Paper ID #50926

Impact of Virtual Reality Intervention on Word-Based Problem-Solving Skills in Chemical Engineering

Syahidah Mohd Khairi, University of Michigan

Syahidah Mohd Khairi is a Ph.D. candidate in the Department of Chemical Engineering, University of Michigan, Ann Arbor. Syahidah received her Bachelor of Science in Chemical Engineering from The Ohio State University, Master of Science in Applied Computing from Universiti Teknologi PETRONAS in Malaysia, and Master of Science and Engineering in Chemical Engineering from University of Michigan, Ann Arbor. Her current research interests are colloidal self-assembly for crystals and gels application, and educational technology for engineering and nanoscience education.

Riley Michael Garliauskas, University of Michigan

Riley Garliauskas is a Startup Engineer at Membrane System Specialists in Wisconsin Rapids, Wisconsin. He earned both his Bachelor of Science (2023) and Master of Science (2024) in Chemical Engineering from the University of Michigan. During his time at the university, he served for three years as a teaching staff member in both the Department of Chemical Engineering and the Department of Mathematics. Riley is passionate about communicating complex concepts through innovative and engaging methods to inspire the next generation of engineers. His work in education and mentorship has included volunteer teaching, coordination of youth STEM outreach in his community, and leadership of organizations promoting integrity and excellence.

Prof. Albert Tianxiang Liu, University of Michigan

Albert Liu is an Assistant Professor of Chemical Engineering, Materials Science and Engineering, and Macromolecular Science and Engineering at the University of Michigan, Ann Arbor. He is a member of the Biointerfaces Institute and the Rogel Cancer Center. Albert received his B.S. in Chemical Engineering from Caltech in 2014 and his Ph.D. in Chemical Engineering from MIT in 2020, working with Michael Strano. Before joining Michigan, he completed postdoctoral training in the Stanford University School of Medicine with Steven Chu in 2022. His research group, the Laboratory of Emergent and Galvanically Orchestrated Systems, focuses on colloidal and surface engineering to control nanoscale interfacial transport processes in micro-confined materials.

©American Society for Engineering Education, 2026

Impact of Virtual Reality Intervention on Word-Based Problem-Solving Skills in Chemical Engineering

Abstract

Spatial ability is central to engineering problem solving. When problems are presented only through written or verbal descriptions, learners must construct an internal representation of the system without the aid of a diagram, which can be especially challenging for students with lower spatial ability. From the cognitive load theory perspective, these students may experience greater cognitive load while solving a problem because of working memory constraints. Virtual reality (VR) may mitigate this burden by externalizing key spatial features and supporting mental visualization. Here, we examine whether VR interventions improve word-based problem-solving performance in an undergraduate chemical engineering heat and mass transfer course, with spatial ability treated as a moderating factor. We developed supplemental VR modules for head-mounted displays (VR-HMDs) that enable students to visualize three-dimensional transient temperature and concentration profiles. The modules were assigned after the corresponding topics were introduced in lecture. To compare VR integration with a non-VR alternative, students chose between completing the VR modules or parallel programming assignments covering the same systems. Using a pre-test-post-test design, we assessed students' ability to solve word-based heat and mass transfer problems using two tests, one focusing on the problem representation stage (multiple-choice format) and one on the solving stage (constructed-response format). We compared learning gains between VR and non-VR groups while accounting for spatial ability. Spatial ability was measured before the pre-tests and after the post-tests using the Revised Purdue Spatial Visualization Test: Rotations (Revised PSVT:R), with items customized to reflect heat and mass transfer contexts. Our preliminary results suggest that VR improved word-based engineering problem-solving for students with low and high spatial ability in distinct ways. We observed that students with low spatial ability who used VR performed better in problem-representation-focused questions compared to those who did not use VR. This trend suggests the potential reduction of cognitive load during the problem representation stage, which was not observed for students with high spatial ability. Meanwhile, for the solving-stage-focused test, students with high spatial ability were associated with increased performance when they used VR. Surprisingly, students with low spatial ability in this study who completed the VR assignment appeared to perform worse in this test than those who completed the programming assignment, indicating a possible increase in cognitive loads that warrants further investigation. These findings can guide instructors integrate VR strategically into engineering curricula and guide students in selecting learning tools aligned with their needs.

1 Introduction

Engineering problem-solving processes typically involve mental visualization of the studied system. Often, engineers are presented with verbal or written problem descriptions and are expected to diagnose and solve issues in real time. To solve a problem, problem definition is a key step, in addition to the technical expertise and mathematical skills [1]. This step requires rapid visualization of the presented systems to provide prompt solutions, which is related to spatial ability. Studies have shown that spatial visualization is an important skill to predict the performance of students in science, technology, engineering and mathematics (STEM), where students with low spatial ability typically dropped out of the program [2]. One example of low spatial ability is the aphantasia condition, which is a neurological condition where the person does not have the ability to construct mental images [3]. Students without an accurate mental representation of abstractly proposed problems may struggle in solving practical engineering problems they receive via word-based descriptions, inhibiting success in their engineering career trajectory. In order to decrease the barrier for students with low spatial ability to succeed in engineering, educational interventions should help providing the tools to allow for mental visualizations for better conceptual understanding and enhance the subsequent problem-solving skills. Virtual reality (VR) has been shown to improve spatial ability [4], [5], [6] and benefit students with low spatial ability [7], [8], [9]. The sense of presence and agency afforded through immersive VR as compared to screen-based VR further enhances the factors (e.g. embodied cognition) to achieve the learning outcomes [10]. Because VR implementation could be time-consuming and costly for both students and instructors, an informed strategy (e.g. prioritization based on spatial ability) is important to ensure that the implementation is beneficial in achieving the learning outcomes for students of diverse backgrounds.

2 Literature Review and Theories

2.1 Cognitive Load Theory and Spatial Ability

Cognitive load theory suggests that humans have limited working memory capacity that affects problem-solving skills [11]. Working memory is a short-term memory that receives inputs from the sensory register and also retrieves information from the long-term memory storage [12], serving as storage and processing capacity [13], [14]. Both time and memory resources serve as limitations in working memory [14], [15]. For a limited processing time such as solving problems promptly, the needed working memory capacity is higher. Baddeley's working memory model proposes that the working memory consists of the central executive as the main control, phonological loop for processing auditory and verbal contents, the visuospatial sketchpad for processing visual contents, and the episodic buffer to connect between the phonological loop and the visuospatial sketchpad [13], [14]. The visuospatial sketchpad (including visual, spatial and possibly kinesthetic components) includes both the storing and manipulating of visual images [13]. In the case of the problem-solving process, the steps of recognizing and forming the visuals of the problem statement may compete with other components, such as mental arithmetic and conceptualization, for the limited capacity of working memory [14]. This means that if there is an increase in the cognitive load to form the images, the capacity to process and solve related problems will be lower, affecting the solving speed. Spatial abilities, especially the spatial visualization ability, relate to the visuospatial working memory in that they require the use of memory for visual storage and spatial visualization tasks, which involve manipulation of a visual in multiple steps [16]. Individuals with higher spatial ability experience less cognitive load as they require less engagement of visuospatial working memory.

2.2 Spatial Ability and Word-Based Problem-Solving

In math-related problems, which typically includes engineering problems, visuospatial working memory is shown to be an important component in addition to intelligence, especially for word-based problems where the problems are represented only using words and no visuals are provided [17]. Individuals with low spatial ability struggle with word-based problems, particularly the problem representational stage that includes the visualization of the system's physical shapes and dimensions and manipulating the system accurately [1]. High spatial ability has been shown to be important in the problem-solving success of mathematical word problems through the correct choice of schema [18]. These students have the mental representations that they can retrieve readily and focus on solving the problems, while students with low spatial ability experience greater cognitive load to mentally represent the problems, reducing the working memory available for the solving stage [19]. Based on these findings, spatial ability is clearly important in solving word-based problems.

2.3 Spatial Ability and Virtual Reality

Educational interventions should accommodate students with low spatial ability. One strategy is to use Immersive Virtual Reality (IVR; using a head-mounted display, HMD), which has been shown to improve the educational experiences both in cognitive and affective factors [10]. A study by Sun et al. shows that for students with lower spatial ability, the use of IVR as a learning tool reduces the cognitive load of students compared to using presentation slides, which improves learning performance. Meanwhile, for students with high spatial ability, there are no cognitive load differences [7]. This learning performance difference between spatial ability levels is also consistent with the use of desktop VR as compared to presentation slides [8], and with internet-based immersive virtual environment compared to traditional 2-D learning [9]. Moreover, several studies suggest that IVR can be used to train students in improving their spatial skills [4], [5], [6], which may explain the benefits for students with low spatial ability. These findings show that IVR-based learning is beneficial for students with low spatial ability, which suggest that IVR intervention will improve their word-based problem-solving skills.

2.4 Applying Virtual Reality in Education

Among the first virtual reality modules in chemical engineering are microscopic views of a catalytic reaction and non-isothermal packed bed reactor processes, designed to accommodate students of different learning styles [20]. Here, IVR helps in chemical plant exploration with reduced time, cost and space, and to manipulate and/or dissect the equipment for better understanding beyond the outer surface [20], [21], [22]. Another important skill for chemical engineering students is the problem-solving skill (identify, formulate and solve), which is one of the learning outcomes in Criterion 3 for Accreditation Board for Engineering and Technology (ABET) accreditation for Baccalaureate Level Programs [23]. To instill problem-solving skills outside of lecture, transfer of knowledge from classroom is important. The study by Araiza-Alba et al. shows that IVR intervention improves the completion rate of the learning activities, as well as improving the transfer of knowledge through embodied interaction [24]. This shows that IVR has the potential in both learning and knowledge transfer phases. However, IVR may distract the students from the actual learning due to increase in emotions and cognitive load [25], decreasing its effectiveness [26]. To mitigate this problem, an introduction to the core concepts prior to using IVR, or pre-training, has been shown to improve knowledge transfer and self-efficacy of the students after using VR [27]. This is because pre-training reduces the cognitive load during the VR intervention by reducing the amount of new knowledge to be processed [27], [28]. One way to incorporate the pre-training is by first conducting the traditional classroom lecture to introduce the core concepts [27]. In addition, it can also help in accommodating students with different learning preferences and accessibility issues (e.g. motion sickness) [21].

2.5 Research Purpose

Since there is a common theme of spatial ability between VR and math-based problem-solving, this study aims to understand the benefits of IVR integration as a supplemental learning tool on engineering word-based problem-solving performance for students with low and high spatial ability. Although previous literature has shown that IVR improves the learning performance of students with low spatial ability, the context of the learning focuses on knowledge that mainly requires memorization [7], [8], [9] and hands-on equipment training [21], [22], [29]. In this study, we focus on the problem-solving aspect which has been shown to be affected by spatial ability [1], [18], [19]. We conducted a pre-test-post-test study in a chemical engineering heat and mass transfer course that requires students to solve temperature and concentration profiles in a given system as explained by word problems. By supplementing the learning experience using IVR, we expect that it will reduce the cognitive load while solving the problems, especially for students with low spatial ability, thus increasing problem-solving performance.

3 Research Hypotheses

The purpose of this study is to examine the effectiveness of VR as a supplemental instructional tool in improving engineering students' word-based problem-solving skills, with spatial ability as a moderating factor. The hypotheses of this study are:

1. Students who use VR-HMD as a supplemental instructional tool demonstrate improved performance in solving word-based chemical engineering transport problems, compared to those using traditional methods.
2. Spatial ability moderates the effect of VR-HMD on word-based problem-solving performance.

4 Theoretical Framework

We used the cognitive load theory to guide our methods, data analysis and interpretation. VR intervention has the potential to reduce students' cognitive load during the learning process by providing the 3D immersive visualization. This supports the construction of schema that students can later recall and apply to solve the word problems. In addition, we also hypothesized that spatial ability acts as a moderating factor. Students with lower spatial ability may experience increased cognitive load that increase the higher demand of visuospatial working memory during the problem-solving tasks. This could reduce their working memory capacity for other conceptual and arithmetic aspects of the tasks. We used spatial rotation tests to assess students' spatial ability at multiple points across the semester, before and after the learning interventions.

5 Methods

We performed a quasi-experimental quantitative pre-test-post-test study in a heat and mass transfer course to study the cognitive improvement and spatial skills change among chemical engineering undergraduate students after the VR educational interventions as shown in Figure 1.

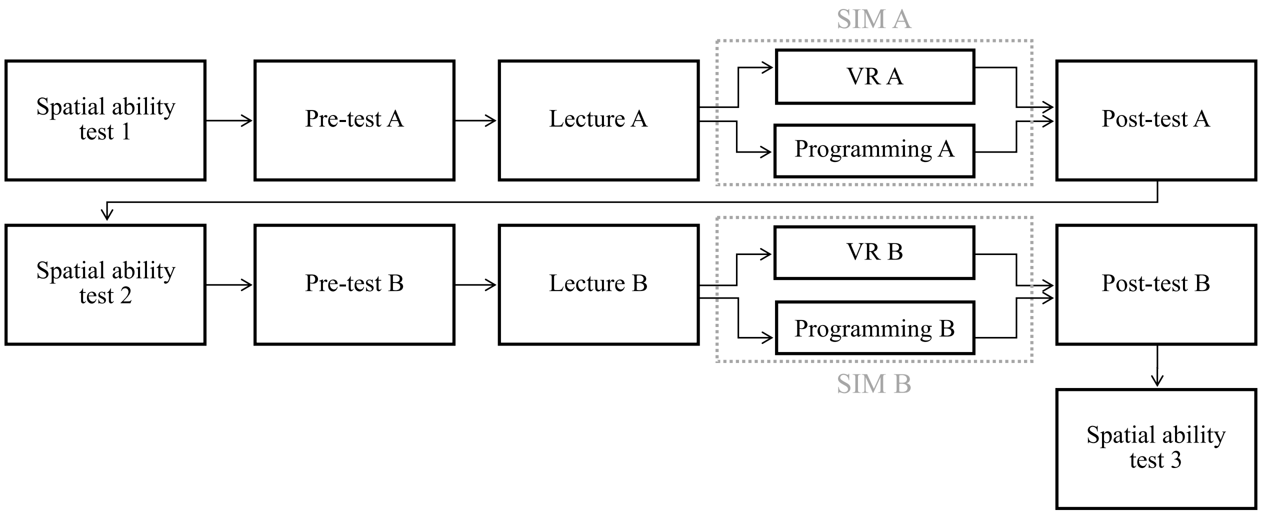

Figure 1: Flow diagram of data collection and learning interventions.

5.1 *Participant recruitment*

This study was conducted at a large public university in the United States. Heat and mass transfer is a core course in the undergraduate chemical engineering program, typically taken by third-year students. All students in the heat and mass transfer course were invited to participate in this study via a consent form. The participation in this study was separate from the course activities and did not affect their grades. Data from students who did not consent to participate in this study were removed from the research data and not reported in this paper. A total of 86 students consented to participate in this research. This study was reviewed and determined as exempt by the Institutional Review Board (HUM00279628).

5.2 *Course structure and intervention groups*

One of the learning objectives of this course is to solve problems involving steady and unsteady state heat transfer and mass transfer. Participation points consist of 5% of the final grade, which include spatial test and pre-test completions, and post-test grades. The homework grade consists of 20% of the final grade, where 10% of this includes two Safety/Industry Module (SIM) assignments on unsteady state heat transfer and unsteady state mass transfer. For each SIM assignment, students were able to choose either a VR assignment or a programming exercise on the finite-difference method of solving for temperature and concentration profiles. These assignments were designed to be similar in content, with the main difference being the direct visualization and measurement in VR versus the derivation and implementation of the finite-difference method using Microsoft Excel. After related lectures, students were able to see both assignments at the same time and choose their preferred assignment. The intervention groups were designated based on their choice: students who choose the VR experience were referred to as “VR group”, while students who choose the programming component were referred to as "non-VR group." To provide the opportunity for all students to complete the VR assignment given the limited number of HMD devices, students had 1 week to complete the SIM (VR and programming) assignments on their own time. Staff members were available during office hours to help with technical issues. For VR assignments, students were encouraged to complete the exercises in pairs, but each student needed to submit their own answers for grading to ensure that they completed the assignments.

5.3 VR-HMD learning activities

Two VR modules (VR A and VR B for unsteady state heat and mass transfer respectively, see Figure 2) were developed using an in-house VR development software and visualized in an immersive scene using the Meta Quest Pro HMD. Since these are time-dependent spatial profiles, the scene is animated where the colors represent the temperature and concentration profiles for heat and mass transfer respectively. Each scene consists of 200 total time frames, and the animations were visualized at 20 frames per second by default. Students were able to speed up/slow down and pause/play the animations using keyboard keys. Another important tool used in these modules is the measuring tool controlled using Meta Quest Pro controllers, where students measured the distance between two points in the 3D virtual space.

VR A was designed to understand unsteady heat transfer problem (conduction and convection) by analyzing the transient 3-dimensional sphere temperature profiles at 3 different Biot numbers (corresponding to 3 columns in Figure 2a). There were 3 different visualizations for each Biot number (corresponding to 3 rows in Figure 2a): a hemisphere, an octant (1/8 of a sphere) and a radial slice that is assumed to be infinitely thin crossing the centroid of the spheres. The octants and radial slices were included to helps students see the internal temperature profiles easily as compared to the hemispheres. The colors correspond to the scaled temperature relative to the initial temperature and the surrounding temperature, where red (value 1 in the color bar) corresponds to the initial uniform temperature of the sphere, and dark blue (value 0 in the color bar) corresponds to the surrounding temperature (assumed as constant). The color changes across time as incorporated in the animation feature. While completing this activity, students were asked to calculate and drag-and-drop the Biot numbers to the corresponding columns, measure and compare the temperature gradients for different Biot numbers and identify the scenarios when the temperature profiles change.

VR B was designed to facilitate understanding of unsteady state mass transfer concepts through visualizing the concentration change of solvent in spherical polymer pellets as they move through a tank of air. Similar to VR A, these pellets start as red (value of 1) corresponding to the initial uniform concentration of solvent in the pellets and become dark blue (value of 0) as the solvents diffuse out of the pellets, the same as the concentration of solvent in the surrounding (assumed as constant). There were two visuals present in the VR scene. First, 1000 pellets as seen as small spheres were dynamically moving from the top to the bottom of the cylinder tank, as shown on the right in Figure 2b. Students were able to track the random walk motion of the pellets using a marker annotation tool using the controller. Three of these 1000 particles were highlighted with the color corresponding to their volume-average concentration value to allow students to focus on them; other particles remained in the dark blue color. On the left of Figure 2b, an octant represents the actual concentration profile in each pellet at the same time. Students were asked to determine the radial location where the concentration is the same as the volume-average concentration seen on the pellets. Moreover, students were also introduced with the concept of residence time and collected data to fit a curve of concentration versus time and location in the tank.

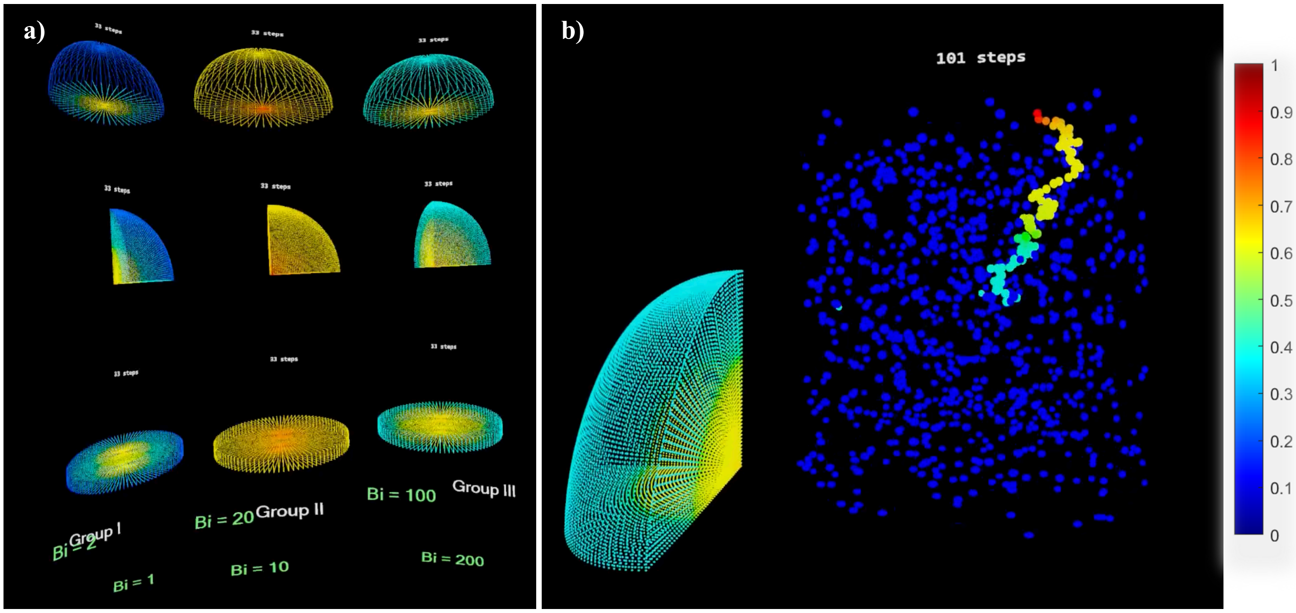

Figure 2: VR visualization scenes for (a) VR A at an intermediate time step (t = 33 or "33 steps" as shown above each temperature profiles) and (b) VR B at an intermediate time step (t = 101) as seen through the VR-HMD.

5.4 Pre-tests and post-tests

To measure the problem-solving performance before and after the learning experiences (lecture and SIM), we administered pre-tests in the beginning of the respective lectures and post-tests after the respective VR experiences were due (5 days for VR A and 7 days for VR B). Pre-tests were graded by completion, while post-tests were graded by correctness, as part of participation grades. In addition, students who answered correctly in pre-tests were able to recover up to 1 missing lecture point. The test solutions were released after the post-tests were administered.

These tests were conducted during the scheduled class time using two different formats to address two different problem-solving stages: a 5-minute, 5-multiple-choice question set (pre- and post-tests A, problem-representation-stage-focused, using Google Form) and a 10-minute constructed-response question set (pre- and post-tests B, solving-stage-focused, paper-based). For pre- and post-tests A regarding heat transfer, students were given scenarios of objects of different shapes and scene changes (i.e. object rotation) and were asked to predict the corresponding temperature profiles or any changes in heat transfer that occurred (e.g. heat transfer rate and system/model considerations). There were three unique systems in the test and two questions involving manipulation of the object orientation. For each question, students needed to select one answer among three given choices. For pre- and post-tests B, students were given a scenario of mass transfer diffusion in a hemispherical object and were asked to define the problem through sketching the system (20%), recognition of initial and boundary conditions (50%), manipulation of the object (10%), and prediction of the concentration profile (20%).

5.5 Spatial ability test

Purdue Spatial Visualization Test: Rotation (PSVT:R) is a spatial ability test used widely in STEM contexts due to its reliability, relevance to STEM, appropriate difficulty and ease of use [30], [31], [32]. The original PSVT:R consists of 30 test items to be conducted within 20 minutes. Based on PSVT:R, we measured students' spatial ability before and after the interventions (3 test instances: 3 days before pre-test A in lecture, 11 days after post-test A in lecture, and 9-14 days after post-test B, which students were allowed to complete on their own time). The questions were contextualized for heat transfer 2D and 3D simple shape factors and radiation view factors (Figure 3). For each test, we administered 5 questions with a set time limit of 90 seconds. The completion of these tests (graded by completion) contributed to the in-class participation grade.

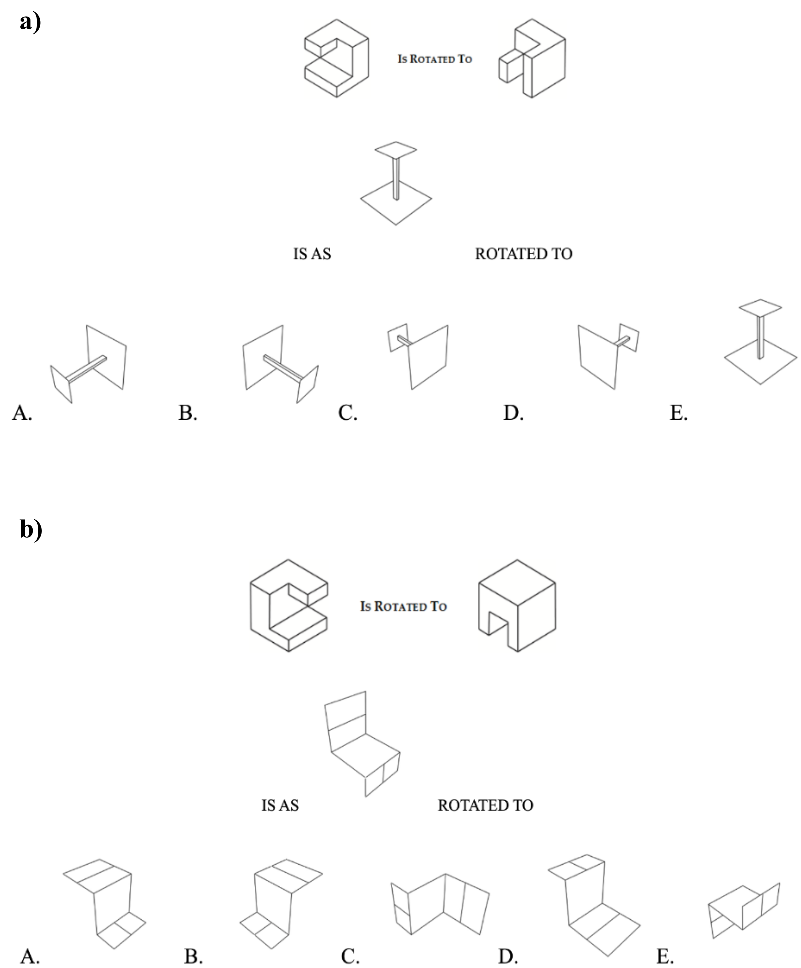

Figure 3: Examples of spatial ability tests contextualized to mass and heat transfer course based on (a) simple shapes covered in 3-dimensional heat transfer and (b) radiation view factors.

5.6 *Data analysis and representation*

Statistical analyses and data visualizations were performed using GraphPad PRISM 10 software. The statistical significance was set to 0.05, where $p < 0.05$ is considered as statistically significant (represented by asterisk, *) and $p \geq 0.05$ is considered not statistically significant (represented as "ns"). We used the Mixed-Effects Model (Restricted Maximum Likelihood, REML) for statistical analysis to account for repeated measures of the same students over the semester. Effective matching was tested using chi-square statistics to compare between the model accounting for repeated measures and the model not accounting for repeated measures. All test scores are represented as percentages, and un-attempted questions were scored as zero. The error bars in the scatter/line plots represent standard error of the mean (SEM), while the box-and-whiskers plot represent the range (whiskers), 25th to 75th percentile (box), median (horizontal line in the box), mean (plus sign, '+') and all data points (scattered points). The predicted least squares means of differences between groups are represented by mean ± SEM. The score differences between pre- and post-tests (increase from pre- to post-tests) only include students who completed both tests.

6 Results

82 students consented for this study, of whom 36 students were in the VR group and 39 students were in the non-VR group. These 75 students (N = 75) were included in the main analysis, while 7 students were excluded due to group classification ambiguity (4 students completed the VR A assignment and the programming B assignment, 1 student completed the programming A assignment and the VR B assignment, and 2 students did not complete one of the assignments). Table 1 shows the descriptive statistics for the test scores for VR and non-VR groups and all students. We note that only 38 students completed all 7 tests, while other 37 students have some missing data. This is because of the limitations of the quasi-experimental study conducted in a classroom, where the tests depend on students' attendance during the non-compulsory testing day. The presence of these missing data was accounted in the Mixed-Effects Model.

Table 1: Descriptive statistics for test scores of different intervention groups (VR and non-VR) and overall students (combined VR and non-VR). (SD represents standard deviation and N represents the number of data present in the sample).

Tests	Non-VR			VR			Overall		
	Mean	SD	N	Mean	SD	N	Mean	SD	N
Spatial Test 1	53	23	31	58	27	28	55	25	59
Spatial Test 2	46	29	31	49	22	32	47	26	63
Spatial Test 3	59	26	36	62	23	35	61	24	71
Pre-test A	52	20	36	55	19	34	54	19	70
Post-test A	52	20	36	62	19	31	56	20	67
Pre-test B	46	14	32	44	10	32	45	12	64
Post-test B	57	16	31	50	13	28	53	15	59

6.1 *Spatial ability comparison between VR users and non-VR users*

We first show that there is no significant difference of spatial test scores between VR and non-VR groups (Figure 4a). We used the Mixed-Effects Model to test this hypothesis (matching is effective with $p < 0.01$, sphericity is not assumed, Geisser-Greenhouse correction is used). The fixed effects are time (within-subject factor) and intervention (between-subjects factor), where time yields a statistically significant effect ($p < 0.01$) while intervention and cross-term are not statistically significant ($p = 0.39$ and 0.88 respectively). Since there is no significant difference of spatial ability scores between VR and non-VR groups, this analysis shows that there is no preference of students on choosing the interventions.

Figure 4b shows the change of spatial test scores within each group. For the VR group, post-hoc analysis using Tukey's test shows that Spatial 2 is significantly lower than Spatial 3 ($p = 0.04$) while Spatial 1 is not significantly different from Spatial 2 and 3 ($p = 0.23$ and 0.72 respectively). For the non-VR group, there are no significant differences between the spatial tests. Mixed-Effects Model and Tukey's post-hoc tests on all students with only time factor as a fixed effect (Figure 4b, overall) confirms that time is a significant factor ($p < 0.01$), with Spatial 2 is significantly lower than Spatial 3 ($p < 0.01$) and Spatial 1 is not significantly different from Spatial 2 and 3 ($p = 0.13$ and 0.22 respectively). One possible reason for the lower Spatial 2 score is that Spatial 2 was administered approximately 1.5 weeks away from the related radiation lecture, where students were probably less motivated with the relevance of the tests. This is different from Spatial 1, where the test was conducted on the same day of the related topic of 2D and 3D heat transfer, while Spatial 3 was completed by students on their own time.

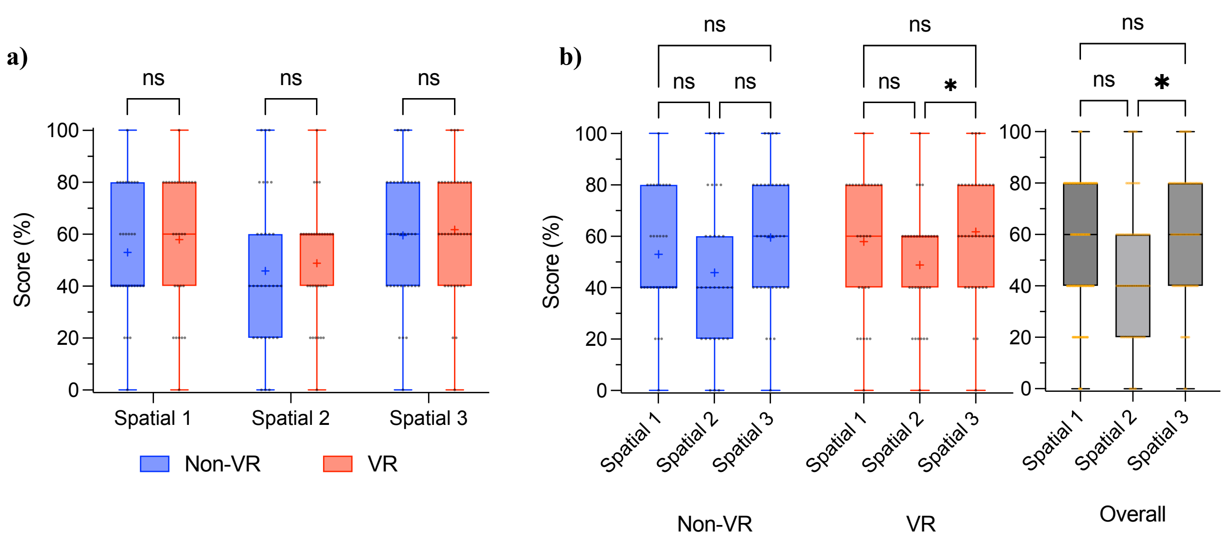

Figure 4: Comparison of spatial ability between VR and non-VR groups. (a) Spatial test scores comparison between groups. (a) Spatial test scores across time for each group.

We categorized the students into two categories of spatial ability, LOW and HIGH, based on the mean spatial test scores. A mean score of greater than 50% is represented by HIGH, while 50% and lower is represented by LOW. The number of students in each spatial category is shown in Table 2. We also separately considered the effect of very low scores for Spatial 2. We found that only 3 students had their spatial category change if we exclude Spatial 2: 1 student in the VR group changes from LOW (all included) to HIGH (Spatial 2 excluded) and 2 students (1 student from each of VR and non-VR group) change from HIGH (all included) to LOW (Spatial 2 excluded). Since this change is minor, we proceeded with using the overall mean to categorize the students into HIGH and LOW spatial categories. These categories were used to understand the differences of effects on the problem-solving performance (pre- and post-tests).

Table 2: Number of students in each spatial category and intervention groups overall and those who completed the pre-tests and post-tests.

Intervention	**Non-VR**		**VR**	
Spatial Category	LOW	HIGH	LOW	HIGH
Pre-test A	19	17	11	23
Post-test A	17	19	10	21
Both pre- and post-tests A	17	17	9	20
Pre-test B	16	16	10	22
Post-test B	15	16	8	20
Both pre- and post-tests B	14	15	8	19
Overall students	**20**	**19**	**12**	**24**

6.2 Pre-tests and post-tests comparison between VR users and non-VR users

Figure 5 shows the pre-test and post-test scores for VR and non-VR groups. Since the test formats were different for A and B (multiple-choice and constructed-response formats respectively), we performed the analyses separately. For both tests, we used the Mixed-Effects Model with time (pre- and post-tests) and intervention groups as fixed effects, and uncorrected Fisher's Least Significant Difference (LSD) test for the post-hoc multiple comparisons test. Additionally, we also performed Welch's unpaired t-tests to compare test scores difference (increase of post-test scores from pre-test scores) between the VR and non-VR groups (Figure 5b).

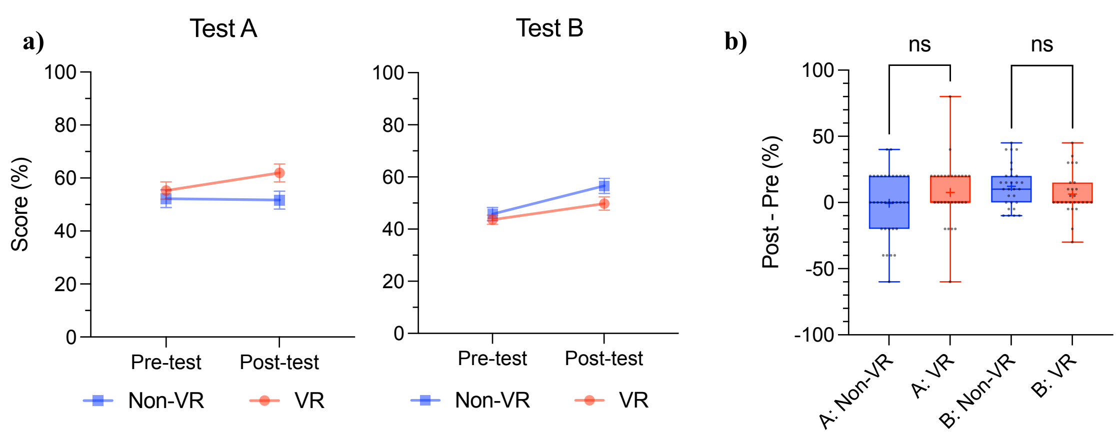

Figure 5: Comparison of pre- and post-test scores between VR and non-VR groups. (a) Pre- and post-tests scores between groups. (b) Score difference between groups.

For test A, the model yields p = 0.29, 0.07 and 0.21 for time, intervention and cross-term respectively, with significant matching at p = 0.04. Post-hoc analysis between intervention groups shows that while there is no significant difference between VR and non-VR groups for the pre-test (p = 0.51, absolute difference = 3.1% ± 4.7%), the VR group has significantly higher post-test scores compared to the non-VR group ($p = 0.03$, absolute difference = 10% ± 4.8%). Meanwhile, the scores for the VR group increase ($p = 0.11$, absolute difference = 6.9% ± 4.3%) while the scores for non-VR group slightly decrease ($p = 0.89$, absolute difference = 0.56% ± 4.0%). The test scores differences (Figure 5b) show that the VR group improves slightly better than the non-VR group after the intervention (Welch's t-test $p = 0.19$, absolute difference = 8.2% ± 6.1%).

For test B, the model yields $p < 0.01$, = 0.12 and 0.26 for time, intervention and cross-term respectively, with matching at $p = 0.05$. Post-hoc analysis between intervention groups shows that although non-VR group scores are consistently higher than VR group scores, there is no significant difference between the groups for both pre- and post-tests (pre-test: $p = 0.59$, absolute difference = 1.8% ± 3.4%; post-test: $p = 0.06$, absolute difference 6.7% ± 3.5%). However, both groups have a significant score increase from pre-test to post-test (VR: $p = 0.04$, absolute difference = 6.3% ± 3.1%; non-VR: $p < 0.01$, absolute difference = 11% ± 3.0%). Figure 5b shows that the test score increase for the non-VR group is slightly higher than for the VR group (Welch's t-test $p = 0.19$, absolute difference = 5.8% ± 4.3%).

These results show that the effects of educational interventions differ between test formats. For the multiple-choice questions format (test A), the VR group has significantly higher post-test scores of 10% ± 4.8% compared to the non-VR group, demonstrating the benefits of VR intervention, although the increase in the scores between pre- and post-test are not statistically different. However, test B results indicate that there is no significant difference between intervention groups. The difference between these tests is the type of questions. For test A, all five questions required students to implicitly visualize the system for each word-based multiple-choice question and compare with multiple given possible answers. They needed to differentiate between provided options to choose the best answer. For test B, we explicitly asked students to draw the system for that set of questions without any answer reference, while the remaining questions involved specifying the equations of boundary conditions, changing the system, and drawing the concentration profile. Thus, the number of visualizations in the test questions may explain the difference in the effect of interventions on the scores between the different test sets, where more visualizations may increase the significance in the intervention difference.

We acknowledge that test A and VR A were conducted first, followed by test B and VR B. This timing effect may also explain the difference, where some students might have higher motivation for the first VR than the second VR, resulting in more learning for the first one and diminishing return for the second.

6.3 Analysis of spatial ability as a mediating effect for pre-post-tests

We also sought to analyze the potential of spatial ability as a mediating factor based on the two spatial categories (HIGH and LOW). Figure 6 shows the pre- and post-test scores for different intervention groups and spatial categories. We used the Mixed-Effects Model on the pre- and post-test scores A and B (Figure 6a) with time (pre- and post-tests), intervention group (VR and non-VR) and spatial category (LOW and HIGH) as fixed effects, followed by post-hoc Šídák's multiple comparisons test. The full result is shown in Appendix A.

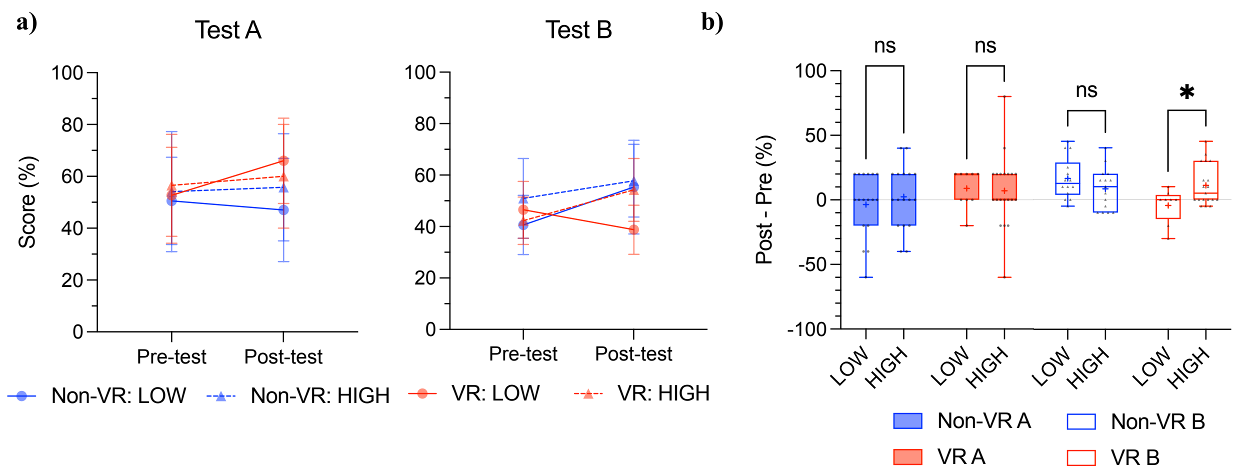

Figure 6: Comparison of pre- and post-test scores between intervention groups and spatial categories. (a) Pre- and post-test scores between groups. (b) Score difference between groups.

For test A, all fixed effects and the resulting post-hoc tests are not statistically significant. Similarly, Analysis of Variance (ANOVA) on the difference of test scores (Figure 6b) are also not significant ($p = 0.76$, 0.19 and 0.55 for spatial category, intervention group and interaction term respectively). However, the numerical difference between groups demonstrates that the effect of intervention depends on spatial category. For LOW spatial, although the difference between the VR and non-VR groups for the pre-test is small at 2.5% ± 7.5%, the VR group has 18% ± 7.8% higher post-test scores than the non-VR group ($p = 0.22$). On the other hand, for HIGH spatial, the intervention shows numerically similar scores between both pre- and post-tests (2.4% ± 6.3% and 4.8% ± 6.2% respectively). These results demonstrate that VR intervention could help in improving multiple-choice test performance for students with low spatial ability.

For test B, the three-way interaction term is statistically significant ($p < 0.01$), demonstrating the significance of spatial category in determining the effects of VR on the test scores change. Post-hoc analysis shows that there is a significant increase from pre- to post-test scores for non-VR with LOW spatial (15% ± 4.0% increase, $p < 0.01$) and VR with HIGH spatial (12% ± 3.5% increase, $p = 0.02$). For the VR group with LOW spatial, the test score decreases by 6.8% ± 5.3% ($p = 0.94$), while for the non-VR group with HIGH spatial, the test score increases by 7.3% ± 3.9% ($p = 0.58$). Additionally, ANOVA on the test score differences (Figure 6b) with spatial category and intervention group as factors shows that the interaction term is significant ($p = 0.01$), with spatial category not significant ($p = 0.40$) and intervention group significant ($p = 0.04$). A post-hoc uncorrected Fisher's LSD test shows that the test score differences are significantly different for HIGH and LOW spatial for the VR group ($p = 0.02$, absolute difference = 15% ± 6.5%) and for the LOW spatial category between VR and non-VR groups ($p < 0.01$, absolute difference = 21% ± 6.8%). Meanwhile, the score differences are not significantly different for the non-VR group between HIGH and LOW ($p = 0.16$, absolute difference = 8.1% ± 5.7%) and for the HIGH category between non-VR and VR groups ($p = 0.61$, absolute difference = 2.7% ± 5.3%).

In summary, we observed that for students with high spatial ability, VR intervention is significantly effective in improving their performance on constructed-response questions, while the improvement is not significant for the non-VR group. Surprisingly, for students with low spatial ability, the score decreased from pre-test to post-test for the VR group, but increased for the non-VR group. We hypothesize that since test B involves equations and derivations that require skills beyond visualization, the programming assignment completed by the non-VR group could be more helpful for these students as discussed in the next section.

7 Discussion

In this study, we observed that VR interventions improve word-based test performance in chemical engineering heat and mass transfer for all students in general. However, the benefits varied based on their spatial ability and the test format. Here, the two test formats represent two stages of word-based problem-solving processes: problem representation and solving. There are two ways that test A (multiple-choice) demonstrates the more problem-representational stage than test B. First, there were more visualizations needed in test A, where the questions in test A involved the visualization of three systems plus two rotations, while in test B, students needed to represent only one system and one geometry change. Second, since test A is device-based, it might require more working memory capacity to represent and solve the problem simultaneously, compared to test B that explicitly required students to draw the system on paper where they could refer to the problem while in the solving stage. Test B mainly assessed the solving stage: students responded to the constructed-response questions by identifying the equations for boundary conditions and drawing profiles without any given answer choices. Thus, students would need to have a deeper conceptual understanding to answer these constructed-response questions, rather than comparing and recognizing the given answer choices as in the multiple-choice questions.

For students with low spatial ability, we observed that VR intervention improved their performance in solving the problem-representation-focused problems as shown through the results of the multiple-choice tests (test A). Meanwhile, the performance of students with high spatial ability in this study was not significantly different between interventions. These are consistent with the findings in [8], where cognitive load theory could explain the advantage of VR intervention in reducing the cognitive load during the learning process to be processed into the long-term memory. In our study, the learning process involved developing mental images and schemas useful for word-based problem solving. With lower cognitive load, students with low spatial ability achieved a better learning outcome. Students with high spatial ability already have low cognitive load, so there is no significant difference of interventions [8].

On the other hand, the findings for the solving-stage-focused tests (test B) suggest that students with high spatial ability in our study benefited from VR for the solving stage. However, given that VR did not yield significantly different outcomes from the non-VR intervention on the problem-representation-focused tests (test A), further investigation is needed to understand the underlying mechanism by which VR influences the solving stage.

In addition, programming intervention could be beneficial for conceptual and math reasoning. For students with low spatial ability, we observed that the non-VR (programming) group performed better than the VR group in the solving-stage-focused questions. Since word-based problem solving involves both problem representations and conceptual and math reasoning, this preliminary observation suggests that the programming assignment could enhance these students' conceptual and math skills leading to higher performance in solving the problems, which can be confirmed in future studies with a larger sample size.

We also noted that these two VR interventions did not significantly change students' spatial ability over a semester. Although VR was used to improve spatial ability [4], [5], [6], the VR modules used in this study were specific to heat and mass transfer topics with simple shapes rather than the general spatial reasoning that involves more complex shapes and rotation in the spatial tests. However, these spatial test scores were useful in categorizing students' spatial ability level.

8 Implications

This study suggests that VR interventions can be used to improve word-problem solving skills. In this study, one of the learning objectives of this chemical engineering mass and heat transfer course is to be able to solve mass and heat transfer problems, which require students to understand the problem and define the system. The VR modules were designed so that students could have the mental image of the typical system of problems to be solved. VR interventions of a similar design may also be applied to other engineering problem-solving-based courses with similar learning objectives.

However, VR interventions may be limited in terms of budget, equipment availability and time. If it is not possible for all students to use the VR as supplemental learning tools, spatial ability and learning objectives should be considered. Spatial ability can be identified by conducting spatial tests in the beginning of the semester as a baseline to determine the students' needs. The previous finding shows that students with low spatial ability benefit most from VR intervention [8], and our preliminary data indicate that this benefit is particularly relevant for the learning objectives focusing on problem-representation. On the other hand, VR appears to support the students with high spatial ability during the solving stage, although these trends should be validated with larger sample sizes. In addition, programming or numerical-focused assignments may be helpful for students with low spatial ability to improve their solving-stage skills based on our preliminary findings. In general, both VR and programming interventions show promise in improving problem-solving skills, while spatial tests may serve as guidance for both students and instructors to discuss the best learning strategies for the course and future success.

9 Limitations

This is a quasi-experimental study, where a non-randomized control group (non-VR) was recruited based on students' preferences to compare with the VR intervention group. Although the spatial test scores show no significant differences between the groups, other factors such as math self-efficacy may also be a determining factor in their preference (students who have low math self-efficacy may tend to avoid programming assignments), which could relate to their test performance. We also acknowledge that we used a shortened version of PSVT:R (5 items instead of 30 items), which may not measure the spatial ability of students as accurately as the full version. The limited number of test items, in addition to the small sample size (10-25 students per intervention and spatial group, with the presence of missing data), may contribute to low statistical significance (high p-value) in the results. Future work includes larger sample size and randomization, as well as ensuring the accuracy of the spatial test results, to confirm these preliminary findings.

Tests A and B were designed to assess the students' performance in different question types in the two distinct stages. However, the grading also differed. Test A has only one correct and two incorrect answers for each multiple-choice question, providing a binary grading for each question. However, since test B is of constructed-response format, partial credits were also given, in addition to multiple accepted answers. There was also more tendency for students to leave their responses blank (graded as zero) rather than guessing if they do not have enough time to solve the problems. Thus, test B scores are more subjective from the students' responses and grading perspectives. Although multiple-choice questions are more reliable, constructed-response questions are perceived as closer to real-world problems, and these different formats may provide different scores based on the levels that they are testing [33], [34], demonstrating the importance of both formats despite the challenges for constructed-response tests.

Moreover, although we seek to measure the direct effect of the related VR interventions through pre-tests and post-tests, knowledge retention and recency effects should be considered in comparing between tests A and B, and between students, because of timing variations between tests and interventions. A future research direction should also include the potential long-term effects on the problem-solving performance through transfer of learning to apply the skills to broader engineering problems.

Finally, it is also important to gauge students' acceptance of these recently developed VR modules for the heat and mass transfer course. It was found that students had more positive attitude towards learning when they positively perceived the modules' usefulness and ease of use [35], depending on various class dynamics and structures [36]. In addition, some students in our study reported that they suffered from motion sickness and some did not. It was determined that comfort and well-being is important in educational technology acceptance [36]. Further study is needed to understand the effects of acceptance and well-being on learning performance.

10 Conclusions

This study provides preliminary evidence that VR may serve as a useful supplemental tool to improve problem-solving skills in chemical engineering in the context of a heat and mass transfer course. Our findings suggest that these benefits help students with different spatial ability in distinct ways. For students with low spatial ability, VR appeared to improve their performance in problem-representation-focused tests. This finding may be consistent with the cognitive load theory, where VR provides visual support during learning to help them develop mental models during the problem representation stage of the problem-solving process. This will lessen the demand on their visuospatial working memory and provide more capacity to apply the concepts they learned in class into the presented problems and solve them. For students with high spatial ability, VR was associated with the increase in their performance in the solving-stage-focused tests, highlighting the need for further study to understand the mechanism behind VR benefits on the solving stage. In designing their courses, instructors can work together with students to develop targeted strategies, tailored to factors such as spatial ability, to help them improve their problem-solving skills. By having better problem-solving skills, students are better prepared to solve problems efficiently as future engineers.

Acknowledgments

This work was supported by the National Science Foundation Center for Complex Particle Systems (COMPASS) Science and Technology Center (award no. 2243104) and University of Michigan College of Engineering (Engineering Innovation grant). We would like to thank Rhonda K. Holton, Patricia Jaimes, Daniel Fessahazion, Shanna R. Daly and anonymous reviewers for helpful discussions. We also would like to thank Stephanie O'Malley and Ted Hall from the Duderstadt Center at University of Michigan for VR modules development and support.

Appendix A: Statistical analysis results for Mixed-Effects Model with time, spatial category and intervention as fixed effects

Effects	Predicted mean difference	SEM of difference	p-value
Test A: Fixed-effects tests (matching p = 0.06)			
Time			0.23
Spatial category			0.07
Intervention			0.49
Time x Spatial category			0.14
Time x Intervention			0.82
Spatial category x Intervention			0.38
Time x Spatial category x Intervention			0.28
Test A: Post-hoc Šídák's multiple comparisons test			
Pre-test: Non-VR A: LOW vs. Post-test: Non-VR A: LOW	3.5	5.8	> 0.99
Pre-test: Non-VR A: HIGH vs. Post-test: Non-VR A: HIGH	-1.8	5.8	> 0.99
Pre-test: VR A: LOW vs. Post-test: VR A: LOW	-12	7.6	0.76
Pre-test: VR A: HIGH vs. Post-test: VR A: HIGH	-4.3	5.2	> 0.99
Pre-test: Non-VR A: LOW vs. Pre-test: Non-VR A: HIGH	-3.4	6.6	> 0.99
Pre-test: VR A: LOW vs. Pre-test: VR A: HIGH	-3.3	7.2	> 0.99
Post-test: Non-VR A: LOW vs. Post-test: Non-VR A: HIGH	-8.7	6.6	0.91
Post-test: VR A: LOW vs. Post-test: VR A: HIGH	4.8	7.5	> 0.99
Pre-test: Non-VR A: LOW vs. Pre-test: VR A: LOW	-2.5	7.5	> 0.99
Pre-test: Non-VR A: HIGH vs. Pre-test: VR A: HIGH	-2.4	6.3	> 0.99
Post-test: Non-VR A: LOW vs. Post-test: VR A: LOW	-18	7.8	0.22
Post-test: Non-VR A: HIGH vs. Post-test: VR A: HIGH	-4.8	6.2	> 0.99
Test B: Fixed-effects tests (matching p = 0.04)			
Time			< 0.01*
Spatial category			0.05
Intervention			0.04*
Time x Spatial category			0.04*
Time x Intervention			0.22
Spatial category x Intervention			0.78
Time x Spatial category x Intervention			< 0.01*
Test B: Post-hoc Šídák's multiple comparisons test			

Pre-test: Non-VR B: LOW vs. Post-test: Non-VR B: LOW	-15	4.0	< 0.01*
Pre-test: Non-VR B: HIGH vs. Post-test: Non-VR B: HIGH	-7.3	3.9	0.58
Pre-test: VR B: LOW vs. Post-test: VR B: LOW	6.8	5.3	0.94
Pre-test: VR B: HIGH vs. Post-test: VR B: HIGH	-12	3.5	0.02*
Pre-test: Non-VR B: LOW vs. Pre-test: Non-VR B: HIGH	-11	4.6	0.24
Pre-test: VR B: LOW vs. Pre-test: VR B: HIGH	4.2	5.0	> 0.99
Post-test: Non-VR B: LOW vs. Post-test: Non-VR B: HIGH	-2.7	4.7	>0.99
Post-test: VR B: LOW vs. Post-test: VR B: HIGH	-14	5.4	0.11
Pre-test: Non-VR B: LOW vs. Pre-test: VR B: LOW	-6.3	5.3	0.96
Pre-test: Non-VR B: HIGH vs. Pre-test: VR B: HIGH	8.4	4.3	0.47
Post-test: Non-VR B: LOW vs. Post-test: VR B: LOW	16	5.7	0.08
Post-test: Non-VR B: HIGH vs. Post-test: VR B: HIGH	4.0	4.4	> 0.99

References

[1] G. Duffy, S. Sorby, and B. Bowe, "Exploring the role of spatial ability in the mental representation of word problems in mathematics," *Front. Educ.*, vol. 9, Jun. 2024, doi: 10.3389/feduc.2024.1346474.

[2] D. H. Uttal and C. A. cohen, "Spatial Thinking and STEM Education: When, Why, and How?," in *Psychology of Learning and Motivation*, vol. 57, Academic Press, 2012, pp. 147–181. doi: 10.1016/B978-0-12-394293-7.00004-2.

[3] R. Keogh, J. Pearson, and A. Zeman, "Aphantasia: The science of visual imagery extremes," in *Handbook of Clinical Neurology*, vol. 178, Elsevier, 2021, pp. 277–296. doi: 10.1016/B978-0-12-821377-3.00012-X.

[4] H. C. Gomez-Tone, J. Martin-Gutierrez, and B. K. Valencia-Anci, "Chapter 2 - Augmented and immersive virtual reality to train spatial skills in STEAM university students," in *Innovations in Artificial Intelligence and Human-Computer Interaction in the Digital Era*, S. Bhatia Khan, S. Namasudra, S. Chandna, A. Mashat, and F. Xhafa, Eds., in Intelligent Data-Centric Systems. , Academic Press, 2023, pp. 7–30. doi: 10.1016/B978-0-323-99891-8.00002-4.

[5] C. Roca-González, J. M. Gutierrez, M. GarcÍa-Dominguez, and M. del C. M. Carrodeguas, "Virtual Technologies to Develop Visual-Spatial Ability in Engineering Students," *EURASIA J Math Sci Tech Ed*, vol. 13, no. 2, pp. 441–468, Dec. 2016, doi: 10.12973/eurasia.2017.00625a.

[6] C.-W. Chang, J. Heo, S.-C. Yeh, H.-Y. Han, and M. Li, "The Effects of Immersion and Interactivity on College Students' Acceptance of a Novel VR-Supported Educational Technology for Mental Rotation," *IEEE Access*, vol. 6, pp. 66590–66599, 2018, doi: 10.1109/ACCESS.2018.2878270.

[7] R. Sun, Y. J. Wu, and Q. Cai, "The effect of a virtual reality learning environment on learners' spatial ability," *Virtual Reality*, vol. 23, no. 4, pp. 385–398, Dec. 2019, doi: 10.1007/s10055-018-0355-2.

[8] E. A.-L. Lee and K. W. Wong, "Learning with desktop virtual reality: Low spatial ability learners are more positively affected," *Computers & Education*, vol. 79, pp. 49–58, Oct. 2014, doi: 10.1016/j.compedu.2014.07.010.

[9] Z. Merchant, E. t. Goetz, W. Keeney-Kennicutt, L. Cifuentes, O. Kwok, and T. j. Davis, "Exploring 3-D virtual reality technology for spatial ability and chemistry achievement,"

Journal of Computer Assisted Learning, vol. 29, no. 6, pp. 579–590, 2013, doi: 10.1111/jcal.12018.
[10] G. Makransky and G. B. Petersen, "The Cognitive Affective Model of Immersive Learning (CAMIL): a Theoretical Research-Based Model of Learning in Immersive Virtual Reality," *Educ Psychol Rev*, vol. 33, no. 3, pp. 937–958, Sep. 2021, doi: 10.1007/s10648-020-09586-2.
[11] J. Sweller, "Cognitive load during problem solving: Effects on learning," *Cognitive Science*, vol. 12, no. 2, pp. 257–285, Apr. 1988, doi: 10.1016/0364-0213(88)90023-7.
[12] R. C. Atkinson and R. M. Shiffrin, "Human Memory: A Proposed System and its Control Processes," in *Psychology of Learning and Motivation*, vol. 2, Academic Press, 1968, pp. 89–195. doi: 10.1016/S0079-7421(08)60422-3.
[13] A. Baddeley, "The episodic buffer: a new component of working memory?," *Trends in Cognitive Sciences*, vol. 4, no. 11, pp. 417–423, Nov. 2000, doi: 10.1016/S1364-6613(00)01538-2.
[14] A. D. Baddeley and G. Hitch, "Working Memory," in *Psychology of Learning and Motivation*, vol. 8, Academic Press, 1974, pp. 47–89. doi: 10.1016/S0079-7421(08)60452-1.
[15] P. Barrouillet and V. Camos, "Developmental Increase in Working Memory Span: Resource Sharing or Temporal Decay?," *Journal of Memory and Language*, vol. 45, no. 1, pp. 1–20, Jul. 2001, doi: 10.1006/jmla.2001.2767.
[16] A. Miyake, N. P. Friedman, D. A. Rettinger, P. Shah, and M. Hegarty, "How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis," *J Exp Psychol Gen*, vol. 130, no. 4, pp. 621–640, Dec. 2001, doi: 10.1037//0096-3445.130.4.621.
[17] M. Kyttälä and J. E. Lehto, "Some factors underlying mathematical performance: The role of visuospatial working memory and non-verbal intelligence," *Eur J Psychol Educ*, vol. 23, no. 1, pp. 77–94, Mar. 2008, doi: 10.1007/BF03173141.
[18] G. Duffy, S. Sorby, and B. Bowe, "An investigation of the role of spatial ability in representing and solving word problems among engineering students," *Journal of Engineering Education*, vol. 109, no. 3, pp. 424–442, 2020, doi: 10.1002/jee.20349.
[19] S. A. Sorby, G. Duffy, and N. Loney, "An examination of the role of spatial ability in the process of problem solving in chemical engineering," *Australasian Journal of Engineering Education*, vol. 25, no. 1, pp. 55–65, Jan. 2020, doi: 10.1080/22054952.2020.1785653.
[20] J. T. Bell and H. S. Fogler, "Vicher: A virtual reality based educational module for chemical reaction engineering," *Computer Applications in Engineering Education*, vol. 4, no. 4, pp. 285–296, 1996, doi: 10.1002/(SICI)1099-0542(1996)4:4<285::AID-CAE4>3.0.CO;2-9.
[21] C. Udeozor, R. Toyoda, F. Russo Abegão, and J. Glassey, "Perceptions of the use of virtual reality games for chemical engineering education and professional training," *Higher Education Pedagogies*, vol. 6, no. 1, pp. 175–194, Jan. 2021, doi: 10.1080/23752696.2021.1951615.
[22] V. V. Kumar, D. Carberry, C. Beenfeldt, M. P. Andersson, S. S. Mansouri, and F. Gallucci, "Virtual reality in chemical and biochemical engineering education and training," *Education for Chemical Engineers*, vol. 36, pp. 143–153, Jul. 2021, doi: 10.1016/j.ece.2021.05.002.

[23] "Criteria for Accrediting Engineering Programs, 2025 - 2026." Accessed: Dec. 20, 2025. [Online]. Available: https://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-engineering-programs-2025-2026/
[24] P. Araiza-Alba, T. Keane, W. S. Chen, and J. Kaufman, "Immersive virtual reality as a tool to learn problem-solving skills," *Computers & Education*, vol. 164, p. 104121, Apr. 2021, doi: 10.1016/j.compedu.2020.104121.
[25] J. Parong and R. E. Mayer, "Cognitive and affective processes for learning science in immersive virtual reality," *Journal of Computer Assisted Learning*, vol. 37, no. 1, pp. 226–241, 2021, doi: 10.1111/jcal.12482.
[26] C. Norton *et al.*, "Development and deployment of an immersive learning environment for enhancing process systems engineering concepts," *Education for Chemical Engineers*, vol. 3, no. 2, pp. e75–e83, Dec. 2008, doi: 10.1016/j.ece.2008.04.001.
[27] O. A. Meyer, M. K. Omdahl, and G. Makransky, "Investigating the effect of pre-training when learning through immersive virtual reality and video: A media and methods experiment," *Computers & Education*, vol. 140, p. 103603, Oct. 2019, doi: 10.1016/j.compedu.2019.103603.
[28] R. E. Mayer, *The Cambridge Handbook of Multimedia Learning*. Cambridge University Press, 2014.
[29] M. J. Díaz, C. Mantell, I. Caro, I. de Ory, J. Sánchez, and J. R. Portela, "Creation of Immersive Resources Based on Virtual Reality for Dissemination and Teaching in Chemical Engineering," *Education Sciences*, vol. 12, no. 8, Art. no. 8, Aug. 2022, doi: 10.3390/educsci12080572.
[30] G. M. BODNER and R. B. GUAY, "The Purdue Visualization of Rotations Test," *Chem. Educator*, vol. 2, no. 4, pp. 1–17, Oct. 1997, doi: 10.1007/s00897970138a.
[31] S. Y. Yoon, "Psychometric properties of the Revised Purdue Spatial Visualization Tests: Visualization of Rotations (the Revised PSVT:R)," *Theses and Dissertations Available from ProQuest*, pp. 1–180, Jan. 2011.
[32] Y. Maeda and S. Y. Yoon, "A Meta-Analysis on Gender Differences in Mental Rotation Ability Measured by the Purdue Spatial Visualization Tests: Visualization of Rotations (PSVT:R)," *Educ Psychol Rev*, vol. 25, no. 1, pp. 69–94, Mar. 2013, doi: 10.1007/s10648-012-9215-x.
[33] W. L. Kuechler and M. G. Simkin, "Why Is Performance on Multiple-Choice Tests and Constructed-Response Tests Not More Closely Related? Theory and an Empirical Test," *Decision Sciences Journal of Innovative Education*, vol. 8, no. 1, pp. 55–73, 2010, doi: 10.1111/j.1540-4609.2009.00243.x.
[34] D. R. Bacon, "Assessing Learning Outcomes: A Comparison of Multiple-Choice and Short-Answer Questions in a Marketing Context," *Journal of Marketing Education*, vol. 25, no. 1, pp. 31–36, Apr. 2003, doi: 10.1177/0273475302250570.
[35] J. A. M. García, C. G. Gómez, A. T. López, and M. J. Schlosser, "Applying the technology acceptance model to online self-learning: A multigroup analysis," *Journal of Innovation & Knowledge*, vol. 9, no. 4, p. 100571, Oct. 2024, doi: 10.1016/j.jik.2024.100571.
[36] A. Kemp, E. Palmer, P. Strelan, and H. (Mery) Thompson, "Testing a novel extended educational technology acceptance model using student attitudes towards virtual classrooms," *British Journal of Educational Technology*, vol. 55, no. 5, pp. 2110–2131, 2024, doi: 10.1111/bjet.13440.