AC 2010-546: SPATIAL ABILITY TESTING WITH AUGMENTED REALITY

Patrick Connolly, Purdue University

James Beeler, Purdue University

Pat Connaughton, Purdue University

Jared Price, Purdue University

Ben Trefz, Purdue University

Spatial Ability Testing with Augmented Reality

Abstract

This paper reports on research in spatial ability testing comparing traditional 2D test instrument images with hybrid testing involving augmented virtual reality images. Augmented reality involves the combination of virtual reality displays with real world environments. Non-immersive virtual reality systems are becoming popular in many industries and across many applications such as scientific visualization, product development, and interactive entertainment. The goal of this research was to examine the potential impact of augmented reality technology on the successful completion of a standardized visualization test instrument. Participants in the study were students enrolled in an introductory engineering design graphics course at a Midwestern university. Two test groups were organized from the students in the class, one control and one treatment. The control group completed the standard Vandenberg Mental Rotations Test, with static images displayed on computer workstation monitors. The participants received 90 seconds to answer each question on the test. The treatment group received identical test questions but viewed the object as a 3D augmented reality image that was slowly rotated through one revolution. The participants in this group also received 90 seconds to answer each question on the test. Quantitative and qualitative results were recorded for the study. Future research plans are discussed as well as lessons learned from this augmented reality application.

Introduction

Spatial skills, sometimes referred to as spatial ability, are increasingly important in a workplace that is dependent on collaboration and communication. These spatial skills are vital in numerous fields including engineering, medicine, and manufacturing. Historically, there has been a great deal of interest in methods of instruction and technology that could potentially increase the spatial skills of its users. In recent years, the rise in virtual reality, be it immersive, augmented, or desktop, has fueled renewed research in spatial ability development.

During the Spring 2009 semester at ________ University, a research group sought to investigate the potential of augmented reality (AR) as a solution to this growing desire for a technology to aid in spatial ability development[1]. In their study, they found qualitative support that augmented reality was a potential means of achieving greater spatial skill development. This initial research featured a control group and an experimental group, each taking a standard visualization test. The control group took the test without any technology to assist them, while the experimental group was allowed to use augmented virtual reality displays to assist in completing the visualization test. It was found that a majority of the students that used the augmented reality tool found it to be a useful resource in their taking of the test. Many participants also felt that it would be a good educational tool to help develop spatial skills.

This paper reports on a follow-on study attempting to validate quantitatively the qualitative results of the first study. According to Firestone[2], there is a significant advantage to using both qualitative and quantitative studies in combination to further qualify a conclusion. Qualitative studies work to create a series of potential avenues for further research while quantitative studies choose one of these routes, singles it out, and attempts to validate it.

Spatial Ability

Spatial ability has been researched extensively for more than a century. There is significant evidence that the ability to mentally envision, retain, and manipulate three dimensional images[3] is a crucial cognitive ability to be successful in many tasks in both academic and professional applications.[4, 5, 6, 7] Research into spatial ability and constructs branches into many fields, such as psychology, early-age development, and physical and mental activities.

Spatial skills are critical in today's technological and computer-dependent society. According to Contero et al.[8], spatial ability is an increasingly prevalent skill required in the workplace, and many employees are lacking in this area. Because of this growing need for spatial ability development, there is an accompanying demand for any tool that can aid in the development of spatial skill.

Virtual/Augmented Reality

Virtual reality (VR) can be defined many ways according to a broad range of applications[1]. Augmented reality is considered to be a form of virtual reality that incorporates both virtual and real world environments simultaneously.[9, 10] Augmented reality has been somewhat limited in application until relatively recently due to excessive cost and technology limitations. However, augmented reality is growing in popularity as these challenges are overcome. It has been embraced by a variety of industries. It has already been incorporated into the medical industry as a tool for doctors to practice before complicated and delicate surgeries,[11] by psychiatrists as a potential way to cure phobias,[12] and by universities as a means to create interactive tours of their campuses for visitors.[13]

The technology required for simple augmented reality applications is easily attainable.[1] For this study, an augmented reality software program was utilized called BuildAR, provided by HITLabNZ. The program utilizes the computer-based camera to create augmented reality images on the computer screen. The user creates 'markers,' black squares with white patterns inside of them, that are associated with specific models. Examples of these markers are shown in Figure 1.

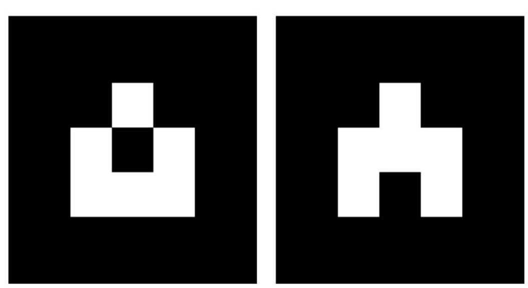

Figure 1. Examples of BuildAR markers

The pattern shown on the marker is linked by the BuildAR software to a specific 3D model file. When one of the markers with an associated model file comes into the camera's viewpoint, the model appears floating above the marker within the camera viewport. A picture representation of what would be shown in the camera's viewport is shown in Figure 2.

Figure 2: Augmented reality image using BuildAR

Methodology

The participants in this study were students in a freshman level computer graphics technology course. This introductory course focuses on basic 3D modeling techniques in a variety of software packages, and enhancing spatial ability. Because of this, the students were assumed to have at least some exposure to spatial ability tests throughout the duration of the class. The majority of the students in the computer graphics technology program tend to have average to above average spatial skills.

A random number generator was used to separate the participants into the control or experimental group. The students were taken two at a time to either the control test location or the experimental test location and their results were recorded. The students were not told whether they were in the control or the experimental group.

Both the experimental and control forms of the test were based on the Vandenberg Mental Rotation Test. This test, based on a 1971 instrument developed by Steven Vandenberg, is a two-part test. Both parts consist of ten questions and have a time limit of five minutes. They feature a picture of one model and four potential answers. Each answer is an image of a model rotated to a different angle. Two of the four answers feature an identical model to the original while the other two are different. The person taking the test has to correctly choose the two that are the same.[14] An example question is shown below in Figure 3.

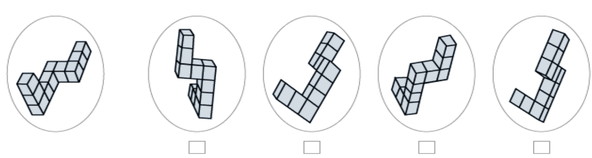

Figure 3: An example of a Vandenberg Mental Rotation Test Problem

The control group of this study completed the standard paper-based version of the Vandenberg Test. They had five minutes to complete each section. The control group consisted of 37 participants. The experimental group participants were presented with twenty markers, each corresponding to one of the questions on the test. Using the markers and the provided laptop with the BuildAR interface displayed, they were allowed to use the augmented reality image of the object to aid them in taking the test. Before they were given the test, the experiment group participants were given a brief explanation of how the BuildAR interface works and a short demonstration on how to use the markers. Similar to the control group, the experiment group participants were given five minutes to complete each ten-question section. This group consisted of 41 total students. Figure 4 shows an example of one of the augmented reality images.

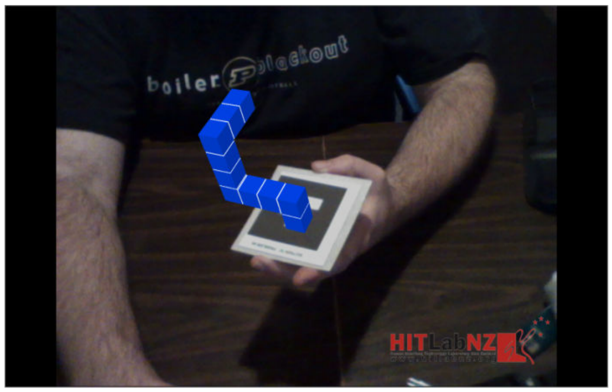

Figure 4: An example augmented reality object used in the study

The main focus of the research was to see if the test scores would increase with the addition of the AR resource. The test administrators also tracked how long it took the students to complete the test.

After the test was completed, the participants were asked to fill out a demographic survey. The survey asked them for their major field of study, age, gender, and year in school (freshman, sophomore, junior, or senior).

Results

The data from this study were classified by the overall experiment group test scores vs. the control group test scores, and also examined by major, age, gender, and year in school, as gathered by the post-test demographic survey. Once the results were compiled, they were evaluated in terms of test average and time to completion, to see if there was a statistically significant increase in either.

It was noted during the testing, that a large number of students in the experimental group used the augmented reality feature for part 1 of the test but did not use it for part 2. Furthermore, a large majority of students in the experimental group (28 of 33), did not finish part 1 of the test in the allotted time. Despite the assurances of the test administrators in between parts 1 and 2 that there was no penalty for not finishing the test, nearly all of the students rushed through part 2 in order to finish the test. This rushing was evident in the students' large drop in time per question, decrease in accuracy, and the general observations of the testers. Because of this, the administrators felt that the validity of the results for part 2 of the test was compromised. Therefore, the averages that were primarily used in the statistical analysis were the averages from each group's part 1 portion of the test. Figure 5 shows the statistical results from part 1 of the test in both accuracy and time, separated by demographic criteria.

PART 1 SCORES (in percent)

	Without AR	With AR	Difference	StanDev	Z Score	P-Value
Overall	84.5	90.8	6.3	15.5	0.40645	0.34221
Males	92.4	94.8	2.4	9.8	0.24490	0.40327
Females	73.3	78.6	5.3	17.7	0.29944	0.38230
CGT majors	84.8	89.5	4.7	16.4	0.28659	0.38722
Other majors	84.0	96.6	12.6	14.2	0.88732	0.18745
Freshmen	85.4	89.8	4.4	16.2	0.27161	0.39296
Upperclassmen	83.1	91.7	8.6	18.1	0.47514	0.31734

PART 1 TIMES (in seconds)

	Without AR	With AR	Difference	StanDev	Z Score	P-Value
Overall	12.5	21.6	9.1	3.7	2.45946	0.00696
Males	12.3	20.8	8.5	4.1	2.07317	0.01908
Females	13.2	24.1	10.9	2.8	3.89286	0.00004
CGT majors	13.2	21.2	8.0	3.8	2.10526	0.01763
Other majors	11.5	21.9	10.4	3.2	3.25000	0.00057
Freshmen	13.1	23.9	10.8	3.5	3.08571	0.00102
Upperclassmen	11.4	19.8	8.4	3.2	2.62500	0.00433

Figure 5: Statistical results by demographic (Test Part 1 Only)

For both test parts combined, the accuracy improved from 85.9% in the control group to 88.9% in the experimental group. While this is appeared to be a noticeable improvement, the results were not statistically significant. This may be due in part to the large standard deviation of the scores. The spread ranged from 8/20 on the test (40%) to 20/20 (100%). For part 1 only of the test, the addition of augmented reality to the test caused the average to rise from 84.5% to 90.8%. While this is a larger jump than the overall averages, it was also not statistically significant.

In terms of time of completion, the increased accuracy in test results also came at the expense of an increase in time. While it took the control group an average of just 24.5 seconds to complete each question for both parts of the test, it took the experimental group an average of 40.88 seconds. This increase in time, like many of the individual demographic time differences, was statistically significant. The overall time increase with the augmented reality group had a z score of 2.34 and a corresponding p-value of .01. In part 1 of the test, the time difference was even more pronounced. It took the control group an average of 12.5 seconds to complete each question in Part 1 of the test, while it took the experimental group 21.6 seconds to complete each question in this part. This result was also statistically significant.

There was a noticeable difference in the returned test results of the people by each demographical category. The first examined was that by major. 41 of the test takers were Computer Graphics Technology majors. 25 of the remaining students were comprised of other majors, including 24 Industrial Design students and 1 Computer Science student. One student listed their major as both CGT and CS and was counted in both groups. For both groups, their control group averages were very similar. The CGT students averaged 84.8% on the control group test while the other majors scored an average of 84.0%. The AR scores of the two groups were similar as well. The CGT majors scored an 89.5%, while the non-CGT students scored 96.6%. This jump by the non-CGT students was the largest of any demographic, although not statistically significant. They had a z score of .887 and a p-value of .188. A graph of these results is shown in Figure 6.

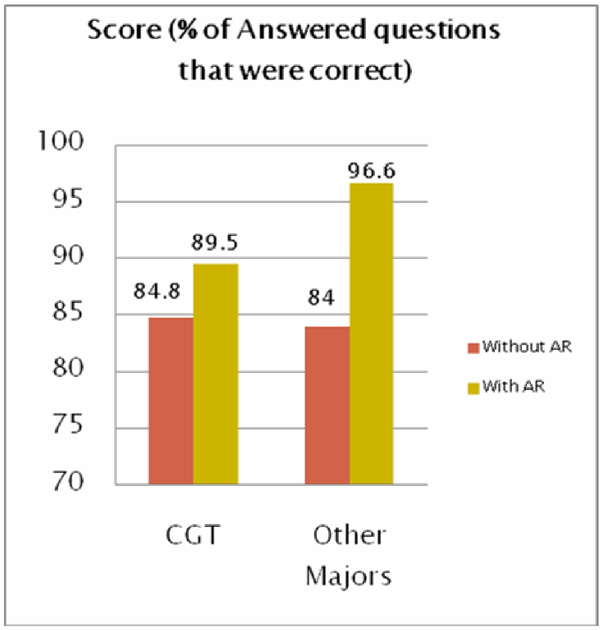

Figure 6: Test results by major

One of the most noticeable differences in our results was that of the difference in results by gender. Many studies have shown that women tend to test lower on mental rotation spatial abilities tests than men.[15] This result was confirmed in this study. With the 49 men and 21 women who took the test, the men scored higher than women in both groups. In the control group, men scored over 20 percentage points higher than the women, 90.7% to 70.6%. In the experimental group, the males scored higher than their female counterparts by a margin of 94.8% to 78.6%. However, women showed a much larger improvement with the use of augmented reality. Their increase in percentage points from the control to the experimental group was double that of the men's. The standard deviation being so high (17.7) again led to this not being statistically significant. A graph of these results is shown in Figure 7.

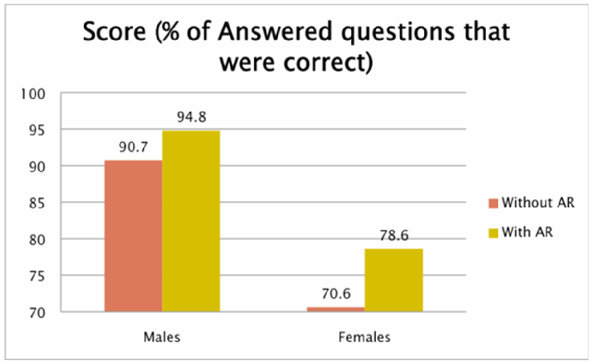

Figure 7 Results by gender

There are several reasons for a potentially larger rise in female scores. First, it appears that augmented reality is a highly beneficial remediation tool for people with low initial spatial ability skills. It may be that the higher the initial spatial ability level is, the lesser the impact of the tool. Also, there also existed the possibility of a plateau effect of the scores. Augmented reality worked well for helping those with initially low spatial ability scores, raising their percentages by large degrees. However, its impact was lessened as the original scores increased. Because the women's initial score was so much lower, the augmented reality addition may have had a larger effect.

There were noticeable differences in the class years of the students taking the test as well. There were 40 freshmen and 29 non-freshmen (sophomores to seniors) that completed the study. The freshmen scored minimally higher than their upper classmen counterparts in the control group, scoring an average of 85.4% versus the upper classmen's 83.1%.. However, those roles were reversed in the augmented reality test results. The upper classmen were able to score 91.7% on section 1 of the test while the freshmen scored 89.8%. While not statistically significant, it appeared that the upper classmen experienced a larger boost with augmented reality. A graph of this data is shown in Figure 8.

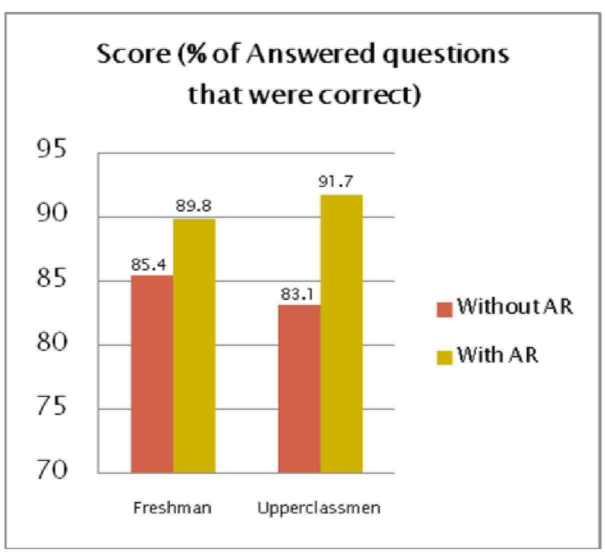

Figure 8: Results by class year.

A potential explanation for these numbers lies with the upper classmen's assumed increased experience with new technology. It can be assumed that part of the college experience in high technology fields involves learning how to deal with and overcome the challenges inherent in dealing with novel interfaces. Some freshmen participants seemed confused or overwhelmed by the new technology and could have potentially been overly frustrated with early struggles to use the augmented reality tool effectively.

There were seven students who were placed in the experimental group who stopped using the augmented reality tool after section 1 of the test. Two of these participants completed eight of the ten questions of part1, and the other five participants completed five or less of the ten questions in Part 1. Because of this, they felt that the augmented reality tool was hindering their speed to a significant degree, and decided to take the test without the tool for Part 2.

Their Part 1 results were consistent with the overall augmented reality results. They scored an average of 89.6 versus the overall augmented reality Part 1 average of 90.8. However, they were slower than the average for the augmented reality group in Part 1. The average was 43.2 seconds per question while this sub-group averaged 54.6 seconds per question. However, the Part 2 mean score for these participants was 95.3%, compared to the 87.3% average of the control group on Part 2. Neither used the augmented reality tool, but the experimental group did noticeably better. One potential recommendation from this result may be that the augmented reality tool may be best utilized as a reference or learning aide as opposed to a mainstream educational tool. The time-consuming nature of the augmented reality tool could be difficult to incorporate into a busy curriculum and could become frustrating to the user. However, when it is used as a tool in the initial stages of spatial ability exercises, or as a remediation tool, it may be beneficial in spatial skill development.

Recommendations for Future Research

There are many potential avenues to pursue to further validate the concept of augmented reality applications in spatial skill development. It is noted that eliminating the test time limit would most likely provide vastly different results for Part 2 of the test. It was felt that the inclination of the test participants to rush through Part 2 of the test negatively impacted the accuracy and usefulness of the test results. The ability was lost to compare a potential increase in accuracy and reduction in time with a longer exposure to the augmented reality tools. It would be interesting to research if the participants increased at the same rate as the control group from Part 1 to Part 2 without the time constraint.

A second recommendation would be to increase the number of test-takers. Many of the statistical calculations were impacted by the broad result spreads. An increased number of participants would help to provide more statistical power to results data.

Another future recommendation would be to involve other academic majors in similar research. Due to a restricted participation of the participant population, there was a limit on discovering the variation from major to major. It is anticipated that students from different majors would have different spatial ability norms. In general, this research project has suggested that the lower the initial spatial ability skill level, the higher the impact of the augmented reality tool. This analysis may not be consistent across all majors and with those that have a significantly lower starting spatial skill level.

Similar research could involve age groups other than college students, such as elementary school children and more mature populations. The ability to analyze the results based on age was extremely limited due to the tiny variance in the ages of the participants in this test. More research exposure in this realm may prove valuable in identifying future applications of augmented reality in other areas outside of education.

Bibliography

1. Connolly, P., Hartman, N., Bertoline, G., Takahashi, G., Lynam, J., Tittle, D., & Rose, D. (2009). Augmented Reality in Spatial Ability Development: A Concept Study. Proceedings of the 64th Mid-Year Conference of the Engineering Design Graphics Division of the American Society for Engineering Education, Erie, PA.
2. Firestone, William A. (1987). Meaning in Method: The Rhetoric of Quantitative and Qualitative Research. Educational Researcher, Vol. 16 No. 7, 16-21
3. Lohman, D. F. (1979). Spatial ability: A review and reanalysis of the correlational literature. Palo Alto, CA: Stanford University.
4. McGee, M. G. (1979). Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences. Psychological Bulletin, 86(5), 889-918.
5. Miller, C. L. (1992). Enhancing visual literacy of engineering students through the use of real and computer generated models. Engineering Design Graphics Journal, 56(1), 27-38.
6. Smith, I. M. (1964). Spatial ability: Its educational and social significance. San Diego, CA: Robert R. Knapp.
7. Sorby, S. A. (2000). Spatial abilities and their relationship to effective learning of 3-D solid modeling software. Engineering Design Graphics Journal, 64(3), 30-35.
8. Contero, M., Company, P., Saorin, J. L., & Naya, F. (2006). Learning support tools for developing spatial abilities in engineering design. International Journal of Engineering Education, 22(3), 470-477.
9. Billinghurst, M. (2002). Augmented Reality in Education. Retrieved September 18, 2009, from http://www.newhorizons.org/strategies/technology/billinghurst.htm
10. Kaufmann, H. (2003). Collaborative augmented reality in education. Paper for keynote speech at Imagina 2003 conference, Monaco.
11. Stadie, Axel Thomas & Kockro, Ralf Alfons & Reisch, Robert & Tropine, Andrei & Boor, Stephan & Stoeter, Peter & Perneczky, Axel (2008). Virtual reality system for planning minimally invasive neurosurgery. Journal of Neurosurgery, Vol. 108, 382-394.
12. Botella, C.M. & Juan, M.C. & Banos, R.M. & Alcaniz, M. & Guillen, V. (2005). Mixing Realities? An Application of Augmented Reality for the Treatment of Cockroach Phobia. Cyber Psychology & Behavior. Vol. 8, Num 2, 162-171.
13. Feiner, Steven K. (2002). Augmented Reality: A New Way of Seeing. Scientific American, April 2002. 50-55.
14. Miller, C. L. (1996). A historical review of applied and theoretical spatial visualization publications in engineering graphics. Engineering Design Graphics Journal, 60(3), 12-33.
15. Smith, S. S. (2003). A design-based engineering graphics course for first-year students. Engineering Design Graphics Journal, 67(2), 33-42.