Spatial representations of virtual mazes: the role of visual fidelity and individual differences

DAVID WALLER[1], DAVID KNAPP, and EARL HUNT, University of Washington, Seattle, Washington

Twenty-four people were presented three versions of a large-scale maze: a wireframe desktop virtual environment (VE), a normally rendered desktop VE and a real-world maze. Differences between the mental representations formed from each environment were measured with pointing and distance estimation tasks in a real-world version of each maze. Participants also constructed maps of each environment. People were more accurate pointing after having learned the wireframe VE maze than the rendered VE maze, however this effect was small compared to the effect of individual differences. Differences in gender, spatial ability (assessed by the Guilford-Zimmerman test of Spatial Orientation), and prior computer experience were all significantly related to the ability to acquire spatial information from the desktop VE. There was a high correlation between spatial knowledge when it was measured in the VE and spatial knowledge measured in the real-world. Gender significantly influenced the accuracy of bearing estimations when they were measured in the VE.

[1] Address inquiries to David Waller, Department of Psychology, University of Santa Barbara, Santa Barbara, CA 93106. Email can be sent to waller@psych.ucsb.edu.

Running title: SPATIAL REPRESENTATIONS OF VIRTUAL MAZES

Key words: Virtual environments, spatial cognition, individual differences, simulation fidelity, computer-based training.

Spatial representations of virtual mazes: the role of visual fidelity and individual differences

In the last few years, several studies have shown that computer-generated desktop virtual environments (VE's) can be effective for training knowledge of large-scale spaces (Peruch, Vercher, & Gauthier, 1995; Rossano & Moak, 1999; Ruddle, Payne, & Jones, 1997; Waller, Hunt, Knapp, 1998; Wilson, Foreman, & Tlauka, 1997). Because VE's are able to depict three-dimensional spaces interactively, they offer a promising medium for teaching people about the spatial characteristics of places and situations that are rare, remote, or dangerous. However, the psychological processes by which the spatial properties of computer-generated environments become mentally represented are poorly understood. Although exposure to a desktop VE can result in accurate spatial knowledge, it not known what aspects of either the VE or its user are most important for enabling this knowledge. In this paper, we examine the role of the VE's visual fidelity and the role of several cognitive characteristics of the user in enabling configurational (or "survey") knowledge of a computer-simulated large-scale maze. Configurational spatial knowledge is characterized by knowledge of the overall pattern of spatial relationships in an environment and is operationally defined here as skill at pointing and estimating distances to unseen locations from any place in an environment.

It is generally assumed that increasing the fidelity of a VE will result in improvements in its training effectiveness. When VE's are used to train perceptual and motor skills, this principle is likely to be true (Hunt & Waller, 1999). For example, Witmer and his colleagues showed that a relatively high-fidelity VE (with a head-tracked stereoscopic immersive display) resulted in more accurate and faster traversal of a complex real-world route than did training with a map (Witmer, Bailey, Knerr, & Parsons, 1996). Similarly, Loomis and his colleagues showed that increasing the fidelity of the VE interface (by including kinesthetic and vestibular input) significantly improved the accuracy of people's memory about short paths that they had recently traversed (Chance, Gaunet, Beall, & Loomis, 1998; Klatzky, Loomis, Beall, Chance, & Golledge, 1998). It is important to realize that route memory of the type studied by Witmer et al., and by Loomis et al. is largely mediated by perceptual and motor processes and generally does not require one to form a flexible mental representation of the global characteristics of an environment (i.e., configurational knowledge). The acquisition of configurational knowledge is generally considered to be a controlled process, requiring conscious attention (Linberg & Garling, 1983). We suggest that the fidelity of a VE is much less important when it is used to train tasks that primarily require these more controlled cognitive processes. For example, in the Witmer et al. study cited above, researchers found no difference in performance on pointing and distance estimation tasks between participants who had learned in a VE and those who had learned it with a map. Perhaps the fact that these pointing tasks required more conscious processing made the difference between the fidelity of the map and the VE less relevant. Indeed, there is some evidence that maps – which are clearly very low-fidelity representations of environments – can be more effective than VE's for training configurational knowledge of an environment (Darken & Banker, 1998; Goerger et al, 1998). In fact, maps can be more effective than actual experience in an environment at teaching people about an environment's spatial characteristics (Lloyd, 1989).

The effectiveness of a map for teaching people configurational knowledge undoubtedly derives from the fact that maps typically represent critical spatial information (such as relative distances and bearings) very accurately, and fail to represent irrelevant information (such as color and texture). In the case of environmental simulations such as VE’s, a similar principle probably applies; however, it is not clear what aspects of the VE are essential for conveying spatial information. We explored this issue by exposing people to three different environments: a real-world maze, a relatively high-fidelity desktop VE maze, and a low-fidelity wireframe desktop VE maze (see Figure 1). Several studies have suggested that linear perspective is especially important for conveying depth information in VE's (Jaeger, 1998; Surdick, David, King, & Hodges, 1997). By allowing the user to focus on motion and perspective cues, wireframe VE’s may allow users sufficient (and perhaps superior) information to form an accurate mental representation of the environment. We were interested in investigating whether configurational knowledge acquired in a wireframe VE would be at least as accurate as that acquired in a normally rendered VE. In addition to clarifying the role of visual fidelity in VE spatial knowledge training, such a finding may help inform researchers about the relative efficacy of the cues to which people attend while acquiring spatial knowledge of an environment.

If the fidelity of the VE is less influential for cognitively controlled tasks than it is for perceptually and motor-driven tasks, what factors might enable us to predict and understand how mental representations of space are formed from experience with VE’s? Obvious candidates for these variables are individual differences in cognitive abilities and prior experience with computers. There is some evidence that spatial ability, as assessed by paper-and-pencil psychometric tests, has significant (though moderate) predictive validity for spatial knowledge acquisition in VE’s (Bailey, 1994; Darken, 1996; Waller, 1999). In general, tests of spatial visualization (the ability to manipulate figures mentally) and spatial orientation (the ability to account for changes in viewpoint) are more predictive of the ability to acquire spatial information in a VE than are tests of visual memory (the ability to remember the configuration, location, and orientation of figures) or spatial scanning (the ability to explore visually a complex spatial array) (see Waller, 1999). In the present study, we assessed participants' spatial orientation ability with the Guilford Zimmerman test of Spatial Orientation (GZ-SO) (Guilford & Zimmerman, 1981). The 60 items on this test require participants to determine the change in orientation implied by two views from the prow of a boat. To assess the role of general intelligence and to provide discriminant validity for the GZ-SO, we also gave participants a test of fluid intelligence designed by Salthouse (1993). Administered on the computer, the Salthouse test is very similar to Raven's matrices. It requires examinees to study a 3 x 3 array of figures that has one missing cell in the lower corner. Examinees must decide which of several possibilities completes the pattern formed by the array. Finally, we assessed participants' prior experience and attitudes toward computers with a short questionnaire (see Appendix). We were interested in whether scores on these individual difference measures would be more predictive of participants' ability to acquire an accurate spatial representation of a VE than would be the visual fidelity of the VE.

A final issue that we address involves how to measure spatial knowledge acquired in a VE. In addition to their potential for training, VE's also show promise as tools for understanding human spatial cognition. This is primarily because VE's: (a) enable participants to explore (simulated) large spaces within the confines of the laboratory, and (b) allow experimenters more control over stimulus characteristics than they typically have in the real world. In fact, several recent neurophysiological studies of spatial cognition have used VE's as a research tool, in order to draw conclusions about real-world spatial cognition (Aguirre & D'Esposito, 1997; Maguire et al., 1998). Training and research applications that use VE's as substitutes for real-world spaces raise an important question about measurement: When spatial knowledge is acquired in a VE, what is the degree to which measurements taken in the VE can substitute for similar ones in the real-world? This question is especially important for training applications in which assessment of the trainee's capabilities must be made before transfer.

A recent study by Ruddle et al. (1997) focused our attention on this question. Ruddle et al. (1997) trained people to learn the spatial layout of a virtual office building and later measured their knowledge of it while they were in the VE. They then compared these results directly with those from a widely-cited study of spatial cognition in a similar but real office building (Thorndyke & Hayes-Roth, 1982). The authors suggested that measurements made in the VE are generally more indicative of a person's spatial knowledge than are measurements taken afterward or in another medium. However, it is possible that VE's introduce additional error or variance into people's spatial judgments and that performance measures acquired in a VE may not be as indicative of a person's mental representation as other types of measures. Before making direct comparisons between measurements acquired in a VE and those acquired in the real-world, we feel that it is necessary to assess empirically the relationship between measurements made in either location. In the present study, we addressed this issue by measuring people's knowledge of a VE maze when they were in the VE and then again when they were in a real-world version of it. The accuracy of participants' spatial representation was assessed by pointing, distance estimation, and map construction tasks.

Method

Participants

The participants were 24 students (12 men) enrolled in an introductory Psychology course at the University of Washington. Mean age of the participants was 20.17 years (SD = 3.21). Eighteen of the participants received extra credit for their participation. The remaining six were paid $10 per hour.

Materials and apparatus

The real-world environments were three 4.88 m x 4.88 m mazes constructed from 2.13 m black curtains hanging from an overhead grid of cables. The system of cables and curtains allowed the experimenter to reconfigure the mazes rapidly between conditions of the experiment. The ceiling of each maze was a white fabric that helped to reduce the amount of directed light entering the maze. Four prominent landmarks were placed at fixed locations in each maze. To enhance the memorability of each type of maze, this set of landmarks varied depending on which maze had been learned in the study phase of the experiment. (Real-world maze landmarks were large cardboard letters A, B, and C, and a cardboard box. Rendered maze landmarks were a gun, a sword, a violin, and a blue ball. Wireframe maze landmarks were a frog, a duck, the bust of a popular cartoon character, and a red cylinder.) Figure 1 shows ground-level images from the same location in each type of maze. Figure 2 illustrates the three maze configurations and landmark locations.

For the virtual conditions of the experiment, two versions (rendered and wireframe) of each of the three maze configurations were constructed using World Up® by SENSE8®. These VE’s were run on a Pentium® Pro 200 using a Diamond FireGL 3000 graphics accelerator board. Mazes in the rendered condition used color and fog effects (but no texture mapping) to enhance their interpretability. Wireframe mazes used no such cues; however, they were programmed so that landmarks were not visible when a maze wall lied between them and the user's viewpoint.

In addition to these VE mazes, a similar practice VE maze (and its real-world counterpart) was constructed to teach participants the experimental procedures and tasks. Approximately every four seconds, the practice VE maze switched between being fully rendered and being wireframe.

Participants viewed the VE's while sitting 38 cm from a 35 cm x 26.5 cm monitor with a resolution of 1152 x 900 (32K colors, 76 Hz refresh). The speed of this system was approximately 12.64 frames per second. Navigation in the VE was controlled with a Thrustmaster PFCS™ joystick that provided three degrees of freedom of movement: translation in either dimension along the ground as well as the ability to pan the viewpoint around the vertical axis (yaw). A button on the joystick was programmed to transport the virtual viewpoint to the entrance to the maze. Joystick sensitivity was altered between rendered and wireframe conditions to keep constant rates of translation and angular rotation (maximums were approximately 3 modeled feet per second and 180 degrees per second).

Procedure

Each participant was run individually through two separate two-hour sessions conducted within four to seven days of each other. The first session consisted of administration of the psychometric tests and training the participant on the tasks that they would perform in the second session. After taking the computer use questionnaire (see Appendix), the GZ-SO, and the matrix completion test (Salthouse, 1993), the experimenter instructed the participant about the navigational features of the joystick. Participants then practiced with the joystick until they could complete a "virtual obstacle course" in under five minutes. They were then introduced to the VE practice maze and were instructed to explore it for as much time as they felt necessary to perform the pointing and distance estimation tasks described below.

After learning the practice maze, participants were tested on their knowledge of it in three ways. They were first asked to point and estimate distances from within the VE, providing data subsequently referred to as “virtual measures.” Participants then performed pointing and distance estimation tasks in the real-world version of the maze, providing data that will be called “transfer measures.” Finally, they were asked to construct a map of the maze. During the practice session of the experiment, participants received error-corrective feedback about the accuracy of all of their responses. This feedback was withheld during the actual experiment.

When measured in the VE, each participant provided nine bearing and six distance estimations. For bearing estimations, the participant’s virtual viewpoint was placed directly in front of one of the maze objects and the participant was instructed to use the joystick to rotate his or her viewpoint so that it was facing in the direction of one of the three other targets. A set of cross-hairs superimposed on the monitor screen helped participants align the direction of their viewpoint on the target. After each bearing estimation, participants estimated the distance to the target in feet. The three symmetric distance estimations (e.g., from A to B after having already estimated the distance from B to A) were not used in subsequent analyses.

Next, the participant was escorted into the real-world version of the maze and asked to point and estimate distances from another set of locations. These data comprise participants’ “transfer measures.” Participants made eight bearing and eight distance estimations from four fixed locations to various unseen (and un-passed) objects in the maze. Bearing estimations in the real world were measured with a dial mounted on a 1.04 m stand. Participants rotated the dial to point in the direction of each target, and the direction was recorded to the nearest degree. After each bearing estimation, participants estimated the distance to the target in feet.

The first session ended with the participant constructing a map of the practice maze. Each participant was given a piece of paper containing a square (106 cm^2) coordinate grid and four small cardboard pieces, each containing a picture of one of the maze locations. Participants were asked to configure the pieces on the paper as if they were making a map, so that the pieces were arranged in the correct positions relative to each other.

When participants returned for the second session of the experiment, they were allowed a few minutes in the VE practice maze to regain their familiarity with the interface and with the experimental tasks. They then learned and were tested on the three experimental mazes: a real-world maze, a rendered desktop VE maze, and a wireframe desktop VE maze. The order of presentation of these conditions was determined randomly, and the three possible maze configurations were counterbalanced across the three environmental fidelity conditions. Participants were given as much time as they wanted (minimum of seven minutes in VE mazes and a minimum of four minutes in the real world) to explore and learn the relative locations of the landmarks. When the participant indicated that he or she had adequately learned each maze, we tested his or her knowledge of it in the real-world maze with the "transfer measure" procedures described above. An additional test of the participant's knowledge of the rendered VE maze was administered in the VE itself using the "virtual measure" procedures described above. Figure 2 shows the sighting locations and target locations for the transfer and virtual pointing tasks in each maze. Table 1 summarizes the four main types of measurements that we obtained.

After pointing and estimating distances in each maze, participants constructed a map of the maze, configuring pieces that represented the landmarks on a blank sheet of grid paper. The x-y coordinates of their map placements were recorded for further analysis (described below). After learning, pointing, estimating distances, and constructing maps of the three mazes, participants were asked which maze type they though was the most difficult to learn.

Scoring

Measures of spatial knowledge. Each individual's set of bearing estimations in a given maze was summarized by two statistics. First, the mean signed directional error was calculated to represent a measure of constant-error (or pointing bias). Secondly, a variable-error score for bearing estimations was calculated for each participant as the standard deviation of his or her set of signed bearing errors. Defined this way, variable errors correlated very highly (ca. .93) with absolute (or unsigned) bearing errors. Moreover, variable errors and unsigned bearing errors were generally within 15% of the same magnitude. Thus, the results and conclusions obtained below for variable errors do not change when we used unsigned bearing errors as an alternate measure of orientation (see Rieser, Hill, Talor, Bradfield, & Rosen, 1992 for a discussion of constant and variable errors).

We also used two measures of participants' accuracy at estimating distances. First, we computed signed distance estimation errors for each participant's distance judgments by subtracting actual distances from their estimations. Positive values for these distance errors are thus indicative of overestimation of distances. Negative values represent distance underestimation. Secondly, we used a measure of relative distance estimation accuracy by correlating participants' distance estimations with the actual values. These correlations were transformed (Fisher's r-to-z) before submitting them to statistical analyses, and were re-transformed to their original scale in order to present the results of these analyses. Note that unlike the preceding measures based on errors, higher values of relative accuracy are associated with better performance.

Maps were scored by comparing each participant's estimated configuration of landmarks with the actual configuration in the maze they had learned. The sum of the distances between corresponding locations in the actual and estimated configuration (called E) was used as a measure of map error. In order for E to be independent of the overall position, orientation, and scale of the participant's map placements, a set of geometric transformations were applied to the estimated configuration. Intuitively, each participant's map was rigidly shifted, rotated, and scaled in a way that minimized the value of E. More formally, given two n-dimensional vectors x and y that represent the x and y coordinates of the participant's n map placement locations (and vectors ξ and ψ, the coordinates of the actual locations), we determine least-squares parameter estimates $\hat{s}$ (a scaling factor) and $\hat{a}$ (a rotation) that minimize the quantity:

$$E = \sum_{i=1}^{n} \sqrt{\left(x_i - x_i'\right)^2 + \left(y_i - y_i'\right)^2}$$

where the vectors x' and y' are the rotated and scaled versions of the participant's estimated locations:

$$x_i' = s \cdot \left[x_{ci} \cdot \cos(-a) - y_{ci} \cdot \sin(-a)\right] \quad \text{and} \qquad y_i' = s \cdot \left[x_{ci} \cdot \sin(-a) + y_{ci} \cdot \cos(-a)\right]$$

with x_{ci} and y_{ci} being the centered coordinates, namely:

$$x_{ci} = x_i - \left(\frac{\sum_{j=1}^{n} x_j}{n} \right) \quad \text{and} \quad y_{ci} = y_i - \left(\frac{\sum_{j=1}^{n} y_j}{n} \right)$$

One caveat to this general procedure was observed: If the estimated scaling parameter ($\hat{s}$) was negative, the algorithm was restarted with a starting value of 180^{o} for the rotation parameter. Conceptually, a map with a negative scale is one which has been rotated, and this slight modification makes parameter estimates more consistently interpretable. It is worth noting that this procedure is, with very minor caveats, the functional equivalent of bi-dimensional regression (see Kitchin, 1996; Tobler, 1976) and serves the same function as the CONGRU algorithm.

Measures of individual differences. Scores on the GZ-SO were calculated as the percentage of correct responses, corrected for guessing. The matrix completion test was scored by calculating the percent correct over all items. Items on the computer use questionnaire assess both a person's attitude toward computers (items 3, 5, 7, 9, and 10) and prior experience with computers (items 2, 4, 6, and 8). Previous analyses have shown that these scales represent two associated factors of computer use (see Waller, 1999). Responses on items 3 and 8 were reflected so that high answers indicated either a more positive attitude toward or more experience with computers. Attitude towards computers and prior experience with computers were then measured as the average score on items from their respective scales.

Results

Possible confounds

In general, the order of presentation of the three learning environments did not have a statistically reliable effect on participants' spatial knowledge. Order effects for each dependent variable were tested in several oneway repeated measures ANOVA's. The only significant effect of order was on the map scores. Map scores derived from the last maze learned tended to be worse than those of previous mazes. The mean map error of 13.07 mm (SD = 8.18) on the last maze was significantly higher than the first (M = 8.17, SD = 5.34) or the second (M = 8.63, SD = 6.95).

Leaving the time of exposure to each maze uncontrolled represented another possible confound to the effects of interest. Significantly less time was spent learning the real-world maze (M = 423.42 s; SD = 156.22) than either the rendered VE maze (M = 740.75 s; SD = 283.39) or the wireframe VE maze (M = 651.79; SD = 268.79) ($F(2,22) = 15.75$, $p < .001$). Yet there was very little evidence that exposure time affected the accuracy of participants' spatial knowledge. Correlations between the time spent learning each maze and the measures of spatial knowledge were generally low (M = .24; SD = .17) and nonsignificant. The sign of all of these correlations indicated that more learning time was associated with greater error, indicating that there was no tradeoff between speed and accuracy.

The effect of visual fidelity and individual differences

We examined the effects of individual differences and visual fidelity in several 3 (maze type – real, rendered, wireframe) x 2 (gender) repeated measures ANCOVA's that used computer experience, attitude towards computers, scores on the GZ-SO and scores on the matrix completion test as covariates. These models were tested separately for each of the five major dependent variables: constant bearing error, variable bearing error, mean signed distance error, relative distance estimation accuracy, and map error. For two of these dependent variables – constant bearing error and map error – no effect or interaction represented in the ANCOVA was significant. In particular, averaged over all participants, mean signed (constant) bearing errors for the three measures acquired in the real world ($\underline{M}_{rendered}$ = 0.21°, SD = 23.65; $\underline{M}_{wireframe}$ = -3.99°, SD = 24.38; $\underline{M}_{real}$ = -0.70°, SD = 14.11) were not significantly different from zero, nor were they significantly different from each other ($\underline{F}(2,22) = 1.44$, $p = .26$). Visual examination of the signed bearing errors illustrates the wide range of individual differences in pointing ability. Figure 3 illustrates participants' signed bearing errors made in the real maze (transfer measures) for each of the three maze types. In addition, the figure illustrates participants' mean signed bearing errors when measured in the VE immediately after participants had learned the fully rendered maze. Signed bearing errors had an estimated reliability (Chronbach's alpha) of .41 for the virtual measure, .67 for the rendered transfer, .86 for the wireframe transfer, and .71 in the real world.

Variable bearing estimation errors were more sensitive to the effects of environmental fidelity and individual differences. Figure 4 illustrates the effect of environmental fidelity and gender on participants' variable errors in bearing estimations Averaged over all three maze types, variable errors were greater for women ($\underline{M} = 31.32^{o}$) than for men ($\underline{M} = 20.47^{o}$). Averaged over gender, variable errors were greater after having learned the rendered VE ($\underline{M}$ = 33.60^{o}, SD = 19.18) than the wireframe VE ($\underline{M} = 23.66^{o}$, SD = 13.22). Variable errors for bearing estimations were lowest after having learned the real-world maze ($\underline{M} = 20.42^{o}$, SD = 8.82). When the effect of visual fidelity and gender on variable errors was tested with a 3 (maze type) x 2 (gender) ANOVA, the effect of gender ($F(1,22) = 5.74$, $p = .026$) and environment ($F(2,21) = 14.28$, $p < .001$) were both significant. The contrast comparing variable error between the rendered and wireframe transfer conditions ($F(1,22) = 9.89$, $p = .005$) accounts for most of this effect (compared to $F(1,22) = 3.25$, $p = .085$ for the contrast comparing variable bearing error made after learning the real maze with that after learning the wireframe maze). Importantly, however, when the covariates (matrix completion and GZ-SO scores, computer attitude and experience) are included as predictors in this model, the effect of environment becomes nonsignficant ($F(2,12) = 2.72$, $p = .11$) evidently accounted for by prior computer experience ($F(1,13) = 18.42$, $p = .001$), which is highly significant. The measure of prior computer experience was particularly predictive of variable errors, correlating on average at -.55 with variable errors across the three environmental conditions. Effect sizes (η^2) for environmental fidelity, gender, and computer experience were estimated at .31, .42, and .59, respectively.

Analysis of signed distance errors showed that on average, distances were overestimated. Overestimation was greatest after having learned the wireframe maze ($M = 1.00$ feet, $SD = 1.50$), followed by the real-world maze ($M = 0.57$ feet, $SD = 1.02$) and the rendered maze ($M = 0.49$ feet, $SD = 1.02$). One-sample t-tests confirmed that signed distance errors were significantly greater than zero in each of the three environmental conditions. When the same ANCOVA model as above was applied to signed distance estimation errors, we found that scores on the GZ-SO ($F(1,13) = 4.64$, $p = .05$) and prior computer experience ($F(1,13) = 15.05$, $p = .002$) were significantly associated with distance estimation errors However, there was no significant difference in signed distance errors between the three learning environments ($F(2, 12) = 2.49$, $p = .12$) or between genders ($F(1,13) = 0.001$, $p = .98$). Signed distance errors had an estimated reliability (Chronbach's alpha) of .25 for the virtual measure, .28 for the rendered transfer measure, .72 for the wireframe transfer measure, and .51 in the real world.

The effects of visual fidelity and individual differences were similar for relative distance estimation accuracy as they were for variable bearing errors, however, gender was generally not a statistically significant predictor of relative distance estimation accuracy. A 3 (fidelity) x 2 (gender) repeated measures ANOVA showed a significant effect of environmental fidelity ($F(2,21) = 4.29$, $p = .027$), no significant gender effect ($F(1,22) = 1.67$, $p = .209$), and no significant interaction between environment and gender ($F(2,21) = 2.89$, $p = .078$). This is illustrated in Figure 5. However, when the four covariates were entered as predictors in this model, the effect of visual fidelity became nonsignificant ($F(2,12) = 0.24$, $p = .790$) while computer experience ($F(1,13) = 20.00$, $p = .001$), matrix reasoning ($F(1,13) = 16.82$, $p = .001$) and gender ($F(1,13) = 5.47$, $p = .036$) were significantly related to relative distance estimation accuracy. Effect sizes (η^2) of environmental fidelity, gender, computer experience, and matrix reasoning were estimated at .04, .30, .61, and .56, respectively.

Of those participants who expressed an opinion, seventeen out of twenty-two (approximately 77%) chose the fully rendered maze as being the most difficult of the two VE mazes to learn. A chi-squared test confirmed that this is a significantly larger proportion of responses than the 50% that would be expected by chance ($\chi^2(1) = 6.55$, $p = .01$).

Zero-order correlations between the individual difference variables, variable bearing errors, and relative distance estimation accuracy are presented in Table 2. It is clear from the table that gender (mean absolute point-biserial correlation = .36) and prior computer experience (mean absolute correlation = .43) are the most powerful single predictors of both bearing and distance estimation ability. Even after controlling for spatial ability (GZ-SO), the partial (point-biserial) correlations between gender and bearing errors from both the virtual (pr = .64 , p = .001) and transfer$_{rendered}$ (pr = .45, p = .033) measurements are statistically significant. Table 2 also shows that the GZ-SO tended to be more closely associated with pointing errors (mean absolute correlation = .21) than with distance estimation accuracy (mean absolute correlation = .10) although none of these correlations were statistically significant. The matrix completion test tended to be more closely associated with distance estimation accuracy (mean absolute correlation = .31) than with errors in bearing estimations (mean absolute correlation = .17).

Effect of measurement location

Measures of spatial knowledge that were obtained in the VE were compared with those obtained in the real world in four separate 2 (measurement location – real or virtual) x 2 (gender) ANCOVA's that used the individual difference variables as covariates. Mean signed distance errors, relative distance estimation accuracy, and mean signed (constant) bearing errors were not significantly affected by any of these factors or their interactions. However, mean signed bearing errors for measurements made in the VE ($M_{virtual}$ = -8.24°, SD = 19.24) were significantly less than zero ($t(23)$ = 2.10, p=.05), indicating a slight bias to respond too far counterclockwise when using the joystick to indicate directions.

Variable bearing errors were significantly different between men and women. Regardless of the environment in which they were tested, on average, men (M_{VE} = 20.52°, SD = 13.56; M_{real} = 24.32°, SD = 10.14) always pointed significantly more accurately than did women (M_{VE} = 66.58°, SD = 33.70; M_{real} = 42.88°, SD = 21.88) ($F(1,13) = 15.61$, $p = .002$). Men were actually more accurate pointing with the joystick in the VE than they were pointing with the dial in the real maze, while the opposite trend occurred with women. This difference made the interaction between measuring location and gender significant ($F(1,13) = 7.80$, $p = .015$). The effects of gender and measuring location on variable bearing errors are illustrated in Figure 6. Variable bearing errors were also significantly associated with prior computer experience ($F(1,13) = 7.35$, $p = .018$), with more experienced participants pointing more accurately.

Relationships between measurement types

In general, correlations between variable bearing errors in the different maze types (M= .63, SD =.15) were higher than those for relative distance accuracy (M = .28, SD = .22). This difference in mean correlations is significant ($t(10) = 3.18$, $p = .010$). Bearing errors that were measured in the VE were highly correlated with those that were measured in the real-world version of the same maze ($r = .73$, $p < .001$). Distance estimation accuracy in the VE is also highly predictive of distance estimation accuracy in the real world ($r = .63$, $p = .001$). Table 3 presents the intercorrelations, means, and standard deviations of the bearing and distance errors for the four measurement types.

Discussion

It is clear from this experiment that differences between individuals on characteristics such as gender, prior computer use and cognitive ability accounted for more variance in performance on tasks requiring spatial knowledge acquisition from a VE than did major differences in the visual fidelity of the VE. Effect sizes for individual difference variables were routinely higher than those for differences in the visual fidelity of the learning environment. Very few studies have directly compared the relative contributions of VE design variables and individual difference variables in accounting for people's ability to learn spatial information. Darken (1996) examined the use of navigational aids in VE's and found that the addition of features such as maps and grids had a significant effect on the accuracy and amount of spatial information that people were able to acquire from the VE. In general, Darken found that differences between these VE designs had a greater impact on people's performance than did cognitive differences on tests of spatial ability. Darken's study provides an interesting contrast to the present one, in which the predictive power of individual differences dominated the effects of system design variables. The difference between our findings and those of Darken probably stems from the fact that Darken was not primarily concerned with creating high-fidelity simulations of the real world. He was more interested in the utility of additional features that could be used as navigational aids. It is likely that when features such as maps are introduced into VE's specifically to augment spatial understanding, they may have a substantial impact on spatial learning regardless of the users' prior abilities. In other words, if one is interested in designing a VE that maximizes the user's ability to learn its spatial characteristics, one should focus on including aids that specifically enhance spatial knowledge – not on creating a VE that looks exactly like the real world.

Given the extreme range of visual fidelity in this experiment – from a real-world maze environment to a wireframe computer simulation – the failure to find a strong effect of visual fidelity supports our contention that for some tasks, the appearance of the simulated environment is not particularly relevant. The tasks used in the present experiment required participants to form a mental representation of a maze, and then to use it in flexible and unpredictable ways. Distances and bearings that participants estimated were never directly perceived, and had to be inferred from their prior experience. This study supports the hypothesis that such tasks are influenced more by user characteristics and cognitive variables than by variables associated with the VE system. A corollary to this hypothesis suggests that lower-fidelity desktop VE's may be just as effective as more expensive "immersed" VE's in enabling knowledge that requires conscious effort to acquire.

The effect of gender on VE spatial knowledge acquisition was particularly strong in this study. It is important to note that disorientation in virtual mazes was, on average, quite severe for women. Several (7) women – and only one man – had average bearing errors in excess of 30° for both VE transfer phases of the experiment. Moreover, gender differences in the ability to point to objects in the VE by using a joystick showed enormous gender differences (see Figure 6). The interaction between gender and measurement location is an important finding and suggests that women may be especially error-prone when measured in the VE. Interestingly, men and women performed comparably after learning the real maze (see Figs. 4 and 5). These results corroborate earlier findings that understanding the spatial characteristics of VE's may be more challenging for women than for men (Astur, Ortiz, & Sutherland, 1998; Waller et al., 1998) and that a gender-related difference in proficiency with the VE's navigational interface is a particularly important determinant of people's ability to acquire spatial information from a VE (Waller, 1999).

People generally thought that the wireframe maze was easier to learn than the rendered maze, and performance on spatial tasks was generally superior after having learned the wireframe maze than after the rendered maze. This may be because wireframe forces the user to focus on perspective cues which are generally more effective in determining depth (Surdick et al., 1997) because wireframe allows users to see spaces that would normally be occluded by walls, or because a wireframe VE forces users to attend consciously to their training. We suspect that all of these reasons influenced people's ability to acquire spatial information more accurately from the wireframe VE, although further research into this effect is clearly warranted.

One of the aims of this experiment was to address the degree to which measuring spatial knowledge in a VE adequately predicts subsequent performance in the real world. Our results suggest that a person's ability to point to unseen objects in a VE is quite predictive (r = .73) of his or her ability to do so in the real world. Similarly, the accuracy of relative distance judgments in a VE transfers fairly well to the real world (r = .63). Despite the high correlation between errors made in a VE and those made in the real world, it is important to realize that the magnitude of errors is not necessarily similar between these conditions. For example, when using the joystick to estimate directions, women on average erred about 67^{o}; however, women's pointing errors when measured in the real world averaged only about 43^{o}. It would be prudent, therefore, for investigators to account for gender differences and interface proficiency before using measurements acquired in a VE to draw conclusions about people's knowledge of a real-world space.

In general, computer-simulated environments can be effective tools for teaching people about the spatial characteristics of real-world places. Most research in this field has concentrated on examining aspects of the VE's that influence their training effectiveness and relatively little has focused on the role of user abilities and characteristics. We have found that for some tasks (particularly those involving controlled mental effort) individual differences can be much more influential on performance than very large differences in the appearance of the training system. Our message is not new, but it is one that in the renewed interest in computer-based training deserves repeating: Those who are interested in training complex cognitive tasks will do well to control for differences between people in cognitive abilities and experience with computers.

References

Aguirre, G. K., & Desposito, M. (1997). Environmental knowledge is subserved by separable dorsal/ventral neural areas. Journal of Neuroscience, 17(7), 2512-2518.

Astur, R. S., Ortiz, M. L., & Sutherland, R. J. (1998). A characterization of performance by men and women in a virtual Morris water task: A large and reliable sex difference. Behavioural Brain Research, 93(1-2), 185-190.

Bailey, J. H. (1994). Spatial knowledge acquisition in a virtual environment (Doctoral dissertation, University of Central Florida). Dissertation Abstracts International, 55(6-B), 2421.

Chance, S. S., Gaunet, F., Beall, A. C., & Loomis, J. M., (1998). Locomotion mode affects the updating of objects encountered during travel: The contribution of vestibular and proprioceptive inputs to path integration. Presence: Teleoperators and Virtual Environments, 7(2), 168 - 178.

Darken, R. P. (1996). Wayfinding in large-scale virtual worlds. The George Washington University department of Electrical Engineering and Computer Science, Doctoral dissertation.

Darken, R. P., & Banker, W. P. (1998). Navigating in natural environments: A virtual environment training transfer study. In Proceedings of the Virtual Reality Annual International Symposium, (VRAIS), pp. 12-19.

Goerger, S., Darken, R., Boyd, M., Gangnon, T., Liles, S., Sullivan, J., & Lawson, J. (1998). Spatial knowledge acquisition from maps and virtual environemtns in complex architectural spaces. In Proceedings of the 16th Applied Behavioral Sciences Symposium (pp. 6-10). U.S. Air Force Academy, Colorado Springs, CO.

Guilford, J. P., & Zimmerman, W. S. (1981). The Guilford-Zimmerman Aptitude Survey manual of instruction and interpretation. Palo Alto, CA: Consulting Psychologists Press.

Hunt, E., & Waller, D. (1999). Orientation and wayfinding: A review (ONR technical report N00014-96-0380). Arlington, VA: Office of Naval Research.

Jaeger, B. K. (1998). The effects of training and visual detail on accuracy of movement production in virtual and real-world environments. In Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting, (pp. 1486 - 1490). Santa Monica, CA: Human Factors and Ergonomics Society.

Kitchin, R. M. (1996). Methodological convergence in cognitive mapping research: Investigating configurational knowledge. Journal of Environmental Psychology, 16, 163-185.

Klatzky, R. L., Loomis, J. M., Beall, A. C., Chance, S. S., & Golledge, R. G. (1998). Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. Psychological Science, 9(4), 293-298.

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

Lloyd, R. (1989). Cognitive maps: Encoding and decoding information. Annals of the Association of American Geographers, 79(1), 101 - 124.

Maguire, E. A., Burgess, N., Donnett, J. G., Frackowiak, R. S. J., Frith, C. D., & Okeefe, J. (1998). Knowing where and getting there: A human navigation network. Science, 280(5365), 921-924.

Peruch, P., Vercher, J., & Gauthier, G. M. (1995). Acquisition of spatial knowledge through visual exploration of simulated environments, Ecological Psychology, 7, 1 - 20.

Rieser, J. J., Hill, E. W., Talor, C. R., Bradfield, A., & Rosen, S. (1992). Visual experience, visual field size, and the development of nonvisual sensitivity to the spatial structure of outdoor neighborhoods explored by walking. Journal of Experimental Psychology: General, 121(2), 210-221.

Rosanno, M. J., & Moak, J. (1998). Spatial representations acquired from computer models: Cognitive load, orientation specificity, and the acquisition of survey knowledge. British Journal of Psychology, 89, 481-497.

Ruddle, R. A., Payne, S. J., & Jones, D. M. (1997). Navigating buildings in “desk-top” virtual environments: Experimental investigations using extended navigational experience. Journal of Experimental Psychology: Applied, 3, 143-159.

Salthouse, T. A. (1993). Influence of working-memory on adult age-differences in matrix reasoning. British Journal of Psychology, 84, 171-199.

Surdick, R. T., Davis, E. T., King, R. A., & Hodges, L. F. (1997). The perception of distance in simulated visual displays: A comparison of the effectiveness and accuracy of multiple depth cues across viewing distances. Presence: Teleoperators and Virtual Environments, 6(5), 513 - 531.

Thorndyke, P. W., & Hayes-Roth, B. (1982). Differences in spatial knowledge acquired from maps and navigation. Cognitive Psychology, 14, 560-589.

Tobler, W. R. (1976). The geometry of mental maps. In R. G. Golledge & G. Rushton, (Eds.), Spatial Choice and Spatial Behavior: Geographic essays on the analysis of preferences and perception. Columbus, OH: Ohio State University Press. (pp. 69 - 82).

Waller, D. (1999). An assessment of individual differences in spatial knowledge of real and virtual environments. Doctoral dissertation, University of Washington.

Waller, D., Hunt, E., & Knapp, D. (1998). The transfer of spatial knowledge in virtual environment training. Presence: Teleoperators and Virtual environments, 7(2), 129-143.

Wilson, P. N., Foreman, N. & Tlauka, M. (1997). Transfer of spatial information from a virtual to a real environment. Human Factors, 39(4), 526 - 531.

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.

Appendix

Computer Use Questionnaire

This appendix contains the stems from the 10 items on the computer use questionnaire. Participants used a seven-point scale to rate the degree to which each statement applied to them (1 = Completely disagree; 7 = Completely agree).

1. Computers dehumanize society by treating everyone as a number.
2. I am able to learn about computers very quickly.
3. Computers are beyond the understanding of the typical person.
4. I know how to program computers.
5. I feel at ease when I am around computers.
6. I have played a lot of computer games.
7. I feel comfortable when a conversation turns to computers.
8. Kids these days know more about computers than I do.
9. I have a lot of self-confidence when it comes to computers.
10. I think working with computers would be enjoyable and stimulating.

Author Note

This research was supported by the Office of Naval Research grant N00014-96-0380 to E. H.

Correspondence concerning this article should be addressed to David Waller, Department of Psychology, University of California, Santa Barbara, CA 93106. Electronic mail may be sent to waller@psych.ucsb.edu

Table 1

Summary of the four measurement types taken from each participant. The effect of environmental (visual) fidelity was tested by comparing measurement types 2, 3, and 4. The effect of measurement location was examined by comparing types 1 and 2.

Measure type	Where is knowledge acquired	Where is knowledge measured
1. *Virtual*	Rendered VE	Rendered VE
2. $Transfer_{rendered}$	Rendered VE	Real world
3. $Transfer_{wire}$	Wireframe VE	Real world
4. *Real*	Real world	Real world

Table 2

Pearson's r correlations between measures of individual differences and measures of spatial knowledge.

		1	2	3	4	5
	1. Gender[a]					
Individual differences	2. Computer experience	-.125				
	3. Computer attitude	-.385	.670			
	4. GZ-SO	-.381	.401	.267		
	5. Matrix completion	.214	.109	.131	-.044	
Variable bearing error	Virtual	.684	-.385	-.400	-.333	-.012
	Transfer_{ren}	.494	-.435	-.532	-.252	-.190
	Transfer_{wire}	.308	-.593	-.467	-.011	-.317
	Real	.348	-.638	-.503	-.250	-.178
Relative distance accuracy	Virtual	-.291	.370	.420	.028	-.120
	Transfer_{ren}	-.438	.309	.347	.142	.243
	Transfer_{wire}	-.211	.361	.444	-.026	.423
	Real	.138	.329	-.047	.207	.463
	Mean	1.50	4.56	3.63	39.62%	83.37%
	Standard Deviation	0.51	1.02	0.99	17.50	7.20

[a] Men coded as 1; women coded as 2.

Table 3

Intercorrelations and descriptive statistics for bearing and distance errors of each measurement type.

		1	2	3	4	5	6	7	8
Variable	1. Virtual								
Bearing	2. $Transfer_{ren}$	0.726							
Error	3. $Transfer_{wire}$	0.367	0.560						
	4. Real	0.599	0.749	0.761					
Relative	5. Virtual	-.489	-.367	-.461	-.479				
Distance	6. $Transfer_{ren}$	-.591	-.570	-.481	-.608	.631			
Accuracy	7. $Transfer_{wire}$	-.204	-.438	-.669	-.442	.395	.320		
	8. Real	-.287	-.317	-.285	-.551	.166	.183	.011	
Mean (deg or r)		43.55	33.60	23.66	20.42	.57	.69	.80	.86
Std. Dev.		34.42	19.18	13.22	8.82	.36	.29	.16	.12

Figure captions

Figure 1. Views of the three types of maze used in this experiment. The same maze configuration (and observation point) is shown in the three panels. Participants explored a real-world maze (left), a rendered desktop virtual maze (center) and a wireframe desktop virtual maze (right). In each maze, participants learned the relative locations of a different set of objects.

Figure 2. The three maze configurations used in the experiment. Participants learned each of these mazes in one of three visual fidelity conditions (see Figure 1). After exploring each maze, participants were escorted through a real-world version of the maze that they learned and were asked to point to several landmarks (circles) from several sighting locations (transparent triangles). Participants also made these judgments while they were still in the rendered VE maze (solid triangles).

Figure 3. Circular plots of participants bearing errors for each type of measurement. Upright on the plot represents the true bearing to the target, with errors plotted as angular deviations from upright. Men’s data are plotted as triangles; women’s data are plotted as circles. Note that disorientation is more common among the women in the sample. Errors of 90^o or more occurs for several women (and in one observation from a man) when using the joystick to indicate directions (virtual measure – far left).

Figure 4. Mean variable bearing error for men and women when measured in the real-world version of each of the three types of mazes. Variable error is defined as the standard deviation of signed bearing errors (and corresponds closely with unsigned bearing error). Error bars represent one standard error of the mean.

Figure 5. Mean relative distance accuracy for men and women across the three visual fidelity conditions. Relative accuracy is defined as the correlation between judged and actual distances. Error bars represent one standard error of the mean.

Figure 6. Mean variable bearing error for men and women when measured in either the virtual environment or the real world. Error bars represent one standard error of the mean.

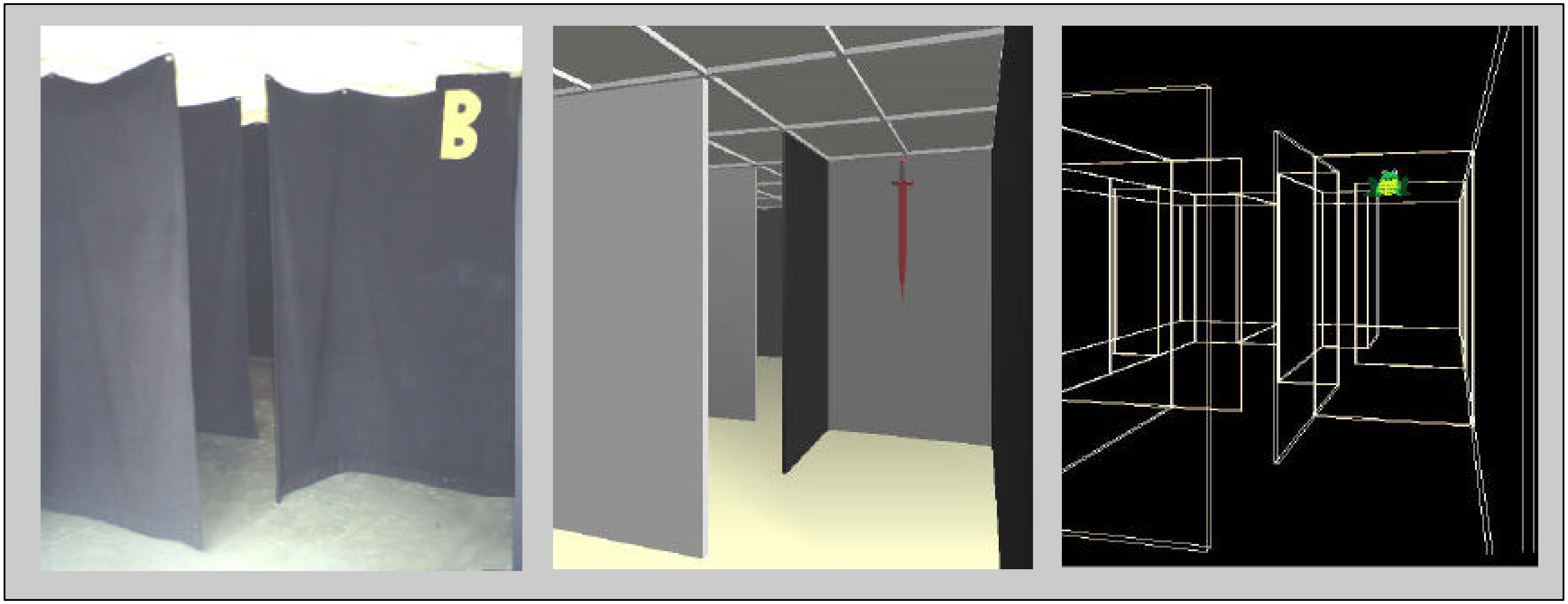


Figure 1

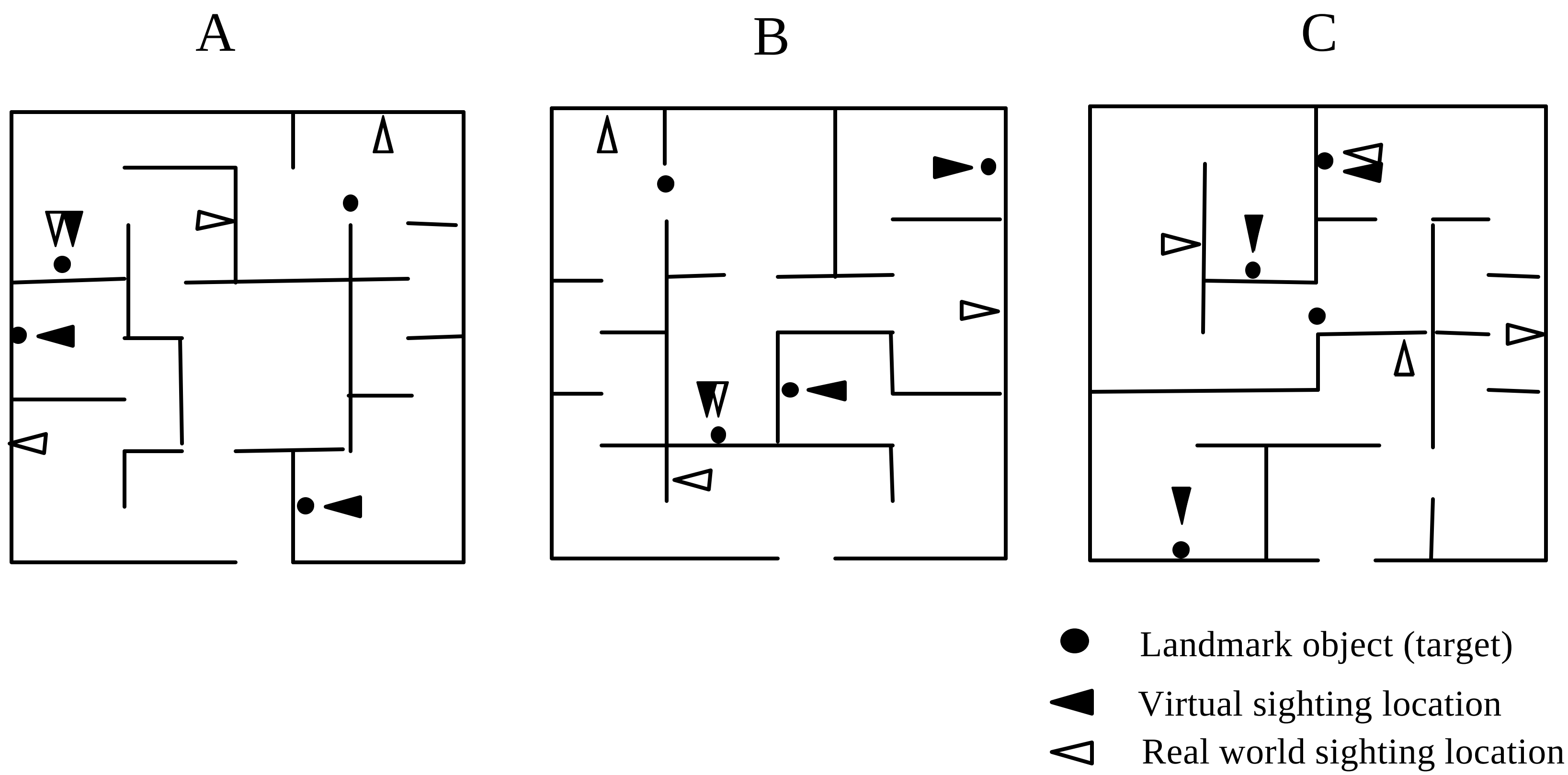

A
B
C
Landmark object (target)
Virtual sighting location
Real world sighting location

Figure 2

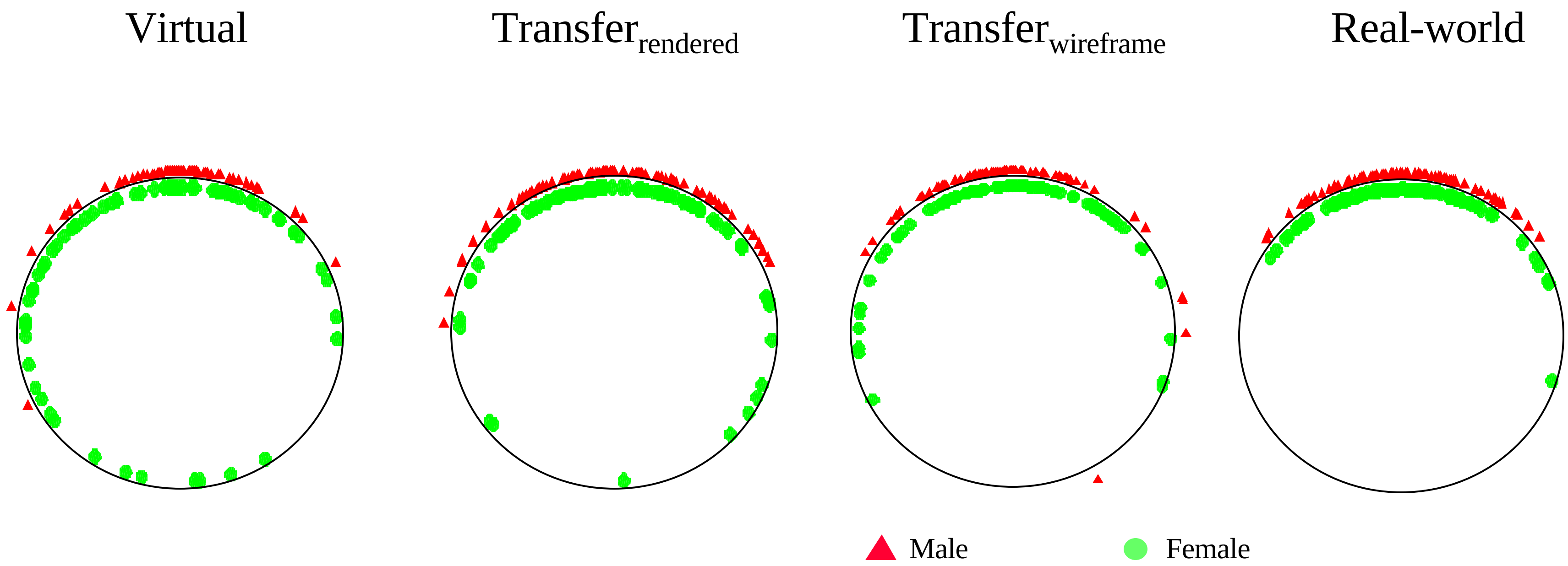

Virtual
Transfer$_{rendered}$
Transfer$_{wireframe}$
Real-world
Male
Female

Figure 3

Figure 4

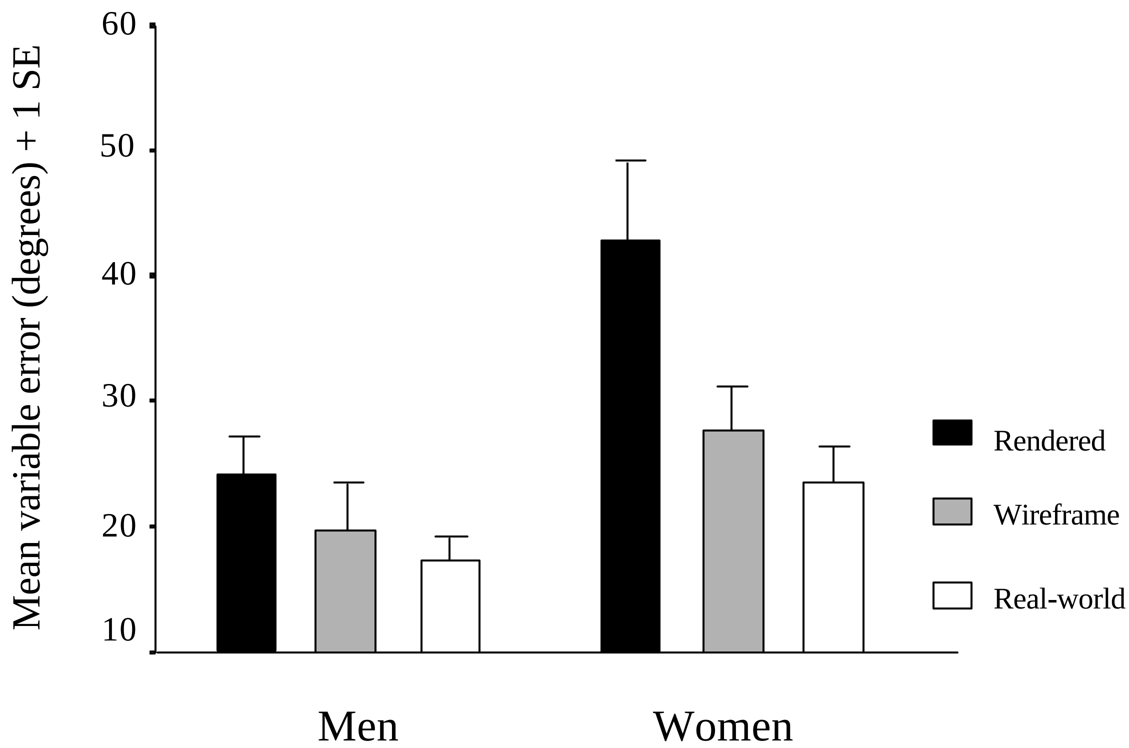

Mean variable error (degrees) + 1 SE
60
50
40
30
20
10
Men
Women
Rendered
Wireframe
Real-world

Figure 5

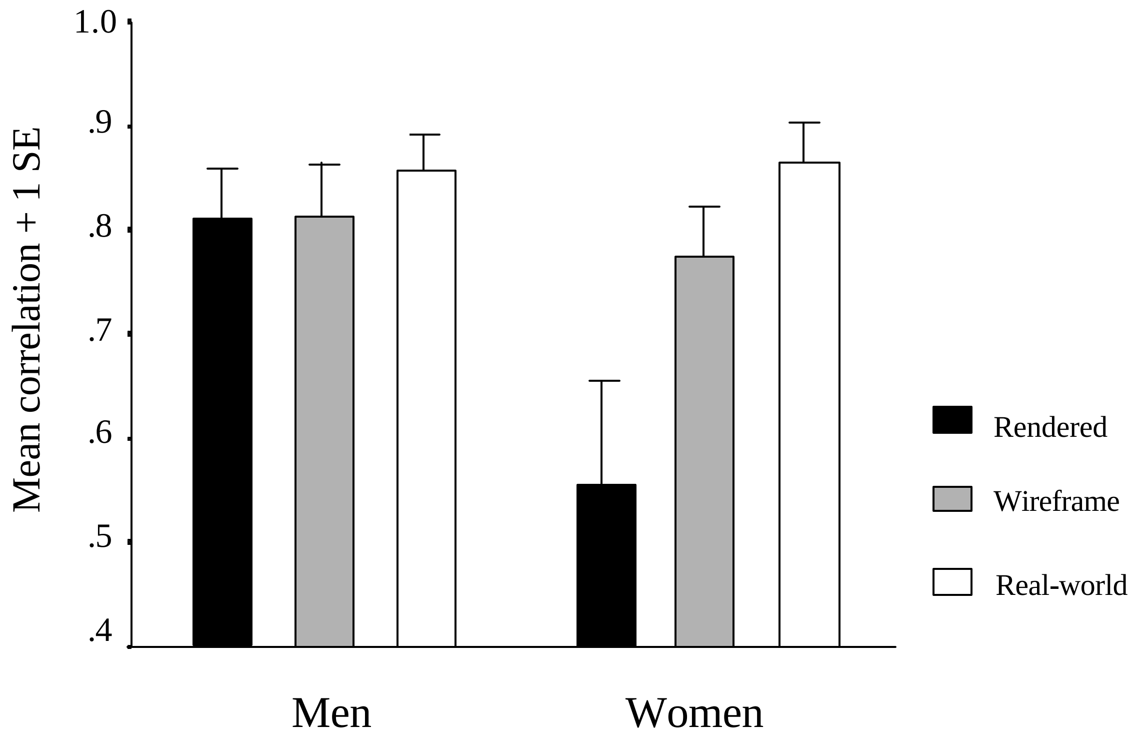
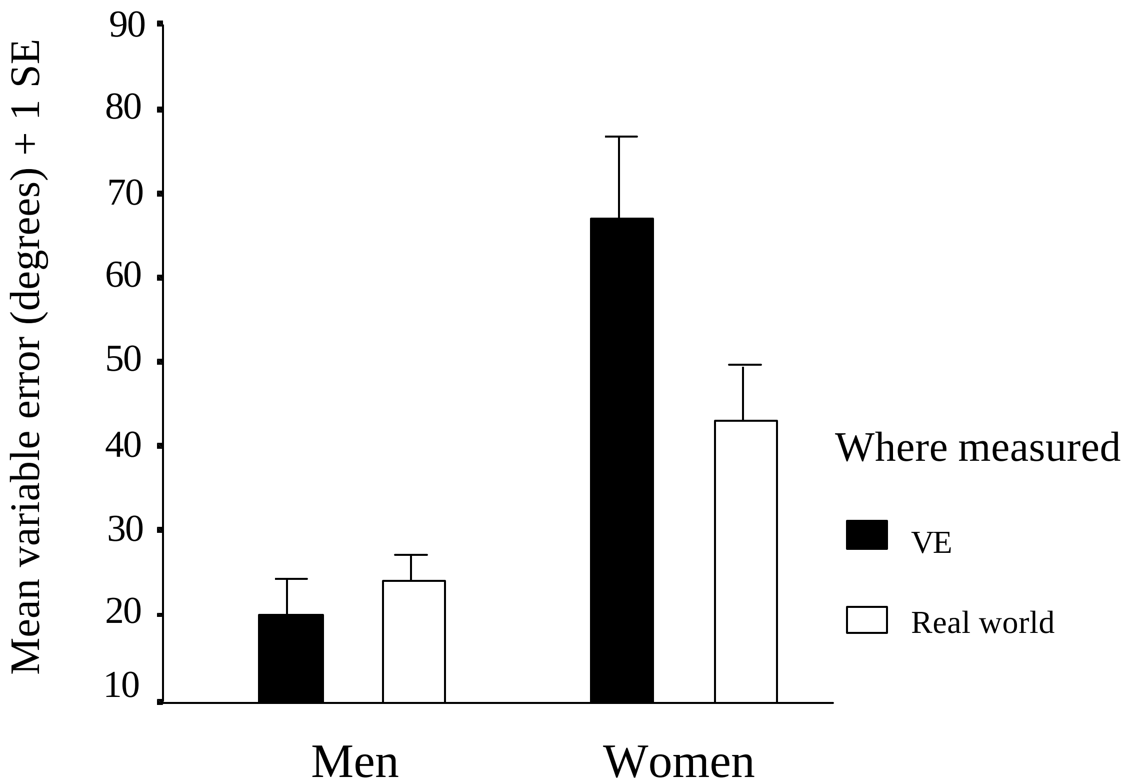

1.0
.9
.8
.7
.6
.5
.4
Mean correlation + 1 SE
Men
Women
Rendered
Wireframe
Real-world

Figure 6


Mean variable error (degrees) + 1 SE
90
80
70
60
50
40
30
20
10
Men
Women
Where measured
VE
Real world

Author biographies

David Waller is a visiting postdoctoral researcher at the University of California, Santa Barbara. He received his Ph.D. in cognitive psychology from the University of Washington in 1999.

David Knapp is a human-factors and usability consultant in Seattle, WA. He received his B.A. in psychology from the University of Washington in 1998.

Earl Hunt is a professor of psychology and an adjunct professor of computer science at the University of Washington. He received his Ph.D. from Yale University in 1960.