Educational Application of 3D Interactive, Dynamic Multimodal Visualizations in Virtual Learning Environments

Phuong T. Doa*, Dennis P. Korcheckb, and John R. Morelandc
aDepartment of Behavioral Sciences/Cognitive Psychology; bDepartment of Construction Science and Organizational Leadership; cThe Center for Innovation through Visualization and Simulation (CIVS)
Purdue University Calumet, Hammond, IN 46323-2094, USA

*Correspondence concerning this manuscript should be addressed to Phuong T. Do, PhD, Department of Behavioral Sciences, Purdue University Calumet, 2200 169th Street, Hammond, IN 46323-2094. Email: Phuong.Do@purduecal.edu. Office: 219/989-2576. Fax: 219/989-2008.

Abstract

The present study is a joined effort of our interdisciplinary collaboration in the fields of cognitive psychology, mechanical engineering, and computer sciences. The goal is to implement multimodal training methods and evaluate the effectiveness of our visualization designs on virtual learning acquisition. In a series of two experiments, participants were randomly assigned to different training conditions and rehearsal processes to learn a list of 14 terms associated with construction of a wood-frame house. They then completed a memory test determining their cognitive ability to free recall the definitions of the 14 studied terms immediately after learning. The audiovisual modality training condition displayed the highest accuracy rate, while the auditory-modality condition showing lower accuracy. The visual-modality and no-training conditions revealed similar performance by exhibiting little or no learning acquisition. An increase in performance accuracy with training repetition for the audiovisual condition suggested the relative importance of rehearsal in learning in multimodal visualizations. Findings revealed the potential for advancing knowledge and understanding of the practical use of interactive, dynamic visualizations on learning and teaching effectiveness.

Key words: 3D, multimodal visualizations, learning and memory, virtual reality, e-learning

1 Introduction

It is important from both theoretical and practical perspectives to examine the potential effects immersive 3D virtual reality may have on learning performance. Three-dimensional (3D) dynamic information visualization is a relatively new field whose interdisciplinary nature incorporates computer and behavioral sciences.9 The extent to which human cognitive abilities are being modulated by computer simulations and interactive interfaces is a fundamental component of our learning adaptability to recent advances in information technology.10-12

The categorization field has recently embraced the possibility that learning and transfer are mediated by an integration of information from multiple sensory modalities.1, 2 The primary findings on
multimodal perceptual effect focus on the combination of narration and two-dimensional animations.\textsuperscript{3} The question of whether knowledge derived from stereoscopic 3D dynamic visualization is mediated by an integration of information from multiple sensory modalities has been largely overlooked in the literature.\textsuperscript{34}

The new movement towards web-based and multimedia learning has recently placed great emphases on interactive, dynamic visualizations.\textsuperscript{13, 14} Dynamic visualizations have been shown to facilitate learning and teaching effectiveness.\textsuperscript{15, 16} Interactive 3D virtual learning environments not only display a stimulating and informative framework for distance learners, but provide educators a communicative tool that can be utilized to foster students’ experiential learning and constructive collaboration.\textsuperscript{17}

Despite the potential benefits of embedding stereoscopic 3D content into our e-learning instructional curriculum, the reciprocal relationship between 3D human-computer interfaces and distance education is unclear. The current research explores the pragmatic application of 3D information visualization in traditional classroom environments. Undergraduate students completed a memory test determining their cognitive ability to free recall the definitions of 14 studied terms associated with the construction of a wood-frame house after going through a multimodal visualization and/or rehearsal learning phase. Two research questions are proposed: (1) which types of dynamic visualizations enable learners to gain experiential learning experience without inducing cognitive overload, and (2) what is the role of rehearsal on learning acquisition as a function of multimodal visualizations? The theoretical goal of this work is to suggest how the present findings provide implications for theories of sensory integration, and perhaps propose a theoretical framework for testing dynamic visualization with recent technologies such as haptic user interfaces and stereoscopic 3D computer monitors.

The practical application of immersive virtual reality as a training tool in web-based education has not yet been fully explored. In the combined effort of interdisciplinary collaboration across campuses, the proposed project demonstrates the viable usefulness of immersive 3D visualizations in the exchange of theoretical concepts and human-computer user interfaces. Our research presents strong support for meritorious, innovative, and interdisciplinary research by addressing a critical barrier to progress in the fields of psychological and computer sciences. Findings present innovative strategies to successfully design interactive, dynamic visualizations that optimize performance accuracy and learning acquisition.

2 Method

In a series of two empirical studies, we examined the influence of multimodal visualizations and rehearsal processes on learning acquisition through the application of immersive virtual reality.

2.1 Participants

One hundred and fifty four ($N = 154$) Purdue University Calumet undergraduate students (68 males and 86 females) were randomly recruited to participate in the current study. Student
participation was voluntary, and participation or nonparticipation did not affect their grades. All of the participants ranged in age from 18 to 25 years and had normal or corrected-to-normal visual acuity, and were naïve to the purpose of the experiment. A total of twenty to thirty participants were randomly assigned to one of the eight learning-testing conditions. Probability sampling strategies and random selection were applied during the participant recruitment process to eliminate the potential for selection bias. Of the 146 participants who provided demographic information, 51% of the students were identified as Caucasians, 15% as African Americans, 15% as Asians, and 19% as other ethnic minorities.

2.2 Materials, Stimuli, and Apparatus

The learning stimuli consisted of a vocabulary list of 14 technical terms associated with the construction of a wood-frame house: beams, ceiling joist, columns, floor joist, floor sheathing, footing, formwork, foundation wall, header, mud sill, reinforcing, ribbon joist, top/bottom plate, and wall studs. These terms were presented in different contexts based on the training condition (i.e., visual-, auditory-, audiovisual-modality, or no learning) as a function of rehearsal process (i.e., repetition or no repetition).

For the audiovisual and visual condition, a virtual walkthrough was used to assist participants in their comprehension of the wood-frame house construction and design while allowing them to interact with objects in 3D. The walkthrough was presented in an immersive virtual environment and was developed using a mixture of commercial, OpenSource, and custom-built software. The wood-frame house was constructed and animated using Autodesk 3dsMax (Figure 1).

This artifact was then exported as a 3D mesh model that was loaded into a custom-built application that utilized the OpenSource VRJuggler & OpenSceneGraph libraries. Once loaded into the immersive environment, the different phases of construction were shown as a series of interactive stereoscopic 3D animations.

Figure 1. A brief demonstration of the 3D, dynamic visualization process of designing the
immersive virtual reality environment of a wood-frame construction.

The immersive virtual environment was displayed using a virtual reality system that allowed participants to see and interact with data and objects in a three-dimensional space. The display consists of two large screens: a 7’6” x 10’ rear projected wall, and a floor screen of the same size that participants stand on (Figure 2). Combined with 3D glasses, this configuration allows users to stand inside the visualization, with imagery coming out of both the screen in front of them and the floor beneath them.

![Figure 2](image.png)

Figure 2. An illustration of the immersive virtual reality system used in the study, by VisBox Inc.

The system uses four projectors (two per screen) to provide the 3D imagery (called Stereoscopic 3D) and creates an illusion of depth. Filters built into the projectors correspond with the special lenses in the 3D glasses so that each eye only sees an image from one projector in a stereoscopic pair. The 3D images are sent to the projectors from a high-end PC with four video outputs. While the system is also capable of tracking a user’s head and hand movements within the space, this functionality was not used in the present study. Following the learning phase, participants were given a written test asking them to match each term with its corresponding definition.

### 2.3 Research Design and Procedure

#### 2.3.1 Experiment 1

Experiment 1 examined the influence of multimodal 3D visualizations (i.e., visual-, auditory-, audiovisual-modality, or no training) on learning performance accuracy. An immersive virtual environment was developed to assist participants in their comprehension of the wood-framed construction design while allowing them to interact with objects in a 3D environment (Figures 1-2). The stereoscopic 3D system tracks an observer’s head and hand movements as they navigate in the 3D space. The system uses these measurements to modify the virtual objects so they appear at the...
correct scale, allowing the observer to view 3D objects at the size they would appear in real life. Following their interaction with the virtual environment (i.e., the learning phase), the participants completed a memory recall test.

The primary objective of Experiment I was to evaluate (1) the effectiveness and usability of our 3D virtual construction design of a wood-framed single family house, (2) whether or not participants could successfully consolidate and transfer information in immersive 3D virtual environment, and (3) the extent to which multimodal visualizations play in learning and transfer of abstract concepts such as architectural principles.

A significant effect of multimodal visualizations on learning performance was found. Despite the medium flooring effect in performance accuracy, evidence suggested that multimodal visualization training facilitated learning performance accuracy as compared to no training. The audiovisual- and visual-modality training conditions displayed comparable accuracy rates, while the auditory-modality training condition revealed lower accuracy, and the no-training condition exhibited little or no learning acquisition. The process of simultaneously exposing learners to interactive dynamic visualizations and prompting them to attend to information through the pragmatic use of audio cues reduced memory load, and in turn facilitated memory recall.

Nonetheless, the flooring effect in learning performance found in the present study might suggest that participants were not able to easily comprehend the construction terms with just one learning trial. Whether this effect could be remedied using multiple learning trials was the primary rationale behind Experiment 2.

2.3.2 Experiment 2

Experiment 2 was a direct replication and extension of Experiment 1 coupled with the rehearsal manipulation. A cohort of five participants were randomly assigned to one of the four counterbalanced learning conditions. The four cohorts of participants for each condition were counterbalanced and randomly rotated across the four training conditions (e.g., ABCD, BCDA, CDAB, DABC). A Greco-Latin square was used to assign participants to the four training conditions, with O1, O2, O3, O4; and Visual, Auditory, Audiovisual, No Learning representing the order of learning and training conditions, respectively. A total of five subjects in each condition were randomly assigned to each row of the resulting square. As a consequence, each participant was tested equally often by each training condition and learning order.

Prior to the experiment, standard learning instructions were read to inform participants that all 14 technical terms were associated with wood-frame constructions. Participants were then presented a series of 14 terms and their definitions in the different contexts already described over a duration of 6 minutes to help them familiarize with the technical usage. Based on our pilot data and prior evidence that multimodal visualization20 had been shown to display a much faster processing time than visual or auditory training alone, the present experiment equated learning time intervals for each respective training condition. For the visual training condition, each learning word was carefully presented in concurrence with the demonstration of the virtual
construction of a single-family wood framed house. For the auditory training condition, the recorded narrative was the primary learning source. For the audiovisual training condition, participants were visually exposed to the virtual depiction of the construction and they simultaneously heard the narrative as the wood framed house was being built in virtual reality. Thus, information was presented simultaneously in bimodal situations. It should be noted that the visual-auditory training modality condition had exposure of neither visual nor auditory response modality dominance. An imposition of a memory demand on vision or audio might bias learning judgments, due to the possibility of modality differences in memory.21 Participants in the present study were instructed to rely on neither modality, but to use their cognitive ability to recalibrate multimodal information during learning.

Following the first learning phase, participants were given a 5-minute break for questions and answers. A second learning phase was soon followed after the brief intermission. The participants were then given standard testing instructions for the transfer test, which occurred immediately after the second learning trial. The amount of time it took to complete the research activity ranged from 30 to 60 minutes.

3 Results

3.1. Comparative Analyses

Figure 3a demonstrated the results obtained from Experiment 1 without learning rehearsal. The results of Experiment 1 were obtained from our previous study.34 Data for Experiment 2 with learning rehearsal revealed a significant effect of multimodal visualizations on learning performance, $F(3, 70) = 194.75, p < .001$ (Figure 3b). For a detailed description of Experiment 1’s results, please refer to Do, Moreland, & Korche (2010).34 The present article only shows the overall graph of Experiment 1 to compare learning performance against Experiment 2.

The audiovisual condition displayed the highest mean proportion ($M = 8.74, SD = 1.83, \sigma_M = .34$) as compared to the auditory condition ($M = 4.96, SD = 1.47, \sigma_M = .32$), the no learning condition ($M = 1.55, SD = 1.12, \sigma_M = .32$), and the visual condition ($M = 1.43, SD = .89, \sigma_M = .27$). Such a difference is particularly large, $\eta^2 = .89$. It should be noted that the assumption of homogeneity of variance was violated. Hence, the Brown-Forsythe and the Welch $F$-ratios are reported, $F(3, 68.97) = 257.21, p < .001$ and $F(3, 33.28) = 224.55, p < .001$, respectively.
Figure 3. Learning performance accuracy as a function of multimodal visualizations without and with rehearsal, respectively.

Table 2a shows the Bonferroni post hoc analysis for Experiment 1 obtained from our previous study. Table 2b presents the Bonferroni post hoc analysis for the multimodal visualization training conditions with rehearsal, i.e., Experiment 2. When participants were exposed to the auditory or visual training condition, they displayed different learning performance accuracy ($p < .001$). Similar findings were found for the visual and audiovisual training comparison which revealed significantly different performance accuracy ($p < .001$). The visual and no learning conditions revealed similar performance accuracy such that the probability of obtaining any differences between these two group means was not probable ($p = 1.00$). It was unclear why participants performed equally well whether or not they learned the construction concepts visually or received no learning at all. Moreover, participants showed different rates of learning acquisition when they were assigned to the auditory versus audiovisual, no learning versus auditory, or no learning versus audiovisual training condition ($ps < .001$). With the exception of the visual training condition, it is interesting to note that the mean difference of performance accuracy between the no learning condition and the different multimodal visualization training conditions was significantly different ($ps < .001$).
Table 2. Bonferroni post hoc analyses of multimodal visualization performance accuracy without and with rehearsal, respectively.

<table>
<thead>
<tr>
<th>Bonferroni Analysis</th>
<th>Experiment 1: Without Rehearsal</th>
<th>Significance*</th>
<th>Bonferroni Analysis</th>
<th>Experiment 2: With Rehearsal</th>
<th>Significance*</th>
</tr>
</thead>
<tbody>
<tr>
<td>5. No Learning vs. Visual</td>
<td>6.10</td>
<td>.001 – SIG</td>
<td>5. No Learning vs. Visual</td>
<td>0.17</td>
<td>1.00 – NS</td>
</tr>
</tbody>
</table>

*The mean difference is significant at the .05 alpha level.

4 Discussion

Our previous study suggested that participants were not able to easily comprehend the construction terms with just one learning trial, as a result of the flooring effect in learning performance. Whether or not this flooring effect could be remedied using learning rehearsal was the primary objective of our second experiment. The present study investigated the effects of 3D interactive, dynamic multimodal visualizations and training repetition (i.e., rehearsal strategy) on learning acquisition. Evidence suggested that multimodal visualization training facilitated learning performance accuracy as compared to no training. Participants who received training repetition performed better than those who did not. Results overall indicated the relative importance of both multimodal visualizations and rehearsal strategy on performance accuracy. Findings provided strong support for meritorious, innovative, and interdisciplinary research by addressing a critical barrier to progress in the fields of behavioral sciences and visualization development of human-computer user interfaces.

It was unclear why the visual-modality training condition with learning rehearsal produced such low performance accuracy in the present study. One possibility was that the randomly selected cohort for this condition comprised mainly of international students whose English was not their native language. Another possibility might simply be due to learning fatigue and boredom as the participants were instructed to learn the architectural concepts twice. The relatively low performance accuracy rate for these visual learners might also be resulted from motion sickness of interactively immersing their physical presence in 3D virtual environments. We planned to rerun the visual condition with a different set of cohort to fully examine whether or not the low performance accuracy was a result of systematic and/or random testing errors.

Despite the low accurate rate for the visual condition, findings overall suggested the important role multimodal learning (i.e., audiovisual-modality) and training repetition together plays in
learning acquisition. The process of simultaneously exposing learners to interactive dynamic visualizations and prompting them to attend to information through the pragmatic use of audio cues reduced memory load, and in turn facilitated memory recall. The possibility of both instructors and students to utilize digital media in which avatars interact in immerse virtual reality to facilitate experiential learning across interdisciplinary fields is promising.

Future studies should also examine the role domain-specific knowledge play on learning acquisition. Inconsistent evidence has been reported regarding the effect of audiovisual perception in the design of dynamic visualizations with respect to information specificity. For example, perceptually integrated dynamic visualizations facilitated learning performance only when presented information was domain-specific, i.e., interactive learning environments were designed to complement the individual learners’ cognitive processes, such as their different levels of knowledge expertise.\textsuperscript{28, 29} Without taking into account the learner’s knowledge base, dynamic visualizations in audiovisual modality are expected to attenuate learning compared to visual or audio alone.\textsuperscript{30} The incorporation of domain knowledge into the design of dynamic visualizations has been largely overlooked. If not properly designed, interactive dynamic visualizations may result in a learning decrement due to memory overload and redundant information.\textsuperscript{31-33}

Our visualization and simulation laboratory is currently exploring the extent to which domain-specific knowledge impact learning performance in web-based distance education. The purpose is to implement both the theoretical applications and pragmatic use of 3D interactive, dynamic visualizations on instructional curriculum. It is imperative to develop 3D visualizations effective for individual PC/MAC computers using a mouse and/or a joy stick as vehicles for navigation in virtual reality. Additionally, though stereoscopic 3D displays are available to the general public, 2D monitors are currently the standard and will likely continue to be used for the immediate future and relevant comparisons between 2D & 3D display should be factored into future work.

Acknowledgements

This research was supported in parts by the Purdue Research Foundation (PRF) and Northwest Indiana Computational Grid (NWICG) Project at Purdue University Calumet.

The authors wish to express their gratitude to the faculty members in the Department of Behavioral Sciences, the CIVS undergraduate/graduate research assistants, and many others who have contributed to our research project.

Bibliography


of multimodal visualizations, domain-specific knowledge, and rehearsal process on learning acquisition.


