Joint Virtual Reality Conference of EGVE - ICAT - EuroVR (2009)
M. Hirose, D. Schmalstieg, C. A. Wingrave, and K. Nishimura (Editors)

Higher Levels of Immersion Improve Procedure Memorization Performance

D. A. Bowman, A. Sowndararajan, E. D. Ragan, R. Kopper

Virginia Center for Human-Computer Interaction and Dept. of Computer Science, Virginia Tech

ABSTRACT

Researchers have proposed that immersion could have advantages for tasks involving abstract mental activities, such as conceptual learning; however, there are few empirical results that support this idea. We hypothesized that higher levels of immersion would benefit such tasks if the mental activity can be mapped to objects or locations in a 3D environment. To investigate this hypothesis, we performed an experiment in which participants memorized procedures in a virtual environment and then attempted to recall those procedures. We aimed to understand the effects of three components of immersion on performance. Results demonstrate that a matched software field of view (SFOV), a higher physical FOV, and a higher field of regard (FOR) all contributed to more effective memorization. The best performance was achieved with a matched SFOV and either a high FOV or a high FOR, or both. In addition, our experiment demonstrated that memorization in a virtual environment could be transferred to the real world. The results suggest that, for procedure memorization tasks, increasing the level of immersion even to moderate levels, such as those found in head-mounted displays (HMDs) and display walls, can improve performance significantly compared to lower levels of immersion.

Categories and Subject Descriptors: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—*artificial, augmented, and virtual realities*; H.5.2 [Information Interfaces]: User Interfaces—*evaluation/methodology*; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—*virtual reality*.

1. Introduction

Virtual reality (VR) technologies have been used successfully for a variety of applications to facilitate learning of real-world activities and procedures. Such applications, including many of those used for vehicular operation training, military simulations, and medical operations training, often employ immersive VR systems in which the virtual environment (VE) appears to surround the user in space. Applications in these domains take advantage of the physical, "whole-body" interactions provided by such systems. For example, flight simulators make use of a real physical cockpit so that pilots-in-training can use the actual controls to fly the simulated airplane [BroF99]. Similarly, laparoscopic surgery simulators use high-fidelity haptic devices to help physicians learn the necessary motor skills before operating on a real patients [BBSJ07].

Other types of applications take advantage of immersive VR's higher-quality and more realistic spatial cues (e.g., stereoscopy, motion parallax), which makes it possible to provide users with higher levels of spatial understanding than could be achieved using traditional displays. For instance, vehicle designers have long used immersive VR systems to better understand their designs before they are built [BroF99]. Scientists use immersive technologies to visualize complex 3D structures and data sets [vDFL*00]. Engineers plan underground features, such as oil wells, using immersive VR [LLG*07].

While the reasons for the success of these two sets of VR applications are understood, other proposed applications, such as educational applications, do not fit within these categories. Educational VR systems have been developed for the purpose of helping students to learn conceptual information and principles. For example, researchers have prototyped immersive VR systems for mathematics education [KSW00, ROS06] and for learning complex principles of physics [DSL96]. We can characterize these applications as interactive visualizations for the purpose of conceptual learning, in which abstract concepts or very large- or small-scale phenomena are mapped to human-scale visual representations. But it is not known if immersive VR technology is necessary or beneficial for such learning-based applications or if standard, non-immersive displays would work just as well.

This is a difficult problem to attack directly, particularly because measurement of conceptual learning is not well understood and is subject to many potential biases. Furthermore, different educational objectives are met with different pedagogical approaches, and it is not clearly

DOI: 10.2312/EGVE/JVRC09/121-128

understood what features of VR are beneficial for what educational purposes [DSL96, SDLC99]. Greater knowledge of what features of immersive VR support different levels of cognitive processing is needed to understand how to effectively design VR applications that are conducive to learning activities.

Rather than attempt the unwieldy evaluation of conceptual learning directly, we use a memorization task as a more manageable example of a mental activity that still requires the transfer of information from a VE to a user. As knowledge and recollection of facts is considered to be a simple, foundational learning process [BKM56, Kra02], supporting such activities can reinforce the deeper levels of learning that are desired in educational applications. We have conducted a study of the use of VR technology for a procedure memorization activity. In this task, a user in a VE is shown a procedure involving several steps/actions and is asked to rehearse and memorize the procedure. This task requires perception and interpretation of abstract information; thus it is a simple approximation of conceptual learning.

We performed this study within a theoretical framework centered on the concept of *immersion*. Following Slater [Sla03], we define immersion as “the objective level of fidelity of the sensory stimuli produced by a VR system.” In other words, immersion depends only on the technology used to produce the VE, and is not necessarily related to the user’s experience of the VE (the sense of *presence*). Immersion can be modified, controlled and used as an independent variable for empirical studies. With this definition, we can speak of *levels* of immersion, rather than using terms such as “non-immersive” and “immersive” VR. Furthermore, we note that the overall level of immersion is made up of many components, such as field of view, resolution, and stereoscopy [BM07].

We hypothesized that higher levels of immersion would lead to better performance on the procedure memorization task if the procedure could be mapped to spatial locations in the VE; that is, if the VE could be used as a cognitive aid during learning and recall. Our research supports this hypothesis, showing that particular components of immersion (or combinations of components) are particularly beneficial. This is a first step in demonstrating the benefits of immersion for abstract mental activities such as conceptual learning, and in determining which VR technologies should be used for such applications.

2. Related Work

Many researchers have explored procedural training and conceptual learning applications of VR. For example, several projects have explored whether users can learn a procedure through interaction with a virtual agent in a VE [JDR*05, JR97, PHM*03]. Others have hypothesized that content will be more memorable if students experience it firsthand in an immersive VE [AH00, SDLC99]. In some cases, researchers have attempted to measure the effectiveness of the training/learning [BWHA99, JMOL01, SDLC99], but it has proven difficult to quantify the effects of VR systems on learning.

Despite the many educational applications that take advantage of VR technology, few projects have attempted to formally evaluate the benefits. One such project, ScienceSpace [DSL96], studied the benefits of groups of features of VR for three specific applications. While this was an important step in evaluation, it was not possible to determine the values of the individual components of immersive technologies.

Several VR researchers have performed controlled studies investigating the effects of one or two components of immersion on user performance in other tasks [e.g., Art00, BHB97]. In our study, we use a memorization activity as an approximation of conceptual learning and vary three components of immersion independently to determine their effects on performance.

In a study related to the memorization of object information, Mania, Robinson, and Brandt [MRB05] found evidence that object recognition was significantly better with higher rendering quality. While several researchers have investigated the effects of various components of immersive VR [AHC97, MC04], as well as interaction techniques [e.g., BroB99], on memorization of spatial layouts of objects, these studies focused on the effects on memorization of spatial location rather than additional information.

Placing greater emphasis on learning new information that is not bound to the specifics of the VE, our previous research [SWB08] found that users performed significantly better in a procedural memorization task when they used a more immersive VE. The experiment compared a laptop display (low immersion) to a large two-wall projection display (high immersion). Users were shown a medical treatment procedure consisting of multiple steps and asked to view, rehearse, and memorize the procedure before recalling it in the VE. Such a mental activity is a simplified version of conceptual processing involving perception and interpretation, but not necessarily understanding.

Higher levels of immersion are known to provide stronger spatial cues that can help improve spatial understanding and spatial memory [SB07, WM05]. While our previous study in procedure memorization [SWB08] did not evaluate participant strategy, we hypothesized that better performance resulted from an increased ability to use a spatial memorization strategy in the high-immersion condition, since it is still unknown what components of immersion effectively improve procedure memorization, whether higher levels of immersion also improve memorization of more abstract, non-spatial procedures, and whether such learning transfers to the real world. We addressed these questions with the experiment presented in the following session.

3. Experiment

We conducted a controlled study to further investigate the effects of immersion on procedure memorization, and to determine which components of immersion were responsible for any effects observed in [SWB08].

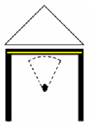
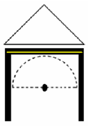
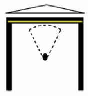
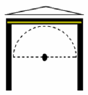
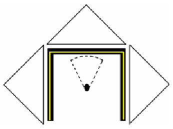
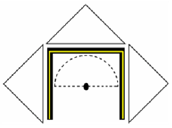
	Matched SFOV		Unmatched SFOV	
	Low FOV	High FOV	Low FOV	High FOV
Low FOR				
High FOR				

Table 1: *Levels of immersion tested in experiment II. The figures are top-down views of the CAVE display. The yellow highlighted sides represent screens that were turned on in each condition (to control FOR). The dotted arcs in the center represent the user's physical FOV. The solid triangles represent the SFOV of each virtual camera.*

3.1 Hypotheses

Our overall hypothesis was that a learning environment with a higher level of immersion would produce better performance in the procedure memorization task. The rationale for this hypothesis comes from the enhanced spatial cues provided by higher levels of immersion—cues resulting from display characteristics such as high field of view (FOV), allowing the user to see more of the environment at any one time, and high field of regard (FOR), allowing the user to make use of natural head and body movements to view other parts of the environment.

We know that these enhanced spatial cues in higher levels of immersion can lead to improved spatial understanding and spatial memory [SB07, WM05]. Our hypothesis was motivated by the idea that *spatial memory* could be used as a substitute for *procedural memory* during memorization of a procedure in a VE. In other words, if the steps of the procedure can be mapped to objects or spatial locations, the learner can associate the procedure with these locations in spatial memory. The steps can then be recalled by referencing the spatial locations. Thus, the VE acts as a cognitive aid for the learner. It follows, then, that a learning environment with better spatial cues (higher level of immersion) should result in better recall performance than a learning environment with impoverished spatial cues (lower level of immersion).

This idea is not new; in fact, it has been used as a memorization technique since classical times. In the "method of loci," one memorizes a speech, story, or list by associating each element with a physical or imagined location in a large space, and rehearsing the items while physically or mentally walking through this space [Eri03, Yat01]. Our contribution to this idea is to use a VE as a replacement for the physical or imagined space used in the classical method.

From our previous study [SWB08] alone, it was not possible to deduce which component(s) of immersion resulted in the observed difference between the two conditions, which differed in at least the following ways:

- Field of view (FOV; the angular area in the physical world within which the user can see the virtual world at any instant in time).
- Software FOV (SFOV; the angular area in the virtual world that the user can see at any instant in time, or the FOV of the virtual camera).
- Field of regard (FOR; the angular area surrounding the user within which the virtual world is displayed).

In our study, we investigate which of these components had a positive effect on procedure memorization. We hypothesized that an SFOV that is matched to the physical world, in combination with a high FOV and a high FOR, would result in a high level of spatial understanding, and thus facilitating memorization. Further, unlike the medical procedure used in [SWB08], which was concrete and easily mapped to a virtual world, our experiment was designed to use a more abstract procedure type. We also tested the transfer of learning from a VE to the real world.

3.2 Experimental Design

We varied FOV, SFOV, and FOR as between-subjects independent variables in this experiment. FOV had two levels: low (60 degrees), and high (nearly 180 degrees). We used physical blinders on goggles to restrict the FOV in the low FOV conditions, and non-blinded glasses for the high FOV conditions. SFOV had two levels: matched (virtual camera has the same FOV as the user, 90 degrees for each screen), and unmatched (virtual camera has an FOV of 135 degrees for a screen). FOR also had two levels: low (90 degrees, using one projection screen), and high (270 degrees, using three projection screens surrounding the user).

Overall, then, there were eight possible between-subjects conditions. However, we did not test conditions with an unmatched SFOV and a high FOR, as this would result in severe distortions across the three screens. This left us with six between-subjects conditions (Table 1).

In practical terms, the components that we varied enabled us to simulate the conditions of widely-used VR systems, such as the CAVE (high FOR, high FOV, matched SFOV), large-size display walls (low FOR, high FOV, matched SFOV) HMDs (high FOR, low FOV, matched FOV), and desktop displays (low FOR, low FOV, unmatched SFOV). Thus, our results can guide the choice of display system to use for educational applications that involve procedure memorization.

We also had two within-subjects independent variables: assessment environment (AE) and object consistency (OC). AE refers to the setting in which the assessment (or recall) phase was performed: in the virtual world, or in the physical world. In the virtual world conditions, the highest level of immersion (high FOV, matched SFOV, high FOR) was always used for assessment. AE was used to determine whether learning transferred to the real world. OC refers to the spatial location of the objects in the environment during the assessment: objects could be in the same or different locations. OC was used to test whether participants relied on the exact spatial location of the objects for recall, or whether they could remember the procedure accurately even when the objects were moved. We hypothesized that neither AE nor OC would have a significant effect on results, because we believed that participants' recall would be based on their memory of the VE in which they learned a procedure, rather than on their surroundings during recall.

Thus, there were four within-subjects conditions. Each participant was placed in one of the six between-subjects groups and performed one trial in each of the four within-subjects conditions. The four trials required participants to learn different procedures, but each procedure had an equivalent level of complexity; thus we do not consider procedure as an independent variable.

As in [SWB08], the dependent variables were the time to complete the assessment phase and the number of errors in the assessment phase. An error was counted every time the participant specified the next step of the procedure incorrectly, up to a maximum of ten errors per step. After ten errors on a step, the experimenter provided that step of the procedure to the participant.

3.3 Participants

Forty-one voluntary, unpaid participants (twenty-five males) took part in the experiment. The mean age was 22. Eight of them had used immersive VR previously, while 18 had video game experience. Each participant was screened with an initial memory test (described below) in the real world; five participants scored below a predefined threshold on this test and did not complete the remainder of the experiment. The remaining 36 participants were assigned to the six between-subjects groups so that each group had six participants and the groups had approximately equal average scores on the initial memory test.

3.4 Experimental Procedure

Before beginning the experiment, participants gave their informed consent and answered a demographic questionnaire. They then performed the initial memory test, which involved the memorization of an eight-step procedure (similar to those used in the main experiment), with both learning and assessment taking place in a real-world setting.

As in [SWB08], the participant's task was to memorize a multi-step procedure. Each step consisted of an object, a source location, and a destination. For example, a step might be: "move the yellow sphere from the left table to position number 6" (the destination was denoted by a numbered position on a 4x4 grid (Figure 3), or by an object that was already located on the grid).

Each trial consisted of learning, practice, and assessment phases. In the *Learning phase*, the experimenter first identified the objects that would be used in the procedure, and then explained each step of the procedure in detail. The participant was taught to describe the steps of the procedure in specific terms. For example, a step might be described as: "move the red sphere from the center table to position number eight." While the experimenter described each step, that step was shown visually in the VE (the object would be moved automatically from its source location to its destination).

The *Practice phase* allowed participants to rehearse, with the experimenter's assistance, the procedure from the learning phase. In this phase we asked the participant to verbally describe the procedure, following the protocol from the learning phase. As the participant described each step correctly, that step was shown visually in the VE. If the participant made a mistake or could not recall the next step, the experimenter helped him/her to remember the correct step in the procedure.

In the *Assessment phase*, participants were asked to recall the entire procedure in the assigned assessment environment (real or virtual), without any assistance from the experimenter. When the participant provided the current step correctly, the experimenter showed the next step visually (automatically in the VE and manually in the real world), and the scenario moved on to the next step.

The learning and practice phases were always conducted in the virtual world, with the level of immersion determined by the participant's group. The assessment phase was conducted in either the virtual world or the real world, and with objects in the same or different locations as compared to the learning and practice phases, depending on the values of AE and OC for the trial in question.

In two practice trials, the procedures consisted of four steps, while in the four main trials the procedures consisted of eight steps. In the main trials, participants always encountered the four procedures in the same order, while the order of the four within-subjects conditions was counterbalanced using a Latin Square.

Before each trial, participants were informed of the values of AE and OC for that trial. Participants were asked to concentrate on the task of memorization and to refrain from asking questions during a trial. We allowed participants to rotate the virtual world around its vertical axis during the learning and practice phases, but no other virtual navigation was allowed. Rotation was not necessary in the assessment phase, since it always used either the VE with the highest level of immersion, or the real world, when all objects were visible without rotation.

Figure 1: *Real room used for initial memory test and real-world assessment conditions*

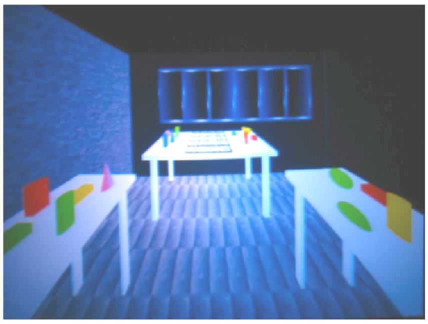


Figure 2: *Virtual room (shown here with distorted SFOV)*

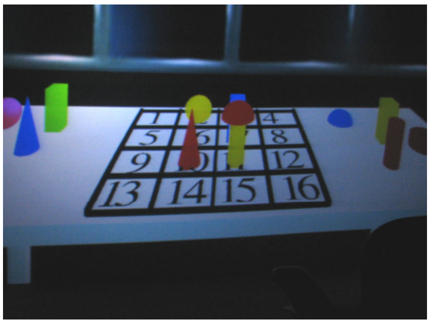


Figure 3: *Target square in the virtual environment*

3.5 Apparatus

We used a three-screen (front wall and two side walls) Fakespace CAVE™ to implement all six between-subjects conditions. Each screen was 10’ wide and 9’ high. Screens were rear-projected, using 1280x1024 Electrohome CRT projectors. In the high FOR conditions, all three screens were used, while the low FOR condition used only the front screen. Physical blinders attached to lab goggles were used to restrict the FOV to 60 degrees in the low FOV conditions; participants wore the goggles without blinders in the high FOV conditions. Participants held an Intersense IS-900 wand, using the analog joystick to rotate the virtual world around its vertical axis. We did not track the participants’ head or hands (since participants were stationary and could still look to the left or right in the high FOR condition without head tracking, and since no direct interaction with the environment was needed), nor did we use stereoscopic graphics (the stereo glasses would have limited the range of physical FOV we could test). The environment was rendered using DIVERSE [KSA*03].

Participants were seated on a chair in the center of the CAVE, and we varied the height of the chair so that each participant’s head was at the same level.

The virtual world was modeled to look like a real room that was adjacent to the CAVE (Figures 1 and 2). Both the virtual and real environments contained three tables holding a variety of geometric objects. There were seven different object shapes and six different object colors. The front table also held a white 4x4 grid with numbered squares to serve as the target area (Figure 3).

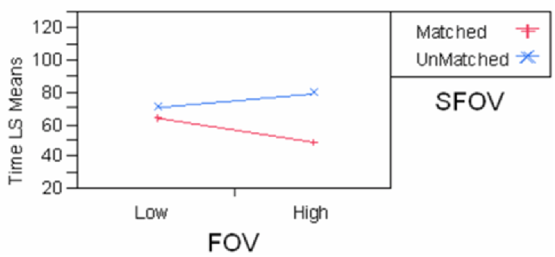


Figure 4: *Interaction between SFOV and FOV for time. The combination of matched SFOV with high FOV was significantly faster than the other three combinations.*

3.6 Results

We performed a mixed analysis of variance (ANOVA) on both time and errors. Note that p-values for the error analysis are only approximate, since the number of errors is not a continuous variable or necessarily normally distributed. Even after consulting with statisticians, we were not able to identify any non-parametric tests that could do the analysis due to our complex experimental design. In cases like these, ANOVA is considered the best approximation. As we will show, our results for the error metric are nearly identical to the results for the time metric, increasing our confidence that our analysis of the error metric is reasonable. More importantly, our effect sizes are large relative to variability, and are clearly meaningful in this context.

As we hypothesized, the within-subjects factors had no significant effect. AE was neither significant for time ($F(1,103) = 0.037$, $p = 0.849$) nor errors ($F(1,103) = 0.862$, $p = 0.355$). Similarly, OC was neither significant for time ($F(1,103) = 0.228$, $p = 0.634$) nor errors ($F(1,103) = 0.364$, $p = 0.547$).

Table 2 shows the least squares means for time in all six between-subjects conditions, while Table 3 gives the same information for errors. We found main effects of SFOV for both time ($F(1,30) = 15.85$, $p < 0.001$) and errors ($F(1,30) = 123.81$, $p < 0.0001$). Matched SFOV resulted in less recall time (mean = 59.99s) and fewer errors (mean = 1.56) than unmatched SFOV (71.05s and 5.40 errors).

Our analysis also found main effects of FOR for both time (F(1,30) = 17.09, $p < 0.001$) and errors (F(1,30) = 13.35, $p < 0.001$). The high FOR conditions had lower assessment time (58.35s) and fewer errors (2.10) than the low FOR conditions (72.69s and 4.86 errors).

We did not find a main effect of FOV on time (F(1,30) = 0.589, $p = 0.449$), although high FOV conditions did have a faster average time (60.06s) than low FOV (70.99s). There was a main effect of FOV on errors (F(1,30) = 4.31, $p < 0.05$), with high FOV resulting in fewer errors (2.72) than low FOV (4.21).

We found a significant interaction between SFOV and FOV for the time metric (F(1,30) = 6.982, $p < 0.02$), shown in Figure 4. A post-hoc analysis using a Tukey HSD test showed that the combination of matched SFOV with high FOV resulted in significantly faster recall than the other three combinations of these two variables.

There was also a significant interaction between FOV and FOR for both time (F(1,30) = 6.24, $p < 0.02$) and errors (F(1,30) = 4.31, $p < 0.05$), as shown in Figures 5 and 6. Tukey HSD tests revealed that the combination of low FOV and low FOR was significantly slower and resulted in significantly more errors than the other three combinations of these variables.

We also ran post-hoc Tukey tests to compare the six between-subject conditions with one another. Table 2 shows the results for time, while Table 3 shows the results for errors. In both cases, the same three conditions formed a separate group with better performance than the other conditions; these conditions all had a matched SFOV and either a high FOR or a high FOV or both.

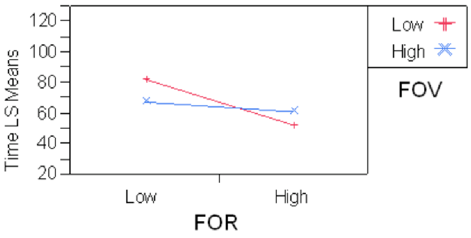


Figure 5: *Interaction between FOV and FOR for time. The combination of low FOV and low FOR was significantly slower than the other three combinations.*

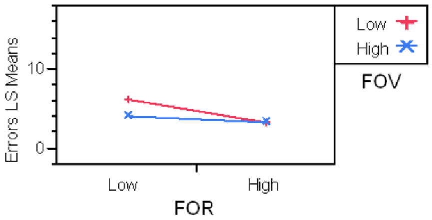


Figure 6: *Interaction between FOV and FOR for errors. The combination of low FOV and low FOR resulted in significantly more errors than the other three combinations.*

	Matched SFOV		Unmatched SFOV	
	Low FOV	High FOV	Low FOV	High FOV
Low FOR	78.76	52.73	84.66	81.91
High FOR	49.54	45.53		

Table 2: *Results for time, given as least square means. Times are in seconds. Conditions in shaded cells are significantly faster than other conditions.*

	Matched SFOV		Unmatched SFOV	
	Low FOV	High FOV	Low FOV	High FOV
Low FOR	3.46	1.00	8.67	6.96
High FOR	0.58	0.21		

Table 3: *Results for errors, given as least square means. Conditions in shaded cells are significantly more accurate than all other conditions.*

3.7 Discussion

As we hypothesized, assessment environment had no statistically significant effect on recall time or accuracy in our procedure memorization task. This means that the procedure, memorized in the virtual world, could be recalled in the real world just as quickly and with the same level of accuracy. Thus, learning transferred from the virtual to the real world. This has important implications for the use of immersive VR technologies for conceptual learning, since the learning would be useless unless it could be used outside the virtual world. Our experiment does not show, however, that learning in a virtual environment is as effective as learning in the real world; we leave this for future work. Regardless, the advantages of virtual environments (flexibility, control, the ability to display scenes not possible in the real world) would still make them attractive for learning applications.

We also found that object consistency (whether objects were in the same positions during assessment as they were in training and practice) had no effect on either recall time or accuracy. This means that participants did not rely (exclusively) on the presence of the same cognitive aid used during training/practice when recalling the memorized procedure. In other words, they appear to have been able to recall the procedure in its abstract form, rather than simply remembering the spatial locations or necessary movements of the objects. We hypothesize that participants were referencing their memories of the spatial locations of the objects in the VE in order to recall the colors and shapes of the objects and their final positions on the grid. Recalling the procedure in exactly the same environment did not increase performance; or, said another way, recalling the procedure in a jumbled version of that environment did not decrease performance. This result does not necessarily imply that performance would be similar if the recall took place in a sterile environment with no cognitive aids; this neeeds to be verified in a follow-up study.

Higher levels of immersion during learning and rehearsal (i.e., matched SFOV, high FOV, and high FOR) all significantly improved recall accuracy, while both matched SFOV and high FOR significantly reduced recall time.

These results also matched our hypotheses, with the one exception that high FOV did not reduce recall time significantly. As we described in section 3, we believe that these increases in participants' ability to memorize procedures relates to the richness and quality of the spatial cues provided by the VE. The higher levels of immersion provided richer and better spatial cues, leading to increased spatial understanding, and allowing participants to use a spatial memory strategy (similar to the method of loci) for memorizing the procedure. Our experiment does not prove this assertion, but our results are consistent with this overall hypothesis.

Furthermore, we saw that various combinations of components of immersion produced better results than others. The significant interaction between SFOV and FOV (Figure 4) showed that both a matched SFOV and a high (unrestricted) FOV were necessary to achieve a significant decrease in recall time. The significant interactions between FOV and FOR (Figures 5 and 6) reveal that performance was significantly reduced when both FOV and FOR were at low levels.

These findings were reinforced by the post-hoc comparison of all six levels of immersion (Tables 2 and 3). For both time and accuracy, the best conditions were those that had a matched SFOV and either a high FOV or a high FOR, or both. Matched SFOV seems, therefore, to be the most important component (among those we tested) for producing good recall performance, which partially explains the experiment results of [SWB08]. On the laptop screen, we were forced to use an unmatched SFOV to allow the user to see more of the virtual world. We claim that the unmatched SFOV, because it distorts the user's view of the environment, hinders accurate spatial understanding of the virtual scene, and therefore makes it more difficult to use a spatial memory strategy to memorize the procedure. But matched SFOV by itself was not enough to achieve the best performance in our experiment; either high FOV or high FOR were also needed to provide sufficient spatial cues.

Although the highest level of immersion (matched SFOV, high FOV, high FOR) produced the lowest recall times and highest accuracy rates in absolute terms, the post-hoc tests reveal that performance in this condition was not significantly different than two other conditions (the matched SFOV, high FOV, and low FOR condition with $p = 0.30$; and matched SFOV, low FOV, and high FOR condition with $p = 0.14$). With a greater number of subjects, the difference might have been significant, but this result still has important implications for real-world systems. In terms of widely-used VR displays, the condition with matched SFOV, high FOV, and high FOR corresponds to a CAVE-like system. The condition with matched SFOV, high FOV, and low FOR corresponds roughly to a large display wall. And the condition with matched SFOV, low FOV, and high FOR has characteristics similar to most HMDs. From a practical point of view, our experiment seems to suggest that while high-end VR systems (CAVE-like displays) may provide the best overall performance, lower-cost VR systems with moderate levels of immersion (display walls and HMDs) can still result in significant performance gains over less immersive displays for the task of procedure memorization.

4. Conclusions and Future Work

Applications of immersive VR to conceptual learning and training applications have been proposed, but there has been little evidence to support the assertion that immersive VR systems can produce better learning. Our experiment on the effects of level of immersion on procedure memorization does not fully answer the question, but it does provide empirical evidence that higher levels of immersion can produce a measurable improvement in the performance of an abstract mental activity.

In addition, we have shown that a finer-grained view of immersion as a multidimensional continuum can result in a deeper understanding of its effects. In our experiment, because we studied three independent components of immersion, we were able to say not only that higher levels of immersion resulted in better performance, but also that these benefits are due to increased FOV and FOR and to matched SFOV. Further, conditions corresponding to lower-cost VR systems offered statistically significant performance improvements over conditions with lower levels of immersion. Significant benefits for procedure memorization can be obtained even without using the highest possible level of immersion.

Clearly, much future work is needed. We noted in the previous section the need for follow-up studies to compare learning in a VE vs. learning in the real world, and to investigate how well learners recall a procedure in the absence of any (physical or virtual) cognitive aid. Additional experimentation is also required to determine whether participants are indeed using a spatial memorization strategy, as we surmise, and whether they use this strategy more often when they learn in higher levels of immersion. We are also interested in exploring the importance of utilizing a spatial layout for information presentation, as we hypothesize that the effectiveness of memorization techniques may be enhanced through additional spatial cues. We could also study whether participants with higher spatial ability, or those who gain more spatial understanding of the VE, perform better in the procedure memorization task.

Beyond procedure memorization, empirical evidence of the effects of immersion (and its components) is needed for other abstract mental activities, and for higher-level conceptual learning processes. Finding appropriate measures and procedures for such experiments, however, will be a difficult challenge.

References

[AH00] ALLLISON, D., HODGES, L. F.: Virtual reality for education?. *ACM Symposium on Virtual Reality Software and Technology* Seoul, Korea (2000), pp.160-165.

[ACC99] ARNS, L.,. COOK, D., CRUZ-NEIRA, C.: The Benefits of Statistical Visualization in an Immersive Environment. *Proceedings of IEEE Virtual Reality* (1999), pp.88-95.

[AHC97] Arthur, E. J., Hancock, P. A., Chrysler, S. T.: The perception of spatial layout in real and virtual worlds. *Ergonomics*, *40*, 1 (1997), pp. 69.

[Art00] Arthur, K.: Effects of Field of View on Performance with Head-Mounted Displays. *Computer Science*. Chapel Hill, NC: University of North Carolina (2000).

[BHB97] Barfield, W., Hendrix, C., Bystrom, K.: Visualizing the Structure of Virtual Objects Using Head Tracked Stereoscopic Displays. Proceedings of the Virtual Reality Annual International Symposium (1997), pp. 114-120.

[BKM56] Bloom, B. S., Krathwohl, D. R., Masia, B. B. *Taxonomy of educational objectives: the classification of educational goals: handbook*. Longmans. (1956).

[BBSJ07] Botden, S. M. B. I.,. Buzink, S. N, Schijven, M. P., Jakimowicz, J. J.: Augmented versus Virtual Reality Laparoscopic Simulation: What Is the Difference? *World Journal of Surgery* 31 (2007), pp. 764–772.

[BM07] Bowman, D.,.McMahan, R.: Virtual Reality: How Much Immersion is Enough? *IEEE Computer* 40 (2007), pp. 36-43.

[BWHA99] Bowman, D., Wineman, J., Hodges, L, Allison, D.: The Educational Value of an Information-Rich Virtual Environment. *Presence: Teleoperators and Virtual Environments* 8 (1999), pp. 317-331.

[BroB99] Brooks, B. M.: The Specificity of Memory Enhancement During Interaction with a Virtual Environment. *Memory*, *7*,1 (1999), pp. 65.

[BroF99] Brooks, F.: What's Real About Virtual Reality? *IEEE Computer Graphics & Applications* 19 (1999), pp. 16-27.

[DSL96] Dede, C, Salzman, M., Loftin, R.: ScienceSpace: Virtual Realities for Learning Complex and Abstract Scientific Concepts. *Proceedings of the Virtual Reality Annual International Symposium* (1996), pp. 246-252.

[Eri03] Ericsson, K. A: Exceptional memorizers: made, not born. *Trends in Cognitive Sciences* 7(2003), pp. 233-235.

[JDR*05] Johnsen, K., Dickerson, R., Raij, A., Lok, B., Jackson, J., Shin, M., Hernandez, J., Stevens, A., Lind, D. S.: Experiences in Using Immersive Virtual Characters to Educate Medical Communication Skills. *IEEE Virtual Reality* 324 IEEE Computer Society (2005), pp. 179-186.

[JMOL01] Johnson, A., Moher, T., Ohlsson, S., Leigh, J.: Exploring Multiple Representations in Elementary School Science Education. *Proceedings of IEEE Virtual Reality*, Yokohama, Japan (2001), pp. 201-208.

[JR97] Johnson, W. L., Rickel, J.: Steve: an animated pedagogical agent for procedural training in virtual environments. *SIGART Bull* 8 (1997) , pp. 16-21.

[KSW00] Kaufmann, H., Schmalstieg, D., Wagner, M.: Construct3D: A Virtual Reality Application for Mathematics and Geometry Education. *Education and Information Technologies* 5 (2000), pp. 263-276.

[KSA*03] Kelso, J., Satterfield, S. G., Arsenault, L. E., Ketchan, P. M., Kriz, R. D.: DIVERSE: a framework for building extensible and reconfigurable device-independent virtual environments and distributed asynchronous simulations. *Presence: Teleoper. Virtual Environ.* 12 (2003), pp. 19-36.

[Kra02] Krathwohl, D. R.: A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice*, *41*, 4 (2002), pp. 212.

[LLG*07] Lidal, E., Langeland, T., Giertsen, C., Grimsgaard, J., Helland, R.: A Decade of Increased Oil Recovery in Virtual Reality. *IEEE Computer Graphics & Applications* 27 (2007), pp. 94-97.

[MC04] Mania, K., Chalmers, A.: The Effects of Levels of Immersion on Memory and Presence in Virtual Environments: A Reality Centered Approach. (2004, July 5).

[MRB05] Mania, K., Robinson, A., Brandt, K. R. : The Effect of Memory Schemas on Object Recognition in Virtual Environments. *Presence: Teleoperators & Virtual Environments*, *14*,5 (2005), pp.606-615.

[PHM*03] Ponder M., Herbelin B., Molet T., Schertenlieb S., Ulicny B., Papagiannakis G., Magnenat-Thalmann, N., Thalmann, D.: Immersive VR decision training: telling interactive stories featuring advanced virtual human simulation technologies. *Workshop on Virtual Environments 2003*. Zurich, Switzerland: ACM (2003), pp. 97-106.

[ROS06] Roussou, M., Oliver, M., and Slater, M.: The virtual playground: an educational virtual reality environment for evaluating interactivity and conceptual learning. Virtual Real. 10 (2006), pp. 227-240.

[SDLC99] Salzman, M., Dede, C., Loftin, R., Chen, J.: A Model for Understanding How Virtual Reality Aids Complex Conceptual Learning. *Presence: Teleoperators and Virtual Environments* 8 (1999), pp. 293-216.

[SB07] Schuchardt, P., Bowman, D.: The Benefits of Immersion for Spatial Understanding of Complex Underground Cave Systems. *ACM Virtual Reality Software and Technology*. Newport Beach, California (2007), pp. 121-124.

[Sla03] Slater, M.: A Note on Presence Terminology. *Presence-Connect* 3 (2003).

[SWB08] Sowndararajan, A., Wang, R., Bowman, D.: Quantifying the Benefits of Immersion for Procedural Training. *Workshop on Immersive Projection Technologies/Emerging Display Technologies.* Los Angeles: ACM Press (2008), pp. 4 pages.

[vDFL*00] van Dam, A., Forsberg, A., Laidlaw, D., LaViola, J.,Simpson, R.: Immersive VR for Scientific Visualization: A Progress Report. *IEEE Computer Graphics & Applications* 20 (2000), pp. 26-52.

[WM05] Ware, C., Mitchell, P.: Reevaluating Stereo and Motion Cues for Visualizing Graphs in Three Dimensions. Proceedings of Symposium on Applied Perception in Graphics and Visualization, 2005, pp. 51-58.

[Yat01] Yates, F.: *The Art of Memory*. Chicago, Illinois: University of Chicago Press (2001).