant associations between scores on the OSIVQ subscales and the CCI at the start or end of the term (all  $ps > .05$ ).

## Study 2

At the start of the fall term, the men in the sample identified mostly, on average, as spatial learners ( $M = 50.00$ ,  $SD = 7.43$ ) and, unlike in Study 1, least as object learners ( $M = 43.63$ ,  $SD = 13.36$ ). Men's average scores remained highest on the spatial subscale at the end of the term ( $M = 51.13$ ,  $SD = 9.05$ ; see Table 2). However, their lowest score, on average, was on the verbal subscale at the end of the term, rather than the object subscale ( $M = 46.63$ ,  $SD = 4.44$ , and  $M = 47.25$ ,  $SD = 9.02$ , respectively). Although men's scores increased on each subscale over time, the increases were not statistically significant (all  $ps > .05$ ).

The highest average score among the three cognitive styles for women in the sample was on the object subscale at the start of term ( $M = 44.80$ ,  $SD = 8.70$ ). However, average scores on this subscale decreased at the end of term ( $M = 40.20$ ,  $SD = 15.12$ ). The lowest average score self-reported by women was on the spatial subscale at the start of term ( $M = 40.20$ ,  $SD = 12.52$ ) and remained the lowest at the end of term ( $M = 36.00$ ,  $SD = 14.42$ ). And, women's scores on the verbal subscale increased during the term ( $M = 42.60$ ,  $SD = 7.37$ , and  $M = 43.20$ ,  $SD = 9.09$ , respectively). None of these changes in subscale scores were statistically significant (all  $ps > .05$ ).

Table 2: *Descriptive Statistics for Test Variables Per Gender Type in Study 2—Treatment Group*

Variable	Gender	Mean	Standard Deviation
		Start of Term	Start of Term
CCI (scored out of 22)	Male	7.38	5.26
	Female	5.60	2.61
PSVT:R (scored out of 15)	Male	11.00	2.88
	Female	9.40	3.78
OSIVQ: Spatial (scored out of 75)	Male	50.00	7.43
	Female	40.20	12.52
OSIVQ: Object (scored out of 75)	Male	43.63	13.36
	Female	44.80	8.70
OSIVQ: Verbal (scored out of 75)	Male	46.38	5.40
	Female	42.60	7.37
		End of Term	End of Term
CCI (scored out of 22)	Male	8.88	4.76
	Female	8.60	5.03
PSVT:R (scored out of 15)	Male	10.75	3.81
	Female	10.20	2.05
OSIVQ: Spatial (scored out of 75)	Male	51.13	9.05
	Female	36.00	14.42
OSIVQ: Object (scored out of 75)	Male	47.25	9.02
	Female	40.20	15.12
OSIVQ: Verbal (scored out of 75)	Male	46.63	4.44
	Female	43.20	9.09

Just as in Study 1, after the first round of data collection at the start of the term, independent samples t-tests revealed no significant differences between male and female participants' scores on the three OSIVQ subscales (all  $ps > .05$ ). However, unlike Study 1, at the end of the term, a significant difference between men and women's scores on the spatial subscale of the OSIVQ was revealed,  $t(11) = 2.35$ ,  $p < .05$ ,  $d = 1.26$ . No other significant differences were revealed between genders on the OSIVQ subscales after the term was complete (all  $ps > .05$ ).

No significant correlations were revealed between scores on the PSVT:R and subscales of the OSIVQ, nor were there any significant associations between scores on the OSIVQ subscales and the CCI at the start or end of the term.

With respect to the control sample, unlike in the treatment group, and the results of Study 1, men in the control group, at the start of term, mostly identified, on average, as verbal learners ( $M = 51.75$ ,  $SD = 7.14$ ) but, just as in the treatment group (and unlike in Study 1), they identified least as object learners ( $M = 44.25$ ,  $SD = 2.87$ ) at the start of term. This pattern continued for men at the end of term: their average scores remained highest on the verbal subscale ( $M = 47.00$ ,  $SD = 3.16$ ; see Table 2) and lowest on the object subscale ( $M = 43.75$ ,  $SD = 13.72$ ). Although men's average scores decreased on each of the three subscales during the term, none of these differences were statistically significant (all  $ps > .05$ ).

Just as in the treatment group, the highest average score among the three cognitive styles for women in the control group was on the object subscale at the start of term ( $M = 44.00$ ,  $SD = 2.65$ ). However, at the end of the term, women's predominant subscale was no longer object but verbal ( $M = 46.33$ ,  $SD = 10.97$ ). Similar to the results from the treatment group, the lowest average score self-reported by women was on the spatial subscale at the start of term ( $M = 33.33$ ,  $SD = 11.37$ ) and remained the lowest at the end of term ( $M = 35.67$ ,  $SD = 8.08$ ).

While women's scores on the spatial and verbal subscales increased over the term, and scores on the object subscale decreased, none of the differences were statistically significant (all  $ps > .05$ ).

Table 3: *Descriptive Statistics for Test Variables Per Gender Type in Study 2—Control Group*

Variable	Gender	Mean Start of Term	Standard Deviation Start of Term
CCI (scored out of 22)	Male	7.75	6.29
	Female	13.33	4.51
PSVT:R (scored out of 15)	Male	12.25	1.89
	Female	10.67	2.52
OSIVQ: Spatial (scored out of 75)	Male	48.25	12.58
	Female	33.33	11.37
OSIVQ: Object (scored out of 75)	Male	44.25	2.87
	Female	44.00	2.65
OSIVQ: Verbal (scored out of 75)	Male	51.75	7.14
	Female	42.00	8.19
		End of Term	End of Term
CCI (scored out of 22)	Male	11.25	2.87
	Female	4.33	2.52
PSVT:R (scored out of 15)	Male	11.25	0.96
	Female	3.67	3.79
OSIVQ: Spatial (scored out of 75)	Male	45.25	10.84
	Female	35.67	8.08
OSIVQ: Object (scored out of 75)	Male	43.75	13.72
	Female	40.67	12.34
OSIVQ: Verbal (scored out of 75)	Male	47.00	3.16
	Female	4	

Just as in Study 1, and in the treatment group in Study 2, independent samples *t*-tests revealed no significant differences between male and female participants' scores on the three OSIVQ subscales (all *ps* > .05) at the start of term. No significant differences were revealed between genders on the OSIVQ subscales after the term was complete (all *ps* > .05).

No significant correlations were revealed between control group participant's scores on the PSVT:R and the subscales of the OSIVQ at the start or end of term, nor were there any significant associations between scores on the OSIVQ subscales and the CCI at the end of the term (all *ps* > .05). However, while no significant correlations between the spatial and object subscales on the OSIVQ and the CCI were revealed at the start of term, a significant positive correlation between scores on the verbal subscale and scores on the CCI ( $r = .85$ ,  $p = .01$ , Fisher's  $Z_r = 1.26$ ) was found.

## Discussion and Teaching Implications

This study found contradicting results to Sorby et al. (2013) concerning the positive impact spatial training was expected to have on students' calculus and rotation ability, on average. Although average scores increased on both of these tasks during the term, the increases were statistically insignificant and occurred with or without spatial training. These results may be because of differences present in how spatial training was implemented in the courses. For example, Sorby et al. afforded spatial training to students in a separate, one-credit course that students were required to enroll in if their spatial skills were deemed insufficient (based on their results on a placement test). In the present research, all students took part in the spatial training (and only for a limited time during class).

Unlike Hawes et al. (2017), we did not find that spatial training affected mathematics performance. However, our results are consistent with those of Rodán et al. (2019) in that spatial abilities were found to increase after spatial training, but not mathematics ability. Our study may be distinct insofar as it was not done with young learners, as were the two studies noted above and focused on participants' self-reported perceptions of their own cognitive style which may influence confidence levels in calculus courses. While our work used a method of spatial training appropriate for undergraduate learners, both types of interventions included aspects of spatial visualization and mental rotation—the shared cognitive skills of forming and manipulating mental images (Mix & Cheng, 2012).

In addition, students' lack of increased performance in calculus after experiencing the spatial training may relate to the nature of the problems asked on the CCI. One study done with sixth-grade students found that spatial skills and representations were more likely to be drawn upon when students are asked to solve new or challenging mathematics problems, compared to routine problems (Mix, 2019). Because the CCI tests introductory calculus concepts (and all students in this study were enrolled in second- or third-year calculus courses), the problems on the CCI may not have seemed very novel to students and, thus, they did not strongly engage their spatial reasoning skills during the task.

In Study 1, spatial training did not alter students' self-reported cognitive style to be more spatial. On average, students perceived themselves as predominantly object learners before and after spatial training. And, in Study 2, while the spatial style was predominant for students before and after the term, spatial training did not significantly increase scores. We found no statistically significant differences between the treatment and control groups with respect to average scores on the three subscales of the OSIVQ before or after the term. It would seem that, on average, spatial training does not significantly influence students' self-reported cognitive styles toward a spatial preference. Perhaps, spatial training ought to be discipline-specific for large effects to be evident. Indeed, the workbook developed by Sorby et al. (2013) was intended for engineering students using calculus.

However, one point of evidence in the present study that spatial training associates with cognitive style and spatial ability may be the significant positive correlation in Study 1 between students' preference for the spatial cognitive style and their scores on the rotation task after spatial training was involved in their curriculum. Because this correlation was not present in Study 2, and because a similar significant correlation was revealed between CCI scores and preference for the verbal style in the control group, further investigation with larger samples sizes must be done to better understand relations between these variables. It should also be mentioned that a limitation of the study is the size of each sample. Given our ability to access only one section of a calculus course from which to collect data (and the circumstance whereby only those participants who completed all study instruments were included in the data set), a larger study with more diverse samples, over longer periods of time, will be prudent to undertake in the future. Despite this limitation, this study lays the groundwork for a promising interdisciplinary line of enquiry.

It is worth noting that in both Study 1, and in the treatment group of Study 2, males reported a predominant spatial cognitive style and women did not. Women reported being predominantly object visualizers—and this perception, on average, did not change after spatial training. In general, no significant differences were found between men's and women's self-perceptions of their cognitive styles. However, in Study 2, men self-scored significantly higher on the spatial subscale compared to women after receiving spatial training. This may indicate that men benefited from the spatial training in a psychological way and that the training may not have affected female students' confidence about their spatial abilities as we had hoped.

## Limitations and Conclusion

Although this research has valuable implications for the teaching practices of mathematics educators, some limitations are present. This work is a series of case studies that focus on data from undergraduate students at a single university. Thus, results are initial and external validity must be bolstered in the future by replicating results and utilizing the methodology with more diverse samples in additional locations.

This research is innovative in its blending of variables at the intersection of cognitive psychology and mathematics education. However, with this novelty comes a scarcity of literature combining the measurement of self-reported cognitive style with spatial training techniques and calculus performance. It seems that interdisciplinary research is needed if we are to better understand whether these variables affect participation with, and success in, the STEM fields for both men and women at various stages in their education. Indeed, earlier research suggests that students who are predominately object visualizers tend to see graphs as static pictures. Thus, they tend to experience difficulties in interpreting science graphs as abstract schematic representations because of this literal interpretation of graphs (e.g., Heagarty & Kozhevnikov, 1999; Kozhevnikov et al., 2002). This tendency contrasts with that of individuals who score themselves as predominantly spatial who tend to understand graphs as dynamic encapsulations of information—an essential skill to perform well in undergraduate-level calculus courses.

Our study offers mixed results in relation to these findings: a significant positive correlation between students' preference for a spatial cognitive style and their scores on the rotation task after spatial training was only revealed in one of the two studies. As additional research with larger sample sizes is put forward, it may be prudent to encourage teaching practices that offer students the opportunity to pair physical events or words with graphical representations moving in real time. This may afford students better connections between calculus concepts and how they naturally perceive the world as individuals with preferred cognitive styles, as Kozhevnikov and Thornton (2006) suggest.

In conclusion, our study is one of very few in the body of mathematics education literature to incorporate the psychological construct of cognitive style into research concerning spatial training. Now that associations between these variables have been shown to exist, future research can investigate other ways to measure and enhance the impact spatial training may have on student learning, confidence, and performance in the STEM fields.



## References

- Acredolo, L. P. (1988). From signal to "symbol": The development of landmark knowledge from 9 to 13 months. *British Journal of Developmental Psychology*, *6*, 369–372.
- Baenninger, M., & Newcombe, N. S. (1989). The role of experience in spatial test performance: A meta-analysis. *Sex Roles*, *20*, 327–344.
- Blajenkova, O., Kozhevnikov, M., & Motes, M. A. (2006). Object-spatial imagery: A new self-report imagery questionnaire. *Applied Cognitive Psychology*, *20*, 239–263.
- Blajenkova, O., & Kozhevnikov, M. (2008). The new object-spatial-verbal cognitive style model: Theory and measurement. *Applied Cognitive Psychology*, *23*, 638–663.
- Blazhenkova, O., & Kozhevnikov, M. (2009). The new object-spatial-verbal cognitive style model: Theory and measurement. *Applied Cognitive Psychology*, *23*(5), 638–663. <https://doi.org/10.1002/acp.1473>
- Bressoud, D., V. Mesa, and C. Rasmussen (Eds.) (2009). *Insights and recommendations from the MAA national study of college calculus*. MAA Press.
- Brown, L. N., Lahar, C. J., & Mosley, J. L. (1998). Age and gender-related differences in strategy use for route information: A "map-present" direction-giving paradigm. *Environment and Behavior*, *30*, 123–143.
- Burte, H., Gardony, A., Hutton, A., & Taylor, H. (2017). Think3d!: Improving mathematics learning through embodied spatial training. *Cognitive Research: Principles and Implications*, *2*. <https://doi.org/10.1186/S41235-017-0052-9>
- Casey, M. B., Nuttall, R., & Pezaris, E. (2001). Spatial-mechanical reasoning skills versus mathematics self-confidence as mediators of gender differences on mathematics subtests using cross-national gender-based items. *Journal for Research in Mathematics Education*, *32*, 28–57. DOI: 10.2307/749620
- Casey, M. B. (2013). Individual and group differences in spatial ability. In D. Waller, & L. Nadel (Eds.), *Handbook of spatial cognition* (pp. 117–134). American Psychological Association.
- Cooper, S. E., & Robinson, D. A. G. (1991). The relationship of mathematics self-efficacy beliefs to mathematics anxiety and performance. *Measurement and Evaluation in Counseling and Development*, *24*, 4–11.
- Cutmore, T. R. H., Hine, T. J., Maberly, K. J., Langford, N. M., & Hawgood, G. (2000). Cognitive and gender factors influencing navigation in a virtual environment. *International Journal of Human Computer Studies*, *53*, 223–249. <https://doi.org/10.1006/ijhc.2000.0389>
- Dawson, C. (2019). Tackling limited spatial ability: Lowering one barrier into STEM? *European Journal of Science and Mathematics Education*, *7*(1), 14–31.
- de Winter, J. C. F., & Dodou, D. (2010). Five-point Likert items: t-test versus Mann-Whitney-Wilcoxon. *Practical Assessment, Research and Evaluation*, *15*, 1–16.
- Dweck, C. (2006). Is math a gift? Beliefs that put females at risk. In S. J. Ceci & W. M. Williams (Eds.), *Why aren't more women in science? Top researchers debate the evidence* (pp. 47–55). American Psychological Association.
- Ellis, J., Fosdick, B. K., & Ramussen, C. (2016). Women 1.5 times more likely to leave STEM pipeline after calculus compared to men: Lack of mathematical confidence a potential culprit. *PLoS ONE*, *11*, e0157447. <https://doi.org/10.1371/journal.pone.0157447>
- Epstein, J. (2013). The calculus concept inventory—Measurement of the effect of teaching methodology in mathematics. *Notices of the AMS*, *60*(8), 1018–1026. <http://dx.doi.org/10.1090/noti1033>

- Ferrini-Mundy, J. (1987). Spatial training for calculus students: Sex differences in achievement and in visualization ability. *Journal for Research in Mathematics Education* 18(2), 126–140.
- Gärling, T., Book, A., Lindberg, E. & Nilsson, T. (1981). Memory for the spatial layout of the everyday physical environment: Factors affecting rate of acquisition. *Journal of Environmental Psychology*, 1, 263–277.
- Gifford, R. (2007). *Environmental psychology: Principles and practice* (4th edition). Optimal Books.
- Gleason, J., White, D., Thomas, M., Bagley, S., & Rice, L. (2016). The calculus concept inventory: A psychometric analysis and framework for a new instrument. Proceedings of the 19th Annual Conference on Research in Undergraduate Mathematics.
- Guay, R. (1977). Purdue Spatial Visualization Tests. Purdue Research Foundation.
- Hargreaves, M., Homer, M., & Swinnerton, B. (2008) A comparison of performance and attitudes in mathematics amongst the “gifted.” Are boys better at mathematics or do they just think they are?, *Assessment in Education: Principles, Policy & Practice*, 15(1), 19-38. <https://doi.org/10.1080/09695940701876037>
- Hawes, Z., & Ansari, D. (2020). What explains the relationship between spatial and mathematical skills? A review of evidence from brain and behavior. *Psychonomic Bulletin & Review*. doi.org/10.3758/s13423-019-01694-7.
- Hawes, Z., Moss, J., Caswell, B., Naqvi, S., & MacKinnon, S., (2017) Enhancing children’s spatial and numerical skills through a dynamic spatial approach to early geometry instruction: Effects of a 32-week intervention. *Cognition and Instruction*, 35, 236–264.
- Harvard Mental Imagery and Human-Computer Interaction Lab. (2013). Neural correlates of object vs. spatial visualization abilities. [http://www.nmr.mgh.harvard.edu/mkozhevnlab/?page\\_id=663](http://www.nmr.mgh.harvard.edu/mkozhevnlab/?page_id=663)
- Hill, C., Corbett, C., St. Rose, A. (2010). Why so few?: Women in science, technology, engineering, and mathematics. American Association of University Women.
- Johnson, W., & Bouchard, T. J. (2005). The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence*, 33, 393–416. <https://doi.org/10.1016/j.intell.2004.12.002>
- Kirasic, K. C., Allen, G. L., Siegel, A. W. (1984). Expression of configurational knowledge of large-scale environments: Students’ performance of cognitive tasks. *Environment and Behavior*, 16(6), 687–712. <https://doi.org/10.1177/0013916584166002>
- Kirasic, K. C., Allen, G. L., & Haggerty, D. (1992). Age-related differences in adults’ macrospatial cognitive processes. *Experimental Aging Research*, 18, 33–39.
- Kline, R. B. (1997). *Principles and practice of structural equation modeling*. Guilford.
- Kozhevnikov, M., Hegarty, M., & Mayer, R. E. (2002). Revising the visualizer-verbalizer dimension: Evidence for two types of visualizers. *Cognition and Instruction*, 20, 47–78. [https://doi.org/10.1207/S1532690XCI2001\\_3](https://doi.org/10.1207/S1532690XCI2001_3)
- Kozhevnikov, M., Kosslyn, S., & Shephard, J. (2005). Spatial versus object visualizers: A new characterization of visual cognitive style. *Memory and Cognition*, 33, 710–726. <https://doi.org/10.3758/BF03195337>
- Kozhevnikov, M., & Thornton, R. (2006). Real-time data display, spatial visualization ability, and learning force and motion concepts. *Journal of Science Education and Technology*, 15, 113–134. DOI: 10.1007/s10956-006-0361-0
- Lachance, J. A., & Mazocco, M. M. M. (2006). A longitudinal analysis of sex differences in math and spatial skills in primary school age children. *Learning and Individual Differences*, 16, 195–216. <https://doi.org/10.1016/j.lindif.2005.12.001>

- Lehning, M., Leplow, B., Haaland, V. O., Mehdorn, M. & Ferstl, R. (2003). Pointing accuracy is dependent on age, sex, and experience. *Journal of Environmental Psychology*, 23, 419–425. [https://doi.org/10.1016/S0272-4944\(02\)00084-1](https://doi.org/10.1016/S0272-4944(02)00084-1)
- Lent, R. W., Brown, S. D., & Larkin, K. C. (1986). Self-efficacy in the prediction of academic performance and perceived career options. *Journal of Counseling Psychology*, 33, 265–69.
- Levine, S. C., Huttenlocher, J., Taylor, A., & Langrock, A. (1999). Early sex differences in spatial skills. *Developmental Psychology*, 35, 940–949.
- Linn, M. C., & Hyde, J. S. (1989). Gender, mathematics, and science. *Educational Researcher*, 18, 17–27.
- Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of gender differences in spatial abilities: a meta-analysis. *Child Development*, 56, 1479–1498.
- Lubinski, D., & Benbow, C. P. (2006). Study of mathematically precocious youth after 35 years: Uncovering antecedents for the development of math-science expertise. *Perspectives on Psychological Science*, 1, 316–345. <https://doi.org/10.1111/j.1745-6916.2006.00019.x>
- MM Virtual Design. (2016). *Object-Spatial Imagery and Verbal Questionnaire*. [http://www.mmvirtualdesign.com/html/osivq\\_.html](http://www.mmvirtualdesign.com/html/osivq_.html)
- Mix, K. (2019). Why are spatial skill and mathematics related? *Child Development Perspectives*, 13(2), 121–126. <https://doi.org/10.1111/cdep.12323>
- Mix, K. S., & Cheng, Y. L. (2012). The relation between space and math: developmental and educational implications. *Advances in Child Development and Behavior*, 42, 197–243. <https://doi.org/10.1016/B978-0-12-394388-0.00006-X>
- National Science Board. (2010). Preparing the next generation of STEM innovators: Identifying and developing our nation's human capital. NSB-10-33 <https://www.nsf.gov/nsb/publications/2010/nsb1033.pdf>
- Newcombe, N. (2017). Harnessing spatial thinking to support STEM learning. OECD Education Working Papers, No. 161. OECD Publishing. <https://doi.org/10.1787/7d5dcae6-en>.
- Newcombe, N. (2010). Picture this: Increasing math and science learning by improving spatial thinking. *American Educator* 14(4), 171–177. <https://www.aft.org/sites/default/files/periodicals/Newcombe.pdf>
- Nimmessgern, H. (2016). Why are women underrepresented in STEM fields? *Chemistry--A European Journal*, 22, 3529–3530. <https://doi.org/10.1002/chem.201600035>
- Organization for Economic Cooperation and Development (OECD). (2012). *Closing the gender gap: Act now*. OECD Publishing.
- Pajares, F. (1996). Self-efficacy beliefs and mathematical problem-solving of gifted students. *Contemporary Educational Psychology*, 21, 325–44.
- Pajares, F. (2005). Gender differences in mathematics self-efficacy beliefs. In A. M. Gallagher & J. C. Kaufman (Eds.), *Gender differences in mathematics: An integrative psychological approach* (pp. 294–315). Cambridge University Press.
- Redmond, P., & Gutke, H. (2019). STEMming the flow: Supporting females in STEM. *International Journal of Science and Mathematics Education*, 18, 221–237. <https://doi.org/10.1007/s10763-019-09963-6>
- Rodán, A., Gimeno, P., Elosúa, M. R., Montoro, P. R., & Contreras, M. J. (2019). Boys and girls gain in spatial, but not in mathematical ability after mental rotation training in primary education. *Learning and Individual Differences*, 70, 1–11.
- Saucerman, J., & Vasquez, K. (2014). Psychological barriers to STEM participation for women over the course of development. *Adultspan Journal*, 13, 46–64. <https://doi.org/10.1002/j.2161-0029.2014.00025.x>

- Schlenker, E. (2015). The labour supply of women in STEM. *IZA Journal of European Labour Studies*, 4, doi:10.1186/s40174-015-0034-1.
- Skolnik, J. (2015). Why are girls and women underrepresented in STEM and what can be done about it? *Science & Education*, 24, 1301–1306. <https://doi.org/10.1007/s11191-015-9774-6>
- Sorby, S., Casey, B., Veurink, N., & Dulaney, A. (2013). The role of spatial training in improving spatial and calculus performance in engineering students. *Learning and Individual Differences* 26, 20–29. <https://doi.org/10.1016/j.lindif.2013.03.010>
- Spelke, E. S. (2005). Sex differences in intrinsic aptitude for mathematics and science?: A critical review. *American Psychologist*, 60, 950–958.
- Stieff, M. and Uttal, D. (2015). How much can spatial training improve STEM achievement? *Educational Psychology Review* 27(4), 607–615. 10.1007/s10648-015-9304-8
- Stieff, M., Lira, M. E., Scopelitis, S. A. (2016). Gesture supports spatial thinking in STEM. *Cognition and Instruction*, 34, 80–99. <https://doi.org/10.1080/07370008.2016.1145122>
- Taylor, H., & Hutton, A. (2013). Think3d!: Training spatial thinking fundamental to STEM education. *Cognition and Instruction*, 31, 434–455. <https://doi.org/10.1080/07370008.2013.828727>
- Uttal, D. H., (2009, April). Training spatial skills: What works, for whom, and for how long? Paper presented at the meeting of the Society for Research in Child Development, Denver, CO.
- Vasta, R., Knott, J. A., & Gaze, C. E. (1996). Can spatial training erase the gender differences on the water-level task? *Psychology of Women Quarterly*, 20, 549–567.
- Voyer, D., Voyer, S., Saint-Aubin, J. (2017). Sex differences in visual-spatial working memory: A meta-analysis. *Psychonomic Bulletin & Review*, 24, 307–334. <https://doi.org/10.3758/s13423-016-1085-7>
- Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin*, 117, 250–270.
- Wai, J., D. Lubinski, and C. Benbow (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. DOI: 10.1037/a0016127
- Wang, L. (2019). Mediation relationships among gender, spatial ability, math anxiety, and math achievement. *Educational Psychology Review*. <https://doi.org/10.1007/s10648-019-09487-z>
- Ward, S. L., Newcombe, N., & Overton, W. F. (1986). Turn left at the church, or three miles north: A study of direction giving and sex differences. *Environment and Behavior*, 18, 192–213.
- Webley, P., & Whalley, A. (1987). Sex differences in children's environmental cognition. *Journal of Social Psychology*, 127, 223–225.
- Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology*, 82, 51–59.



The *Journal of Educational Research and Practice* is a peer-reviewed journal that provides a forum for studies and dialogue about developments and change in the field of education and learning. The journal includes research and related content that examine current relevant educational issues and processes. The aim is to provide readers with knowledge and with strategies to use that knowledge in educational or learning environments. *JERAP* focuses on education at all levels and in any setting, and includes peer-reviewed research reports, commentaries, book reviews, interviews of prominent individuals, and reports about educational practice. The journal is sponsored by the Richard W. Riley College of Education and Leadership at [Walden University](#), and publication in *JERAP* is always free to authors and readers.