Running Head: Spatial Knowledge Transfer

The Transfer of Spatial Knowledge in Virtual Environment Training

David Waller, Earl Hunt, and David Knapp
University of Washington

Abstract

Many training applications of virtual environments (VE’s) require people to be able to transfer spatial knowledge acquired in a VE to a real world situation. Using the concept of fidelity, we examine the variables that mediate the transfer of spatial knowledge and discuss the form and development of spatial representations in VE training. We report the results of an experiment in which groups were trained in 6 different environments (no training, real world, map, VE desktop, VE immersive, and VE long immersive) and then asked to apply route and configurational knowledge in a real world maze environment. Short periods of VE training were no more effective than map training; however with sufficient exposure to the virtual training environment, VE training eventually surpassed real world training. Robust gender differences in training effectiveness of VE’s were also found.

Introduction

Virtual environment (VE) training is based on the assumption that knowledge or skills acquired in a virtual world will transfer to the real world. In fact, virtual training simulators are effective only to the degree that they enable a user to apply knowledge or skills acquired in the VE to their real world counterparts. This paper considers those applications for which an essential part of the knowledge to be transferred includes the spatial properties of a large-scale environment. As an example, consider a proposed virtual fire-fighting simulator (see Bliss, Tidwell, and Guest, 1997; Egsegian, Pittman, Farmer, and Zobel, 1993). The virtual fire-fighting simulator trains firefighters to move through the spatial layout of buildings or ships before they enter the real environment to put out an actual fire. Training firefighters in a simulator is attractive because it is risk-free, controlled, convenient, and potentially cost-effective. However, before simulators such as this are implemented, it is important to understand the degree to which there are differences between knowledge acquired in the real world and knowledge acquired in a VE.

In this paper we will investigate the ways in which exposure to a computer generated replica of an environment can substitute for actual exploration of the real world. Recently, Witmer and his colleagues have shown that exposure to a VE can be effective in training knowledge about the route through a large and complicated office complex (Witmer, Bailey, Knerr, and Parsons, 1996). Bliss, Tidwell, and Guest (1997) showed further that civilian firefighters can apply route knowledge learned in a VE to a mock rescue situation. These studies are promising and persuasive demonstrations of the efficacy of VE’s in training spatial knowledge. In fact, we feel that researchers now no longer need to question whether VE’s can be effective in training spatial knowledge. Today’s more pressing research questions involve examining the variables that mediate the training effects of VE’s. We have taken this analytic view by controlling our training environment, while varying the fidelity of the interface, the quality of the VE, and training time. We use this approach to consider some of the psychological issues associated with VE training, and discuss whether training in a VE affects the cognitive representation of the corresponding real world environment.

Fidelity

We define fidelity to be the extent to which the VE and interactions with it are indistinguishable from the participant’s observations of and interactions with a real environment. It is generally assumed that if all other factors are held constant, increasing the overall fidelity of a simulator will lead to increases in transfer (Hays & Singer, 1989; Caird, 1996). To some degree, this intuitive principle is undoubtedly true. In the extreme case, training in a futuristic virtual simulator with perfect fidelity would yield transfer equivalent to real world training because the two environments would be indistinguishable. However, the reality of today's technology is that a slight increase in fidelity may be very expensive. Because the effectiveness of VE's always hinges on a trade-off between economic and technological variables, it is important to have an understanding of which technological variables may be most easily sacrificed without degrading trainee performance. For example, Witmer et al. (1996) were able to achieve positive transfer of route knowledge from people trained in a VE despite the fact that their simulator did not provide tactile feedback or aural cues. By today's standards, the simulator used by Witmer et al. was virtually state of the art and the environment was very detailed visually. Perhaps Witmer et al. could have achieved positive results with even less fidelity and cost. Bliss, Tidwell, and Guest (1997), for example, were able to show with relatively low-fidelity interface devices that people can transfer spatial knowledge acquired in a VE.

In addition to using lower-fidelity equipment to help determine the cost-effectiveness of VE's in training, we wish to determine those characteristics of the virtual training environment that are most important for successful transfer of knowledge. There are three information domains in VE training: the real world environment, the training environment, and the trainee’s mental representation of the environment (see figure. 1). In general, information about a real world environment will never be preserved perfectly in either the training environment or the trainee’s mental representation. We know this because there are systematic differences in people’s representations of real environments, even after years of experience in them (Tversky, 1981). On the other hand, some structures are preserved in the mappings between the three domains. Fidelity is concerned with the quality of these mappings.

-- Insert figure 1 about here --

The mapping from the real world environment to the training environment is mediated by what we call environmental fidelity, while the mapping from the virtual environment to the mental environment of the trainee is affected by what we term interface fidelity (see fig 1). We briefly discuss these types of fidelity and then describe an experiment that examines their role in the transfer of spatial knowledge.

Environmental Fidelity

Environmental fidelity depends on the degree to which variables in the training environment resemble those in the real world. This means that environmental fidelity is a psychological rather than an engineering concept, because it depends on a psychological judgment of similarity rather than a mathematical correspondence between values of variables. For instance, although map distances may almost perfectly correspond to distances on the ground, a map is not a high fidelity simulation of geography. In the experiment reported here, we consider three quite disparate levels of environmental fidelity: a real world maze environment, a virtual replica of the environment, or a map of the environment. The degree to which environmental fidelity affects the transfer of spatial knowledge from the learning to the real environment should be reflected in the ordering of the performance of these groups.

Interface fidelity

Interface fidelity deals with the mapping of the variables in the training environment to those in the trainee’s mental representation of the environment. It addresses the degree to which the input and output devices associated with the VE function similarly to the way in which the trainee would interact with the real world. Of course, the trainee’s assessment of an intuitive interface is also a psychological judgment. For example, a VE which uses a mouse as a navigational device represents walking or riding by mouse motion. It is an important question whether this correspondence sufficiently represents interaction with the real world through navigation. In fact, Bliss et al. (1997) have shown that for acquiring certain kinds of spatial knowledge, minimal navigational control with a mouse can be sufficient.

In our experiment, we clustered several interface devices together to form two principal treatment groups: an immersed condition, and a desktop condition. Participants in the immersed condition experienced the virtual training environment through a head-mounted display (Virtual Research's VR4 – 742 x 230 resolution; 60 degree field of view) in conjunction with a six degree of freedom head tracker. This immersion allowed a greater concordance between the user's actual body and their virtual viewpoint and thus represents a higher degree of interface fidelity. On the other hand, the lower fidelity desktop condition consisted solely of a 21-inch flat panel display (subtending approximately 36 degrees of visual angle). Both conditions allowed users four degrees of freedom of navigational control with a joystick (three translational dimensions – although the vertical dimension was limited to realistic heights – plus rotation around the vertical axis, or yaw). The degree to which immersion is an aid in training is an unresolved issue and our experiment was designed to test its value. If the fidelity of the interface is in fact an important aid in the transfer of spatial knowledge, then we would expect participants in our immersed condition to outperform those in the desktop condition.

Developing mental environments

Our remarks to this point have referred mostly to the quality of the mappings between a real world environment, a VE, and a mental representation of an environment. Here we focus on the form and development of a person’s mental representation of either a VE or a real environment. Because such representations develop over time, it is important to ask whether the developmental sequence is the same in the real and virtual environments.

Psychologists have identified three stages of development of an individual’s cognitive representation of a large-scale navigable space (Seigel & White, 1975). During an initial period of familiarization, a person focuses on the important locations in the environment. Knowledge in this stage consists of a disconnected set of landmarks. After more exposure to an environment, people are able to link together important landmarks into routes. Knowledge of this type is said to be a route representation. With additional exposure, some people may develop a more flexible, map-like representation of the environment called a survey representation (also known as configurational knowledge). An individual with a survey representation of a space understands the spatial relationship between the various landmarks in an environment independently of the routes that connect these landmarks. Survey representations facilitate spatial inferences and can allow people access to spatial information regardless of orientation (Sholl, 1987); however, they differ from real environments in well-documented ways (e.g. Tversky, 1981; McNamara, Hardy, & Hirtle 1989; Engebretson & Huttenlocher, 1996). Nevertheless, it is generally assumed that survey knowledge represents a more thorough and flexible understanding of the spatial characteristics of a large scale environment than does route knowledge. Therefore, it is important to study the variables which affect the transition from a route to a survey representation.

The issue of transition is particularly important when learning a VE because people can be in a real environment for more than a year without necessarily acquiring a survey representation of it (Moeser, 1988). Developing a surveyor’s representation requires a conscious effort (Lindberg & Garling, 1983), which implies that people are motivated to learn it and the environment allows them to do so. Just as in a real world environment, people can learn procedures for moving from point to point in a VE (Bliss et al., 1997). However, if a task requires no more than this, a surveyor’s representation may not develop. Although it has been shown that navigation through a VE can lead to orientation-free judgments of relative directions between objects (Tlauka & Wilson, 1996) it is not clear whether these are truly survey representations. Witmer et al. (1996) were not able to establish whether exposure to a VE differentially affects the quality of a person’s survey representation, although their findings are consistent with there being no difference.

The problem of developing a survey representation in a VE is confounded by two important characteristics of VE’s compared to real environments. VE’s typically have restricted fields of view compared to real environments, and restricted fields of view have been shown to interfere with spatial learning in the real world (Sholl, 1993). Additionally, today’s VE’s are, for most people, a non-intuitive experience; wearing a headset and navigating by looking or by mouse movement are not natural activities. To the extent that learning to interact with the VE is a controlled activity, it would be expected to interfere with the cognitive effort required to develop a surveyor’s representation. Therefore when evaluating VE training it is important to assess participant’s representation at different points in their VE experience. We have done this by measuring participants’ spatial knowledge after each of six exposures to the training environment. Additionally, we have included a group that is given much longer exposure to the VE during the training phase of each trial, allowing us to evaluate the effects of increased simulator time on an individual’s performance.

Our primary measure of spatial knowledge involved recording the participants’ behavior in the performance environment while he or she was blindfolded. Fast and accurate navigation in a room while blindfolded clearly requires one to have a good mental representation of the room. Conversely, we feel that efficient blindfolded navigation is indicative of an accurate mental representation. We measured efficient navigation by recording both the time taken to navigate blindfolded and the number of times the participants bumped into the walls of our environment. Unlike most methods of measuring route or survey knowledge (e.g. distance estimations, recognition methods, and graphic and reconstructive methods such as map drawing) the blindfold measure has the advantage of measuring actual spatial behavior in the performance environment. Moreover, denying participants visual information about the performance environment required them to rely more heavily on their training to learn the environment’s spatial characteristics.

Method

Subjects

A total of 125 people (61 men and 64 women) participated in the experiment. 82 of the participants were undergraduates between the ages of 18 and 40 enrolled in an introductory psychology course at the University of Washington. These people participated in the experiment in return for extra credit in their class. For the four weeks of the study during which the Psychology Department’s human subject pool was unavailable, the remaining 43 participants were recruited through an advertisement in the campus paper. These recruited participants were between the ages of 18 and 50 and were paid an hourly rate for their participation.

Materials and apparatus

The real world environment was a 14′ x 18′ maze with movable 7′ tall black curtains, configured as shown in Figure 2. The curtains hung from a rectangular grid of cables each spaced two feet apart. The ceiling of the maze extended approximately eight feet above this grid, allowing ambient light into the maze. At four locations, a large stuffed animal was suspended from the grid, four and a half feet from the ground, along with a large cardboard numeral. The numerals helped indicate to the subject the correct path to take through the maze as well as making communication about the maze’s locations easier. In this way, the numerals and stuffed animals served as prominent landmarks in the maze.

-- Insert figure 2 about here --

For the virtual portions of the experiment, the room described above was modeled using WorldUp™ by Sense8 Corp on a Pentium Pro™ 200 with an Oxygen™ 102 graphics accelerator card (see fig 3). At seven locations in the virtual maze we placed red directional arrows at eye-level. The arrows indicated the correct route to take through the maze and they also helped to make the space more interpretable for inexperienced users. A Thrustmaster PFCS™ joystick provided four degrees of freedom of navigational control (three degrees of translation plus the ability to pan one's viewpoint). Subjects in the immersed condition used a VR4™ HMD from Virtual Research and had additional navigational (and viewpoint) control with a six degree of freedom head tracker (Polhemus Fastrak™). Depending on the location of the user in the VE, the virtual scenes for both desktop and immersed conditions were rendered between 8.0 and 11.1 frames per second with a mean of 10.0. In addition to the virtual maze room, a large maze-like virtual environment was created in which participants learned the rudiments of navigation in a virtual environment. The computer on which the virtual portions of the experiment were conducted was in the same room (though in a separate area) as the actual real world maze.

-- Insert figure 3 about here --

Procedure

All participants were initially given the Guilford Zimmerman standardized test of spatial orientation ability. They were then given a short task that familiarized them with wearing a blindfold while walking around a practice maze. This task gave the experimenter the opportunity to correct the participants of habits such as walking too slowly and taking too small steps while blindfolded. (Pilot studies had shown that a short practice session with a blindfold helped to reduce the variance in scores during the testing phase of the experiment.) The experimenter then reconfigured the maze to the standard configuration (see figure 2) and allowed the participant exposure to a version the maze according to their experimental condition. Twenty participants were randomly assigned to each of the following six exposure conditions:

1. Blind: Participants in this group were given no exposure to the maze room.

2. Real: For each trial, participants in this group were given one minute in which to explore freely the real world maze. At the beginning of the first trial, the experimenter pointed out the appropriate route between each object and thereafter, participants were given no information or advice but were allowed to wander through the maze on their own.

3. Map: At each trial, participants in this group were shown a map of the maze and were asked to study it for one minute. At the beginning of the first study session, the experimenter oriented the map for the participant and pointed out the correct route to take through the maze.

4. VR-Desk: This group was given two minutes of exposure to a virtual replica of the maze at each trial. (Previous experiments had shown that one minute of VE exposure was not enough time to allow a person to navigate through the maze. A two minute exposure period allowed most participants to navigate through the entire maze on their first trial.) These participants were seated 24 inches from a 21-inch color monitor which rendered the virtual scenes at 800 x 600 resolution (true color, 60 Hz refresh). As with the real world group, participants were initially instructed which way to go so that they could get to each location in order. The arrows in the virtual maze also provided path information. The participants’ motion and viewpoint in the virtual environment were controlled by the user with a joystick.

5. VR-Immersive: At the beginning of each trial, this group was given two minutes of exposure to the same virtual maze and were given the same advice as the other groups on the route to take through the maze; however, they experienced it with a VR4

head-mounted display (742 x 230 resolution, 60 degree field of view) and a six degree of freedom tracker. These participants also controlled their motion and gaze with the joystick.

6. VR-Long Immersive: This group was identical to the immersive group; however, at each trial, they were allowed five minutes of exposure time to the virtual maze.

Prior to the maze exposure, all of the participants in the virtual conditions had been given between 30 and 75 minutes of instruction and training in how to use the input devices efficiently. A virtual practice world was used in which the elements of navigation with the joystick (and tracker) were trained and practiced. After learning the basic navigation skills, these participants were timed on a “virtual obstacle course” that required extensive use of the elements of navigation and concentrated on those that would be important for navigating through the maze room. Subjects were not allowed to proceed with the experiment until they could complete the obstacle course in under four minutes. Training time for participants thus varied depending on their abilities. All but four people were able to complete the obstacle course in less than four minutes. It was clear to the experimenter that the four participants who did not complete the obstacle course had difficulty physically moving the joystick because of its relatively high spring tension. These four people were randomly re-assigned to one of the three non-virtual conditions.

After encountering either a virtual, a real, or a map version of the maze, subjects were blindfolded and escorted to the beginning of the real world maze. They were then instructed to touch each stuffed animal in order, as quickly as possible, while minimizing the number of times that they hit the walls of the maze. As participants went through the maze blindfolded, the experimenter timed them and counted how many times they touched the walls. Participants were informed and continually reminded of how they were being scored, and were asked to do their best to minimize their time and touches of the walls (or “bump count”). This process of exposure to the maze followed by a blindfolded walk-through task was repeated six times.

After the sixth exposure to the maze, the experimenter gave the participant a distracting task while he altered two of the curtains in the maze (see Figure 4). This new configuration was identical to the one on which subjects had been trained except that two of the possible three paths between the _first_ and _third_ stuffed animals were now blocked. In the new configuration, both the most familiar and the most direct path were blocked. The participant was instructed that his or her task was no longer to touch each animal. Rather, the task was to go as quickly as possible from the first stuffed animal directly to the third stuffed animal. When the subjects discovered that the typical path between the first and third animals had been blocked, they were forced to rely on their mental representation of the maze and integrate the piece-meal knowledge they had acquired to that point. We refer to this task as the “integration task.” The experimenter recorded how long the participants took to complete the integration task, how many times they touched the maze walls, and the route(s) that they attempted to take.

-- Insert figure 4 about here --

Finally, participants were given a 30 question true/false test in which they identified whether a given map of the room correctly represented a portion of the maze. False items on this test (see figure 5) were wrong because they either showed an incorrect route between locations (see figure 5a) or an incorrect relative position of locations (see figure 5b). Though both types of items required configurational knowledge, we refer to the items showing possible routes between object locations as ‘route items.’ Those items which showed only the relative location of points we refer to as ‘survey items.’

-- Insert figure 5a and 5b about here --

Results

Five participants (four men and one woman) voluntarily withdrew from the experiment complaining of mild symptoms of simulator sickness. Unless otherwise specified, the following analyses are conducted on the remaining 120 participants, using a two-tailed alpha level of .05. Multivariate tests of significance use Wilk's lamda criterion, and significance tests of contrast estimates do not assume equal variance.

Blindfold task

A number of participants (particularly those in the blind condition) took what we felt was an inordinately long time to complete the blindfold portion of some of the trials. In order to allow the entire experiment to be completed in the time allotted, the experimenter stopped the participant if he or she had not traversed the maze in ten minutes. For the 31 subjects and 39 trials on which this occurred (less than 5% of the trials), we assigned a time of ten minutes and a bump count equal to the number of bumps the participant had made in those ten minutes.

Across all trials, the partial correlation (controlling for subject) between the participant's time through the maze and their bump count was quite high ($r(832) = .94$, $p < .0005$). Because of this correlation, all of the subsequent analyses will be univariate, treating time through the maze as the dependent variable. None of the results we obtain differ substantially when we consider bump count as an additional (or sole) measure.

Figure 6 illustrates the effect of repeated exposures on time through the maze for each experimental group. Differences in transfer of spatial knowledge between the non-blind conditions were evaluated using a repeated measures fixed effects ANOVA, treating trial number (1 - 6) as a within subject independent variable, experimental condition (map, real, desk, short immersion, and long immersion) and gender as between subject independent variables. Time through the maze was the dependent variable. In summary, this analysis yielded a significant two-way interaction between condition and trial ($F(20, 286) = 1.84$, $p = .017$) and significant main effects of trial ($F(5, 86) = 31.73$, $p < .0005$), condition ($F(4, 90) = 2.73$, $p = .034$) and gender ($F(1, 90) = 18.75$, $p < .0005$). All other effects were not significant. We defer analysis of the gender related effects to the section on individual differences.

-- Insert figure 6 about here --

Not surprisingly, subjects’ performance in all conditions improved steadily over trials -- the main effect of trial was significant ($F(5, 86) = 31.73$, $p < .0005$). More importantly, the rate of this improvement depended on the type of training the participant had received, probably because participants in the different conditions were converging towards the same asymptote. The interaction of trial and condition was significant (*F*(20, 286) = 1.84, *p* = .017).

Early Learning: effects of immersion and maps

Table 1 shows mean times through the maze on the first two trials for each experimental group. Participants who were allowed only one minute of exposure to the real maze were able to traverse it blindfolded much faster on the first two trials than those participants in the other conditions. On average, subjects in all VR conditions performed worse in the initial trials (*M* = 270.51 s) than those people in either the real world (*M* = 163.32 s) or map (*M* = 242.88 s) conditions. By the second trial, only the group that was given a much longer training time in the immersive VE was able to outperform participants trained on the map ($M_{long\ immerse}$ = 122.05 s; M_{map} = 191.70 s). The lag in performance for the participants who trained in most VE conditions is partly responsible for the significance of the trial by condition interaction. Statistical comparisons between immersed and non-immersed VE groups and between VE training in general and map training are not significant over the first two trials. The only significant Helmert contrast comparing group differences on the first two trials is the one that compares the real world group with all other non-blind groups. The difference between the real world group's mean time and that of the other non-blind groups over the first two trials was, with 95% confidence, estimated to be 150.81 ± 136.05 seconds.

Later learning: asymptotic performance and the effect of long immersion

Figure 7 illustrates the mean times for each group on the blindfold task after the sixth training session. By the sixth trial, participants in the long immersive condition outperformed those in the real world training group ($\underline{M}_{\text{long immerse}} = 40.95$; $\underline{M}_{\text{real}} = 56.5$ seconds), although this difference is not significant. Participants who trained in the other conditions converged on a somewhat worse performance. The contrast comparing mean times for the real and long immersed condition with the two other VE conditions is significant ($\underline{t}(43) = 3.22$, $\underline{p} = .002$).

-- Insert figure 7 about here --

The convergence of performance between the real and virtual groups by the sixth trial cannot be attributed to the learning of the environment that occurs while the participants were blindfolded. By the sixth trial, those people in the blind condition still performed significantly worse than the other groups. The contrast comparing the times of the blind group on the sixth trial with those of the other groups is highly significant (_t_(19) = 4.55, _p_ = .0002).

Representation differences

Differences in mental representations after the sixth trial were measured by combining the results of the integration task and the true-false questionnaire. The integration task forced participants to access their mental representation of the maze by blocking off the familiar path from one maze location to the other. In addition to recording the time to complete the integration task and the number of times the participant ran into the maze walls, the experimenter also recorded which alternate routes the subject attempted to take. Two statistics were derived from the true-false questionnaires: the total percent correct, and a “survey score” which was calculated by subtracting the number of correct ‘route’ items (see figure 5a) from the number of correct ‘survey’ items (see figure 5b).

Representation differences were tested using a fixed effect MANOVA with gender and experimental condition (real world, map, desktop, immersive, and long immersive) as between subjects independent variables. Time on the integration task, bumps in the maze on the integration task, whether the participant initially attempted the shortest route, percent correct on the true-false test, and survey score on the true false test were included as dependent variables. The analysis revealed significant main effects of both condition ($\underline{F}(20, 253) = 1.90$, $\underline{p} = .013$) and gender ($\underline{F}(5, 76) = 2.58$, $\underline{p} = .033$). The interaction of condition and gender was not significant.

The true-false test proved to be a particularly strong measure of both group and gender differences in spatial representations. In all training conditions, men averaged higher scores on the true-false test than women ($\underline{M}_{men}$ = 69% correct; $\underline{M}_{women}$ = 64% correct). For both genders, map training ($\underline{M}$ = 70% correct) yielded the best performance and the two-minute immersed group yielded the worst ($\underline{M}$ = 61% correct). A 4 (training condition) x 2 (gender) univariate ANOVA indicated that both training condition ($\underline{F}(4, 85) = 2.78$, $\underline{p} = .032$) and gender ($\underline{F}(1, 85) = 7.78$, $\underline{p} = .007$) were significant predictors of participants' test performance. Figure 8 illustrates these differences.

-- Insert figure 8 about here --

Another particularly sensitive measure of differences between training conditions was whether the participant initially attempted to take the shortest route in the integration task. Follow-up analysis on this variable revealed that much of the effect of training condition was accounted for by the lower scores of participants in the two immersed conditions. On average, participants in the immersed conditions chose to take the shortest route on the integration task only 35% of the time, whereas, people in the other conditions averaged 63%. This difference is significant ($t(79) = 2.90$, $p = .005$).

Individual differences

Most of our dependent measures varied reliably with gender. On average, in all non-blind experimental groups, men outperformed women at the blindfold task ($F(1, 90) = 18.75$, $p < .0005$). A gender effect was particularly strong for women who trained in the three VE conditions. This trend is illustrated in figure 9. VE-trained women performed significantly worse than men in the VE conditions (the 95% confidence interval for the mean difference between VE men and VE women [assuming equal variances] was 225.21 ± 141.45 s). They also performed significantly worse than women trained in the real world (the 95% CI for the mean difference between these groups [assuming equal variance] was 271.65 ± 173.24 s). Moreover, there was not a significant difference between women and men who trained in the real world.

-- Insert figure 9 about here --

Men also outperformed women on four of the five measures of spatial representation (all but the survey score from the true-false questionnaire), and the MANOVA conducted on our representation measures showed a significant effect of gender ($F(5, 76) = 2.58$, $p = .033$). T-tests confirmed that men took less time to complete the integration task ($t(65) = 2.70$, $p = .009$), touched the walls of the maze less frequently ($t(65) = 2.40$, $p = .020$), and scored higher on the true-false test ($t(91) = 2.39$, $p = .019$).

The Guilford Zimmerman test of spatial orientation was moderately predictive of a participant’s overall performance on the true-false test ($r(95) = .44$, $p = < .0005$); however, it was not predictive of the survey score derived from the true-false test, nor was it predictive of any of the behavioral measures of spatial knowledge.

Discussion

We have shown that training in a VE of relatively low fidelity allows people to develop useful representations of a large scale navigable space; however, when allowed only short exposures, VE training may be no more effective than training with a map, and immersive VE training may be no more effective than desktop VE. On the other hand, if one is willing to spend enough time, then training in an immersive VE on tasks requiring route knowledge may surpass map training and be indistinguishable from training in the real world. However, there is evidence that immersive VE training does not allow survey understanding. Perhaps this is because conscious effort (Lindberg & Garling, 1983) and a wide field of view (Sholl, 1993) are necessary for acquiring survey knowledge. It is possible that for our participants, the novelty of the HMD and tracker diverted some of the cognitive effort required for survey understanding. A follow-up to this experiment will involve training participants extensively in a different VE so that the interface devices are less novel and require less cognitive effort. We are also currently examining the question of whether VE’s enable better long-term retention of spatial skills.

It is important to realize that the VE used for training in this experiment did not make use of many possible properties of virtual worlds, and these properties may have a great impact on acquiring spatial knowledge. We did not allow subjects to ‘fly’ above the maze and look down on it; nor did we allow them to pass through walls or transport themselves to distant locations. Besides the directional arrows, there were no navigational aids in the maze. It is likely that a virtual compass or translucent walls could be used effectively to train spatial knowledge. While this and other research has shown that these characteristics of VE’s are not necessary for training route knowledge, they may be helpful in training survey knowledge.

Darken and Sibert (1996) have shown that environmental cues such as maps and grids placed in a VE decrease directional errors on hand-drawn maps. Participants trained on a map in our experiment also exhibited better configurational knowledge of the real environment. It is natural to hypothesize that providing a map or other navigational aids within the VE during training can enhance the acquisition of survey knowledge.

There has been little evidence for gender effects in VE training, and we were surprised to have found such robust differences between men and women. It is clear that for our population of participants, women’s performance on virtually all tasks tended to lag behind that of men. Importantly, though, women who trained in the real world environment performed nearly as well as similarly trained men. This means that performance differences between genders are not because of differences in the acquisition of spatial knowledge. They are more likely due to gender differences in the effectiveness of VE’s for training. Psychologists have shown that in general, men have more experience with video games (Philips, Rolls, Rouse, & Griffiths, 1995) and report more comfort and confidence with computers (Temple & Lips, 1989). There is also evidence that computer games have differential training effects on men and women (Tirre & Raouf, 1994). Future research in this field will need to determine the extent to which prior computer experience impacts the effectiveness of VE training differently for men and women.

At the very least, future related experiments will need to control for gender.

With a few caveats, VE’s can be an effective medium in which to train spatial knowledge. Because of the expense associated with high fidelity VE training, though, it is essential to know the exact nature of these caveats. One that we have found is that environmental and interface fidelity have little effect on the acquisition of route knowledge. As our research continues, we hope to delineate more clearly the capabilities and limitations of VE training.

References

Bliss, J.P., Tidwell, P.D, & Guest, M.A. (1997). The effectiveness of virtual reality for administering spatial navigation training to firefighters. Presence, 6, 73 - 86.

Caird, J.K. (1996). Persistent issues in the application of virtual environment systems to training. IEEE Computer Society: Human Interaction with Complex Systems 3, 124-132.

Darken, R. P. & Sibert, J. L. (1996). Navigating large virtual spaces. International Journal of Human-computer Interaction 8(1), 47 - 71.

Hays, R. T. & Singer, M. J. (1989). Simulation fidelity in training system design: Bridging the gap between reality and training, New York : Springer-Verlag.

Egsegian, R., Pittman, K., Farmer, K. and Zobel, R. (1993). Practical applications of virtual reality to firefighter training. In Proceedings of the 1993 Simulations Multiconference on the International Emergency Management and Engineering Conference, 155-160. San Diego, CA: Society of Computer Simulation.

Engebretson, P. H. & Huttenlocher, J. (1996). Bias in spatial location due to categorization: Comment on Tversky and Schiano. Journal of Experimental Psychology: General, 125(1) 96-108.

Linberg, E., & Garling, T. (1983). Acquisition of different types of locational information in cognitive maps: Automatic or effortful processing? Psychological Research, 45, 19 - 38.

McNamara, T. P., Hardy, J. K. & Hirtle, S. C. (1989). Subjective hierarchies in spatial memory. Journal of Experimental Psychology: Learning, Memory, and Cognition, 15(2), 211 - 227.

Moeser, S. D. (1988). Cognitive mapping in a complex building. Environment and Behavior, 20, 21-49.

Phillips, C. A., Rolls, S., Rouse, A. & Griffiths, M. D. (1995). Home video game playing in schoolchildren: A study of incidence and patterns of play. Journal of Adolescence, 18, 687 - 691.

Sholl, M. J. (1987). Cognitive maps as orienting schemata. Journal of Experimental Psychology: Learning Memory, and Cognition, 13(4) 615 - 628.

Sholl, M. J. (1993). The effect of visual field restriction on spatial knowledge acquisition, Paper presented at the 34th Annual Meeting of the Psychonomic Society, Washington, D. C.

Siegel, A. W. & White, S. H. (1975). The development of spatial representations of large-scale environments. In H. Reese, (Ed.) Advances in Child Development and Behavior, Vol. 10, (pp 10 - 55). New York: Academic Press.

Temple, L. & Lips, H. M. (1989). Gender differences and similarities in attitudes toward computers. Computers in Human Behavior, 5, 215 - 226.

Tirre, W. C. & Raouf, K. K. (1994). Gender differences in perceptual-motor performance. Aviation, Space, and Environmental Medicine, 65(5), A49 - A53.

Tlauka, M. & Wilson, P. N. (1996). Orientation-free representations from navigation through a computer-simulated environment. Environment and Behavior, 28(5), 647 - 664.

Tversky, B. (1981). Distortions in memory for maps. Cognitive Psychology, 13(3), 407 - 433

Witmer, B. G., Bailey, J. H., Knerr, B. W. & Parsons, K. C. (1996). Virtual spaces and real world places: transfer of route knowledge. International Journal of Human-Computer Studies, 45, 413 - 428.

Author Note

This research was supported by the Office of Naval Research grant N00014-96-1-0380, Earl Hunt, Principal Investigator. We thank Max Wells, Toni Emerson, and Barry Peterson at the Human Interface Technology laboratory at the University of Washington for helpful comments and for help tracking relevant literature. We appreciate Jon Miller's help with building the virtual version of the maze. We also thank Nathaniel G. Auvil at Sense8 technical support for his tireless help with WorldUp scripting questions. Correspondence concerning this article should be addressed to David Waller, Department of Psychology Box 351525, University of Washington, Seattle, WA 98195. Electronic mail may be sent to dwaller@u.washington.edu.

Table 1

Means and standard deviations of times on the blindfold task for the first two trials

Figure Captions

Figure 1. The role of fidelity in virtual environment training.

Figure 2. A schematic map of the maze configuration used in the experiment.

Figure 3. The virtual maze environment.

Figure 4. A schematic map of the maze used for the “integration task” in which the participant’s task was to go from location 1 to location 3 as quickly as possible.

Figure 5a. Example of an incorrect multiple choice item which emphasized the route between maze locations.

Figure 5b. Example of an incorrect multiple choice item which emphasized the relative locations of objects in the maze.

Figure 6. Mean time through the maze blindfolded for each trial in each condition of the experiment.

Figure 7. Mean scores with 95% confidence intervals for the last trial of the blindfold task.

Figure 8. Mean percent correct on the true/false maze identification test for men and women in the different training conditions.

Figure 9. Mean time through the maze blindfolded on each trial for women and men in the VE training conditions and in the real world training condition.

Waller, Hunt & Knapp: Fig. 1

Waller, Hunt & Knapp: Fig. 2

Waller, Hunt & Knapp: Fig. 4

Waller, Hunt & Knapp: Fig. 5a & b

Waller, Hunt & Knapp: Fig. 6

Waller, Hunt & Knapp: Fig. 7

Waller, Hunt & Knapp: Fig. 8

Waller, Hunt & Knapp: Fig 9

Waller, Hunt & Knapp: Table 1