ELSEVIER

Journal of Environmental Psychology 24 (2004) 105–116

Journal of ENVIRONMENTAL PSYCHOLOGY

www.elsevier.com/locate/yjevp

Using virtual environments to assess directional knowledge

David Waller[a,*], Andrew C. Beall[b], Jack M. Loomis[b]

[a] Department of Psychology, Miami University, Oxford, OH 45056, USA
[b] University of California, Santa Barbara, USA

Abstract

Computer-simulated (virtual) environments offer several potential advantages over traditional means (e.g. paper and pencil tests) of assessing people's spatial knowledge of large-scale environments. We examined the relative accuracy and precision with which people estimated directions among unseen landmarks in a large familiar environment using five different methods. People estimated directions (1) in the real environment, (2) in an immersive virtual environment, (3) in a desktop virtual environment, (4) using a static visual image of the environment, and (5) using a traditional paper and pencil assessment technique. Direction estimates were more accurate and more precise in the first three conditions than in the other conditions. Errors and variances were highly correlated among the first three conditions but were not correlated among the other conditions. Direction estimates in the paper and pencil condition were affected by a response bias that has not been adequately addressed in prior literature.

1. Introduction

Assessing people's knowledge of directional relationships has been a fundamental aspect of spatial cognition research since its inception. In one of the first scientific investigations into how people mentally represent their surroundings, Trowbridge (1913) asked people to indicate the directions to several geographical features by drawing lines from the center to the perimeter of a circular piece of paper. Trowbridge examined participants' accuracy in this task and made conclusions about persistent distortions in their mental representations of space. In the nine decades of spatial cognition research following Trowbridge's study, investigators have learned a great deal about people's internal representations of space. Perhaps surprisingly, though, much of this understanding has been gained from methods that are essentially no different from those used by Trowbridge. Although today there are many methods of assessing directional knowledge (see, for example, Montello, Richardson, Hegarty, & Provenza, 1999), it is still common for investigators to assess people's directional knowledge by means of tests that require participants to manipulate or construct a line on a circle that represents a direction to a target (Kozlowski & Bryant, 1977; Gärling, Book, Lindberg, & Nilsson, 1981; Siegel, 1981; Tversky, 1981; Gärling, Lindberg, & Mäntylä, 1983; Moar & Bower, 1983; Bryant, 1984; McNamara, 1986; MacEachren, 1992; Golledge, Ruggles, Pellegrino, & Gale, 1993; Kitchin, 1996; Witmer, Bailey, Knerr, & Parsons, 1996; Rossano & Moak, 1998; Rossano, West, Robertson, Wayne, & Chase, 1999; Kozhevnikov & Hegarty, 2001). In the last decade, this method of testing has been adapted to computers (see for example Shelton & McNamara, 2001). Yet even in these computer applications, the task of indicating a direction by means of a line and a circle is essentially unchanged since Trowbridge's time. After briefly discussing the strengths of this traditional means of assessing directional knowledge, we will discuss several important weaknesses. We will then offer a new method for assessing spatial knowledge (including, primarily, knowledge of directional relationships) that involves computer-simulated environments, and provide experimental evidence that this method can be a more valid and reliable means of testing people on their knowledge of directions in a large familiar environment than traditional assessments in which directions are indicated by lines drawn in circles.

1.1. Traditional assessments of directional knowledge

Convenience and experimental control are probably the two main advantages to using traditional assessments of directional knowledge. Testing people at different locations in an actual environment can be

*Corresponding author. Tel.: +1-513-529-4929; fax: +1-513-529-2420.
E-mail address: wallerda@muohio.edu (D. Waller).

0272-4944/$ - see front matter
doi:10.1016/S0272-4944(03)00051-3

prohibitively difficult for experimenters, particularly when investigating knowledge of large-scale geographic spaces. For example, Moar and Bower (1983) were interested in the degree to which people's knowledge of directions between landmarks in their town contained mutual inconsistencies. Instead of taking their participants to these locations and observing their knowledge of directions in situ, Moar and Bower used a paper and pencil test (that we call the "direction circle") that assessed knowledge of relative directions (see Fig. 1). For each pointing response, participants were shown a circle. They were to imagine being at the location labeled at the center of the circle, facing in the direction of the location labeled at the top of the circle. Participants then drew a line from the center to the perimeter of the circle that represented the proper direction to a target location. As most of the citations in the previous paragraph attest, these direction circles are a fairly common method of assessment in the spatial cognition literature. In general, this method has the advantage of allowing investigators to examine people's behavior in the laboratory, where observations are relatively easy to acquire. These assessments may also allow the experimenter a convenient means of controlling possible confounds involved with exposure to the testing environment. If one is examining spatial learning, for instance, one may not want to allow people to navigate between testing locations within the environment during testing. Another advantage of paper and pencil assessments is that they allow researchers to test large groups of participants at once.

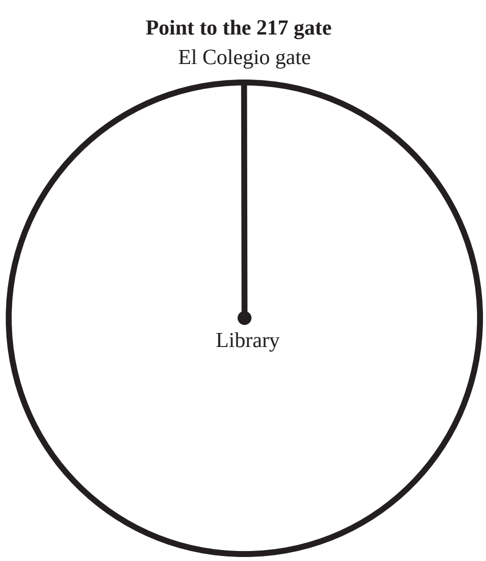

Fig. 1. Example of the "Direction circle" method for assessing knowledge of directions. The user imagines that he or she is at the location labeled at the center of the circle, facing in the direction of the location labeled at the top of the circle (the "orienting location"). He or she then draws a line from the center to the perimeter of the circle that represents the direction of the target location given above, in the instructions. In some applications of this method, the vertical line is not present.

Yet paper and pencil tests of directional relationships are not without problems. One general problem with testing participants' knowledge of an environment while they are in the lab is that such testing places heavy demands on participants' ability to imagine the testing space. In many cases, the mental abilities or processes associated with imagining a known environment are probably quite unrelated to the *spatial* processes or behavior in which the investigator is most interested. Indeed, people may perform quite differently on tests that require imagining an environment and those in which they have visual access to the testing site. For example, Rossano and Moak (1998) exposed people to a computer-simulation of a large campus environment and then asked them to point from various testing locations to several target locations in the environment. One group of participants was shown images of the testing location prior to pointing. The other group was not shown these images, and were asked to imagine each testing location prior to pointing. Rossano and Moak found that the people who were shown images of the testing locations pointed to targets with significantly greater accuracy than those who were required to imagine the locations (see also Thorndyke & Hayes-Roth, 1982).

Another problem with paper and pencil assessments concerns the possibility of response biases that arise from how the testing stimulus is perceived. For example, Huttenlocher and her colleagues have shown that memory for locations inside a circle tends to be biased toward the centers of the quadrants that can be formed by superimposing horizontal and vertical axes through the center of the circular stimulus (Huttenlocher, Hedges, & Duncan, 1991). An alternative possibility for a response bias associated with pointing methods that use depictions of circles suggests that responses tend to be drawn toward the horizontal axis of the response circle. For example, in the study cited above by Moar and Bower, the investigators found that participants' directional estimates were typically biased toward 90° (see also Sadalla & Montello, 1989). Although the authors regarded this bias as indicating that their participants had an inaccurate mental representation of space, it is possible that, to some degree, it was an artifact of a response method that required people to interact with a circular stimulus. A secondary issue with the present research will be to determine whether responses on traditional paper and pencil assessments of directions are subject to consistent biases, and if so, to determine the nature of these biases.

1.2. Virtual environments

In the last few years, we have developed a method of assessing our participants' knowledge of directional relationships that we feel offers several advantages over paper and pencil assessments. Our technique involves testing in the laboratory, with photo-realistic simulations of the actual testing locations that are displayed on a "virtual environment" (VE) computer system. These environmental simulations are created from 360° panoramic photographs (see Fig. 2) that are applied to the inside surface of a computer-modeled cylinder. The participant's viewpoint is fixed at the center of the cylinder (see Fig. 3). In most of our experiments, participants view these panoramas while wearing a head-mounted display (HMD) on which is mounted a small lightweight sensor that keeps track of the participant's facing direction. The system enables participants to interact with these simulated environments much as they would in the real world—by turning their heads to view different directions. Unlike other panoramic techniques (e.g. Wiley & Proctor, 1997) that do not use computer-simulated environments, our system enables experimenters to record with great precision how and where people look in these simulations, as well as measuring participants' latencies to perform pointing tasks. We feel that this system offers a high degree of ecological validity in the task of estimating directions, yet also provides the experimenter the convenience and control associated with testing in the laboratory.

It is important to note that, because our panorama assessment technique provides participants with a visual simulation of the testing location, there is another fundamental difference between it and the more traditional paper and pencil methods. With our system, the visual information present during testing serves to orient the participant at the simulated testing location. It is thus meaningful to instruct participants simply to turn their heads and face in the direction of a target. As if we measured direction estimates in the real world, our system provides the investigator with participants' estimates of *absolute* bearings to target locations. On the other hand, paper and pencil assessments commonly offer the investigator only estimates of *relative* bearings. The reason for this is that paper and pencil assessments must typically orient the participant at the testing location by asking them to imagine facing another known location in the environment (e.g. "El Colegio gate" in Fig. 1). In this case, bearing estimates can only be made relative to the imagined heading.

In our study, we compared tests of directional knowledge that use computer-simulated environments with those that use more traditional paper-and-pencil methods. To compare the bearing estimates of the different methods, we needed to convert the relative bearings of the Direction-circle test into absolute bearings. This was accomplished by adding the participant's estimated relative bearing (to the target location) to the true absolute bearing (from the viewing location to the orienting location). For example, if the estimated relative bearing to the target location obtained from the circle diagram was 187° and the true bearing of the orienting direction from the viewing location was 270°, the estimated absolute bearing to the target location was 97°, which is equal to $187+270$ (mod 360). The estimated absolute bearing obtained in this way reflects two sources of error that are present in the participant's estimate: Error in imagining the direction from the viewing location to the orienting direction and error in imagining the relative bearing.

We were interested in comparing the relative accuracy and precision of our method of assessing directional knowledge with that of the more traditional paper and pencil approach. Five groups of participants were tested on their knowledge of the directional relationships between six locations in a large, familiar campus environment. One group (IMMERSED) was tested with our immersive VE system, while another (DIRECTION-CIRCLE) was administered a traditional paper and pencil test. If we find a large performance difference

Fig. 3. Illustration of how the computer simulates the location shown in Fig. 2. The panoramic photograph is applied to the inside surface of a computer-modeled cylinder. The user's viewpoint is then placed in the center of the cylinder.

Fig. 2. Sample 360° panoramic photograph used in the present experiment.

between these two methods, we will want to understand why. For example, accuracy differences between pointing responses in our system and those from paper and pencil assessments may derive solely from the increased visual information available to people tested in the VE. If this is true, then performance may be quite similar between paper versions of these panoramas (e.g. Fig. 2) and those viewed in the VE. To examine this hypothesis, we tested a third group of participants (PAPER-PANORAMA) using paper versions of these panoramas. Alternatively, if we find that our VE system allows more accurate responses than do traditional methods, it may be because our method immerses participants in a simulated environment, giving them the impression that they are surrounded by its features. Perhaps the realism of this immersion, which requires users to turn their heads in order to experience the environment, is important in allowing participants to be more accurate or precise. To assess this possibility, we added a fourth condition (DESKTOP) in which participants made their direction estimations using the same VE system; however, they viewed the environments on a desktop version of the VE, rather than with an HMD. Finally, a fifth group of people (REAL) were tested on their knowledge of directions at selected locations in the actual environment. This condition provided us with an ecologically valid baseline against which we could compare the other assessment techniques.

2. Method

2.1. Participants

Participants were 240 undergraduate students (120 men and 120 women) at the University of California in Santa Barbara. Their average age was 20.7 (SD = 2.4). Four gender-balanced groups, each with 30 participants, were assigned to the four main conditions of the experiment (IMMERSED, DIRECTION-CIRCLE, PAPER-PANORAMA, and DESKTOP). Most of these 120 students (106) participated in the experiment in return for credit in their introductory Psychology course. The remaining 14 were paid $10 for their participation. Another group of 120 students provided data for the REAL condition of the experiment. These students composed four gender-balanced groups, each with 30 participants, tested at each of four real-world locations. Most of the participants (98) in the REAL condition were passers-by at three popular campus locations. The remaining 22 participants in the REAL condition were students in an Introductory Psychology class that participated in order to receive credit in their class. These participants were all taken to the fourth pointing location (that typically had few willing and eligible passers-by). Participants in the HMD group were run through the experiment approximately six months before those in the other groups.

2.2. Environment and materials

The environment that was tested consisted of six locations on or near the campus of the University of California in Santa Barbara (see Fig. 4). Each location was near a prominent landmark (Recreation Center [A], Student Center [B], Library [C], a local restaurant [D], and two campus entrances [E and F]), and no location was visible from the others. Positions of these locations (and hence their absolute pairwise bearings relative to true North) were determined with a Garmin GPS (Global Positioning System) receiver (model 12XL), accurate to within 15 m.

Fig. 4. Map of the campus and six locations tested in the present experiment.

Participants in the IMMERSED and DESKTOP groups were tested with computer-generated simulations of each of the six locations. These simulations were created from 360° panoramic photographs of each location that were applied to the inside surface of a computer-modeled cylinder. The user's viewpoint was placed in the center of the cylinder. The technique is adapted from Chen and Williams (see Chen, 1995), and it effectively creates a correct perspective view of each location's vista. Participants in the IMMERSED condition viewed these panoramas using a V8 head-mounted display (HMD) from Virtual Research. The display provided 640×480 pixel resolution with a 50° horizontal by 38° vertical field of view and identical images displayed to each eye. Mounted on the HMD was an inertially based orientation tracker (Intersense model IS-300). The tracker was used to update the orientation of the visual image as the participant moved his or her head, as well as to record the facing directions in the pointing task described below. Participants in the DESKTOP condition viewed the panoramas on a $32\,\text{cm} \times 24\,\text{cm}$ monitor, at a resolution of 640×480 pixels. These participants used mouse commands to change their viewing direction. Left-clicking the mouse on the right half of the screen rotated the viewpoint clockwise; left-clicking on the left half of the screen rotated the viewpoint counterclockwise. The computer rendered these scenes using a Pentium III chipset and an NVIDIA GeForce2 MX graphics card, updating the graphics and display at 72 Hz. Randomization and presentation of the stimuli as well as the collection of pointing estimates was controlled through a scripting facility in the Python programming language, supplemented with a utility module specifically for virtual environment applications.

Participants in the PAPER-PANORAMA group were shown $24.5\,\text{cm} \times 3.0\,\text{cm}$ high-resolution color panoramic photographs of the testing locations. One such panorama is illustrated in Fig. 2. For each testing location, two panoramic photographs were shown on the same sheet of $27.9\,\text{cm} \times 21.6\,\text{cm}$ paper. These two photographs presented the same image of the testing location; however the information presented at the center of one photograph was presented at the edges of the other (i.e. the two pictures differed by 180°). A $27.9\,\text{cm} \times 21.6\,\text{cm}$ sheet of transparent acetate was aligned over the two photographs, and participants were instructed to mark their direction estimates on a numerical scale that was printed on the acetate.

We assembled 30 test booklets, each containing the same 24 items, for participants in the DIRECTION CIRCLE group. Each item presented the name of a testing location in the center of a 5 cm diameter circle (see Fig. 1). An orienting location (e.g. "El Colegio Gate" in Fig. 1) was written at the top of each circle, and a target location (e.g. "217 Gate" in Fig. 1) was written above the orienting location. Each of the six testing locations was always paired with the same orienting location. The remaining four locations—those that were neither testing nor orienting locations—served as targets. For each testing location, the orientating location was chosen so that mean absolute pointing error from the testing to the orienting location in the IMMERSED condition was the least across all of the targets. (Participants' mean error in the IMMERSED condition in pointing to each of the orienting locations was uniformly extremely low, averaging just 1.83°). Across all of the testing locations, each of the six locations served exactly once as an orienting location.

At the four testing locations that were most accessible to pedestrians (Library, Recreation Center, University Center, and the restaurant), we measured pointing estimations from participants in the REAL condition with a Raytheon personal digital handbearing compass (model Z074) with an accuracy of 3° RMS and a display precision of 1°.

2.3. Procedure

Participants were run individually through two parts of the experiment. In the first part, the experimenter assessed the participant's familiarity with the to-be-tested locations by showing them the six panoramas in a fixed order in the HMD. (Participants in the REAL condition were shown photographs of each of the target locations.) When the participant identified each location (which never failed to happen) the experimenter gave it a short memorable name that was used as a label for the rest of the experiment.

In the second part of the experiment, participants were asked to estimate directions to and from the six campus locations. In general, testing procedures were similar in each of the experimental conditions; however, differences in the testing media necessitated slight differences between the conditions.

Participants in the REAL condition performed their pointing task at only one testing location. At each testing location, the experimenter asked these participants to hold the compass level and at arm's length, and to point in turn to each of the five target locations. The experimenter recorded each bearing estimate as it was made. The order of testing locations was randomized separately for each participant in the REAL condition. To provide us with an estimate of measurement error in the REAL condition, participants were also asked at one location to point to the same distant visible landmark (a light post). All compass bearings (measured with respect to magnetic north) were converted to true north bearings prior to analysis.

Participants in the IMMERSED and DESKTOP conditions were shown the appropriate panoramas during testing. For each panorama, they were told to pretend that they were at the depicted location and asked to point to each of the target locations by either facing them (IMMERSED) or turning their viewpoint using the mouse (DESKTOP). While participants were pointing, a red arrow appeared in the center of the screen to assist them with lining-up their estimated direction. Direction data (north-referenced bearings) for these conditions were automatically written to an external computer file for later analysis.

Participants in the PAPER-PANORAMA condition marked their pointing response on a sheet of transparent acetate that covered the panorama. Participants were instructed that they were free to use either panorama on the underlying page, and that if their answer was close to the edge of the panorama, that they should use the other image to mark their response. The experimenter recorded each of the participants' direction estimates immediately after it was made. The experimenter then cleaned the acetate and gave it back to the participant for the next estimate.

Participants in the DIRECTION-CIRCLE condition made direction estimates by drawing, for each item, a line from the center to the perimeter of a circle printed in their answer booklet. These direction estimates were subsequently measured and converted to absolute bearings by assuming that the participant had on each item, imagined their orientation with no error.

In each of the four non-Real-world conditions (IMMERSED, DESKTOP, PAPER-PANORMA, and DIRECTION-CIRCLE), participants estimated directions between all pairs of locations. In the IMMERSED, DESKTOP, and PAPER-PANORAMA conditions, 30 pointing questions were blocked into six sets of five—one set for each testing location. In the DIRECTION-CIRCLE condition, items were also blocked by testing location; however, because these items used one potential target as an orienting location, they were grouped into six groups of four questions. The order of testing locations, as well as the order of targets for each testing location, was randomized separately for each participant.

3. Results

The results were clear: Pointing precision and pointing accuracy were relatively high and were extremely similar among the REAL, IMMERSED, and DESKTOP groups. The PAPER-PANORAMA and DIRECTION-CIRCLE groups tended to be more variable and error-prone, and their errors were not closely associated with those made in the other conditions. Fig. 5 illustrates pointing responses in each of the conditions for two representative location/target pairs. Pointing error for each participant is plotted as the angular deviation from the vertical reference line. The relatively low intersubject variability in the REAL, IMMERSED, and DESKTOP conditions is clear from the figures. The right panel of Fig. 5 also illustrates one particular item (at A pointing to F) for which participants exhibited a left (counterclockwise) bias in the REAL, IMMERSED, and DESKTOP conditions. (Significant left- or right-pointing biases were found on approximately 30% of the real-world items. Thus, the figure is not representa-

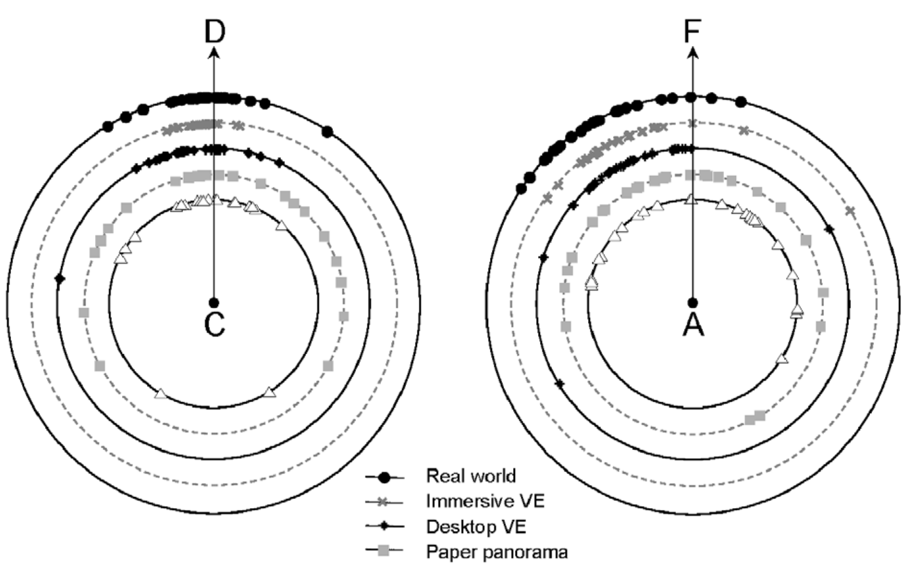

Fig. 5. Pointing errors for each condition at two sample location/target pairs. The testing location is labeled in the center of the circle, and the target is shown above, at a bearing of 0°. Pointing errors are represented as angular deviations from the vertical line. Left. When pointing from location C (see Fig. 4) to location D, performance was relatively unbiased. Right. A left pointing bias can be seen in the Real world, Immersive, and Desktop conditions when pointing from location A to location F. This bias is not as clear in the other conditions.

tive of whatever biases there were for other items.) The bias on this item is not so apparent in the other conditions. The generality of these observations was confirmed with the statistical tests reported below.

3.1. Accuracy

For each testing condition and at each location/target pair, we computed mean pointing error using circular statistics (see Batschelet, 1981) as the average deviation of participants' pointing direction from the true bearing. As described above, the pointing directions of participants in the DIRECTION-CIRCLE condition were computed based on the assumption that the orienting direction was imagined without error. Absolute pointing error, which represents the group's overall accuracy regardless of a left- or right-bias, was then computed as the absolute value of the mean pointing error. Table 1 presents mean absolute pointing error for each group. A oneway ANOVA confirmed that the differences in mean absolute pointing error (averaged over all location/target pairs) among the testing methods were significant ($F(4, 133) = 8.06$, $p < 0.001$). Planned comparisons (using Welch's correction for inhomogeneity of variance) indicated that absolute pointing error for neither the IMMERSED ($t(39) = 0.13$, $p = 0.90$) nor the DESKTOP ($t(43) = 0.35$, $p = 0.73$) conditions was significantly different from that of the Real-world group. However, absolute pointing error for the PAPER PANORAMA ($t(40) = 3.58$, $p = 0.001$) and the DIRECTION-CIRCLE ($t(42) = 2.35$, $p = 0.024$) groups were significantly different from that of the REAL group. Similarly, absolute pointing error for the IMMERSED condition was not significantly different than that of the DESKTOP condition ($t(56) = 0.54$, $p = 0.593$), but it was significantly less than that of the PAPER-PANORAMA condition ($t(36) = 3.80$, $p = 0.001$).

The lower diagonal of the correlation matrix in Table 1 presents associations among the five testing conditions for mean (signed) pointing error. Pointing questions that resulted in left/right biases among participants in the real world also tended to result in left/right biases among participants in the IMMERSED and DESKTOP conditions. Errors made in the REAL, IMMERSIVE, and DESKTOP conditions are highly associated with each other, correlating on average at 0.85. Errors made in the PAPER-PANORAMA and DIRECTION-CIRCLE conditions generally did not correlate highly with each other. Nor did they correlate highly with the other three conditions. Scatterplots representing the data in the first two columns of Table 1 are presented in Fig. 6. The plotting symbols in Fig. 6 are ellipses that extend to the limits of the 95% confidence intervals of the mean error in each dimension. The figure makes evident the close linear association among the REAL, IMMERSED, and DESKTOP conditions, as well as the relatively high precision in these conditions. Moreover, the ellipses in these conditions generally lie near the 45° line, indicating that errors among these conditions were not only linearly related, but that they were, in addition, nearly equivalent.

3.2. Precision

We also examined between-participant variability in pointing errors as a measure of precision among participants. For each condition and at each location/target pair, variability was computed as the standard error of signed pointing errors. High standard errors indicate greater variability among participants, and may thus represent overall uncertainty or imprecision in people's knowledge of the true direction to the target. Table 1 presents mean standard errors for each condition. As with the previous analysis, standard errors were lower in the REAL, IMMERSED, and DESKTOP conditions. (To put these quantities in

Table 1
Intercorrelations and descriptive statistics (mean and SE) for pointing errors in each condition

	1	2	3	4	5
1. Real world	—	0.80**	0.78**	0.74**	0.24
2. Immersive VE	0.87**	—	0.66**	0.43*	0.36
3. Desktop VE	0.86**	0.82**	—	0.49**	0.02
4. Paper panorama	0.24	−0.04	0.20	—	0.19
5. Direction circle	0.06	0.13	0.26	−0.29	—
Mean absolute pointing error (degrees)	7.30	7.03	8.07	20.93	12.79
Mean SE of pointing errors (degrees)	3.64	2.67	4.54	9.98	7.84

Note. Numbers in the lower diagonal of the correlation matrix represent associations among the conditions in signed pointing error. Numbers in the upper diagonal represent associations of standard errors.

*$p < 0.05$.

**$p < 0.01$.

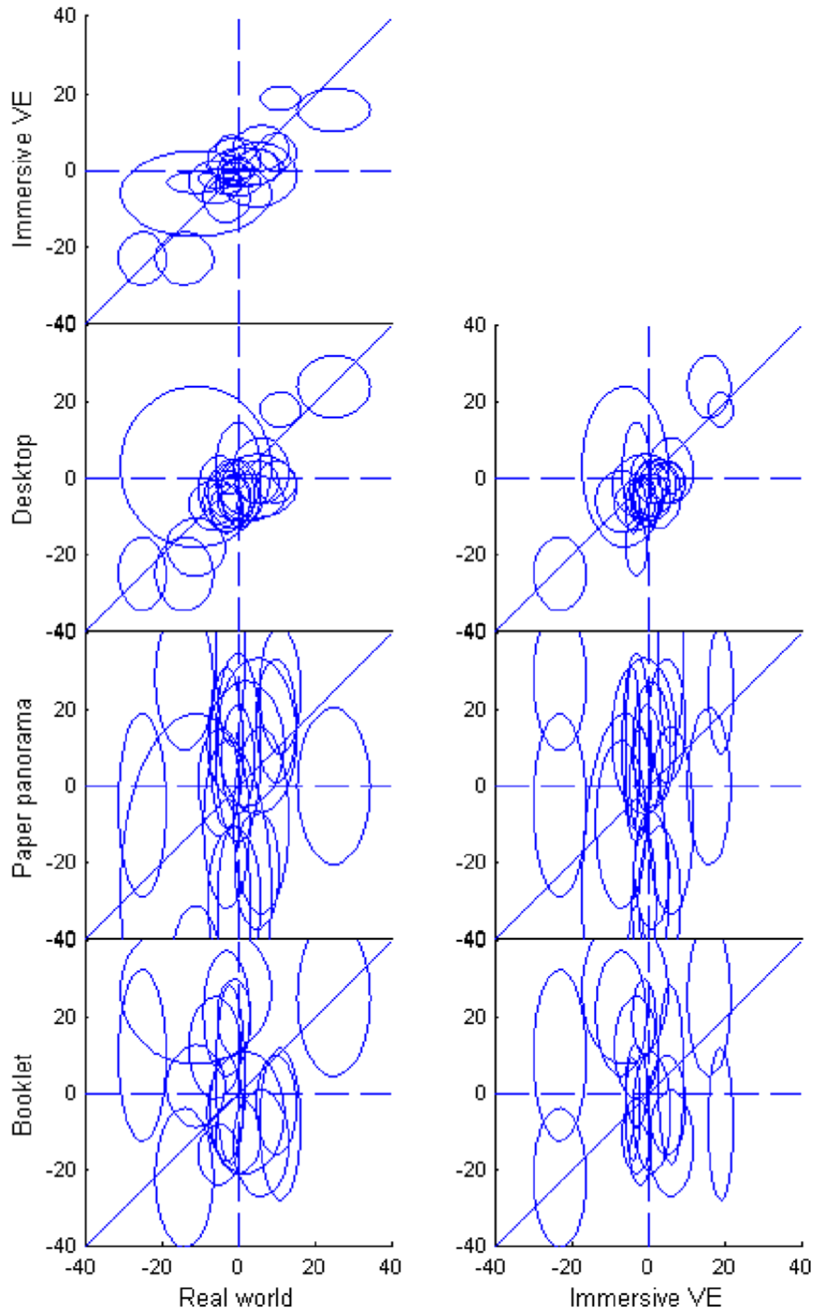

Fig. 6. Scatterplots illustrating the relationships in signed pointing error between Real world condition and the others (left column) and between the Immersive condition and the others (right column). Ellipses are centered at the mean pointing error for the conditions, and extend in both dimensions to the 95% confidence intervals of the mean. Shown in each panel is line of slope 1, representing equal errors between the two conditions.

perspective, the standard error of pointing errors to a visible target in the real world was 1.27°.) A one-way ANOVA on standard errors confirmed that there was a significant effect of testing method ($F(4, 133) = 54.82$, $p < 0.001$), much of which was accounted for by significant differences between the REAL and PAPER-PANORAMA ($t(46) = 8.66$, $p < 0.001$) conditions and between the REAL and DIRECTION CIRCLE conditions ($t(42) = 6.79$, $p < 0.001$). Planned contrasts also confirmed that standard error was significantly lower in the IMMERSED condition than in either the REAL ($t(27) = 2.19$, $p = 0.038$), DESKTOP ($t(48) = 5.24$, $p < 0.001$), or PAPER-PANORAMA ($t(34) = 11.40$, $p < 0.001$) conditions.

The upper diagonal of the matrix in Table 1 presents the intercorrelations among the five testing conditions of the intersubject variability (standard error) of pointing errors. Pointing questions that resulted in high/low dispersion among participants in the real world also tended to result in high/low dispersion among participants in the IMMERSED, DESKTOP, and PAPER-PANORAMA conditions. Thus, standard errors in the REAL, IMMERSED, DESKTOP, and PAPER-PANORAMA conditions were all significantly intercorrelated. On the other hand, intersubject variability for pointing questions in the DIRECTION-CIRCLE condition was not significantly related to that of the other conditions.

3.3. Response bias with direction-circles

We were interested in whether this lack of a relationship between participants' precision in the DIRECTION-CIRCLE condition and that of participants in the other conditions could be partly attributed to a response bias created by the circular stimulus. For these analyses, we considered bias as participants' mean pointing responses in the DIRECTION-CIRCLE condition, relative to the mean indicated direction in the REAL condition. For those location/target pairs for which we did not collect real-world data, we used the indicated directions from the IMMERSED condition. Fig. 7 presents these data. It can be seen from this figure that many of the errors in the DIRECTION-CIRCLE condition appear to be biased toward either the horizontal line (representing indicated directions that were drawn to be more horizontal than those implied by Real-world participants) or toward the diagonal lines (representing a bias toward the center of the quadrants formed by horizontal and vertical axes). To analyse the significance of these biases, we classified each mean response in the DIRECTION-CIRCLE condition as either being representative of a given bias or not. For example, the mean indicated direction from location C to location B in the REAL condition was a bearing of approximately 230° (clockwise from true north), which is 71° counterclockwise from the 301° bearing between location C and location F (see Fig. 4). In the DIRECTION-CIRCLE condition, one item required participants to imagine themselves at C facing F. For this item, the mean indicated direction to B was 59° counterclockwise from upright. Thus, the 71° relative bearing estimated in the REAL condition was estimated at 59° in the DIRECTION-CIRCLE condition. This particular item was then classified as indicating a bias toward the center of the response circle's orthogonal axes (toward ±45° and ±135°) and was also classified as not indicating a bias toward the response circle's horizontal axis (±90°). Eighteen of the 24 items in the DIRECTION-CIRCLE condition (75%) were classified as indicating a bias toward the center of the quadrants (i.e. towards the dashed lines in Fig. 7). This was significantly more than the 50% that would be expected by chance ($\chi^2(1, N = 24) = 6$, $p = 0.014$). Similarly, 71% of the items were classified as showing a bias toward the horizontal axis. This was also a higher percent than would be expected by chance ($\chi^2(1, N = 24) = 4.17$, $p = 0.041$).

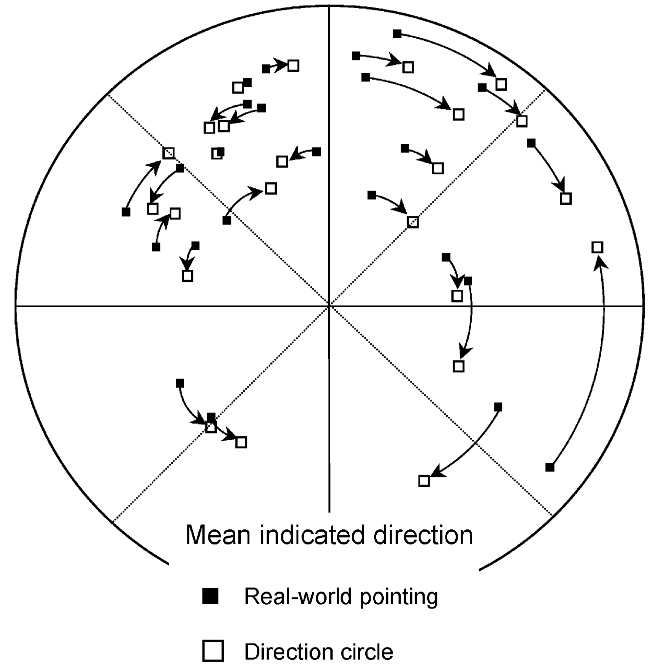

Fig. 7. Indicated directions to targets in the Direction circle condition, relative to those in the Real world condition (For 10 judgments for which real-world data were not available, the indicated direction in the Immersive condition is shown). Radial distance of response pairs was assigned arbitrarily. Many responses in the Direction circle condition are biased toward either the center of the quadrants (dotted lines) or toward the horizontal axis (solid line).

4. Discussion

Our results clearly indicate that assessing people's directional knowledge in a computer simulation of the testing environment provides more accurate and more precise estimates than those acquired from a more traditional assessment method that uses direction circles. Estimates made during computer testing were more accurate than those made from paper and pencil assessments in two ways. First, they were more veridical—closer to the actual value in the real environment. On average, people pointed with only about 7.5° of error in both the IMMERSED and DESKTOP computer simulation conditions, whereas mean pointing errors made with direction circles were slightly, though significantly higher, at about 13°. This finding indicates that the common method of assessing directional knowledge by means of direction circles (at least as they were employed in the present study) underestimates people's knowledge of their environment.

Second, and perhaps more importantly, estimates made during computer testing were more accurate than those in other conditions in the sense that they were more similar to the directions that people indicated in the actual environment. The degree of correspondence between the REAL, IMMERSED, and DESKTOP conditions was striking. Errors were highly correlated and nearly equal among these conditions, whereas in the PAPER-PANORAMA and DIRECTION-CIRCLE conditions, errors were not associated with those made in the actual environment. These results provide compelling evidence that assessing directional knowledge of a familiar area by means of computer-simulated environments can be more ecologically valid than the more common paper and pencil methods. Importantly, this gain in ecological validity is not matched by a loss of experimental control. On the contrary, a computerized assessment such as ours allows investigators to measure more dimensions of participants' behavior—and to measure them more precisely and accurately—than traditional paper and pencil assessment techniques. Although not used in the present study, a computerized system such as ours offers investigators the opportunity to measure participants' head movements over time (see for example Smart, Stoffregen, & Bardy, 2002), eye fixations (Wilder, Hung, Tremaine, & Kaur, 2002), and latencies (see for example Waller, Montello, Richardson, & Hegarty, 2002) while they perform these tasks.

It is also important to note that our computerized methods of assessing people's knowledge yielded significantly reduced variability among participants. Individual differences in spatial and environmental cognition are generally quite pronounced (for a review, see Hegarty & Waller, 2003). These differences often lead to intersubject variability that can render other effects—those in which experimental psychologists and behavioral geographers may be more interested—extremely difficult to detect. Assessment methods and tools that reduce variability in performance among people thus offer an important advance in the field of spatial cognition.

We believe that much of the improved performance of participants who were tested in computer simulations, relative to those who responded with direction circles, derives from the rich, photorealistic visual features provided by the simulation.[1] Having an abundance of accurate visual information available during testing relieved participants in the computer conditions from having to rely on their imagination and their memory of the testing locations. It is likely that visual cues such as landmarks and streets offered computer-tested participants directional information that Direction-circle participants did not imagine (or could not imagine accurately). As a result, our method measured participants' knowledge of the layout of the environment independently of their ability to imagine or recall the appearance of the environment. For many researchers, our method may thus result in greater construct validity, providing a more focused assessment of spatial knowledge, unencumbered by the variability associated with other mental processes.

Yet if the presence of detailed visual information during testing was necessary for participants to achieve relatively high pointing accuracy and precision, it clearly was not sufficient. Participants in the PAPER-PANORAMA condition had access to the same visual information that was available to the IMMERSED and DESKTOP participants; however, their performance was the most error-prone and the most variable of any group in the experiment. The high variability for errors in the PAPER-PANORAMA condition suggests that many participants were confused by the stimuli. Indeed, it was not uncommon for pointing errors in this condition to approach 180°—representing a completely reversed response and complete disorientation. Errors of this magnitude were extremely rare in the IMMERSED and DESKTOP conditions. Informal interviews with the participants in the PAPER-PANORAMA condition confirmed the idea that for many people, the stimuli required effort to interpret. Evidently, presenting 360° of directional information simultaneously in a person's forward field of view can be quite confusing for many people. Perhaps alternate ways of presenting panoramic information on paper could be devised that allow more accurate performance than what we observed in the PAPER-PANORAMA condition and DIRECTION-CIRCLE conditions (see, for example, Kelly, Beall, & Loomis, 2002). Although the paper and pencil testing medium per se may not be responsible for the relatively greater inaccuracy and imprecision of participants tested with it, the medium certainly does not facilitate presentation of ground-level panoramic information. These considerations lead us to conclude that ultimately, computer-simulated environments are valuable because they easily offer both photorealistic visual information as well as the ability to interact with this visual information naturally, such that parts of the environment are revealed progressively, as one actively controls his or her viewing direction.

Our data do not support the idea that the added kinesthetic and vestibular information that is available from viewing an environment in a head-tracked head-mounted display allows people to indicate directions more accurately than when such kinesthetic and vestibular information is not available. Participants in the IMMERSED condition interacted with the simulated environments very naturally—by moving their heads to change their viewing direction. However, performance in this condition was generally no more accurate than that in the DESKTOP condition, in which

[1] As pointed out by an anonymous reviewer, it should be noted that, in addition to its relatively greater fidelity, VE testing may be more effective than paper and pencil assessments because it is a more interesting and motivational medium with which participants can engage.

participants changed their viewing direction using mouse commands. Although the IMMERSED condition showed less variability among direction estimates than the DESKTOP condition, this effect was not particularly large. Our results suggest that during testing in familiar, landmark-rich environments, the addition of proprioceptive and vestibular feedback in computer-simulated environments has little effect on performance accuracy.

Our finding that responses in the DIRECTION-CIRCLE condition were significantly biased complements several studies that have found similar patterns of responses Most of the responses drawn by participants in the DIRECTION-CIRCLE condition were oriented (relative to the direction indicated by participants who pointed in the Real-world condition) closer to a 45° or 135° angle than would have been expected by chance. Huttenlocher and her colleagues have previously documented this same pattern of bias in people's reports of memories for the location of objects presented within a circular display (Huttenlocher et al., 1991). These investigators have provided solid evidence that the existence of such biases derive from the combination of two distinct levels of coding in memory. Coding the location of a stimulus in terms of a category (e.g. a particular quadrant) affects metric reports of location, producing biases in the direction of the category's prototype (e.g. the center of the quadrant). In addition to this ''categorical'' bias, data from the DIRECTION-CIRCLE condition also showed evidence of another bias: More of the responses were drawn closer to horizontal than would have been expected by chance. In other words, relative bearings were generally estimated as being closer to right angles than they were estimated in the REAL condition. As we mentioned in the Introduction, a similar pattern of results was found by Moar and Bower (1983) (see also Chase & Chi, 1981; Tversky, 1981; Sadalla & Montello, 1989) who interpreted this bias as indicating that participants had an inconsistent and distorted mental representation of directional relationships in their environment. We should note that our stimuli were not designed to separate the effects or to examine the relative strengths of the ''categorical'' bias toward 45° and the ''distortion'' bias toward 90°. Most of the stimuli were non-diagnostic of differences between these two biases. However, it is clear from our data that, in general, participants' performance in the DIRECTION-CIRCLE condition was biased in ways that have been previously documented.

Importantly, though, we feel that neither Huttenlocher et al.'s nor Moar and Bower's account of these biases provides a satisfactory explanation of the present results. Whether bias results from combination rules used in memory or from a distorted mental representation implies that bias is due solely to memory processes or structures. However, in our experiments, participants in the computer conditions did not show these biases, and presumably they were accessing similar memories of the environment as participants in the DIRECTION-CIRCLE condition. Indeed, it is somewhat nonsensical to suggest that pointing performance in the computer-simulated conditions would be governed by a bias related to imposed orthogonal axes, for no such stable, identifiable axes were present in these conditions. Why, though, did participants in the DIRECTION-CIRCLE condition show these biases? We suggest that, in addition to biases that arise as a result of multiple levels of coding or as a result of a distorted representation, pointing performance can be biased by the nature of the testing stimulus itself. Tasks that require people to respond to a circular stimulus may be more likely to evoke a response bias because such stimuli can easily be organized in terms of orthogonal axes. Additionally, it is possible that tasks such as direction circles require people to employ relatively greater working memory resources, such as imagining the learned environment and imagining the appropriate heading. Having to use these resources may tax working memory and may thus cause people to rely more on heuristics that result in biased responses. Other tasks that require pointing in a natural environment may be less taxing of cognitive resources, and hence less governed by a stimulus-driven response bias.

Several investigators have noted the potential of computer-simulated environments for conducting basic research in human spatial cognition and navigation as well as their potential for training people on tasks that require spatial knowledge (Wilson, 1997; Peruch & Gaunet, 1998; Loomis, Blascovich, & Beall, 1999). The present paper added to this list of potential benefits of VE technology by showing that computer-simulated environments also can be used effectively for measuring human behavior and assessing human abilities.

Acknowledgements

This research was supported by NSF grant #9873432 to the third author. We thank Eyal Aharoni, Daniel Haun, and Tijana Vujovic for help with data collection. We thank Yvonne Lippa for helpful comments on previous versions of this manuscript.

References

Batschelet, E. (1981). *Circular statistics in biology*. Academic Press: London.

Bryant, K. J. (1984). Methodological convergence as an issue within environmental cognition research. *Journal of Environmental Psychology*, *4*, 43–60.

Chase, W. G., & Chi, M. T. H. (1981). Cognitive skill: Implications for spatial skill in large-scale environments. In J. H. Harvey (Ed.), *Cognition, social behavior, and the environment* (pp. 111–135). Hillsdale, NJ: Erlbaum.

Chen, S. E. (1995). Quicktime VR—an image-based approach to virtual environment navigation. *Proceedings of SIGGRAPH'95*, Los Angeles, CA (pp. 29–38).

Gärling, T., Book, A., Lindberg, E., & Nilsson, T. (1981). Memory for the spatial layout of the everyday physical environment: Factors affecting rate of acquisition. *Journal of Environmental Psychology*, *1*, 263–277.

Gärling, T., Lindberg, E., & Mäntylä, T. (1983). Orientation in buildings: Effects of familiarity, visual access, and orientation aids. *Journal of Applied Psychology*, *68*, 177–186.

Golledge, R. G., Ruggles, A. J., Pellegrino, J. W., & Gale, N. D. (1993). Integrating route knowledge in an unfamiliar neighborhood: Along and across route experiments. *Journal of Environmental Psychology*, *13*, 293–307.

Hegarty, M., & Waller, D. (in press). Individual differences in spatial abilities. In P. Shah, & A. Miyake (Eds.), *Handbook of Visuospatial Thinking*. New York: Cambridge University Press.

Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review*, *98*, 352–376.

Kelly, J. W., Beall, A. C., & Loomis, J. M. (2002). *Accurate judgments of exocentric direction in large-scale space*. Paper presented at the meeting of the Vision Sciences Society, Sarasota, FL.

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

Kozhevnikov, M., & Hegarty, M. (2001). A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory and Cognition*, *29*, 745–756.

Kozlowski, L. T., & Bryant, K. J. (1977). Sense of direction, spatial orientation, and cognitive maps. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 590–598.

Loomis, J. M., Blascovich, J. J., & Beall, A. C. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behavior Research Methods, Instruments, & Computers*, *31*, 557–564.

MacEachren, A. M. (1992). Learning spatial information from maps: Can orientation specificity be overcome. *Professional Geographer*, *44*, 431–443.

McNamara, T. P. (1986). Mental representations of spatial relations. *Cognitive Psychology*, *18*, 87–121.

Moar, I., & Bower, G. H. (1983). Inconsistency in spatial knowledge. *Memory and Cognition*, *11*, 107–113.

Montello, D. R., Richardson, A. E., Hegarty, M., & Provenza, M. (1999). A comparison of methods for estimating directions in egocentric space. *Perception*, *28*, 981–1000.

Peruch, P., & Gaunet, F. (1998). Virtual environments as a promising tool for investigating human spatial cognition. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, *17*, 881–899.

Rossano, 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.

Rossano, M. J., West, S. O., Robertson, T. J., Wayne, M. C., & Chase, R. B. (1999). The acquisition of route and survey knowledge from computer models. *Journal of Environmental Psychology*, *19*, 101–115.

Sadalla, E. K., & Montello, D. R. (1989). Remembering changes in direction. *Environment and Behavior*, *21*, 346–363.

Shelton, A. L., & McNamara, T. P. (2001). Systems of spatial reference in human memory. *Cognitive Psychology*, *43*, 274–310.

Siegel, A. W. (1981). The externalization of cognitive maps by children and adults: In search of ways to ask better questions. In L. S. Liben, A. H. Patterson, & N. Newcombe (Eds.), *Spatial representation and behavior across the life span: Theory and application* (pp. 167–194). New York: Academic Press.

Smart, J. L., Stoffregen, T. A., & Bardy, B. G. (2002). Visually induced motion sickness predicted by postural instability. *Human Factors*, *44*, 451–465.

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

Trowbridge, C. C. (1913). On fundamental methods of orientation and 'imaginary maps'. *Science*, *38*, 888–897.

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

Waller, D., Montello, D. R., Richardson, A. E., & Hegarty, M. (2002). Orientation specificity and spatial updating of memories for layouts. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *28*, 1051–1063.

Wilder, J., Hung, G. K., Tremaine, M. M., & Kaur, M. (2002). Eye tracking in virtual environments. In K. M. Stanney (Ed.), *Handbook of virtual environments: Design, implementation, and applications* (pp. 211–222). Mahwah, NJ: Erlbaum.

Wiley, L. P., & Proctor, R. W. (1997). Self-localization within a simulated environment. *International Journal of Cognitive Ergonomics*, *1*, 43–62.

Wilson, P. N. (1997). Use of virtual reality computing in spatial learning research. In N. Foreman, & R. Gillet (Eds.), *A handbook of spatial research paradigms and methodologies*, Vol. 1: *Spatial cognition in the child and adult* (pp. 181–206). Hove, England, UK: Psychology Press/Erlbaum.

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.