



A Study of Augmented Reality for the Development of Spatial Reasoning Ability

Dr. John E. Bell, Michigan State University

JOHN BELL Professor, Educational Technology, College of Education. John Bell earned his B.S. in Computer Science from Michigan State University, and then his M.S. and Ph.D. in Computer Science from the University of California, Berkeley. His research considered various user interfaces for human-computer interaction among users with a wide range of technology skills. Bell later completed a post doc at UC Berkeley focused on teaching programming to non-computer science majors, and the development of spatial reasoning abilities for engineering students. Bell has worked at Michigan State University since 1995. His work focused on the development of K-12 teacher abilities to use technology for teaching and learning. His recent research has focused on distance learning and collaboration through telepresence. One key aspect of this work is the study of embodied content for learning and collaboration. Embodied content includes collaborative textual environments as well as augmented/mixed reality. Other research includes idea-centered teaching and learning.

Cui Cheng, Michigan State University

Cui Cheng is a doctoral candidate in the Educational Psychology and Educational Technology program at Michigan State University. Her research focuses on (a) new literacies of online reading comprehension, particularly in complex and ill-structured learning domains, (b) teaching and learning in synchronous hybrid learning environments, where physically present and remote participants interact in real time through such technologies as video conferencing tools and robots, and (c) the use of augmented reality in STEM education.

Dr. Hannah Klautke, Michigan State University

Hannah Klautke is a User Experience Research Associate with Usability/Accessibility Research and Consulting (Michigan State University Outreach and Engagement). She is involved in usability evaluations, focus groups, and information architecture projects for MSU and external clients. Her research areas include effects of cooperative online learning, interventions based on cognitive flexibility theory for reading to learn on the web, and student motivation and achievement in flipped classrooms. Hannah holds a B.A. in Psychology from the University of Bonn, a M.A. in Communication from the University of Missouri, and a Ph.D. in both Communication and Educational Psychology and Educational Technology from Michigan State University.

Dr. William Cain, Michigan State University

Dr. William Cain is Assistant Director of CEPSE/COE Design Studio in the College of Education at Michigan State University. William's research focuses on how people teach, learn, create, and collaborate with new technologies and in technology-rich environments.

Dr. Daniel Joseph Freer, Michigan State University

Daniel Freer is Graduate Students studying Educational Psychology and Educational Technology. His focus is on how students learn, specifically the STEM fields.

Mr. Timothy J. Hinds, Michigan State University

TIMOTHY J. HINDS is the Director of the Michigan State University College of Engineering CoRe (Cornerstone Engineering and Residential) Experience program. His administrative responsibilities include management of the 1500-student first-year combined academic and co-curricular program. His teaching includes development, delivery and management of CoRe Experience courses in engineering design, modeling/computation and spatial visualization. He has also taught courses in machine design, manufacturing processes, mechanics, computational tools and international product design as well as graduate-level



courses in engineering innovation and technology management. He has conducted research in the areas of environmentally-responsible manufacturing, globally-distributed engineering teaming and early engineering education development and has over 30 years of combined academic and industrial management experience. He received his BSME and MSME degrees from Michigan Technological University.

A Study of Augmented Reality for the Development of Spatial Reasoning Ability

Abstract

Spatial ability has been reported to be an important predictor for student success in STEM fields [1]. Accordingly, various studies have looked at multiple strategies to help students develop these skills [2].

This study tested the use of augmented reality (AR) on smartphones for developing spatial reasoning in the context of a spatial reasoning skills course. The mobile AR app gave students the ability to view digital three-dimensional objects by pointing their phones at a printed marker. They could either move their phones around the marker, or they could move the markers, in order to view the objects from multiple angles. The app allowed them to manipulate the objects via commands within the app, such as rotations and folding. Various games were implemented in the app to support the development of mental rotation abilities while using the motivational aspects of gamification to increase engagement.

To test the effects of this app, 56 freshman early-engineering students at a major Midwestern university who performed poorly on the PSVT:R spatial abilities test had the opportunity to work with this app in the context of a course designed to develop their spatial reasoning abilities. They were compared with 34 students who took the same course minus the augmented reality app. Both sets of students experienced traditional means for teaching spatial reasoning (see [1]).

The results of this study demonstrate the efficacy of using the augmented reality app in conjunction with traditional classroom activities for developing spatial reasoning abilities.

Introduction

Many have noted the critical value of spatial abilities for success in STEM fields [3], [1]. Spatial abilities include one's "skill to generate, retain, retrieve, and transform well-structured visual images" [3], and one of those key skills is mental rotation [4]. Many researchers have studied a variety of strategies for helping students develop these skills [2], [5]. Ha & Fang [2] identified two broad categories of training: tangible models and virtual models. Tangible models use physical objects, such as blocks, to scaffold the development of mental rotation [6]. In contrast, virtual models employ computer-based animations and simulations, augmented reality tools, and virtual reality tools [7], [8], [9], [10], [11].

Sometimes both tangible models and virtual models are used in tandem [3]. They describe this approach as using "virtual and physical manipulatives" (VPM). Students "touch and play with real-world PMs [physical manipulatives] in their hands but also students can see how the corresponding VMs [virtual manipulatives] (i.e., computer graphics) simultaneously change in real time" [3].

We posit that augmented reality (AR) provides another kind of middle ground between virtual manipulatives and physical manipulatives in which there is an illusion of deep integration. In

AR, virtual 3-D objects appear as a part of a physical 3-D environment [12]. AR has been implemented using multiple means of integrating the display device with the physical experience of the user. [13] describes integration by means of a fixed monitor, a head-mounted device, and a handheld device.

This study involved the use of handheld AR (see Figure 1). The primary reason for this approach to AR was the nearly universal availability of smartphones among the target population for our study. Any other approach would have required providing additional technologies to the subjects, involving considerable expense and logistical challenges.

In comparison to the VPM approach, we reasoned that with AR there is the potential of a closer integration between the 3-D object and the abstractions that are a part of engineering manipulation, such as the orientation of the X-Y-Z axes and true 90 degree rotations. In addition, we had greater flexibility to create objects for which we did not have a scale physical representation (e.g., a house or a bicycle). Finally, we had the ability to create interactions that would be harder to create with physical objects, like having students identify a sequence of moves that would not begin until after locking in those steps, thus requiring a plan for mental rotations prior to beginning the moves.



Figure 1. Student using the augmented reality app to rotate a virtual object.

This approach of handheld AR was implemented using an app that students installed on their devices. They then pointed their devices at printed targets, and the appropriate objects would appear on their displays, integrated with the physical environment seen by their cameras. A series of lessons involved a variety of games with objects of varying complexity.

Bell, et al., [13] applied the “gulf of execution” and “gulf of evaluation” [14], [15] to understand the significance of applying AR to these kinds of interactions. They reasoned that increasing the “feeling of directness” (i.e., reducing the gap) for understanding the digital representation (the gulf of evaluation) and increasing the “feeling of directness” for instructing the digital representation what to do (the gulf of execution) would increase the realism of the simulation and thus increase focus on mental rotation of the objects, unencumbered by distractions of the interface or the artificiality of controlling a computer rather than working with physical objects.

Our Approach

This study tested the effectiveness of a handheld augmented reality app for developing mental rotation ability. The app worked by pointing the smartphone or tablet camera at a printed target, which then made a series of virtual objects and games appear. Users could move their camera around the target in order to view the objects from other perspectives, or they could move the printed target to achieve the same effect.

The app contains a series of lessons. Students were to complete one lesson per week, and each lesson contained four to nine activities. New lessons became available each week on the same day that class met. As students completed each activity, a checkmark would appear by that activity, and the next activity would become unlocked. Some activities simply required students to do the work, while other activities would not be considered completed until students achieve a minimum score.

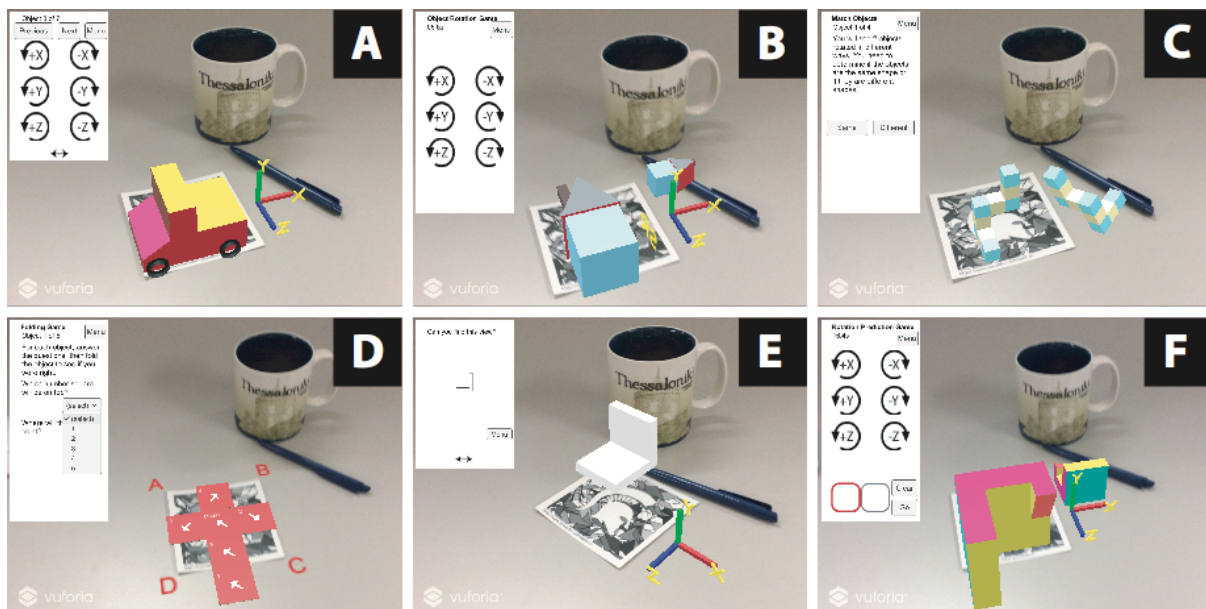


Figure 2. Example activities in the augmented reality app.

As seen in Figure 2, the activities included:

- A. **Rotation practice:** A series of objects were to be rotated using typical engineering symbols for positive and negative rotations on each of the X, Y, and Z axes.
- B. **Rotation matching game:** Two instances of the same object would appear, although with different rotations. Users rotated one of the objects in order to match the rotation of the other object.
- C. **Compare objects game:** Two different stacked objects would appear, and students were to indicate whether they were identical objects or not.
- D. **Folding game:** An unfolded cube would appear as a flat object, with different numbers and arrows in each part of that flat object. Users were to predict which number would appear on top and toward which letter the arrow would point once the object was folded

up into a cube. Users would then step through the folding of the object to see if they were right.

- E. **Finding orthographic views:** Users were shown an object next to a specific orthographic projection of that object. They were then to change their view of the object as needed so that the specified orthographic projection was facing toward them.
- F. **Prediction:** Like the rotation matching game, users saw two instances of the same object, with different rotations. And like the rotation matching game, users were required to rotate the one object to match the rotation of the target object. Unlike the rotation matching game, students had to specify the two required moves before any of them would take place. If their moves were not correct, the rotations would be undone and they would try again.

Student scores were private, so they did not have opportunity to compare their scores with other students. Earlier work with our app showed that for some students, this competition was motivating, but for the students who struggled more on these problems, the competition was discouraging. Accordingly, we chose not to implement competition within this class.

Research Questions

Our research questions were as follows:

Does the addition of student use of the handheld augmented reality app to an existing class based on “Developing Spatial Thinking” [16] show greater pre-post course improvements on a mental rotation assessment than a control group who took the course as guided by that text and using physical objects as well as paper-and-pencil activities but not augmented reality? If so, what factors may influence or moderate the effect of the app on students’ spatial reasoning performance?

To answer these questions, we employed both quantitative and qualitative approaches, including in-app data collection, pretest and posttest performance, and a student survey. The pretest and posttest data are reported here.

Context and Method

Applicants to the university in this study who are accepted and who indicated an interest in engineering all took the PSVT:R assessment of mental rotation prior to coming to campus. Students who scored below 60% were encouraged to take a 1-credit course designed to help them develop their spatial reasoning ability. Roughly half of these students actually took this 1-credit class [17]. Of four sections in this class, two were chosen to be the control group and two were chosen to be experimental. The choice was made for practical reasons while ensuring that the choice was unlikely to result in a biased sample, which was confirmed by the descriptive statistics of these groups.

The control group followed very closely the strategy described by [1] and based on the Sorby textbook [16]. The experimental group did all of the same coursework as the control group, with the addition of having assignments each week based on the AR app. The online in-class time

dedicated to the app was during the first two weeks of class when roughly 15-minutes were taken each of those days to ensure that students were able to install and use the app. At the end of the semester, students were given up to 10% extra credit for the course based on completion of the 11 units in the app. Note that students were told at the beginning of the class that use of the app was required, although the syllabus did not state what role it would have in the grade. Students were not told that the app would be for extra credit only until the last week class.

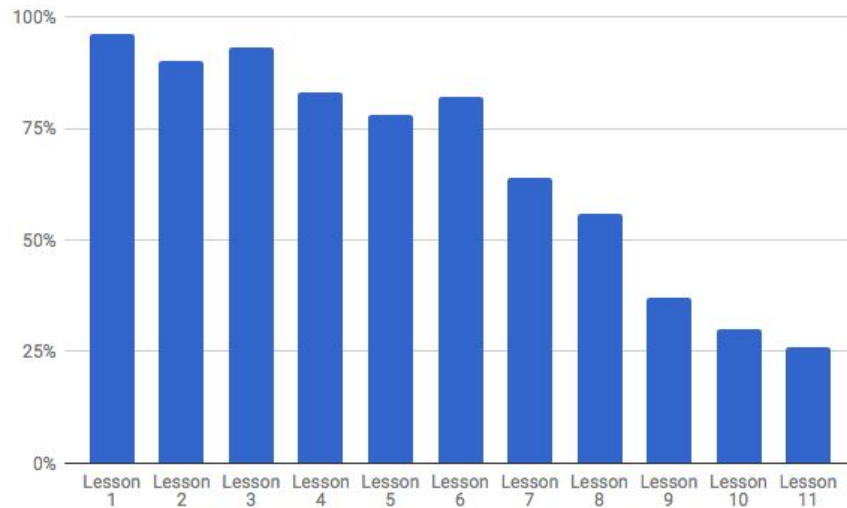


Figure 3: Completion rate of app lessons by week

Of the 102 students who opted in to the study, 10 had actually passed the pretest. To avoid a ceiling effect, we took those 10 students out of our sample. The intervention in this study was the use of the augmented reality app as a supplement to the standard course content. Students in 2 of the 4 sections of the class were invited to use the app. In total, 57 students used the app at least once. To ensure substantive use of the app in the experimental group, we ignored students who completed 60% or less of the in-app tasks. Figure 3 shows the completion rate for each lesson, which was made up of multiple tasks. These ignored students, we reasoned, did not use the app enough to hope for a significant benefit in terms of the final exam.

As a result, 39 students were in the experimental group, and 35 were in the control group. Of the 73 who reported their gender, 24 were women (33%) and 49 were men (67%). In terms of race, we had a diverse population, although race was not intended to be a factor in our research. Of the 72 who answered, we had 9 Asian (12%), 19 Black or African American (26%), 40 White (56%), and 4 Other (6%). Seven reported as being Hispanic or Latino (10%).

This paper focuses on the quantitative data we analyzed. The PSVT:R pre-course scores and the PSVT:R-based final exam served as measures of achievement in mental rotation. A survey administered at the beginning of the course included participant demographics as well as enjoyment of and perceived competence at spatial reasoning activities based on Ryan's [18] longstanding Intrinsic Motivation Inventory (IMI). A post-course survey again assessed enjoyment and perceived competence with regards to spatial reasoning.

Results

Preliminary analysis. To answer our first research question, we compared students' pre-post course improvements on the mental rotation test between the control group and the experimental group. Descriptive statistics for students' scores on the pretest and posttest measures of mental rotation, divided by intervention group membership, are presented in Table 1. A univariate ANCOVA was conducted to determine whether there was a significant difference between the control and experimental groups in terms of students' posttest scores, controlling for their pretest scores. A significant beneficial effect was found for the experimental group ($F(1, 71) = 4.359$, $p = 0.040$, partial eta squared = 0.058). This indicates that the use of the AR app was helpful in improving the spatial reasoning performance of students who completed more than 60% of the in-app lessons.

Table 1. Descriptive Statistics for the Pretest and Posttest Measures in the Control and Experimental Groups

	n	Min	Max	Mean	SD
Control Group					
Pretest	35	6	17	13.40	3.00
Posttest	35	14	28	21.94	3.99
Experimental group					
Pretest	39	7	17	14.03	2.58
Posttest	39	13	30	24.08	3.92
Total					
Pretest	74	6	17	13.73	2.79
Posttest	74	13	30	23.07	4.07

Degree of performance improvement and individual differences in prior spatial reasoning ability. Given the positive effect of the AR app on students' spatial reasoning performance in general, we further analyzed whether students' spatial reasoning ability prior to the course influenced or moderated the effects of the AR app on students' performance improvement. Specifically, we used the median (median=15) of the 74 students' pretest scores to split all students into two groups: low-performing group ($n=36$) and high-performing group ($n=38$). Descriptive statistics for students' scores on the pretest and posttest measures of mental rotation, divided by intervention group membership and prior spatial reasoning ability, are presented in Tables 2 and 3.

Table 2. Descriptive Statistics for the Pretest and Posttest Measures Divided by Intervention Group Membership for High Prior Spatial Reasoning Ability

	n	Min	Max	Mean	SD
Control Group					
Pretest	17	15	17	16.00	0.79
Posttest	17	14	28	23.12	3.87
Experimental group					
Pretest	21	15	17	15.90	0.83
Posttest	21	19	30	24.52	3.31
Total					
Pretest	38	15	17	15.95	0.80
Posttest	38	14	30	23.89	3.59

Table 3. Descriptive Statistics for the Pretest and Posttest Measures Divided by Intervention Group Membership for Low Prior Spatial Reasoning Ability

	n	Min	Max	Mean	SD
Control Group					
Pretest	18	6	14	10.94	2.07
Posttest	18	14	28	20.83	3.87
Experimental group					
Pretest	18	7	14	11.83	2.15
Posttest	18	13	29	23.56	4.58
Total					
Pretest	36	6	14	11.39	2.13
Posttest	36	13	29	22.19	4.40

There was no difference in students' pretest scores either between the two low-performing groups or between the two high-performing groups. Additional univariate ANCOVAs were conducted to determine the group effects (either the control/experimental group effects or the low/high performing group effects) on students' posttest scores by controlling for their pretest scores. Results revealed that for both the high-performing groups and the low-performing groups, there were no significant control/experimental group effects on students' posttest scores after controlling for their pretest scores.

In addition, within the control group, there was no significant difference in students' posttest scores between the low-performing group and the high-performing group controlling for their

pretest scores, which indicates that in the control group, students with lower prior spatial reasoning abilities and students with relatively higher prior spatial reasoning abilities were not significantly different in their performance improvements due to the course. Within the experimental group, however, there was a significant difference in students' posttest scores between the low-performing group and the high-performing group controlling for their pretest scores, meaning that within the experimental group, students with lower prior spatial reasoning abilities made significantly more performance improvement than students with relatively higher spatial reasoning abilities ($F(1, 36)=6.593, p=0.015, \text{partial eta squared}=0.155$). This indicates that the use of the AR app was more helpful for students with lower prior spatial reasoning abilities than for students with relatively higher prior spatial reasoning abilities. We then compared the interaction effect of intervention and prior spatial reasoning ability on students' posttest scores controlling for their pretest scores. No significant interaction effect was found, but the plot (see Figure 4) confirms that the use of the app helped students who had lower prior spatial reasoning abilities improve more on their performance, in comparison to students who had relatively higher prior spatial reasoning abilities.

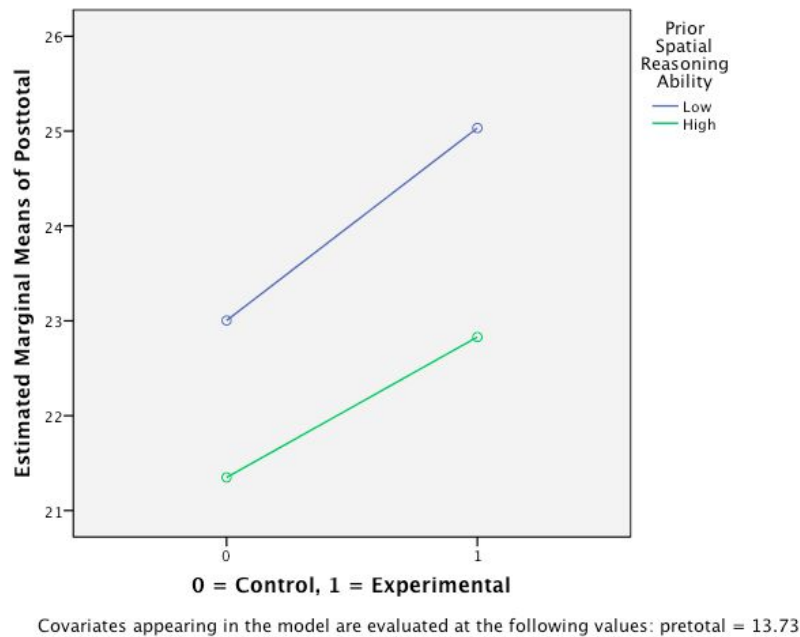


Figure 4. Interaction of Intervention and Prior Spatial Reasoning Ability on Posttest Scores Controlling for Pretest Scores

Degree of performance improvement and individual differences in gender. We further examined whether gender influenced the effects of the app on students' performance improvement. Of the 74 students, 49 were male, 24 were female, and 1 student checked the "prefer not to say" option. Our analysis for this section was based on data from the 73 students who reported their gender. Descriptive statistics for the pretest and posttest measures, divided by intervention group membership and gender, are presented in Table 4 and Table 5.

Table 4. Descriptive Statistics for the Pretest and Posttest Measures of Male Students Divided by Intervention Group Membership

	n	Min	Max	Mean	SD
Control Group					
Pretest	26	6	17	13.12	3.22
Posttest	26	14	28	22.69	3.70
Experimental group					
Pretest	23	7	17	14.57	2.56
Posttest	23	14	30	25.13	3.83
Total					
Pretest	49	6	17	13.80	2.99
Posttest	49	14	30	23.84	3.92

Table 5. Descriptive Statistics for the Pretest and Posttest Measures of Female Students Divided by Intervention Group Membership

	n	Min	Max	Mean	SD
Control Group					
Pretest	8	10	17	14.00	2.27
Posttest	8	14	28	19.63	4.47
Experimental group					
Pretest	16	8	16	13.25	2.49
Posttest	16	13	29	22.56	3.65
Total					
Pretest	24	8	17	13.50	2.40
Posttest	24	13	29	21.58	4.10

Within both the control and experimental groups, there was no significant difference in pretest scores between male and female students. Even though males are reported as having higher spatial ability than females [18], our finding is not surprising because only the students who failed the pretest were included. ANCOVAs were conducted to determine the gender effect on students' posttest scores controlling for their pretest scores. Within the control group, there was significant gender effect on students' posttest scores controlling for their pretest scores ($F(1,31)=5.280$, $p=0.028$, partial eta squared =0.146), indicating that in the control group, male students made significantly more improvement than female students. Within the experimental condition, there was no significant gender effect on students' posttest score controlling for their pretest score ($F(1,36)=3.729$, $p=0.132$, partial eta squared =0.062), which means that in the experimental group, there was no significant difference in performance improvement between

male and female students. We then examined if there was any interaction effect of intervention and gender on students' posttest scores controlling for their pretest scores. No significant interaction effect was found. All these results revealed that the use of the app helped narrow the gender gap in spatial reasoning, although the gap was not significantly narrowed (see Figure 5).

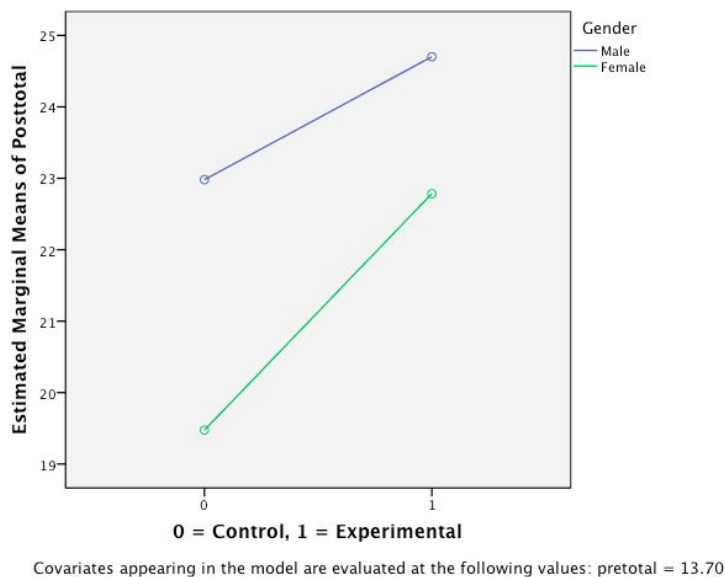


Figure 5. Interaction Intervention and Gender on Posttest Scores Controlling for Pretest Scores

Discussion

Our findings show that the use of the AR app (a) was helpful in improving the spatial reasoning performance of students who originally had poor spatial reasoning abilities, (b) was more helpful for students with lower prior spatial reasoning abilities than for students with relatively higher prior spatial reasoning abilities, and (c) helped narrow the gender gap in spatial reasoning.

The gender gap found in the control condition (males improved significantly more than females) is surprising given the findings of [5] in general and in [19] in particular. We did not study the control condition, so we are unable to account for this advantage for males. On the other hand, it is thus worthy of note that this advantage was not found among students who used the AR app. Further study is required to understand what might account for these differences. This finding may relate to what [3] report with the use of physical manipulatives where females improved more than males.

An important question for further research is a focus on the psychological effects of augmented reality. In what ways is AR like the benefits found when using physical manipulatives, and in what ways is it more like computer simulations?

Next steps for further research include studying an app-only intervention, without taking a class. In this study, students in the experimental condition presumably spent additional time doing

spatial reasoning tasks, so it is possible that this extra time is a contributing factor in the reported gains. On the other hand, students gained substantially simply by taking the class, so the benefits of the app on its own are not clear.

Another direction for further research should focus on additional development of the app. Possibilities include adaptive presentation of lessons based on student progress, further use of gamification to enhance motivation and engagement, and building assessment into the app itself.

References

- [1] S. Sorby, B. Casey, N. Veurink, and A. Dulaney, "The role of spatial training in improving spatial and calculus performance in engineering students," *Learning and Individual Differences*, vol. 26, pp. 20–29, 2013.
- [2] O. Ha and N. Fang, "Spatial Ability in Learning Engineering Mechanics: Critical Review," *Journal of Professional Issues in Engineering Education and Practice*, vol. 142, no. 2, Apr. 2016.
- [3] O. Ha and N. Fang, "Interactive Virtual and Physical Manipulatives for Improving Students' Spatial Skills," *Journal of Educational Computing Research*, vol. 55, no. 8, pp. 1088–1110, 2018.
- [4] M. C. Linn and A. C. Petersen, "Emergence and characterization of sex differences in spatial ability: A meta-analysis," *Child development*, pp. 1479–1498, 1985.
- [5] D. H. Uttal, D. I. Miller, and N. S. Newcombe, "Exploring and enhancing spatial thinking: Links to achievement in science, technology, engineering, and mathematics?," *Current Directions in Psychological Science*, vol. 22, no. 5, pp. 367–373, 2013.
- [6] B. M. Casey, E. E. Pezaris, and J. Bassi, "Adolescent boys' and girls' block constructions differ in structural balance: A block-building characteristic related to math achievement," *Learning and Individual Differences*, vol. 22, no. 1, pp. 25–36, 2012.
- [7] S. Hsi, M. C. Linn, and J. E. Bell, "The role of spatial reasoning in engineering and the design of spatial instruction," *Journal of engineering education*, vol. 86, no. 2, pp. 151–158, 1997.
- [8] M. Contero, J. M. Gomis, F. Naya, F. Albert, and J. Martin-Gutierrez, "Development of an augmented reality based remedial course to improve the spatial ability of engineering students," 2012, pp. 1–5.
- [9] R. T. Azuma, "A survey of augmented reality," *Presence: Teleoperators and virtual environments*, vol. 6, no. 4, pp. 355–385, 1997.
- [10] J. Dorribo-Camba and M. Contero, "Incorporating augmented reality content in Engineering

Design Graphics materials,” in *2013 IEEE Frontiers in Education Conference (FIE)*, 2013, pp. 35–40.

[11] J. Martin-Gutierrez, R. E. Navarro, and M. A. Gonzalez, “Mixed reality for development of spatial skills of first-year engineering students,” 2011, p. T2D–1–T2D–6.

[12] A. Dünser, K. Steinbügl, H. Kaufmann, and J. Glück, “Virtual and augmented reality as spatial ability training tools,” in *Proceedings of the 7th ACM SIGCHI New Zealand chapter’s international conference on Computer-human interaction: design centered HCI*, 2006, pp. 125–132.

[13] J. E. Bell, H. Klautke, C. Cheng, D. Freer, W. Cain, T. Hinds, S.P. Walton, and C. Cugini, “A Study of Augmented Reality for the Development of Spatial Reasoning Ability (Work in Progress),” in *2017 ASEE Annual Conference & Exposition*, 2017.

[14] E. L. Hutchins, J. D. Hollan, and D. A. Norman, “Direct Manipulation Interfaces,” *Human–Computer Interaction*, vol. 1, no. 4, pp. 311–338, Dec. 1985.

[15] D. Kieras, D. Meyer, and J. Ballas, “Towards demystification of direct manipulation: Cognitive modeling charts the gulf of execution,” in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2001, pp. 128–135.

[16] S. A. Sorby and P. Dineen, *Developing Spatial Thinking*. Higher Education Services, 2015.

[17] S. P. Walton, M. Urban-Lurain, A. Idema, T. J. Hinds, and D. Briedis, “Spatial Visualization Skills Intervention for First Year Engineering Students: Everyone’s a Winner!,” *2015 ASEE Annual Conference & Exposition*, 2015.

[18] R. M. Ryan, “Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory.,” *Journal of personality and social psychology*, vol. 43, no. 3, p. 450, 1982.

[19] M. C. Linn and A. C. Petersen, “Emergence and characterization of sex differences in spatial ability: A meta-analysis,” *Child development*, pp. 1479–1498, 1985.