Evaluating the Effectiveness of Virtual Reality Learning in a Mining Context

Lauren Bennett; Phillip Stothard; E. James Kehoe

University of New South Wales, Sydney, Australia
Phillip Stothdard <s3006583@unsw.edu.au>

Abstract. UNSW's Schools of Mining Engineering and Psychology have developed training modules for working at heights in above-ground mines. These modules implement best-available, evidence-based instructional methods combined with a range of immersion. The present paper describes a controlled evaluation of this approach for training novices in the safe operating procedure for a basic maintenance task. All participants received a sequence of instructions using a large-screen, computer-driven visual display accompanied by audio narration in one of three modes: (1) an animated depiction of the target procedure for which the pace of instruction was controlled by the individual participant (Animated + Individual, AI), (2) the same animated depiction but presented to a group with the pace controlled by the trainer (Animated + Group, AG), and (3) a sequence of static slide images presented to a group with trainer pacing (Static + Group, SG). During the training, the participants' active processing of the information was encouraged by preceding each step of the instruction with a challenge question and feedback. Immediately following the module, the participants were given a multiple-choice test, which was repeated after a one-week retention interval. Across all three modes of presentation, the module yielded a high level of acquisition and retention. Among the three modes of presentation, the AI mode produced the highest level of test performance relative to both the AG and SG modes. When the participants were surveyed regarding their immersion in the virtual environment, they generally reported a moderate level of "presence," with the animations (AI, AG) producing higher levels than the static images (SG). These positive outcomes provide a foundation for the further development and testing of additional modules combined with different levels of immersion aimed ultimately at economically producing personnel who can safely and proficiently apply their knowledge and skills in real mines.

1. INTRODUCTION

As is evident in video games and in movies like "Avatar," the development of digital graphics has dramatically increased the ease of producing high-fidelity animations. For instructional purposes, such animations offer an engaging alternative to traditional lectures and demonstrations, especially when combined with the ability of the learner to control the virtual environment created by digital graphics. This "virtual reality" (VR) technology may promote superior learning by fully engaging a learner's cognitive processes through instilling a strong sense of "presence" within the computer-generated environment (Dalgarno, Hedberg, & Harper, 2002; Persky, et al., 2009; Witmer & Singer, 1998). Furthermore, in training for hazardous work places, VR has, like other simulations, the ability to promote learning in safe, well-controlled circumstances in preparation for the real work environment (Stothard, Squelch, et al., 2008; Stothard, Mitra, & Kovalev, 2008).

Nothwithstanding its appeal and technical sophistication, there is no guarantee that VR will be more effective than more traditional lecture-style training using static images on slides to depict the work environment. Although the cost of animation is declining, animation is still a more expensive proposition than taking static photographs. Furthermore, like any technology, VR's effectiveness will depend on how well it is adapted to the capabilities of its human users. For application to training, VR's effectiveness will especially depend on whether its implementation is appropriate to the trainees' cognitive architecture, particularly the features of their two memory stores: working memory and long-term memory.

Working memory filters, organises, and temporarily holds incoming information. Working memory's capacity is very limited. At any one moment, its static holding capacity is seven elements (e.g., holding a phone number), but its dynamic capacity for manipulating elements (e.g., adding digits in mental arithmetic) is limited to three or four elements (Cowan, 2001; Halford, Wilson, & Phillips, 1998). The operations for the manipulations appear themselves to occupy some of working memory's capacity. Thus, working memory can be easily overwhelmed in a dynamic, complex environment, and information can be lost before it is transferred to more permanent storage in long-term memory.

Long-term memory has a huge capacity, but it is not simply a storage closet. Rather, long-term memory contains organised structures – known variously as schemas or mental models – that compile encodings of information obtained in successive, related episodes (Owens & Sweller, 2008; Schnotz & Rasch, 2005; Sweller, van Merrienboer, & Paas, 1998). These compilations can be used either as single elements in working memory or as automated procedures that largely bypass working memory. Thus, as experience is gained and organized, complex information is processed in an increasingly efficient, condensed form.

These features of memory suggest that learning will be most rapid if the training is structured to avoid overburdening working memory and facilitate the formation of useful schemas. With respect to VR-based training methods, there is considerable debate concerning how readily they accommodate these features of memory.

On one side, some theorists contend that the immerse nature of VR will reduce the load on working memory by, among other things, integrating the visual inputs into a coherent single scene and integrating the visual scene with auditory narration (Dalgarno, et al., 2002;

Mayer & Moreno, 2002; Moreno & Mayer, 2000). Additionally, the learner's control of the pace and path through the virtual environment accords with claims that people learn best in minimally-guided environments, in which they are active explorers rather than passive recipients as in instructor-led settings (Christou & Bülthoff, 1999; James, et al., 2002).

On the other side of the argument, immersive and self-guided environments may impede learning, especially for novices who are not familiar with either the learning content or its context. The working memory capacity of novices may be readily overloaded by the abundance of input that is extraneous to the instructional goal and even seductively distracting. In any learning environment, extraneous and distracting information interferes with the processing of the core material (Moreno & Mayer, 2000). Furthermore, Norman (1993) states that dynamic and continually present immersive environments tend to produce event-driven processing. As such, working memory capacity available for organising the core material into long-term memory structures is diminished. With respect to learner control, recent research has revealed that guiding novices through a learning environment helps maintain their orientation to the core material and reduces distraction from extraneous features (Chen, Toh, & Ismail, 2005). The value of self-guidance only begins to appear when learners gain some initial knowledge and skill in the target task (Kirschner, Sweller, & Clark, 2006).

In light of these considerations, the present study evaluated the absolute and relative effectiveness of three versions of one module in a training program developed by UNSW's School of Mining Engineering for working at heights at an above-ground mine. The study had three aims: (1) to determine whether an animated module which the trainee could individually control the speed of instruction (Animation + Individual, AI) would promote greater or lesser acquisition and retention by novices to mining compared to the same animated module presented to a group for whom the instructional pace was controlled by the trainer (Animation + Group, AG; Mayer, 2008); (2) to determine whether the animation in the group setting produced greater or less acquisition and retention relative to the same group presentation illustrated by a series of static slides of the same scenes used in the animated versions (Static + Group, SG), (3) to determine whether the subjective presence in the instructional environment would be enhanced by animation and individual control.

Novices to mining were chosen as the participants for this experiment in order to equate their initial level of relevant knowledge and skill and also to expose as far as possible the differential effects of the alternate methods of presentation independently of any contribution from prior experience in mining. At the same time, the module used in the evaluation did not require technical expertise in mining but concerned the regulations and procedures for the safe replacement of a light globe on an elevated platform.

2. METHOD

To ensure that the participants (N=60, 59% female) were indeed novices to mining, they were all recruited from a first-year university psychology course. None reported being a mining student or former miner.

2.1 Apparatus

The system used for training consisted of a projector, flat screen, PC, and stereo speakers. Images were projected onto an area of 2.5 m x 1.6 m at a resolution of 1280x1024 pixels, and the images were viewed at distance of approximately 3 m. The PC was equipped with an Intel Core 2 Duo running at 2.4 GHz, 2GB of RAM, a NVidia GeForce 7900 GT graphics card, and an inbuilt sound card. Depending on the training condition, a joystick was added to navigate the virtual world. In addition, the joystick's buttons could be used to select and manipulate objects, answer questions, and focus on points of interest in the virtual world.

2.2 Training Content

The module used in this experimental evaluation was selected from a program developed by the School of UNSW Mining Engineering for training miners at an above-ground mine. The specific module concerned the procedure for safely changing a light globe on an elevated platform. The module stepped through the considerations and procedures as if the participants were taken to the mining site to receive a briefing and demonstration. This experiment targeted 17 pieces of knowledge, covering, among other things, required permits, types of equipment to be used, and use of ladders.

2.3 Procedure

After providing consent and demographic information, participants were randomly assigned to one of three training conditions, which were designated as Animation–Individual (AI), Animation–Group (AG), and Static–Group (SG) (*n*s = 20).

In Group AI, each participant was trained individually. After receiving general oral instructions from the experimenter to orient the participant to the apparatus and the training task, the participant was allowed to use the joystick to navigate at their own pace around the animated virtual mine site. Arrows on the image guided the participants along the desired path from point to point for each step that contained a visual demonstration accompanied by a recorded oral narration. After each step, a short written text summarising the key learning point was superimposed on the screen. Depending on the length of the narration, each step was 2-4 min, and the text was displayed until the participant advanced the program to the next step.

Following each step in the instruction, the participants were presented with a relevant multiple-choice question intended to encourage the organisation and storage of the information in long-term memory schemas. Each question was first presented with the four possible answers, which were visually displayed and read aloud by the experimenter. The participants were allowed 3-5 seconds for this task. At the end of this interval, the correct answer was highlighted on the screen. On average, participants took 15 min to complete the module.

In Group AG, participants were trained in clusters of three to five individuals seated together. The researcher acting as trainer controlled the joystick and stepped the participants through the same training module as Group AI. The questions following each step were answered as by a volunteer from the group, as is often the case in class-room instruction. Steps were not repeated. The complete module required 15 min

In Group SG, participants were also trained in clusters of three to five individuals in exactly the same was as Group AG. In Group SG, however, the images were presented as a series of static slides while the narration was read out.

Following the training and a 10-min break, the participants were tested using a 17-item multiple-choice test administered on a PC using a 19-in monitor.

Immediately following this test, the participants were asked to complete a questionnaire concerning their feelings of "presence" in the instructional environment. The Presence Questionnaire contained 33 items, for which the participants used a seven-point scale to rate their perceived control over the instructional environment, the engagement of their visual and auditory senses, their concentration on the instructional environment, and the level of realism in the instructional environment.

One week later, to measure retention, the participants were asked to retake the test via email. However, the order of the multiple-choice options for each question was rearranged. All the participants completed the second test.

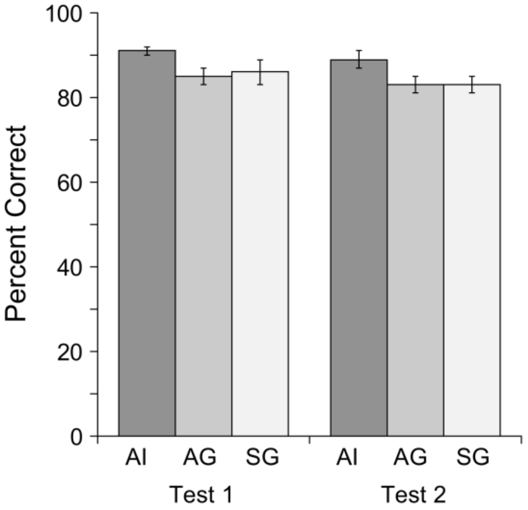

Figure 1: Mean percent correct (± 1 standard error).

3. RESULTS

Figure 1 shows the mean percentage of correct answers in the two tests. Inspection of Figure 1 reveals that the novice participants acquired and retained considerable knowledge of the procedures. Across all three groups and both tests, the scores averaged 86% (SD = 9%). Despite the generally high levels of performance, Group AI showed a significantly higher level than either Group AG or Group SG, smaller $F(1, 57) = 4.25$ $p < .05$, $\eta^2p = 0.069$. There was also a small, but reliable decline over the retention interval from Test 1 to Test 2, $F(1, 57) = 7.31$ $p < .01$, $\eta^2p = 0.114$.

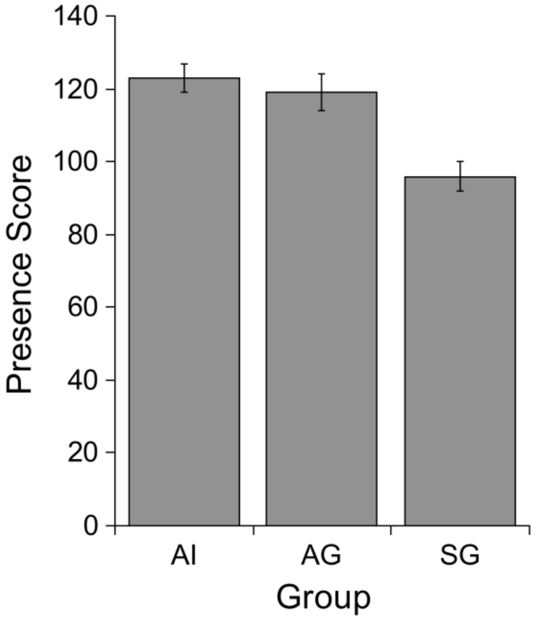

Figure 2: Mean presence score (± 1 standard error) equals sum of ratings for control, sensory engagement, and realism.

Figure 2 displays the means and standard errors ratings for feelings of presence in the instructional environment. Preliminary analyses revealed that the ratings concerning control, sensory engagement, and realism of the simulation were well correlated and comprised a single coherent scale, Cronbach's alpha = .794. Thus, the presence score for each participant was computed by summing together their ratings for those three dimensions (maximum total score = 196). (The fourth set of ratings concerning concentration on the instructional environment versus external distractions was not correlated with the other ratings.)

In relation to the maximum possible score, the participants generally reported ratings clustered above the middle of the scale. The groups that received the animated presentation (AI, AG) showed a significantly higher level of presence than the group that received the static images (SG), $F(1, 57) = 21.74$, $p < .01$, $\eta^2p = .280$. The tiny difference in presence between Groups AI and AG did not even approach statistical significance, $F<1$.

Although both test performance and presence ratings differed across the training conditions, there were no significant correlations between these dependent variables either across all participants or within each group, all *p*s > .05.

4.0 DISCUSSION

In summary, the animated presentation with individual control of the pace (AI) yielded the highest levels of acquisition, retention, and presence. Relative to AI, animation conducted in a group with trainer control of the pace (AG) yielded lower levels of acquisition and retention, but a similar level of presence. The static slide presentation conducted in a group (SG) showed acquisition and retention similar to that of AG but with less presence.

As is readily apparent, the absolute differences among the groups were all relatively small. With respect to the test scores, all three groups achieved high levels of performance that were retained with only a slight loss for a week. These high levels of acquisition and retention were achieved, because the training module was not an artificial laboratory device for testing the relative virtues of animation and learner control. Rather, the module used the best available evidence-based instructional methods to optimize cognitive load and schema formation (Mayer, 2008). The main methods were:

First, the sequence of instruction was segmented into discrete steps containing a single clear learning point within a minimum of extraneous explanation so that, at any one moment, working memory would not be divided between multiple elements (Pollock, Chandler, & Sweller, 2002).

Second, visual input and audio narration were closely coordinated, which both allows learners to use their preferred channel for reception and produces a small, but useful, increase in working memory capacity (Mayer, 2008).

Third, the visual text summary was presented separately from the visual images to prevent an overload of the visual channel (Mayer, 2008)

Fourth, as suitable for novices, there was considerable guidance through the module. Even in the AI version, the individual control was largely limited to self-pacing (Mayer, 2008). Directional arrows were inserted into the animation to encourage the participants in Group AI to follow the same path as the other two groups.

Fifth, active attention and encoding of the training content into long-term memory was encouraged by the inclusion of questions at each step of the instruction (Sweller, et al., 1998).

As a consequence, immediate learning and retention were robust, whether or not animation and individual pacing were available. With these positive results in hand, the way is clear to test these methods using (1) more technically-demanding content, (2) real trainee miners, and (3) tests of transfer from desk-top training to hands-on application in simulated and ultimately real mines. In the case of transfer, the positive value of animation and/or self-pacing may be amplified (Zhang, Stothard, & Kehoe, 2010).

With respect to the immersive value of the training methods, the feelings of presence reported by the participants were generally moderate and not discernibly correlated with the level of acquisition and transfer. Although the animation raised the sense of presence in the virtual environment, all three groups received their training using a 2-D screen in a classroom setting allowing little or no interaction with the virtual environment. Hence, there is room for increases in both the level of immersion provided by, say, a 3-D surround display, and/or amount of interaction with the virtual environment, by enabling the trainee to perform the procedures being taught.

In conclusion, the present results provide a proof-of-concept for the approach taken here to the development of training for the mining industry. By using state-of-the-art, evidence-based instructional methods, a high level of acquisition and retention of knowledge of the target content was achieved based on a single, 15-min training session. Moreover, this outcome was robust across variations in the type of presentation, that is, animated versus static images, learner versus trainer pacing. Thus, there is a solid foundation for further development and testing of the instructional methods used in this study in combination with different levels of immersion with the ultimate aim of producing personnel who can safely and proficiently apply their knowledge and skills in real mines at the greatest possible ratio of benefits to costs in their training.

5.0 INTERIM RECOMMENDATIONS

The present findings combined with increasingly well-established principles of instructional design lead to the following interim recommendations:

5.1 Animated demonstrations should be used to introduce safe operating procedures, even to complete novices. Such demonstrations can provide a consistency in training across sites and instructors, thus helping to ensure compliance with quality standards at modest cost.

5.2 When first introducing a novel procedure, break down the sequence into discrete recipe-like steps. This detailed sequencing will also aid in quality assurance through consistency in process. Do not assume that the learners will fill in missing steps, even if they seem obvious and logical to a proficient person.

5.3 Self-paced, rather than instructor-paced, progression through a new procedure is preferable. However, with attention to the engagement of the learner, instructor-paced progression in group training can be effective.

5.4 Ensure that the learner is engaged at each step. Leading questions, like those used in this study, can focus the learner's attention on the key features of the step. General admonitions to pay attention will not ensure a proper focus.

REFERENCES

Chen, C., Toh, S., & Ismail, W. (2005). Are learning styles relevant to virtual reality? *Journal of Research on Technology in Education, 38*, 123–140.

Christou, C., & Bülthoff, H. H. (1999). View dependence in scene recognition after active learning. *Memory and Cognition, 27*, 996–1007.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87-185.

Dalgarno, B., Hedberg, J., & Harper, B. (2002). The contribution of 3D environments to conceptual understanding. In A.Williamson, C. Gunn, A. Young & T. Clear (Eds.), *Winds of change in the sea of learning: Charting the course of digital education. Proceedings of the 19th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education* (pp. 149-158. Retrieved 148 August 2009 from http://www.ascilite.org.au/ conferences/auckland2002/proceedings/papers/ 2051.pdf). Auckland, NZ: UNITEC Institute of Technology.

Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences, 21*, 803-864.

James, K. H., Humphrey, G. K., Vilis, T., Corrie, B., Baddour, R., & Goodale, M. A. (2002). Active and passive learning of three-dimensional object structure within an immersive virtual reality environment. *Behavior Research Methods, Instruments, & Computers, 34*, 383-390.

Kirschner, P. A., Sweller, J., & Clark, R. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential and inquiry-based teaching. *Educational Psychologist 41*, 75–86.

Mayer, R. E. (2008). Applying the science of learning: Evidence-based principles for the design of multimedia instruction. *American Psychologist, 63*, 760-769.

Mayer, R. E., & Moreno, R. (2002). Animation as an aid to multimedia learning. *Educational Psychology Review, 14*(1), 87-99.

Moreno, R., & Mayer, R. E. (2000). A coherence effect in multimedia learning: The case of minimizing irrelevant sounds in the design of multimedia instructional messages. *Journal of Educational Psychology, 92*, 117-125.

Norman, D. A. (1993). *Things that make us smart: Defending human attributes in the age of the machine*. Cambridge, MA: Perseus Books

Owens, P., & Sweller, J. (2008). Cognitive load theory and music instruction. *Educational Psychology, 28*, 1-29.

Persky, S., Kaphingst, K. A., McCall, C., Lachance, C., Beall, A. C., & Blascovich, J. (2009). Presence relates to distinct outcomes in two virtual environments employing different learning modalities. *CyberPsychology & Behavior, 12*(3), 263-268.

Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and Instruction, 12*, 61-86.

Schnotz, W., & Rasch, T. (2005). Enabling, facilitating, and inhibiting effects of animations in multimedia learning: Why reduction of cognitive load can have negative results on learning. *Educational Technology Research and Development, 53*(3), 47-58.

Stothard, P., Squelch, A., van Wyk, E., Schofield, D., Fowle, K., Caris, C., et al. (2008). *Taxonomy of interactive computer-based visualisation systems and content for the mining industry – Part 1*. Paper presented at the First International Future Mining Conference.

Stothard, P. M., Mitra, R., & Kovalev, A. (2008). *Assessing levels of immersive tendency and presence experienced by mine workers in interactive training simulators developed for the coal mining industry.* Paper presented at the SimTecT 2008, Melbourne, Victoria, Australia.

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251-296.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments, 7*, 225-240.

Zhang, S. Q., Stothard, P., & Kehoe, E. J. (2010). *Evaluation of underground virtual environment training: Is a mining simulation or conventional power point more effective?* Paper presented at SimTecT 2010.