Training in Virtual Environments: Analysis of Task Appropriateness

David A Philbin
William Ribarsky
Neff Walker
C. Ellis Hubbard
Graphics, Visualization and Usability Center
Georgia Institute of Technology
Atlanta GA 30332

ABSTRACT

The need for evaluating task appropriateness in virtual training is discussed. A framework is developed for understanding navigational tasks and spatial cognition and their demands for training. The training abilities of virtual environments are critically examined and compared to the demands for navigational training. An experiment is performed to examine navigational training in a virtual environment and with a map. The results show superior performance for the group trained with the map, yet also show successful training using the virtual environment. A lack of guidelines for conducting virtual training is discussed, and further research is suggested.

INTRODUCTION

Because of its ability to furnish complex interactive visual and auditory stimuli, virtual reality has been lauded as a wonder tool for training. A quick search of a journal database will reveal numerous articles proclaiming the new age of virtual training. Surprisingly, very few articles have been published which critically examine these claims, or which even empirically examine them. The studies which have empirically looked at virtually reality training have produced conflicting results (Kenyon & Afenya, 1995; Kozak, Hancock & Chrysler, 1993; Tate, Sibert & King, 1997). Yet the unabashed exuberance for virtual training is still evident. Kozak et al, after concluding that there was no performance difference between groups receiving training in a virtual environment or receiving no training at all, still concluded that he and his colleagues remain 'enthusiastic about the potential for VR'.

The lack of scrutiny towards claims of VE's training efficacy may have extended to a lack of scrutiny into the tasks and situations in which VEs will be effective. Just as in any training situation, the tasks to be trained must be decomposed into their component cognitive, perceptual and motor demands and these demands must be met in the training environment. In short, to create an effective training environment, one must match up the capabilities of the training environment with the demands of the actual task.

In this paper, we first discuss the characteristics of VEs as they relate to training and the characteristics of tasks which are appropriate for training in a VE. A task with demands that match the characteristics of a VE, specifically navigation, is discussed, and a theoretical framework for understanding navigation is explored. Using that theoretical framework, training methods for developing the locational knowledge required for navigation is developed. Finally, the performance and training demands of navigation are discussed and an experiment is performed.

Training Characteristics of Virtual Environments

A typical virtual environment - with a 60 Hz update rate, head and hand tracking, and head mounted audio and visual displays - excels at providing head coupled visual and audio feedback. This allows one to 'look around' and perceive the surrounding environment, often presenting an impelling sense of being present. However, the typical VE system does not include an impelling tactile display. Kinesthetic feedback, the internal movement and positional feedback of joints and muscles, is not dynamically affected by changes in the VE. Haptic displays, which exert pressure or temperature changes to the skin, are also not in common use. Without the tactile feedback given by haptic displays one cannot tell whether there is pressure or heat against the body. Virtual objects, not having the touch and mass of their real counterparts do not impart the haptic and kinesthetic feedback which can be critical for successful motor tasks.

These characteristics of VEs dictate the types of tasks which can be successfully trained within a VE. Kozak et al (1993) attempted to train subjects in a manual placement task, a task whose performance relies as much on haptic and kinesthetic feedback as on fine visual feedback. Not surprisingly, they found no transfer to the real task. Indeed they found no benefit for the group trained with a VE over a group that had received no training whatsoever. On the other hand, Tate et al (1997) chose to supplement training for a visuospatial perception task with additional VE training. Using a task which matches up well with the characteristics of current VE systems, their results showed a benefit for the groups that were trained in the VE prior to engaging in the real task.

In their evaluation of navigation training, Tate et al showed that a VE with a map provides superior training to a map alone. Yet to date, no one has examined exactly what knowledge and skills you acquire from each of these representations. One question left unanswered by Tate et al is to what extent the benefit provided is merely due to additional training time, as opposed to the use of the VE. In this paper, we are attempting to specify more fully what is gained from each of the types of training, map and VE.

Virtual Environments are currently very weak at haptic displays and kinesthetic feedback. Vestibular feedback, the bodies internal sense of orientation, is also not affected by changes in the virtual environment. Today's VEs simply don't have the display characteristics to easily give the feedback necessary to determine if one's hand is pressed against a target or what the mass is of a virtual object being carried or whether one is flying upside down. VEs are strongest in visual information display and head motion feedback. The tasks which are most likely to benefit from training in a virtual environment are tasks which heavily depend on visual information for success.

Tasks which require interactive 3-D visual displays while minimizing the need for physical interaction, such as inspection tasks and navigation tasks, take advantage of VE's strengths for training purposes. Because of the robust use of the 3-D components of VEs in navigation tasks, we have chosen navigation as the category of tasks to further explore for training in a VE.

Navigational Tasks

Navigating through an environment relies heavily on physical movement and on spatial and locational cognition. Moreover, spatial and locational cognition rely mainly on visual information and spatial working memory. Discounting, for the moment, the physical movement aspects of navigation, the performance demands of navigation map well onto the display characteristics of a VE.

Thorndyke (1980) proposed a framework for understanding spatial knowledge and navigation. Essentially the theory separates spatial knowledge into three types of spatial cognition. As one becomes more familiar with navigating an environment, one may develop each of the three types of knowledge.

The first type is termed *Landmark Knowledge*. Navigation using landmark knowledge, as the term implies, relies exclusively on the recognition of salient landmarks. A landmark may be an identifiable object (e.g. the black skyscraper, or the water fountain), a unique configuration of objects (e.g. 'where the stairs and bathroom meet') or a unique feature of the environment (e.g. the three-way intersection). Persons use the landmarks to orient themselves in the environment. Although navigation between landmarks cannot be performed, the person would know where he or she is when a salient landmark is encountered. Landmark knowledge can be used to communicate locational information and allow collaboration with others in the environment. Landmark knowledge lays the foundation for the second type, *Route Knowledge*.

Navigation using route knowledge depends on cues in the environment and the relationships between those cues. With route knowledge persons begin to be able to navigate from one area to another using the landmarks to make decisions about their route. Unlike landmark knowledge, route knowledge allows the use of cues in the environment to make navigational decisions. Cognitively, the distinction between landmark knowledge and route knowledge is based on the qualitatively different representation that has been built up. Landmark knowledge relies exclusively on recognition of cues in the environment, while route knowledge relies on recognition of the cues in the environment as well as recall of the relationships between those cues. These relationships are thought to be stored in both verbal and spatial code. A person at the route stage of knowledge may navigate using a series of directional statements, such as 'turn left at the statue' or 'if you see the gas station, you've gone too far'. These are thought to be the verbal aspect of route knowledge. However, as noted, route knowledge is not restricted to verbal representation. A spatial representation is also stored. Using this spatial representation, route knowledge can be used to estimate how far the path from one landmark to another is, or how long navigating between them will take. Part of the observable distinction between landmark knowledge and route knowledge is in the person's ability to navigate between landmarks. For instance, the person who says 'I don't know how to get there, but I'll tell you when we are there', is using landmark knowledge. However, a person who says 'I can't tell you how to get there, but I can drive it' is working at the route stage of knowledge.

The third stage of spatial knowledge is *Survey Knowledge*. Survey knowledge is thought to be characterized by the development of a cognitive model of the environment analogous to an internal map of the area. In turn, this 'internal map' is thought to be the mental model equivalent to a regular road map (Thorndyke, 1980; Thorndyke and Hayes-Roth, 1978).

Survey knowledge contains the representations developed during the construction of landmark and route knowledge, but also includes the relationships between the features of the environment in a more global spatial representation. This global spatial representation can be scanned and rotated in the same way one would use a physical map. Using this global representation, subjects would be able to 'fill in' parts of the environment that they had not directly traveled through, thereby creating novel paths to navigate an area. By providing a global spatial representation that can be scanned, survey knowledge also allows better judgments of absolute distances between landmarks.

In summary, Landmark knowledge consists of the recognition of important landmarks in the environment. It does not allow navigation through the environment in any pre-planned manner, because there is no internal representation of any relationship between the landmarks. Route knowledge consists of recognition of landmark items in the environment, as well as a procedural representation of the relationships between them. It allows navigation of previously traveled portions of the environment, and accurate estimates on travel time between items in the environment. Survey knowledge is comprised of a global spatial representation of the environment. This allows novel paths to be generated through the environment, as well as allowing for accurate distance estimates to be made between points in the environment.

Development of Knowledge for Navigation

The most straightforward method to train for navigation is to repeatedly navigate the environment. Repeated exposure to the landmarks and to the navigational cues in the environment will help to develop the route-based representation of the environment, and eventually the survey-based representation. There is a tradeoff between the depth of the subject's knowledge of the environment and the time invested to proceed through the three phases of spatial cognition. Since survey knowledge is the most global form of spatial cognition, it is usually the goal of navigational training. As Thorndyke has noted (1980), survey knowledge provides a global spatial representation of the environment and allows for novel paths to be generated.

Fortunately there is a shortcut to gaining survey knowledge that does not require the extensive navigation time and familiarity with the environment that are developed during the landmark and route stages. This shortcut to survey knowledge is map study. Instead of waiting for the subject to create an internal map of the spatial environment, the subject is presented with a map of the environment and required to study it. This builds the global-spatial representation without relying on the more time consuming development of landmark and route-based knowledge. As a training aid, map study has been used to drastically reduce the time it takes to prepare someone for navigating through an environment (Thorndyke, 1980).

Two of the largest benefits of map study are cost and time. Maps are inexpensive, and take little time to thoroughly study (Williams, Hutchinson & Wickens, 1996). In order to develop route knowledge, one must navigate the environment and learn not only the landmarks but also the spatial relationships between them. If an environment is exceedingly large or complex the time necessary to acquire route knowledge will increase dramatically. And, obviously, if an environment is not available route knowledge cannot be gained at all.

Acquiring Navigational Information in a Virtual Environment

The strengths of visually dominant virtual environments lay in the display of visual information and head motion feedback. The tasks which are most likely to benefit from training in a virtual environment are tasks whose primary components depend on visual information for success. Matching the demands of a task to the interaction aspects of the VE, as previously discussed, is critical to developing successful training using VEs. VEs excel at delivering head-coupled visual information, specifically, they can be used most effectively for presenting dynamic 3-D information regarding the relations of objects in the virtual environment to each other and to the observer. Their main feedback strengths are head-coupled visual feedback, which presents compelling 3-D information (Davis, in press; Padmos and Milders, 1992), and, less commonly, 3-D audio feedback. Their main weaknesses are the lack of haptic or kinesthetic feedback and in vestibular feedback.

To develop route knowledge of an area without actually being able to navigate the area requires a navigable 3-D environment containing the important landmark information from the actual environment, and the ability to follow travel paths essentially identical to the actual environment. In other words, the training requires an interactive visual display with the ability to present a navigable 3-D environment. This is a very good match up for success in training a task in a VE.

For conducting completely novel pathways through an environment, the more global representation of survey knowledge will prove superior. In the research presented here, the focus is on rapid training for navigation in previously untravelled environments of moderate size. The environment size and the time constraints will allow subjects to traverse the entire environment while developing route knowledge. Larger environments, or smaller training times, may make it impossible to fully traverse the environment. In these situations, training for survey knowledge may be the only recourse.

THE EXPERIMENT

Two groups of subjects are trained. Each group is trained in one of the two forms of locational knowledge, either route or survey. The environment used is the second floor of the college of computing building containing the GVU lab. Subjects in the survey group acquired survey knowledge through the study of a map of the environment. Subjects in the route group train for route knowledge by navigating the environment in a VE. Both the map and the VE contained icons representing target objects in the environment. Both groups were given the same verbal descriptions of the target objects. Target objects are salient objects in the environment between which the subjects were to navigate in the real environment.

After the training, subjects were tested for route knowledge and survey knowledge. They were then required to walk through the actual environment. This walk through provided a test of transfer of training.

During the transfer of training task, subjects were instructed to walk to a specified target object. Wrong turns were noted and corrected. Upon reaching a target, subjects were informed of the next target landmark to which they are to proceed. Subjects traveled to a total of 8 targets.

Method

Participants. Participants were students at the Georgia Institute of Technology who had no prior experience with the second floor of the College of Computing. Students were recruited from the Psychology department's undergraduate subject pool.

Apparatus. The experimental environment was displayed using a Virtual Research VR4 head-mounted display with Polhemus Isotrak 3-D tracking hardware. The images were generated with a Silicon Graphics Crimson Reality Engine. The participants interacted with the environment using a plastic mouse, shaped like a pistol grip which was also tracked using the Polhemus Isotrak. During the experiment, they stood within a 1 m by 1 m railed platform.

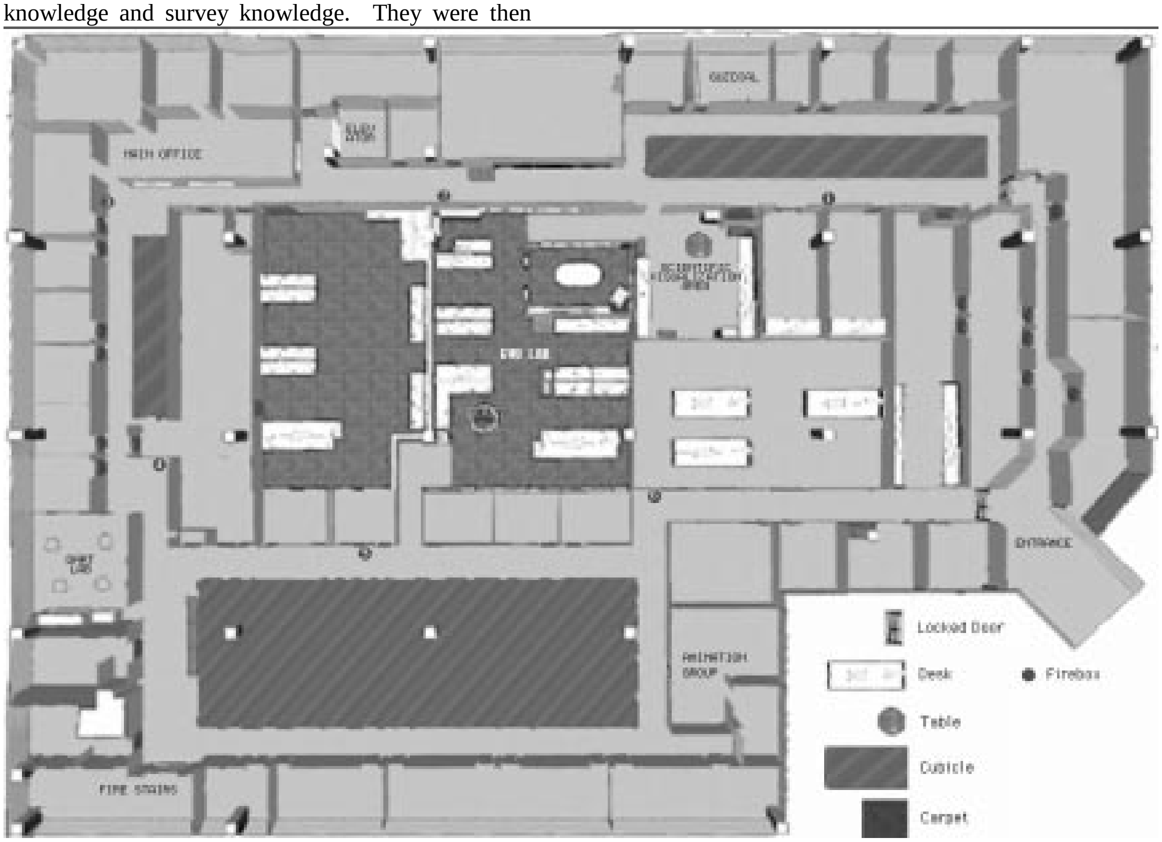

Figure 1. Top down 3D map of Environment

A model was built of the second floor of the College of Computing building. The model uses textures and shading for 3-D information. The environment was marked with landmarks. These landmarks were the most relevant features found in GVU (i.e. the entrance, the fire stairs, the elevator, etc.). Each subject was run individually. Subjects were divided into two groups. The Survey group, and the Route group. To minimize any effects from exposure to the VE, all groups began the study by spending 5 minutes in a VE of a practice environment, a virtual art gallery. This introduced all groups to the VE, although the Survey group did not use the VE again. This familiarized the Route group with the controls in the VE, and equated all groups on experience with the VE. The Survey group then received a 3-D map of the environment, marked with the Landmarks, were given a verbal description of the Landmarks, and were asked to study the map for 10 minutes (See Figure 1).

Subjects were instructed to concentrate on how to get from each landmark to each other landmark. The Route group was given 10 minutes to navigate in the virtual representation of the environment. They were directed on a route to explore by a visible path in the environment. The path was marked with pentagons at which the subject halted and was read a description of the Landmark by the experimenter (see Figure 2).

Figure 1. Path and Landmark at the GVU lab

The path traversed the entire environment once, during the time remaining, subjects were free to explore the environment on their own. All subjects were instructed to focus on learning how to navigate between the landmarks.

At the end of the ten minutes of studying the environment, both sets of subjects were given a brief test. The test was designed to measure the type of knowledge they had gained regarding the environment. Subjects were asked to estimate the absolute distance between Landmarks (referred to as the Distance question) and the time it would take to travel between Landmarks (referred to as the Travel question). The order of the questions asked alternated between Distance first, Travel second, and Travel first, Distance second. These questions reflect the type of knowledge being used to traverse the environment;

Distance questions reflect Survey knowledge while Travel questions reflect Route knowledge.

Each subject was then introduced to the actual environment. Subjects were instructed to travel to a specific landmark. When the subject reached that landmark, they were then instructed to go to another specific Landmark. This procedure was repeated 8 times for each subject.

An experimenter walked behind the subject and scored navigational errors. Upon taking wrong turns, subjects were notified by the experimenter by a tap on the shoulder. After a wrong turn, subjects were required to choose another travel direction.

Results

The design of the study was 30 x 2 (x 8). 30 subjects each were trained with the VE or the map. Data were collected for a total of 8 trials per subject. The dependent variables were: scores on the written test of route and survey knowledge (distance and travel questions); and the number of errors made during the transfer of training walk-through task. In addition we compared the number of errors made in the transfer of training walk-through task to number of errors that could be expected from untrained (no knowledge) participants.

First we compared the subjects in the two training groups on how well they learned Route and Survey knowledge. The Distance question and the Travel question scores were compared using t-tests. The analysis revealed that the two groups did not differ on route knowledge ($t=2.01$, $p > .05$). Because the VE should be stronger than the map in presenting route information, this suggests that either the VE training and Map training were equally effective in presenting Route information, or that effectiveness of the map training was greater than the VE training to the point of bringing the map group's Route knowledge up to the level of the VE group's. As expected with Survey knowledge, the analysis revealed that the map group learned significantly more than the VR group ($t=4.56$, $p<.0001$).

Next we compared how well the participants in the two groups could use the information they had acquired to perform the transfer of training walk-through task. The analysis on number of errors made during the walk-through task revealed that the training with the VE was not as effective as that with the map (mean VE errors = 5.33, mean Map errors = 2.33; $t=4.21$, $p<.001$). This supports the interpretation of the results of the analysis of the Route knowledge that the training time in the VE was not as effective as that with the map.

Finally, to get a measure of the effectiveness of VE training to no training, we compared the mean number of errors made by the VE-trained group to the number of errors that would be made by untrained participants. This analysis revealed that training with the VE resulted in significant improvements in navigation over the expectation of untrained navigation ($t=-13.26$, $p<.0001$).

DISCUSSION

The first and perhaps most important point to make is that the virtual environment training did indeed result in an improvement in the real environment performance. By having subjects perform a virtual walk-through of the building, they were later able to perform at much greater than chance levels when asked to walk through the real building. This result directly examines the efficacy of training in a virtual environment. This shows that virtual environments can be used successfully for training, although it also supports the idea that the tasks must be chosen appropriately.

However, it is also important to note that the group trained with the virtual environment did not perform as well as the group that received the more traditional (and less expensive) training method. The group which trained with the map outperformed the group which trained with a VE, both in actually travelling through the environment, as well as answering questions about the distances between objects in the environment. While this result may not sound encouraging to those interested in VE for training, it should be pointed out that while the VE may not have been foremost, it was successful as a training method on its own. Unlike the work of Tate et al (1997), where the VE was successfully used to supplement more traditional training methods, we have shown here that VE training can stand on its own as the single source of training for the subjects. Training methodologies for using VEs have not been fully defined, and the hardware and software used for VEs have much room to develop. Unlike optimal map study which has been explored extensively, the conditions for optimal VE training are relatively unspecified.

However, another point to be made here is that each group was able to acquire a qualitatively different representation of the environment. The Route knowledge training with the VE was successful. The subjects were able to traverse the environment using a route knowledge representation, while the map subjects relied on a survey knowledge representation. By having the ability to train for different representations, one can take advantage of situations which may better utilize one or the other. The authors are currently working on performing experiment in just such a situation. By disrupting performance with environmental stressors, the differences in the cognitive demands of the two representations should be revealed. This would obviously have implications for training for performance in the presence of environmental stressors, and possibly in other situations. By carefully analyzing the cognitive demands and representations associated with the task and the display characteristics of the task to be trained in the VE, it may be possible to train fundamentally differently and perhaps at times more appropriately than traditional training methods have allowed. However, more research is necessary to realize this potential.

There currently is a dearth of guidelines for constructing successful training environments with VEs. The work that has been done is a start, but, as noted, if VEs are to truly offer the quality of training that people expect of them, much more work is needed. In this paper we have tried to make the initial steps towards specifying what tasks might be appropriately trained in a VE, but a true task taxonomy must be undertaken before real progress can be made.

Acknowledgements
The authors would like to thank Elizabeth Davis, Lori A Rollins, Benjamin Watson and pilot Dabby Phipps. This work was supported in part by the National Science Foundation.

References

1. Davis, E. T. (In press). Visual requirements in HMDs: What can we see and what do we need to see?

2. Kozak, J. J., Hancock, P. A., Arthur, E. J., & Chrysler, S. T. (1993). Transfer of training from virtual reality. *Ergonomics*, Vol. 36, No 7, 774-784

3. Kenyon, R., V., & Afenya, M., B., (1995). Training in Virtual and Real Environments. *Annals of Biomedical Engineering: Starkfest Conference Proceedings*, Biomedical Engineering Society, Vol. 23 , 1995, pp. 445-455.

4. Padmos, P., & Milders, M. V. (1992). Quality criteria for simulator images: A literature review. *Human Factors*, 34(6), 727-748.

5. Tate, D., L., Sibert, L., & King T. (1997). Virtual environments for shipboard firefighting training. *Proceedings of the 1997 Virtual Reality Annual International Symposium*.

6. Thorndyke, P., W. (1980, December). Performance models for spatial and locational cognition. (Technical Report R-2676-ONR). Washington, DC: Rand.

7. Thorndyke, P., W., & Hayes-Roth, B. (1978, November). Spatial knowledge acquisition from maps and navigation. Paper presented at the *Meeting of the Psychonomic Society*, San Antonio, TX.

8. Williams, H., P., Hutchinson, S., & Wickens, C., D., (1996). A comparison of methods for promoting geographic knowledge in simulated aircraft navigation. *Human Factors*, 38, 50-64.