2026 ASEE ANNUAL CONFERENCE June 21–24, 2026 | Charlotte Convention Center, NC
Host Sponsor UNC Charlotte

ASEE AMERICAN SOCIETY FOR ENGINEERING EDUCATION

Paper ID #52757

Spatial Visualization Gains by Mode in Lower-PSVT:R First-Year Engineers

Dr. Prashant Athavale, Clarkson University

Prashant Athavale is an Associate Professor of Mathematics at Clarkson University, where he teaches differential equations, probability and statistics, and computer vision. His research interests span engineering education, applied mathematics, statistics, and image processing. His contributions to engineering education research include the study of student success, retention, and evidence-based interventions aimed at improving outcomes for engineering students, with particular attention to at-risk and underrepresented populations.

Dr. Melissa Carole Richards, Clarkson University

Dr. Melissa C. Richards is an Assistant Professor and the Director of the K12 Robotics Outreach Program at Clarkson University's Institute for STEM Education. She specializes in problem- and project-based STEM education, experiential learning, outreach, and mentorship.

Dr. Richards holds a Ph.D. in Mechanical Engineering. Her research focuses on theoretical rock mechanics and STEM education. She is dedicated to promoting diversity, equity, inclusion, and a sense of belonging for underserved and underrepresented students, particularly in engineering. Furthermore, since 2018, she has served as a Regional Program Delivery Partner for FIRST LEGO League, inspiring young people to explore robotics and engineering.

Dr. Jan DeWaters P.E., Clarkson University

Dr. Jan DeWaters is an Associate Professor in the Institute for STEM Education with a joint appointment in the School of Engineering at Clarkson University, and teaches classes in both areas. Her research focuses on developing and assessing effective, inclusive teaching and learning in a variety of settings. An environmental engineer by training, Dr. DeWaters' work typically integrates environmental topics such as energy and climate into STEM settings.

Emily Houston Monroe , PE, Dartmouth College

Emily H. Monroe, PE is a lecturer at the Thayer School of Engineering at Dartmouth College. She serves as the director of the Cook Engineering Design Center at Dartmouth, which connects industry, government and nonprofit sponsors with Dartmouth Engineering students to collaborate on engineering design projects. Prior to joining Dartmouth, Emily served in manufacturing and product design engineering roles. Emily is a licensed professional engineer (PE) and holds a BS degree in mechanical engineering from MIT and Master of Engineering Management from Duke University.

Dr. Sheryl A. Sorby, University of Cincinnati

Dr. Sheryl Sorby is currently a Professor of STEM Education at the University of Cincinnati and was recently a Fulbright Scholar at the Dublin Institute of Technology in Dublin, Ireland. She is a professor emerita of Mechanical Engineering-Engineering Mec

Ruvarashe Chinyadza, Clarkson University
Rumbidzai Nyaradzo Mushamba, Clarkson University
Dr. Michael W. Ramsdell, Clarkson University

Michael Ramsdell is an Associate Professor of Physics and Director of First Year Physics at Clarkson University. He has over ten years of experience in the design, implementation, and assessment of laboratory curriculum within introductory physics courses

Spatial Visualization Gains by Mode in Lower-PSVT:R First-Year Engineering Students

1. Introduction

Spatial visualization, the ability to mentally manipulate 3D objects, is a crucial skill for success and persistence in STEM disciplines. Strong spatial skills correlate with better performance in core courses such as Calculus, Chemistry, and Physics, as well as higher retention rates in engineering [1], [3]. Although students enter first-year engineering programs with varied preparation, research confirms that spatial thinking is a malleable skill, rather than an inherent trait [1], [2]. Engineering educators who include targeted interventions to strengthen these skills early on may promote an equitable learning environment and student success [3], [4], [5], [6].

Students who demonstrate skills in spatial visualization have been shown to both perform well and persist well in STEM fields [1], [7]. The literature in pre-college age students consistently links high spatial ability to higher performance on the math SAT, higher levels of creativity, and higher performance in STEM college courses [8], [9], [10]. In longitudinal studies to define the role of spatial visualization in engineering education, Sorby demonstrated the potential for an academic intervention consisting of spatial visualization assessment and intervention to impact achievement and persistence in undergraduate engineering programs [1], [3], [4], [11].

The malleability of spatial skills has led to the validation of effective interventions, most notably a multimodal 10-module method [12]. This standardized approach typically combines guided tutorials, handheld manipulatives, workbook sketching exercises, and interactive software. Furthermore, it emphasizes student-centered instruction by providing students with opportunities to actively construct and manipulate representations through gesture, sketching, visualization, and embodied interaction. Through these exercises, they can make explicit connections between physical or representational actions and the spatial information those actions convey [13]. This intentional integration provides students with structured, repeated practice, resulting in significant improvements on standardized assessments such as the Purdue Spatial Visualization Test: Visualization of Rotations (PSVT:R) [14].

The PSVT:R is a 30-question psychometric assessment of mental rotation ability designed to measure an individual's ability to perform specific spatial tasks, including rotating and manipulating three-dimensional objects [14], [15]. With over 50 years of use of the PSVT:R as an assessment for spatial ability, the use of such an assessment in concert with an effective spatial skill training has been offered in various engineering education settings, from pre-college (e.g., [10]) to college (e.g., [3], [5], [16]). Consistent with prior Sorby-based spatial intervention research, students scoring less than or equal to 18 out of 30 on the PSVT:R are typically identified as candidates for spatial skills instruction, in which raw PSVT:R gains are directly linked to enhanced academic performance and increased persistence in engineering (e.g., [1], [3]).

While the efficacy of the Sorby 10-module approach is well-established, a critical gap remains regarding the influence of instructional delivery mode on its outcomes. Recent shifts in higher education, particularly the rise of online, asynchronous, and flipped classroom models (e.g.,

[17]), necessitate a direct comparison. For example, Broadbent and Poon [18] found, in a systematic review of online higher education, that time management, metacognitive regulation, and effort regulation were among the strongest predictors of academic success in online environments. Similarly, Drumm [19] reported that students who actively apply individual learning strategies achieve better outcomes in asynchronous pathways, whereas those without such strategies exhibit greater performance variability. For flipped courses, Blau and Shamir-Inbal [17] emphasize that co-regulation and co-creation processes in redesigned flipped courses can enhance academic outcomes when students actively engage with peers and instructors during application activities.

Research on spatial visualization development has shown success in blended [20], hybrid [21], and web-based [22], [23] implementations. Veurink, Richards, and Sorby demonstrated the effective use of the PSVT:R assessment and intervention during the COVID-19 global pandemic through remote learning in synchronous and asynchronous settings compared to traditional offerings pre-pandemic [24], [25].

While these previous works provide insight into how to deliver and assess the 10-module method, limited research directly contrasts the impact of delivering the 10-module method across three primary formats: Traditional face-to-face, Virtual-Asynchronous, and the Flipped classroom model. Each instructional method offers varied levels of flexibility, structure, challenges, and interaction that may differentially affect the required practice for spatial skill development.

Thus, this research seeks to address this gap by presenting a quantitative comparative study to examine spatial reasoning gains across these three instructional modalities (Traditional face-to-face, Virtual-Asynchronous, and the Flipped classroom). *The guiding question for this work is whether the mode of delivery affects the effectiveness of the established 10-module intervention delivered through a first-year engineering course designed to improve students' spatial reasoning skills.*

This study uses data from three consecutive first-year cohorts, each experiencing the standardized curriculum delivered by the same instructor to ensure consistency. Student outcomes are measured using the PSVT:R as a pre-test and post-test, comparing the magnitude of gains across the delivery modes. Specifically, this study seeks to answer the following research questions (RQ):

- **RQ1**: Do PSVT:R gains (Post − Pre) differ across delivery modes among students with PSVT:R pre-test scores of ≤ 18?
- **RQ2**: Is PSVT:R gain associated with course grade points, and does this association differ by delivery mode?

The findings of this work are intended to inform program-level decisions about scalable spatial-skills instruction that improves overall outcomes while avoiding unintended inequities in early engineering pathways.

2. Methods

2.1 Study Design and Setting

We used a quantitative, quasi-experimental cohort comparison design to examine spatial reasoning gains across three instructional delivery modes in a first-year engineering spatial visualization intervention course. The data comes from a non-randomized, convenience sample consisting of three consecutive cohorts of first-year students enrolled in a spatial reasoning class during 15-week fall semesters, 2019, 2020, and 2021. Each cohort experienced a single delivery format: Traditional face-to-face (2019), Virtual-Asynchronous (2020), or Flipped classroom (2021).

It is important to note that Virtual-Asynchronous (2020) and Flipped classroom (2021) were delivered during the COVID-19 global pandemic. The Virtual-Asynchronous (2020) format was implemented during a period of emergency remote instruction, and the Flipped classroom (2021) was implemented during continued pandemic-related disruptions when students were navigating evolving campus policies, altered social dynamics, and lingering uncertainty [26]. While the course implemented the same 10-module multimodal spatial visualization curriculum across all semesters and was taught by the same instructor to maximize instructional consistency across delivery modes, broader contextual stressors may have influenced student motivation and engagement differentially across cohorts. These external factors represent an inherent limitation of this quasi-experimental cohort design and should be considered when interpreting findings across instructional modalities.

2.2 Participants and Eligibility

First-year students are enrolled in the spatial visualization intervention course based on their performance on the Purdue Spatial Visualization Test: Visualization of Rotations (PSVT:R) [14], which is administered as part of a pre-enrollment assessment through the university's learning management system (LMS) to all incoming engineering and engineering studies majors between May and June prior to matriculation. Students who score 18 or lower on this summer PSVT:R assessment, or who do not complete the assessment, are recommended and enrolled in the optional one-credit spatial visualization intervention course.

Our study sample was selected from a non-randomized, convenience sample of a total of 186 first-year engineering students who enrolled and completed the spatial visualization course across all three instructional delivery modes during the fall semesters of 2019 (Traditional, n = 66), 2020 (Virtual-Asynchronous, n = 63), and 2021 (Flipped, n = 57). Figure 1 illustrates the participant selection process (*further detailed in Section 2.5*), which included only those students with low initial spatial visualization ($PSVT{:}R_{pre} \leq 18$) and valid paired pre- and post-scores who completed the spatial visualization intervention. The final sample consisted of n = 58 students: Traditional (n = 26), Virtual-Asynchronous (n = 13), and Flipped (n = 19).

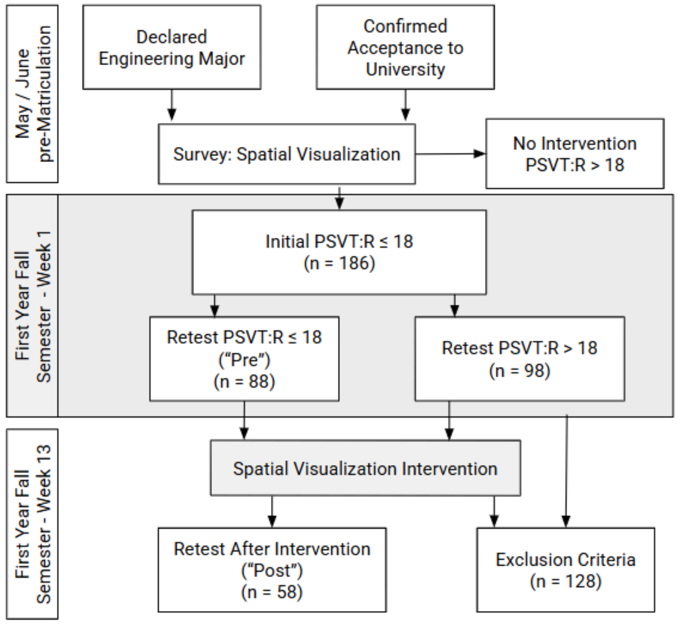

Figure 1. Flowchart of participant selection across multiple academic years of the spatial visualization intervention at Clarkson University, in which first-year engineering majors complete PSVT:R assessments

2.3 Curriculum and Delivery Modes

The intervention used a widely adopted 10-module multimodal approach consisting of guided instruction, structured sketching practice, and hands-on and software-based activities [5]. Students used a sketching workbook and snap cubes/handheld manipulatives; the course also incorporated interactive instructional software [27] and short "getting started" materials aligned with each module [25]. Below, we summarize the implementation by delivery mode.

2.3.1 Traditional:

In the Traditional format, students completed workbook sketching pages and multiple-choice items as homework. In class, the instructor administered short concepts or sketching quizzes after introducing each module. Quiz items were typically drawn from challenging module problems that the instructor partially discussed during class. When quizzes included sketching, students were allowed multiple attempts before submitting their work; if the instructor observed difficulty after an initial attempt, the instructor provided guided support without revealing the answer.

In addition to weekly quizzes and homework, students completed two in-class examinations. Exam 1 covered Modules 1 to 5. Exam 2 covered Modules 6 to 10 and included the PSVT:R post-test as Part 1 (timed), followed by Part 2, which consisted of problems aligned with Modules 6 to 10. In the Traditional face-to-face delivery mode, the actual PSVT:R pre-test score was not included in student course grade points; however, the PSVT:R post-test score was used to calculate their Exam 2 grade. No teaching assistants supported instructional delivery in the Traditional format.

2.3.2 Virtual-Asynchronous:

Implementation details, specific to the pandemic-era asynchronous delivery (the Virtual-Asynchronous format), are presented in [24] and [25]. Nonetheless, for the reader's convenience, we summarize it here. In the Virtual-Asynchronous format, students completed course content independently through the LMS. Students purchased the newly developed sketching workbook and snap cubes/handheld manipulatives. For each module, students were provided pre-recorded mini-lectures, module "getting started" videos, and the instructional software developed by Sorby through the LMS.

Additionally, the instructor recorded and uploaded homework-help videos for selected modules to the LMS. The content covered in the mini-lectures was consistent with the Traditional format; however, the sample exercises used in the asynchronous materials differed from those used during the face-to-face instruction. In this delivery mode, the course emphasized a design project rather than in-class examinations, and optional homework and quizzes were available to replace missed or lower-graded assignments.

The PSVT:R pre- and post-tests in Virtual-Asynchronous were included in the calculation of student course grade points; however, not as the actual PSVT:R score. Instead, all students who attempted the assessment received 100 percent on the corresponding quiz. If a student did not attempt the quiz corresponding to the PSVT:R pre- or post-test, they received a zero for that grade. No teaching assistants supported instructional delivery in the Virtual-Asynchronous format.

2.3.3 Flipped Classroom:

In the Flipped classroom format, students engaged with the same core instructional materials as in the Virtual-Asynchronous format and purchased the same sketching workbook and snap cubes/handheld manipulatives. However, before attending a weekly face-to-face class, students were required to view the pre-recorded mini-lectures, module "getting started" videos, homework-help videos, and try out the instructional software provided through the LMS. At the start of class, during the weekly face-to-face sessions, students completed in-class quizzes on the material covered in the mini-lecture. The instructor then reviewed the quizzes and addressed any misconceptions. Following this review, students worked on individual and group sketching example problems, then started homework individually or in small groups. Similar to the Virtual-Asynchronous format, the course emphasized a design project rather than in-class examinations, and optional homework and quizzes were available to replace missed or lower-graded assignments. Furthermore, as in Virtual-Asynchronous delivery, the actual PSVT:R pre- or post-test scores were not used to calculate grades; however, students received 100 if they attempted it, corresponding to a quiz grade of 100, and zero if no attempt was made.

Founded on the knowledge that peer mentoring and cooperative learning improve student outcomes (e.g., [28]), undergraduate teaching assistant support was incorporated for instructional delivery in the Flipped classroom format. During the weekly face-to-face sessions, in addition to the instructor, one to two undergraduate teaching assistants supported small-group sketching, provided individualized feedback to students on in-class quizzes and homework assignments, and assisted with grading.

2.4 Measures and Outcome Definitions

For first-year, first-semester students who were enrolled in and completed the optional one-credit course in one of the three delivery modes, we assessed spatial visualization using the Purdue Spatial Visualization Test: Visualization of Rotations (PSVT:R) [14], readministered as a pre-test at the beginning (Week 1) of the semester and a post-test after completion of the module sequence (Week 13).

We define spatial visualization (PSVT:R) gain, the raw score change on the PSVT:R points, for each student as:

$$Gain = PSVT{:}R_{post} - PSVT{:}R_{pre}, \tag{1}$$

where:

$PSVT{:}R_{post}$ = PSVT:R post-test scores on the assessment taken at Week 13, and
$PSVT{:}R_{pre}$ = PSVT:R pre-test scores on the assessment taken at Week 1 of the semester.

Recall that the PSVT:R uses a fixed 0–30 scoring scale, so reporting *Gain* in points is directly interpretable. PSVT:R scores are discrete, and so are the gains in these scores. Analyzing *Gain* as raw score change preserves that measurement structure. Prior spatial-skills intervention studies likewise report and analyze pre/post point gains (difference scores) on the PSVT:R, including gain-score analyses by Sorby and Baartmans [29], reporting of PSVT:R point gain by Harris, Harris, and Sadowski [30], and reporting the results in terms of mean gains by Hoople et al. [31]. While normalized gain is widely used in physics education research for percent-based instruments [32], the analogous normalization for PSVT:R would be $Gain / (30 - PSVT{:}R_{pre})$. However, this transformation converts whole-number changes into a continuous quantity and can introduce heteroskedasticity, with greater variability in normalized gains near the upper end of the scale where 30 − Pre is small. Accordingly, we keep *Gain* in points for interpretability and account for differences in starting ability by including $PSVT{:}R_{pre}$ as a covariate in our multiple linear regression model (*see* Section 2.6.3) that relates course performance (*GradePts*) to Gain and delivery mode. Furthermore, we measured course performance using course grade points on a 0.0 to 4.0 scale.

Within the Traditional instructional delivery mode, PSVT:R assessments were administered in class using paper-based test booklets with bubble sheets. However, for both the Virtual-Asynchronous and the Flipped classroom delivery modes, the PSVT:R pre- and post-tests were administered asynchronously through the LMS. The PSVT:R assessments in the Virtual-Asynchronous and Flipped formats awarded credit for completion rather than a score. Thus, the observed gains in these modes should be interpreted cautiously.

2.5 Data Sources, Inclusion Criteria, and Cleaning Procedures

In this quasi-experimental study, 186 first-year engineering students completed the spatial visualization course across all three instructional delivery modes (*Mode*). The students were distributed across three instructional modalities as follows: Traditional (n = 66), Virtual-Asynchronous (n = 63), and Flipped (n = 57). However, data cleaning was required to address the research questions of this study.

Data cleaning procedures included (i) standardizing variable names and formatting; (ii) removing records with missing PSVT:R pre- or post-test scores; (iii) excluding records with implausible PSVT:R completion times (for example, a participant with a completion time of 7 minutes and 40 seconds for the pre-test and 3 minutes for the post-test); and (iv) filtering out unusable cases such as students who did not provide consent to the study. For the present analysis, we further restricted the sample to students with low initial spatial visualization ($PSVT{:}R_{pre} \leq 18$) and valid paired pre/post scores; after applying these criteria, 128 records were excluded from inferential analyses. The final sample consisted of n = 58 students: Traditional (n = 26), Virtual-Asynchronous (n = 13), and Flipped (n = 19). This study extends prior work by [24] and [25]; accordingly, the Traditional (2019) and Virtual-Asynchronous (2020) datasets used in this work reflect the cleaned versions reported in those studies to ensure methodological consistency and comparability across delivery modes.

2.6 Statistical Analysis Plan

2.6.1 Descriptive Statistics and Visualizations:

For this work, we computed overall and mode-specific descriptive statistics (mean and standard deviation) for $PSVT{:}R_{pre}$, $PSVT{:}R_{post}$, and *Gain*. The sample of students with baseline PSVT:R scores ≤ 18 (n = 58) were analyzed. Next, we computed each student's Gain in the PSVT:R scores. The normality of the *Gain* distribution was assessed using Shapiro-Wilk tests overall and within each instructional delivery mode (Traditional, Virtual-Asynchronous, Flipped). Since these tests did not indicate deviation from normality, we used two-sided paired-samples t-tests to evaluate pre–post change in PSVT:R overall and within each delivery mode. We also ran two-sided Wilcoxon signed-rank tests overall and by delivery mode as a nonparametric sensitivity analysis.

2.6.2 RQ1 Analysis - Gain Differences by Delivery Mode ($PSVT{:}R_{pre} \leq 18$):

Since group sizes were unequal and homogeneity of variance could not be assumed, we used Welch's one-way ANOVA (unequal-variance ANOVA) to compare mean *Gain* across the three delivery modes (*Mode*) [33]. Following the omnibus test, we conducted Games-Howell pairwise post-hoc comparisons, which are appropriate for unequal variances and unequal sample sizes, and reported adjusted p-values.

As a nonparametric sensitivity check that does not rely on normality assumptions, we also conducted a Kruskal-Wallis rank-sum test comparing *Gain* across delivery modes. All tests were two-sided with a significance level of $\alpha = 0.05$.

2.6.3 RQ2 Analysis - Relationship Between Gain and Course Performance:

For this study, spatial visualization gain was measured using Equation (1). Furthermore, the measured course performance used was course grade points on a 0.0 to 4.0 scale. Instructional delivery mode, *Mode*, was treated as a three-level categorical factor (Traditional, Virtual-Asynchronous, Flipped).

Mode was coded as a three-level categorical factor and used Traditional as the reference category so that the estimated Mode coefficients represent differences in expected grade points for Virtual-Asynchronous and Flipped relative to Traditional, conditional on the same PSVT:R pre-test score. Traditional was then selected as the reference because it represents the conventional instructional format in the program and provides a substantively meaningful baseline for interpreting whether alternative delivery modes differ from standard practice; this choice affects only the interpretability of the coefficients (i.e., the contrasts reported) and does not change overall model fit or predicted values.

Next, $PSVT{:}R_{pre}$ score (i.e., $Pre = PSVT{:}R_{pre}$) were included as a covariate so that, when estimating the relationship between *Gain* and course grade points, the model compares students who had the same $PSVT{:}R_{pre}$ score (i.e., the same starting spatial visualization level). To thus evaluate whether spatial visualization gains relate to course performance and whether that relationship differs by instructional delivery mode, *Mode*, we fit an ordinary least squares (OLS) linear regression model with an interaction between *Gain* and *Mode*:

$$GradePts_i = \beta_0 + \beta_1\, Gain_i + \beta_2\, Pre_i + \beta_3\, Mode_i + \beta_4\, (Gain_i \times Mode_i) + \varepsilon_i. \tag{2}$$

We also report coefficient estimates, standard errors (SE), *t*-statistics, and *p*-values from the fitted OLS model (*see Table 3*).

Ethics/IRB Statement:

The institutional review board determined this study was exempt from full review (IRB # IRB25-40.1E) because the research was conducted in established educational settings, and all analyses used de-identified records.

3. Results and Discussion

Recall that the dataset for Traditional and Virtual-Asynchronous are those used in [24], and [25]. However, for this work, we provide results and interpretations of the findings related to PSVT:R gains across Traditional, Virtual-Asynchronous and Flipped delivery modes (RQ1), and the association between spatial skill improvement and course performance (RQ2).

3.1 RQ1: Do PSVT:R Gains Differ by Instructional Delivery Mode?

To examine whether students improved and how gains were distributed within each instructional delivery mode, we first describe the PSVT:R *Gain* scores ($PSVT{:}R_{post}$ minus $PSVT{:}R_{pre}$). Figure 2 shows histograms of the frequency of individual gain scores for the three instructional modalities. These visual summaries allow us to distinguish typical improvement from variability and potential outliers.

The histograms (Figure 2) show that gains cluster in the positive range for all three modalities. This indicates that the majority of students demonstrate gains above zero, with relatively few students showing no change or negative gains. The gains within the Traditional and Virtual-Asynchronous instructional modalities appear to have a somewhat more concentrated distribution, whereas the gains for Flipped delivery are more evenly dispersed, with gains spread across a wider range.

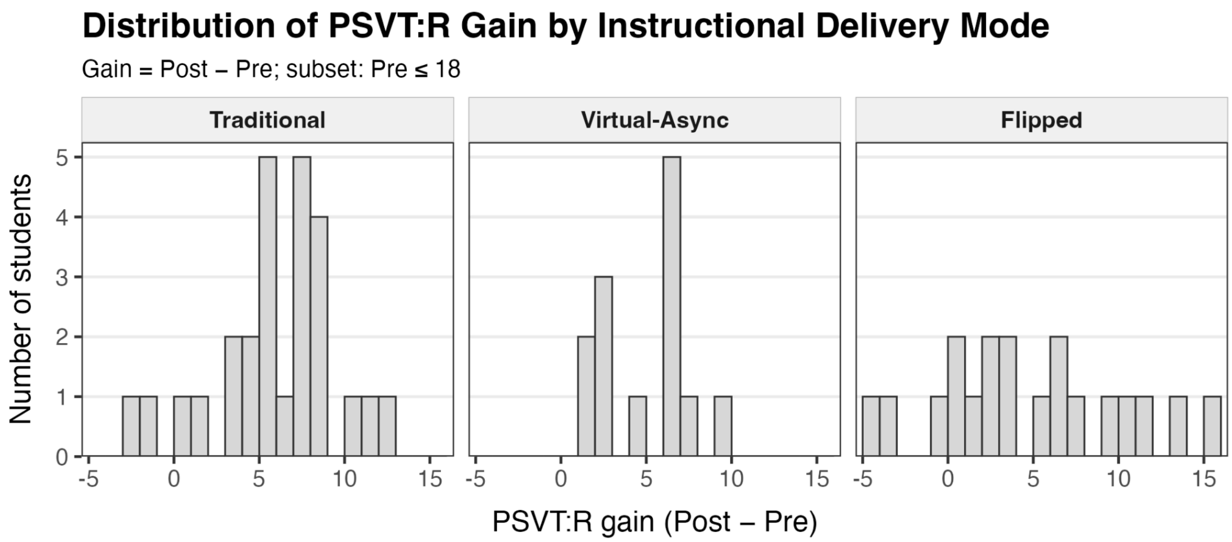

Figure 2. Distribution of PSVT:R Gain by instructional delivery mode for students with low initial scores (*PSVT:R*$_{pre}$ ≤ 18). Histograms show the frequency of individual gain scores, where *Gain* is defined as *PSVT:R*$_{post}$ minus the *PSVT:R*$_{pre}$ score.

A factor that might contribute to variability in gain scores, or to minimal improvements or declines, is the grading structure of the PSVT:R assessments in the Virtual-Asynchronous and Flipped formats. In these formats, students received full credit for attempting the assessment, regardless of their score. Thus, gains observed in the Virtual-Asynchronous and Flipped instructional modalities may not be a true indicator of students' spatial ability, but may instead be a byproduct of decreased effort when assessment outcomes were disconnected from grade implications [24]. While eliminating score-based grading may lower anxiety for some students, it could also diminish motivation for others.

Figure 3 presents boxplots summarizing the distribution of gain scores (median, interquartile range, and whiskers), with the dashed horizontal line indicating zero gain (i.e., no change in PSVT:R scores). Gain scores are generally positive across all three instructional modalities. The dashed zero line in Figure 3 emphasizes that most students improved, with a small number showing little improvement or declines. This finding is consistent with those reported in literature for students who complete the 10-module intervention in Traditional formats [1], [3], [11], [29].

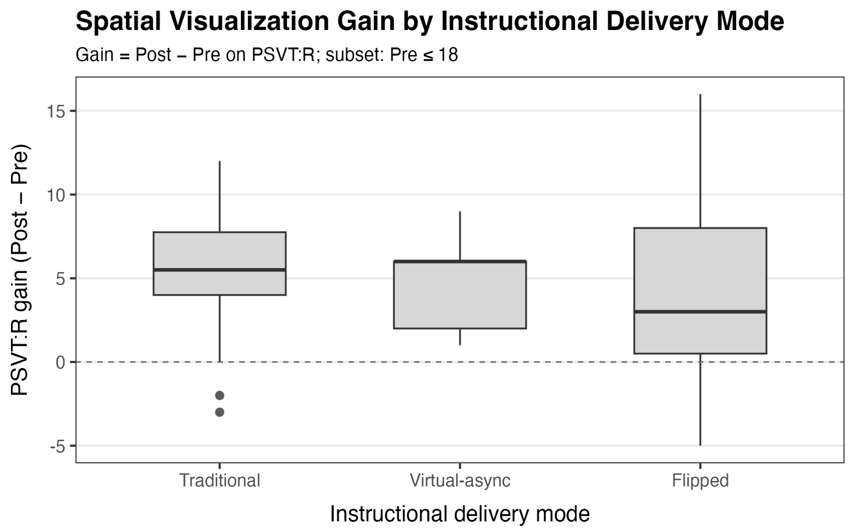

Figure 3. PSVT:R gain by instructional delivery mode for students with low initial scores (*PSVT:R*$_{pre}$ ≤ 18). Boxplots summarize the distribution of gain scores (median, interquartile range, and whiskers), with the dashed horizontal line indicating zero gain (no change from pre- to post-test).

Similar to Figure 2, the boxplot (Figure 3) indicates that the Traditional group has a higher central tendency (median) than the other groups, while the Flipped group shows the widest spread, including both larger positive gains and negative gains.

The descriptive statistics used to quantify the magnitude of pre–post change for the low-baseline subgroup (*PSVT:R*$_{pre}$ ≤ 18) and to establish whether gains are consistently positive across delivery modes are illustrated in Table 1. As reported in [24] and [25] for Traditional and Virtual-Asych., in Table 1, we see that among students who entered with lower spatial visualization scores (*PSVT:R*$_{pre}$ ≤ 18), Flipped instructional delivery is also associated with positive average gains from pre- to post-test.

Table 1. Descriptive statistics for *PSVT:R*$_{pre}$, *PSVT:R*$_{post}$, and *Gain* among students with *PSVT:R*$_{pre}$ ≤ 18

Instructional Mode	***n***	**Mean *PSVT:Rpre*, Mean (*SD*)**	**Mean *PSVT:Rpost*, Mean (*SD*)**	***Gain*, Mean (*SD*)**
Traditional	26	15.85 (2.46)	21.27 (3.88)	5.42 (3.61)
Virtual-Async.	13	15.08 (3.66)	19.54 (3.48)	4.46 (2.60)
Flipped	19	15.21 (2.97)	19.63 (4.81)	4.42 (5.63)
Overall	**58**	**15.47 (2.90)**	**20.35 (4.15)**	**4.88 (4.17)**

Overall, Figures 2 and 3 and Table 1 indicate that most students improved from pre- to post-test. The distributions also appear broadly similar across modalities, suggesting comparable improvement patterns across delivery modes for this low-baseline group.

To test whether gains reflected reliable pre–post improvement, we conducted paired analyses of each student's PSVT:R scores. Because gain scores satisfied normality, we used paired t-tests; we also ran Wilcoxon signed-rank tests to confirm robustness to distributional assumptions. Both approaches showed statistically significant pre–post improvement overall and within each delivery mode ($p < .05$), indicating that the observed gains are not artifacts of modeling assumptions.

As the paired-samples t-test assumes that the distribution of within-student differences is reasonably close to normal, before reporting the pre–post comparisons, it was important to check whether the PSVT:R difference scores (*Gain*) were approximately normally distributed. Normality tests did not detect departures from normality overall ($p = .849$) or within delivery modes (Traditional: $p = .280$; Virtual-Asynchronous: $p = .083$; Flipped: $p = .936$).

The paired *t*-tests showed that PSVT:R scores increased from pre to post in the overall sample (mean *Gain* = 4.88, $p < .001$). The paired *t*-tests also showed significant pre–post increases within each delivery mode: Traditional (mean difference = 5.42, $p < .001$), Virtual-Asynchronous (mean difference = 4.46, $p < .001$), and Flipped (mean difference = 4.42, $p = .003$). Overall, across all three delivery modes, students showed measurable gains in spatial visualization over the semester.

The robustness check with Wilcoxon signed-rank tests (which do not rely on normality assumptions) corroborated these results, showing significant pre–post increases overall ($p < .001$) and within each mode (*Traditional*: $p < .001$; *Virtual-Async.*: $p = .002$; *Flipped*: $p = .005$). These results suggest that students' PSVT:R scores increased, and the improvement was evident across all delivery modes, even when using a nonparametric test.

To address our primary research question, namely whether delivery mode is associated with different levels of improvement in spatial visualization for students entering with low PSVT:R scores, we compared PSVT:R *Gain* across the Traditional, Virtual-Asynchronous, and Flipped modalities using methods appropriate for small, unbalanced groups. Since unequal variances and unequal sample sizes are plausible in this setting, we used Welch's ANOVA as the primary omnibus test. Welch's one-way ANOVA indicated no statistically significant differences in PSVT:R *Gain* among the Traditional ($n = 26$), Virtual-Asynchronous ($n = 13$), and Flipped ($n = 19$) groups, $F(2.000, 32.624) = 0.514$, $p = .603$.

To assess whether any specific pair of delivery modes differed in mean gain, we conducted Games–Howell post hoc tests, which are appropriate under unequal variances. The estimated mean differences in *Gain*, *p*-values, and confidence intervals (CI) are reported in Table 2. All adjusted *p*-values exceeded .60, and all 95% confidence intervals included zero, indicating no statistically significant pairwise differences among the three delivery modes.

Table 2. Games–Howell post hoc comparisons of *Gain* across delivery modes

Pairwise comparison	Estimated mean difference in *Gain*	95% CI	Adjusted *p*-value
Traditional vs Virtual-Async.	0.962	[-3.45, 1.52]	.613
Flipped vs Traditional	1.00	[-2.64, 4.64]	.777
Flipped vs Virtual-Async.	0.0405	[-3.63, 3.71]	1.000
Note: Estimated mean differences in *Gain* are reported as the first-named group minus the second-named group.			

As a nonparametric sensitivity check, we used the Kruskal–Wallis test to compare gain-score distributions across delivery modes without relying on normality assumptions. The Kruskal–Wallis test likewise indicated no differences in gain distributions across modes, $\chi 2(df = 2) = 1.481$, $p = .477$, reinforcing the conclusion that PSVT:R gains were similar across the three delivery modalities.

The agreement between the primary unequal-variance mean-comparison (Welch's ANOVA with Games–Howell) and the nonparametric sensitivity analysis (Kruskal–Wallis) supports the robustness of this conclusion to common concerns in small, unbalanced samples (e.g., heteroscedasticity and departures from normality). Moreover, the Games–Howell confidence intervals were wide relative to the observed mean differences, indicating that any true modality effects, if present, are likely modest compared with the within-modality variability in gains for this low-initial subgroup.

Taken together, these findings indicate that students with lower initial spatial visualization scores improved meaningfully over the semester in each delivery format, and that the typical magnitude of improvement was similar across modes (approximately 4.4–5.4 PSVT:R points). Practically, this suggests that the course supports spatial-skills development for lower-baseline students regardless of whether instruction is delivered in Traditional, Virtual-Asynchronous, or Flipped formats. Across all three instructional delivery modes, students with low initial PSVT:R scores demonstrated statistically significant positive average gains, and the inferential analyses did not provide evidence that gains differed by modality.

3.2 RQ2 Results: Association Between PSVT:R *Gain* and Course Grade Points (*GradePts*)

To test whether improvement on the PSVT:R was associated with course performance, we fit an OLS regression model predicting course grade points, *GradePts,* (0.0–4.0) from PSVT:R *Gain*, baseline *PSVT:Rpre*, *Mode*, and *Gain×Mode* interaction terms, using *Mode:Traditional* as the reference category. We provide the full set of coefficient estimates, uncertainty measures, and *p*-values in Table 3.

As shown in Table 3, the overall model was statistically significant ($F(6, 51) = 4.747$, $p < .001$) and explained a meaningful share of variation in course grade points, *GradePts*, ($R^2 = 0.358$; adjusted $R^2 = 0.283$; residual standard error 0.487). In other words, improvement on the PSVT:R was positively associated with course performance.

We note that *Traditional* does not appear as a separate row in Table 3, because *Mode* was coded with *Traditional* as the reference category (treatment coding). Under this coding, the intercept and the coefficients for *Gain* and *PSVT:Rpre* correspond to the Traditional group, and the *Mode* coefficients shown (*Virtual-Async.* and *Flipped*) represent differences in expected grade points relative to Traditional (with *Gain × Mode* terms indicating differences in the gain–performance slope relative to *Traditional*).

Table 3. OLS regression predicting course grade points (*GradePts*) from PSVT:R *Gain*, baseline PSVT:R, and instructional delivery mode.

Predictor	Coefficients	SE	*t*-score	*p*-value	95% CI
Intercept	2.110***	0.446	4.732	< .001	[1.215, 3.006]
Gain	0.057*	0.027	2.112	.040	[0.003, 0.112]
Mode:Virtual-Async.	- 0.079	0.328	- 0.241	.811	[-0.737, 0.579]
Mode:Flipped	0.766**	0.226	3.389	.001	[0.312, 1.220]
PSVT:Rpre	0.056*	0.025	2.273	.027	[0.007, 0.105]
Gain × Virtual-Async.	0.086	0.061	1.397	.169	[-0.037, 0.209]
Gain × Flipped	- 0.023	0.034	- 0.692	.492	[-0.092, 0.045]

Notes: *Traditional* is the reference category for *Mode*. *Gain = PSVT:Rpost – PSVT:Rpre*. Model fit: $R^2 = 0.358$, adjusted $R^2 = 0.283$, $F(6, 51) = 4.747$, $p < .001$, residual SE = 0.487. $p < .10$. $^*p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Two predictors, namely *Gain* and *PSVT:Rpre*, were positively associated with course grade points, *GradePts*. Furthermore, *Gain* was a statistically significant positive predictor ($\beta = 0.057$, $p = .040$), indicating that, holding baseline *PSVT:Rpre* and delivery mode constant, a one-point larger *Gain* corresponds to an estimated 0.057 increase in *GradePts*. Baseline *PSVT:Rpre* was also positively associated with *GradePts* ($\beta = 0.056$, $p = .027$).

Regarding delivery mode effects (relative to *Traditional*), *Virtual-Async*. was not detectably different from *Traditional* ($\beta = -0.079$, $p = .811$), as reported in [25]. However, relative to *Traditional*, the *Flipped* had higher estimated *GradePts* ($\beta = 0.766$, $p = .001$). Moreover, the *Gain×Mode* interaction terms reported in Table 3 were not statistically significant (*Gain×Virtual-Async.*: $\beta = 0.086$, $p = .169$; *Gain×Flipped*: $\beta = -0.023$, $p = .492$). This indicates that the positive association between *Gain* and *GradePts* did not differ detectably across delivery modes in this sample.

In summary, results show that PSVT:R *Gain* was a positive predictor of course grade points after accounting for baseline *PSVT:Rpre* and delivery mode, indicating that students who improved more in spatial visualization tended to earn higher grades. Furthermore, the findings do not show evidence of an association that differs by delivery mode, as the *Gain*×*Mode* interaction terms were not statistically significant.

These findings suggest that improvements in spatial visualization are linked to stronger course outcomes, even after accounting for initial preparation. They also suggest that this relationship holds across delivery formats, and that the higher grades observed in the Flipped format reflect higher overall performance rather than a different gain–grade relationship.

4. Practical Implications for First-Year Engineering Programs

The results suggest that, regardless of delivery mode, students who begin engineering programs with lower spatial visualization performance can benefit from effective interventions delivered over a 10-module curriculum. Thus, first-year engineering programs can implement an established spatial visualization curriculum in Traditional, Virtual-Asynchronous, or Flipped formats without sacrificing PSVT:R gains for students who enter with lower initial scores. This flexibility allows programs to choose a delivery mode based on local constraints (staffing, classroom space, scheduling, or continuity needs) rather than assuming that one modality is inherently more effective for improving spatial skills.

The positive association between PSVT:R *Gain* and *GradePts* also supports integrating spatial visualization instruction early in the first-year curriculum and tracking student improvement with pre/post assessments. Programs can use these data to identify students who may benefit from additional support and to strengthen implementation features that promote engagement and completion across modes, such as clear pacing, frequent low-stakes practice, structured checkpoints in the LMS, and timely feedback when students struggle.

5. Limitations

This study reports results from a single institutional setting and should be interpreted within that context. The analytic sample for the lower-prepared subgroup is modest (n = 58), and group sizes vary by delivery mode, so estimates are less precise and small between-mode differences may be difficult to detect. Since delivery modes were implemented in different academic years, cohort and contextual factors may have contributed to observed patterns in addition to mode. The fact that two cohorts were experiencing the spatial skills intervention during the global pandemic (Virtual-Asynchronous), or were pandemic-adjacent (Flipped), could also be affecting the results. Lastly, due to the grading structure of the PSVT:R assessments in the Virtual-Asynchronous and Flipped formats, gains observed in these instructional modalities may not be a true indicator of students' spatial ability, but may instead be a byproduct of decreased effort when assessment outcomes were disconnected from grade implications. These factors limit direct generalization to other institutions.

Even so, the findings are likely transferable to similar first-year engineering contexts because the intervention structure was standardized across modes, including a common 10-module curriculum and the same pre/post-test. Future work using larger, multi-institution samples with more balanced group sizes and standardized assessment administration would strengthen the evidence base and clarify the extent of generalizability.

6. Conclusion

This study examined whether instructional delivery mode relates to spatial visualization gains and whether those gains align with course performance in a first-year engineering spatial visualization course delivered in three formats: Traditional face-to-face, Virtual-Asynchronous, and Flipped classroom. We assessed spatial reasoning using PSVT:R pre- and post-test scores and defined *Gain* as $PSVT{:}R_{post} - PSVT{:}R_{pre}$, with primary analyses focusing on students with lower initial scores ($PSVT{:}R_{pre} \leq 18$).

Students exhibited statistically significant positive mean PSVT:R gains across all three delivery modes, yet the inferential analyses did not detect statistically significant differences in gains across modes. These results indicate that, within this instructional context and for lower-prepared students, the 10-module multimodal curriculum produced comparable improvements in spatial visualization across Traditional face-to-face, Virtual-Asynchronous, and Flipped classroom implementations.

Regression analyses showed that PSVT:R gain positively predicted course grade points after controlling for baseline PSVT:R and delivery mode. The $Gain \times Mode$ interaction was not statistically significant, indicating no detectable variation across delivery modes in the relationship between spatial visualization improvement and course performance. Taken together, the findings suggest that spatial visualization gains are meaningfully connected to academic outcomes in the course and that this relationship appears stable across delivery formats in this sample.

Because the study compares non-randomized cohorts at a single institution taught by a single instructor, the findings should be interpreted as context-specific. Future work should replicate these analyses with larger, multi-institution samples and incorporate additional covariates (e.g., prior academic preparation and demographic factors) to better isolate delivery mode effects from cohort-level differences. Even with these constraints, the results provide practical evidence that programs can implement established spatial visualization curricula in multiple delivery formats while maintaining student gains and that improving spatial visualization skills aligns with stronger course performance.

Our future research will explore how assessment context influences effort, score reliability, and student engagement, especially when tests are used for placement or instructional decision-making. We also intend to examine whether differences in effort or completion rates align with testing conditions or PSVT:R pre-test scores.

References

[1] S. A. Sorby, N. Veurink, and S. Streiner, "Does spatial skills instruction improve STEM outcomes? The answer is 'yes,'" *Learn. Individ. Differ.*, vol. 67, pp. 209–222, Oct. 2018, doi: 10.1016/j.lindif.2018.09.001.

[2] D. H. Uttal *et al.*, "The malleability of spatial skills: a meta-analysis of training studies," *Psychol. Bull.*, vol. 139, no. 2, pp. 352–402, Mar. 2013, doi: 10.1037/a0028446.

[3] S. A. Sorby, B. Casey, N. Veurink, and A. Dulaney, "The role of spatial training in improving spatial and calculus performance in engineering students," *Learn. Individ. Differ.*, vol. 26, pp. 20–29, Aug. 2013, doi: 10.1016/j.lindif.2013.03.010.

[4] S. A. Sorby, A. F. Wysocki, and B. Baartmans, *Introduction to 3D spatial visualization: an active approach*. Clifton Park, NY: Delmar Learning, 2003.

[5] J. L. Segil, J. F. Sullivan, B. A. Myers, D. T. Reamon, and M. H. Forbes, "Analysis of multi-modal spatial visualization workshop intervention across gender, nationality, and other engineering student demographics," in *2016 IEEE Frontiers in Education Conference (FIE)*, Oct. 2016, pp. 1–5. doi: 10.1109/FIE.2016.7757525.

[6] E. Atindama, M. Ramsdell, D. P. Wick, S. Mondal, and P. Athavale, "Impact of targeted interventions on success of high-risk engineering students: a focus on historically underrepresented students in STEM," *Front. Educ.*, vol. 10, Feb. 2025, doi: 10.3389/feduc.2025.1435279.

[7] J. Wai, D. Lubinski, and C. P. Benbow, "Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance," *J. Educ. Psychol.*, vol. 101, no. 4, pp. 817–835, 2009, doi: 10.1037/a0016127.

[8] R. M. Webb, D. Lubinski, and C. P. Benbow, "Spatial ability: A neglected dimension in talent searches for intellectually precocious youth," *J. Educ. Psychol.*, vol. 99, no. 2, pp. 397–420, 2007, doi: 10.1037/0022-0663.99.2.397.

[9] H. J. Kell, D. Lubinski, C. P. Benbow, and J. H. Steiger, "Creativity and technical innovation: Spatial ability's unique role," *Psychol. Sci.*, vol. 24, no. 9, pp. 1831–1836, 2013, doi: 10.1177/0956797613478615.

[10] D. L. Shea, D. Lubinski, and C. P. Benbow, "Importance of assessing spatial ability in intellectually talented young adolescents: A 20-year longitudinal study," *J. Educ. Psychol.*, vol. 93, no. 3, pp. 604–614, 2001, doi: 10.1037/0022-0663.93.3.604.

[11] N. Veurink and S. A. Sorby, "Raising the Bar? Longitudinal Study to Determine which Students Would Benefit Most from Spatial Training," in *ASEE 2011*, Jun. 2011, p. 22.1210.1-22.1210.13. doi: 10.18260/1-2--18592.

[12] S. A. Sorby, "Educational Research in Developing 3‐D Spatial Skills for Engineering Students," *Int. J. Sci. Educ.*, vol. 31, no. 3, pp. 459–480, Feb. 2009, doi: 10.1080/09500690802595839.

[13] D. DeSutter and M. Stieff, "Teaching students to think spatially through embodied actions: Design principles for learning environments in science, technology, engineering, and mathematics," *Cogn. Res. Princ. Implic.*, vol. 2, no. 1, p. 22, Mar. 2017, doi: 10.1186/s41235-016-0039-y.

[14] R. B. Guay, *Purdue Spatial Visualization Test*. West Lafayette, IN: Purdue Research Foundation., 1976.

[15] Y. Maeda, S. Y. Yoon, K. Kim-Kang, and P. K. Imbrie, "Psychometric Properties of the Revised PSVT: R for Measuring First Year Engineering Students' Spatial Ability," *Int. J. Eng. Educ.*, vol. 29, Jan. 2013.

[16] S. A. Sorby, "Developing 3D spatial skills for engineering students," *Australas. J. Eng. Educ.*, vol. 13, no. 1, pp. 1–11, Jan. 2007, doi: 10.1080/22054952.2007.11463998.

[17] I. Blau and T. Shamir-Inbal, "Re-designed flipped learning model in an academic course: The role of co-creation and co-regulation," *Comput. Educ.*, vol. 115, pp. 69–81, Dec. 2017, doi: 10.1016/j.compedu.2017.07.014.

[18] J. Broadbent and W. L. Poon, "Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review," *Internet High. Educ.*, vol. 27, pp.

1–13, Oct. 2015, doi: 10.1016/j.iheduc.2015.04.007.
[19] S. Drumm, “Applying individual strategies enhances learning in asynchronous learning paths,” *Comput. Educ. Open*, vol. 8, p. 100257, Jun. 2025, doi: 10.1016/j.caeo.2025.100257.
[20] D. Bairaktarova, D. Williams, and P. J. Katsioloudis, “Analysis of Blended and Multi-Modal Instruction and Its Effects on Spatial Visualization Ability,” *Eng. Des. Graph. J.*, vol. 83, pp. 14–26, 2019.
[21] R. Porat and C. Ceobanu, “Enhancing Spatial Ability: A New Integrated Hybrid Training Approach for Engineering and Architecture Students,” *Educ. Sci.*, vol. 14, no. 6, p. 563, Jun. 2024, doi: 10.3390/educsci14060563.
[22] C. Melgosa Pedrosa, B. R. Barbero, and A. R. Miguel, “Spatial Visualization Learning in Engineering: Traditional Methods vs. a Web-Based Tool,” *Educ. Technol. Soc.*, vol. 17, no. 2, pp. 142–157, 2014.
[23] C. Melgosa Pedrosa, B. Ramos Barbero, and M. E. Baños García, “Interactive learning management system to develop spatial visualization abilities,” *Comput. Appl. Eng. Educ.*, vol. 23, no. 2, pp. 203–216, Mar. 2015, doi: 10.1002/cae.21590.
[24] N. Veurink, M. Richards, and S. A. Sorby, “A Tale of Two Universities: Developing Spatial Skills of Engineering Students during a Global Pandemic,” *Eng. Des. Graph. J.*, vol. 84, pp. 10–27, Feb. 2021.
[25] M. C. Richards, N. L. Veurink, and S. A. Sorby, “Teaching Spatial Skills Online During a Global Pandemic,” presented at the 2021 ASEE Virtual Annual Conference Content Access, Jul. 2021. Accessed: Jan. 20, 2026. [Online]. Available: https://peer.asee.org/teaching-spatial-skills-online-during-a-global-pandemic
[26] J. C. Fruehwirth, S. Biswas, and K. M. Perreira, “The Covid-19 pandemic and mental health of first-year college students: Examining the effect of Covid-19 stressors using longitudinal data,” *PLOS ONE*, vol. 16, no. 3, p. e0247999, Mar. 2021, doi: 10.1371/journal.pone.0247999.
[27] H. B. P. Gerson, S. Sorby, A. F. Wysocki, and B. J. Baartmans, “The Development and Assessment of Multimedia Software for Improving 3-D Spatial Visualization Skills,” *Comput. Appl. Eng. Educ.*, vol. 9, no. 2, pp. 105–113, Jan. 2001, doi: 10.1002/cae.1012.
[28] A. W. Chickering and Z. F. Gamson, “Seven Principles for Good Practice in Undergraduate Education,” Mar. 1987. Accessed: Jan. 20, 2026. [Online]. Available: https://eric.ed.gov/?id=ED282491
[29] S. A. Sorby and B. J. Baartmans, “The Development and Assessment of a Course for Enhancing the 3-D Spatial Visualization Skills of First Year Engineering Students,” *J. Eng. Educ.*, vol. 89, no. 3, pp. 301–307, 2000, doi: 10.1002/j.2168-9830.2000.tb00529.x.
[30] K. S. Harris, L. V. Harris, and M. A. Sadowski, “Measuring Spatial Visualization in a High School Bridge Program,” in *Graphics—in a Changing World*, USA: ASEE, Oct. 2009, pp. 127–131.
[31] G. D. Hoople, E. Cowan, L. D. Van Einde, J. Tara, and N. Delson, “Teaching Spatial Visualization: A Controlled Trial of a Touchscreen App Implemented as Homework,” in *2018 IEEE Frontiers in Education Conference (FIE)*, Oct. 2018, pp. 1–5. doi: 10.1109/FIE.2018.8659068.
[32] R. R. Hake, “Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses,” *Am. J. Phys.*, vol. 66, no. 1, pp. 64–74, Jan. 1998, doi: 10.1119/1.18809.
[33] B. L. Welch, “On the Comparison of Several Mean Values: An Alternative Approach,” *Biometrika*, vol. 38, no. 3–4, pp. 330–336, 1951, doi: 10.1093/biomet/38.3-4.330.