

# *Use of Augmented Reality Interfaces to Enhance Art Student Learning: An Experimental Comparison of Learning Platforms*

Kenny Jones, M.F.A.  
June 1, 2011

## Abstract

Augmented Reality (AR) can be described as layering virtual digital content on top of a real time image, for example as seen through an iPhone. I am interested in exploring whether the AR paradigm can significantly increase learning efficiency and retention as it relates to student perception and their report of what they see. I studied two art classes held in Fall 2010. Both groups were assigned the task to design a piece that would visualize information related to the ACU sculpture, "Jacob's Dream", created by Jack Maxwell. One group created an AR interface that would guide visitors in an informational walking tour. The other group created an informative, non-AR pamphlet. Findings from this study were inconclusive as to the AR affect on improvement of perception. This was due to evidence that the classes were too different before the study began. What we have learned is that it is possible to reliably measure this kind of learning, namely, the perception of visual design forms in far transfer artifacts.

## I. Statement of the Problem

### Can Augmented Reality Enhance Learning for Art and Design Students?

As a professor in the Art and Design department, I have had an experience with freshmen that is probably common to most instructors, a consistent use of vagueness in describing a subject. Furthermore, I have noted a connection between freshmen's use of vocabulary and their ability to accurately perceive relevant features. In my case, the vagueness of terminology involves artworks and the essential elements and principles of visual design. Often freshmen art and design majors, when asked to describe the essential and relevant features of an artwork, will fall back on subject matter labeling, such as, "the artist has painted horses and a race track with spectators." Alternately, some would describe how the artwork makes them feel, usually in equally vague terms of their general liking or disliking of it.

Instead, the foundation of an artwork is first and foremost a composition of visual design elements and principles, e.g. shapes, color, balance, visual rhythm, etc. It is not easy to train students to perceive and describe an artwork in these visual design terms. Over the years of my teaching freshmen level, two-dimensional design courses, I have developed overhead transparency overlays so as to demonstrate how to analyze the shapes and other visual forms present in a projected artwork. As a test I would prompt students to use a list of design terms in order to create diagrams on top of an artwork that highlight the shapes and other forms relevant to a visual design analysis. This was often time consuming and tedious for the students and for me in grading the test. When I first saw the Augmented Reality (AR) concept it appeared to be a perfect extension of this mode of analysis – an overlay of a diagram analysis on top of a scene. Except the potential now is to go beyond simple diagrams to fully realized content comparisons. Augmented Reality, or as it is often referred to AR, can be described as layering virtual digital content on top of a real time image, in this case as seen through an iPhone/iPod touch/iPad 2 camera. I am interested in exploring whether the AR paradigm can

significantly increase learning efficiency and retention of Art and design students as it relates to perception and their report of what they see.

For a long time I have been interested in trying to get some hard data on whether perception is enhanced by digital mobile technology or simply distracting us from the present moment outside of our iPhone. AR's expansion of digital, layered content in real time, on actual objects is one reason I thought that AR might increase student's visual perception of design elements and principles. The interactive engagement afforded students by AR presents a good case to test whether enhancement of perception for visual design components is possible with the iPhone.

## II. Hypothesis

I am hoping to leverage a particular aspect of the iPhone/AR paradigm, that is, a layered context of seeing and interacting with a real-world, real-time artwork via a digital interface. I am hoping that this deeper level of engagement will help students to perceive and understand visual forms better than students who do not engage the AR paradigm.

My Research hypothesis is that students who design an Augmented Reality simulation for an artwork will display a significant increase in perception of visual design forms when compared to students who engage similar projects without the Augmented Reality design context.

## III. Methods

The Field experiment:

The focus and comparison groups were two separate sections of my Two-Dimensional Design class, both held in the regular term semester of Fall 2010. Both groups were assigned the task to design a piece that would envision information related to the campus sculpture, "Jacob's Dream", created by Jack Maxwell.

The focus and comparison groups both engaged in extensive research and analysis of the near transfer artifact, "Jacob's Dream". Both groups used a variety of iPhone apps, writing assignments, digital photography and rough drafts to plan for the final project.

Only the focus group was to use this research to create an AR sim as their final outcome. This AR interface would guide visitors in an informational walking tour of "Jacob's Dream". The students were told that computer programmers could code their AR interface and develop an iPhone app at a later time. Therefore, the students were prompted to make the AR sim as close to an actual AR experience as possible. The comparison group used their research material to create an informative, four-page pamphlet of "Jacob's Dream" that could be printed at a later time and distributed to campus visitors.

I strove to ensure that the only difference in the two groups was the AR sim component. For example, both groups made projects with the same media and tools as each group's final project. Students in the focus AR group were encouraged to produce a rough draft in a pamphlet format similar to the final project of the comparison group. Furthermore, both comparison group and focus group used several iPhone apps, such as Flipbook and PS mobile, though only the focus group used them for designing within an AR context.

## How the data was gathered

Both the comparison group and focus group were prompted to write a four to eight page essay. This essay was administered twice during the semester, once at midterm as a pre-experiment measurement and a final end of term post-experiment measurement. Each essay was comprised of four questions related to two artifacts. The far transfer artifact was changed in the two essays; while the near transfer artifact remained the same in both essay measurements. The near transfer image was of the same sculpture that the students were designing for the project's experiment, "Jacob's Dream". The far transfer artifacts had not been shown or used at any time during the semester and was an older artwork done in oil paint media.

Two essay questions were asked for each artifact. The first essay question asked for a description of the significant visual forms and their role in producing visual cohesion while maintaining diversity (Unity in Variety). And the second essay question asked for the student to report on their affected response to the artifacts – how the artifacts made them feel in a design sense. The first question addressed the research focus response and the second question functioned as a comparison response. A prompt sheet of design terms was provided and each student could refer to it during both essay sessions.

These essays were then coded, copied and mixed so that the research assistants, graduate students in psychology, would not know whether essays were from the pre-experiment or post-experiment. Additionally, the assistants would not know which group any given essay came from, either focus or comparison group.

I then went through a training session with the research assistants, giving them a sample essay from another source. After reviewing their measurement of the essay we then discussed the appropriate and relevant use of the design terms in question. Then the research assistants tallied how many design terms the student reported seeing. The assumption is that each use of a distinct descriptive visual design term will be indicative of how well the student perceived those design elements and principles.

After totaling the number of terms written by the students in all essays, we then entered these scores in a spreadsheet application. A score was recorded for both groups, with scores distributed according to each artifact, each essay question related to design terms, and each measurement session, e.g. pre-experiment or post-experiment. If there was a difference between the raters' scores in any of these categories we adjudicated those scores.

Adjudication was applied in the following situations.

If the difference between the scores was:

1. one point, we took the higher score
2. two points, we took the middle score of the two, e.g. 1 vs. 3 would be rated as 2
3. three or greater, we went back to the essays in question and compared word for word, discussing why different terms were counted or rejected in the first rating. After the raters provided input, I then suggested which terms in question could be added or eliminated from the scores. I could only comment on terms that at least one rater had scored, all other terms were not open for discussion or scoring. There was no time limit to our discussions and we would not move on until consensus was reached.

After adjudication, scores were then entered into SPSS software in order to analyze the data.

#### IV. Analysis and results

I am happy to report that the interrater reliability was very high, even before the adjudication process, the interrater reliability = .98. Results of the analysis show that for both near and far transfer artifacts, the focus group scored significantly higher than the comparison, non-AR group. However, because the groups were not randomly assigned this leaves open the possibility that the groups were also different before the experiment began. In order to test for this possibility, I preformed the same analysis on the pre-experiment test final scores to see if the groups might have been different prior to the experiment. The results of those tests are in the following four sets of tables:

**Group Statistics for  
Post-Experiment Test, Near Transfer table**

<b>Group Final</b>	<b>N</b>	<b>Mean</b>	<b>Std. Deviation</b>	<b>Std. Error Mean</b>
<b>Comparison</b>	<b>20</b>	<b>11.25</b>	<b>9.239</b>	<b>2.066</b>
<b>Focus AR</b>	<b>16</b>	<b>20.00</b>	<b>14.638</b>	<b>3.659</b>

**T-test for Equality of Means - Equal Variances Assumed  
Post-Experiment Test, Near Transfer table**

<b>t</b>	<b>df</b>	<b>Sig. (2-tailed)</b>	<b>Mean Difference Difference</b>	<b>Std. Error Difference</b>
<b>-2.187</b>	<b>34</b>	<b>.036</b>	<b>-8.750</b>	<b>4.000</b>

**Group Statistics for  
Post-Experiment Test, Far Transfer Table**

<b>Group Final</b>	<b>N</b>	<b>Mean</b>	<b>Std. Deviation</b>	<b>Std. Error Mean</b>
<b>Comparison</b>	<b>20</b>	<b>10.25</b>	<b>6.180</b>	<b>1.382</b>
<b>Focus AR</b>	<b>16</b>	<b>17.75</b>	<b>12.477</b>	<b>3.119</b>

**T-test for Equality of Means - Equal Variances Assumed  
Post-Experiment Test, Far Transfer Table**

<b>t</b>	<b>df</b>	<b>Sig. (2-tailed)</b>	<b>Mean Difference Difference</b>	<b>Std. Error Difference</b>
<b>-2.357</b>	<b>34</b>	<b>.024</b>	<b>-7.500</b>	<b>3.182</b>

**Group Statistics for Pre-Experiment Test,  
Near Transfer table**

<b>Group Final</b>	<b>N</b>	<b>Mean</b>	<b>Std. Deviation</b>	<b>Std. Error Mean</b>
<b>Comparison</b>	<b>20</b>	<b>6.05</b>	<b>5.021</b>	<b>1.123</b>
<b>Focus AR</b>	<b>16</b>	<b>15.07</b>	<b>11.504</b>	<b>2.970</b>

**T-test for Equality of Means - Equal Variances Assumed  
Pre-Experiment Test, Near Transfer table**

<b>t</b>	<b>df</b>	<b>Sig. (2-tailed)</b>	<b>Mean Difference Difference</b>	<b>Std. Error</b>
<b>-3.140</b>	<b>33</b>	<b>.004</b>	<b>-9.017</b>	<b>2.871</b>

**Group Statistics for  
Pre-Experiment Test, Far Transfer table**

<b>Group Final</b>	<b>N</b>	<b>Mean</b>	<b>Std. Deviation</b>	<b>Std. Error Mean</b>
<b>Comparison</b>	<b>20</b>	<b>8.14</b>	<b>6.909</b>	<b>1.508</b>
<b>Focus AR</b>	<b>16</b>	<b>15.33</b>	<b>8.934</b>	<b>2.307</b>

**T-test for Equality of Means - Equal Variances Assumed  
Pre-Experiment Test, Far Transfer table**

<b>t</b>	<b>df</b>	<b>Sig. (2-tailed)</b>	<b>Mean Difference Difference</b>	<b>Std. Error Difference</b>
<b>-2.725</b>	<b>34</b>	<b>.010</b>	<b>-7.190</b>	<b>2.639</b>

## V. Conclusion

While there were significant score differences between the groups on both near and far transfer artifacts, no conclusion can be drawn from the current data due to the evidence that the groups were different prior to the first essay. This is often a common problem in doing field research with learning outcomes. These initial conditions could well have been caused by several factors that are not easily controlled or analyzed. For example, the differences in class time, 8am compared to 1:30pm – with the 1:30pm class scoring higher. Also, there was a difference in the classification distribution of the classes with the higher scoring group having more students classified as sophomore or above. The fact that these upper classmen could have been exposed to other courses that presented design terms in a similar manner to this study means that more students in the higher scoring group may have begun the study with better perception skills than the lower scoring group. Unfortunately, the sample size of this study precludes our analyzing these variables with acceptable reliability. One factor I would be especially interested in analyzing would be the teacher's response to the AR group when compared to the non-AR group. Did the study itself create a novelty that subtly engaged the instructor in a deeper way with the AR group when compared to the non-AR group and thus affected student learning? Again the limits of our study exclude this type of analysis.

What we have learned is that it is possible to reliably measure the kind of learning that is key to both this study and student success in art and design - the perception of visual design forms in far transfer artifacts. Moreover, this study provides a valuable foundation for extending this research to future experiments, albeit with greater randomization assignments.

### Thanks and appreciation

I am grateful for the collaborative aid and indispensable help that Dr. Robert McKelvain has given me with this Mobile Learning Fellows opportunity. I am also very grateful to Dr. Scott Perkins, Dr. William Rankin and others who guided me in terms of thinking through the iPhone AR potentials. And I am especially thankful for the generous grant made possible by the ACU Research Center and Dr. Scott Perkins, coordinator of Mobile Learning Research.