Variations in cognitive maps

Understanding individual differences in navigation

Weisberg, Steven M.; Schinazi, Victor R.; Newcombe, Nora S.; Shipley, Thomas F.; Epstein, Russell A.

Published in:
Journal of Experimental Psychology: Learning Memory and Cognition

DOI:
10.1037/a0035261

Variations in Cognitive Maps:

Understanding Individual Differences in Virtual and Real World Navigation

Steven M. Weisberg[1], Victor R. Schinazi[2], Nora S. Newcombe[1],

Thomas F. Shipley[1] and Russell A. Epstein[2]

[1]Temple University

[2]University of Pennsylvania

Author Note

Steven M. Weisberg, Nora S. Newcombe and Thomas F. Shipley, Department of Psychology, Temple University, Philadelphia, PA 19122 USA; Russell A. Epstein and Victor R. Schinazi, Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104.

Work on this project was funded by a grant to the Spatial Intelligence and Learning Center from the National Science Foundation, #SBE-0541957, SBE-1041707.

Correspondence concerning this article should be addressed to Steven M. Weisberg, Department of Psychology, Temple University, Philadelphia, PA 19122 USA. Email: smweis@temple.edu

The final article is available, upon publication, at: https://doi.org/10.1037/a0035261

Abstract

There are marked individual differences in how effectively people construct cognitive maps in both real world and virtual environment settings (e.g., Blajenkova, Motes, & Kozhevnikov, 2005; Chai & Jacobs, 2010; Ishikawa & Montello, 2006; Wen, Ishikawa, & Sato, 2011). However, these individual differences are poorly understood and difficult to assess except by self report. In this study, we studied spatial learning in a virtual environment (VE) consisting of several buildings arrayed along two disconnected routes and investigated how acquisition of spatial knowledge relates to self-reported sense of direction on the Santa Barbara Sense of Direction Scale (SBSOD) and psychometric measures of spatial ability. We also compared the pattern of results in the VE to findings from the real world environment on which the VE was modeled (Schinazi, Nardi, Newcombe, Shipley, & Epstein, in press). Results suggest that more accurate pointing between buildings on different routes correlates with self-reported navigation ability, validating the SBSOD. However, the SBSOD did not distinguish between participants who had more or less difficulty with pointing accurately between pairs of buildings on the same route. Comparing results from the VE to the real world revealed similar patterns of learning, despite higher overall accuracy in the real world. Thus, we confirm the existence of individual differences in the ability to construct a cognitive map of an environment, characterize these differences more completely than is so far possible with self-report, and introduce an objective behavioral measure of navigation ability that can be used as a research tool.

Keywords: Virtual navigation; individual differences; spatial cognition; memory

1. Introduction

Navigation is fundamental for human survival. In evolutionary terms, organisms that could not use spatial representations to re-locate food, shelter, or mates would be unlikely to survive to reproduce. In the present day, techniques and tools have eased much of the burden of navigation, yet for some people finding one's way in complex environments remains a challenging aspect of human life. Past research has demonstrated that people vary along a continuum from expert navigational ability (e.g., Maguire, Woollett, & Spiers, 2006) to serious navigational impairments (Iaria & Barton, 2010), with many gradations in between (e.g., Hegarty, Richardson, Montello, Lovelace, & Subbiah, 2002). Navigation ability relies upon a complex set of skills (e.g., attention, landmark recognition, path integration, etc.), any one or several of which could underlie these individual differences (c.f. Wolbers & Hegarty, 2010).

This study investigates how aspects of navigation ability vary between people by considering how self-report and psychometric measures of spatial ability relate to navigation performance in a virtual environment (VE). Previous research using the same layout of buildings, but conducted in the real world, demonstrated substantial individual differences in both navigation performance and neuroanatomy known to support such performance (Schinazi, Nardi, Newcombe, Shipley, & Epstein, in press).We therefore were able to compare the pattern of results from the VE to the real world and to further explore the characteristics of these individual differences.

There are several reasons to be interested in individual differences in navigation performance. Practical questions focus on improving navigation ability and supporting successful way-finding, and understanding the neural and behavioral-level correlates of individual differences is relevant to those goals. Theoretical interest relates to the controversy concerning the existence of a “cognitive map” proposed by many spatial researchers (e.g., Montello, 1998; Siegel & White, 1975; Tolman, 1948) but critiqued by others (e.g., Foo, Warren, Duchon, & Tarr, 2005). This debate usually revolves around whether cognitive maps do or do not exist. But a recent study suggested that the potential answer could be that there are individual differences in the construction of cognitive maps. Specifically, Ishikawa and Montello (2006) found marked individual differences in the formation of accurate spatial representations, in a study in which they drove participants in a car around a novel environment once a week for 10 weeks. The participants’ task was to learn the locations of buildings along two non-connecting routes. People varied substantially in their ability to learn the environment, as measured by their drawings of locations of buildings (sketch maps) and their accuracy in pointing to locations around the environment. Most participants performed either consistently well or consistently poorly across ten trials, neither improving nor declining, despite also being taken on a connecting route between the two separated routes from the fourth trial on.

One concern about this study, however, is the fact that participants were passively exposed to the environment. To address this issue, Schinazi et al. (in press) led walking participants around a real world environment, once a week for three weeks. Two non-connecting routes were learned the first week and two separate connecting routes were learned each of the two subsequent weeks (see Figure 1). Across the three weeks, Schinazi et al. found that most participants improved, ultimately forming reasonably accurate spatial representations. In particular, subjects with larger posterior hippocampi were significantly better at an off-site pointing task in which they were required to imagine standing next to a learned building, imagine facing down the route, and point to all other buildings. This relation supports findings that

The final article is available, upon publication, at: https://doi.org/10.1037/a0035261

London taxi drivers have larger hippocampi than control subjects (Maguire, Woollett, & Spiers, 2006), and as compared to their own hippocampi before learning (Woollett & Maguire, 2011). Thus, there is strong evidence that individuals differ in their ability to learn new environments, and that there are neural correlates of these differences.

Studying these individual differences is challenging, however, if one must conduct lengthy and demanding spatial learning experiments to assess these differences. Luckily, individual differences in navigation ability can also be explored using self-report measures. Participants who self-report high versus low sense of direction differ in their ability to point to unseen targets, even in studies using just one Likert-scale item (Kozlowski & Bryant, 1977; Sholl, 1988). More recently, two self-report measures of navigation ability have been widely used and validated. Mary Hegarty and her colleagues have developed the Santa Barbara Sense of Direction (SBSOD) scale (Hegarty et al., 2002). Francesca Pazzaglia has also developed a different sense of direction scale (Pazzaglia & DeBeni, 2001). The SBSOD, used in this study, is a unidimensional measure of sense of direction, while Pazzaglia's scale distinguishes between landmark and survey learning preferences. Both scales consist of Likert-scale items that measure the participants' ability and liking for navigation-related tasks.

Both scales are highly reliable and have been shown to be well correlated with tasks that require some form of survey knowledge. Specifically, the sense of direction scales have been shown to correlate with large-scale navigation tasks, but not smaller-scale spatial abilities (Hegarty et al., 2002; Kozhevnikov & Hegarty, 2001), with selecting fewer, but more reliable, landmarks in real world navigation tasks (Ishikawa & Nakamura, 2012), with greater reliance on spatial than verbal working memory in an interference paradigm (Wen, Ishikawa, & Sato, 2011), with better learning in a desktop VE (Pazzaglia & Taylor, 2007), and with greater suppression of

The final article is available, upon publication, at: https://doi.org/10.1037/a0035261

activity in the parahippocampal cortex, a cortical region that supports representations of scenes, upon repeating views of the same building from different viewpoints (Epstein, Higgins, & Thompson-Schill, 2005).

These studies strongly suggest that good and poor navigators may differ in important ways. However, there are several reasons to be wary about exclusive reliance on self-report measures of navigation ability. First, in self-reporting navigation ability, participants may sample from a small number of recent events. Heth, Cornell, and Flood (2002) found that self-report sense of direction was correlated with a route-reversal task when it was administered after, but not before, the navigation task. Second, self-report measures are unlikely to be reliable for measuring improvement or change in navigation ability, because people are likely to regard their sense of direction as a stable trait or be unaware of gradual or relatively small changes. Measuring how navigation ability changes after training protocols requires an objective and consistent form of assessment. One purpose of the present study was to evaluate learning in desktop computer virtual environment navigation as a compromise between the difficulties of gathering real world data and the potential shortcomings of self-report.

How comparable are findings from desktop VE and real world navigation? Non-immersive VEs clearly eliminate many of the sensory and motor cues that occur during actual movement which are important in forming spatial representations (Mittelstaedt & Mittelstaedt, 2001). In a recent review of active and passive navigation, Chrastil & Warren (2011) argue that the bulk of the evidence is consistent with the hypothesis that information deriving from active movement clearly contributes to more accurate spatial knowledge. However, both VE and real world navigation are supported by many of the same spatial processes. While the availability of sensory information is much richer in the real world compared to the virtual environment, and thus performance is typically better, visual information has been shown to be sufficient (Chrastil & Warren, 2011). In addition, individual differences in learning from virtual navigation have been found to be related to individual differences in active navigation (Hegarty, Montello, Richardson, Ishikawa, & Lovelace, 2006; Montello, Waller, Hegarty, & Richardson, 2004).

The VE and paradigm used in this study matched, as closely as possible, a previous real world study (Schinazi et al., in press). Careful attention was taken in order to match the spatial layout and the location of buildings and other objects in the VE to the real world environment. Participants in both studies learned two routes each containing four buildings to be learned, followed by two routes that connected the first two routes. Motion was self-generated in both the virtual environment (using the mouse and keyboard) and real world (by walking). However, while in the real world the full gamut of sensory cues was available, the desktop VE limited cues to visual ones. Thus, sensory cues were richer in the real world, a factor which could have led to more accurate spatial knowledge due to more accurate encoding and because multimodal information is likely more engaging in terms of attention and motivation (Chrastil & Warren, 2011; Witmer & Singer, 1998). An unavoidable difference between the two methodologies required learning to be shorter in the VE compared to the real world. Participants in the VE learned both routes and both connecting routes in about 25 minutes before completing the navigation tasks. Participants in the real world study learned the environment over three separate sessions, each occurring one week apart, for a total of three hours, completing the navigation tasks after each session.[1] Both of these differences led us to predict that learning would be more accurate overall in the real world.

[1] In the first session, participants learned the eight buildings along two disconnected routes. In the second session, participants reviewed the original two routes, and then learned a connecting route between them. In the third

After participants learned the names and locations of buildings around the environment, their spatial knowledge of the environment was tested in two ways: an onsite pointing task (participant was placed next to one of the buildings and cued to point to the other buildings by their names), and a model building task. We expected to find individual differences in overall performance on both tasks, but better performance overall in the real world study. Moreover, we expected the self-report measure of navigation ability, more so than other self-report and psychometric measures, to be sensitive to individual differences in performance. We also examined whether individual differences in navigation as measured in VE went beyond the variability captured by self-report.

2. Method

2.1 Participants

Forty-nine undergraduate students at Temple University (23 male) participated in the experiment in return for course credit or $10.

2.2 Materials

The experiment was administered on an Alienware computer running Windows 7 64 Bit with an Intel Core i7 960 @ 3.20 GHz processor and NVidia GeForce GTX 460 graphics card. The VE was displayed on a 32 cm by 52 cm LCD monitor with a refresh rate of 59 Hz and a field of view of 60°. The viewing distance from the screen was approximately 50 cm and the screen had a resolution of 1920 x 1200. The VE was identical to the spatial layout of the buildings in a real-world college campus (see Schinazi et al., in press), created using Unity3D (www.unity3d.com), and populated with buildings and other objects which were modeled in Google Sketchup and freely available online (http://sketchup.google.com/). While the VE

session, participants reviewed the first two routes, and learned a second connecting route. Participants completed the navigation tasks after each session. See Schinazi et al. (in press) for details.

buildings differed in architectural design compared to the real world, effort was made to match them on saliency and place them in the precise spatial location of the physical buildings in order to keep the relative distance and angles identical. Effort was also made to ensure this was generally true for non-building objects (e.g., signs, trees, trash cans, etc.), which were matched from VE to real world, particularly when those objects could be used as reference points. Care was also taken to match the buildings from VE to real world on prominence and distinctiveness within the environment. For specific comparisons, see Figures 1 and 2, below and Schinazi et al., (in press).

2.3 Procedure

Efforts were made to keep the procedure of the VE study identical to the real world study. In the real world, participants first learned the names and locations of four buildings along two main routes and were then tested using four measures: 1) an onsite pointing task wherein participants stood by one building and pointed to other buildings; 2) an offsite pointing task wherein participants pointed to buildings by imagining the appropriate vantage point in a distinct environment; 3) a distance estimation task wherein participants estimated relative distances between pairs of buildings; and 4) a model building task wherein participants sketched the configuration of buildings. Participants were then brought back a week later, learned one of two connecting paths, and were tested on all four measures again. Finally, participants were brought back a third time, learned a second connecting path, and were tested on the four measures a third time (for detailed methodology see Schinazi et al., in press).

For the VE, pilot data revealed that, using an identical procedure to the real world study, the 3 sessions could be conducted in approximately three hours of testing time, including all psychometric measures. As in the real world study, all participants improved throughout the three sessions, despite variable performance overall. Preliminary data revealed that participants in the VE quickly fatigued after the second session of testing, rendering the second and third sessions extremely unreliable. To improve the reliability of the data, we altered the procedure from the real-world study by front-loading the learning phase. Participants first learned all routes and connecting paths, and then completed the navigation tasks just once. Pilot testing also revealed a very high correlation between onsite and offsite pointing, and extremely high errors for the distance estimation task. The onsite pointing task was chosen over the offsite pointing task because the former transferred better onto a computer interface. Because the offsite pointing and distance estimation tasks took a long time, they were eliminated from the final procedure.

The final procedure used in the VE study took approximately 1 hour. Participants first completed psychometric and self-report measures, either on the computer or on paper. Next, participants familiarized themselves with the VE and learned first the two main routes, then the two connecting routes. Participants learned four buildings along each of the two main routes which were presented in a counter-balanced order. Participants then learned the two connecting routes which were also counter-balanced, but always presented after both main routes (see below and Figure 1 for more detail). Participants then completed the onsite pointing task, the model building task, were debriefed and released. Each of the tasks is described in full below.

2.4 Psychometric Measures

2.4.1 Santa Barbara sense of direction scale. (SBSOD; Hegarty et al., 2002) .The SBSOD consists of 15 items which participants respond to on a 7-point Likert scale (Cronbach's $\alpha = .79$). The scale is designed to measure how strong a navigator a participant feels he or she is, with lower scores indicating lower navigation ability.

2.4.2 Philadelphia spatial ability scale. (PSAS; Hegarty, Crookes, Dara-Abrams, & Shipley, 2010). The PSAS consists of 16 items which participants respond to on a 7-point Likert scale (Cronbach's $\alpha = .77$). This scale is designed to measure how well a participant can perform small-scale spatial tasks such as visualizing and transforming small- or medium-sized objects.

2.4.3 Philadelphia verbal ability scale. (PVAS; Hegarty et al., 2010). The PVAS consists of 10 items which participants respond to on a 7-point Likert scale (Cronbach's $\alpha = .78$). This scale is designed to measure how strong a participant feels his verbal ability is. This scale was administered via paper and pencil.

2.4.4 Mental rotation test. (MRT; Vandenberg & Kuse, 1978; adapted by Peters et al., 1995) The MRT consists of items made up of one target image composed of a number of individual cubes. Participants must choose the two (out of four) objects that correspond to the target after being rigidly rotated. Scoring correcting for guessing was applied such that a participant received 2 points for a correct response, but was penalized 2 points for an incorrect response. No points were awarded or rescinded for omissions. The MRT consists of two parts of 10 items each, with three minutes allotted for each part of the test.

2.4.5 Spatial orientation test (SOT; Kozhevnikov & Hegarty, 2001; we used the revised version by Hegarty & Waller, 2004). The SOT requires viewing an array of objects on a piece of paper, taking the perspective of standing next to one object and facing another, with the task of pointing to a third object. Five minutes are allowed to complete the 12-item measure. The angle between the correct answer and the participant’s response is recorded for each item, and averaged to yield an overall error score. This test was administered via paper and pencil.

2.5 Virtual Environment Learning

The final article is available, upon publication, at: https://doi.org/10.1037/a0035261

After the participant completed all questionnaires, the experimenter explained the navigation tasks. First, the experimenter explained the controls for navigating the VE and provided the participant an opportunity to move around the environment before being required to learn the buildings. Movement through the environment was controlled using the arrow keys (up for forward, down for backward, and left and right for lateral movement or strafing). In addition, participants could look around the environment by moving a standard computer mouse to rotate the camera 360 degrees horizontally, and 60 degrees away from parallel to the ground both up and down. The experimenter explained and demonstrated that the mouse and arrow keys could be used in conjunction to turn (i.e., rotate the mouse while pressing the up arrow). Practice always occurred in the first route in the virtual environment used in the study, but the experimenter instructed the participant not to navigate past the first building.

Once the participant indicated comfort and familiarity with the controls, he was informed that he would be learning the names and locations of eight buildings along two routes in a VE (see Figure 1). These buildings were marked by a blue diamond which floated above the route, and next to a nearby sign with the building’s name. The experimenter then told the participant that two non-connecting routes through the environment would each contain four of the buildings, but to be prepared to be tested on the locations of all eight buildings.

After learning all eight buildings, participants travelled on two paths connecting the first two routes to each other, and were told to pay attention to how the two sets of buildings were positioned in the VE. Participants were clearly instructed that no additional buildings would need to be learned, but that these two routes would provide additional spatial information. The order of experiencing the two connecting routes was counter-balanced across participants, but always occurred after the first two routes. For all routes and paths, participants travelled from the start to the finish and back to the start, but had as much time as they needed.

2.6 Virtual Environment Spatial Tasks

2.6.1 Onsite Pointing. In this task, the participant was placed at the start of one of the two routes (determined randomly), directly adjacent to the first building of that route. By moving the mouse, the participant could point a crosshair in the center of the screen in any direction. A prompt at the top of the screen provided the name of one of the other seven buildings. The participant was instructed to rotate the mouse until the crosshair pointed at the front door of the building in the prompt and to click once to register their answer. Clicking the mouse also changed the name of the building in the prompt and the participant then pointed to that building. Once all seven buildings were completed from the first building, the participant was automatically dropped at the next building along that route, and pointed to the other seven buildings in the same manner. After the participant completed this for the buildings on the first route, the same task was completed for buildings on the second route. The order of the buildings the participant pointed from was fixed to match the order of learning, but the order of buildings pointed to was random. The pointing task was scored by measuring the shortest possible angle between the correct response and the participant's response, yielding the error value in degrees for each trial.

2.6.2 Model Building. In a model-building task, the participant viewed a blank box on the computer screen with aerial images of each of the eight buildings beneath it. The experimenter told the participant the box represented the entire VE. The participant's task was to drag and drop each building using the mouse to where it belonged in the environment. Buildings could be moved as much or as little as necessary and no time limit was given. The orientation of

The final article is available, upon publication, at: https://doi.org/10.1037/a0035261

the buildings was fixed so they could not be rotated, but no instruction was provided about the orientation of the environment. If a participant asked which way the environment was oriented, the experimenter informed him that he could place the buildings in whatever orientation felt most comfortable. Accuracy on the model-building task was measured using a bidimensional regression anal analysis (Friedman & Kohler, 2003; Tobler, 1994).

3. Results

3.1 Virtual Environment Results

3.1.1 Pointing task. For the pointing task, the absolute value of the angular distance from the participant's answer to the correct angle was calculated for each trial, and then averaged across trials to yield the overall error score. For example, if the correct angle for a given trial had the value of 295°, and the participant responded 100°, the absolute value of the difference was calculated and corrected to be below 180° by subtracting the result from 360 (i.e., $|100 - 295| = 195$. $360 - 195 = 165$). Guessing with no knowledge of the environment would yield an average score of 90°. Participants were able to learn the locations of the buildings significantly better than chance, one-sample $t(48) = 25.13$, $p < .001$. No individual participant's pointing error was above the 90° threshold (*Maximum* = 62.90), but there was large variability in performance overall ($M = 42.56°$, $SD = 13.21$, *Range* = 51.73). There were no significant gender differences on the pointing task overall, $t(47) = 1.66$, $p = .10$, $d = .48$.

3.1.2 Between- and within-route pointing task trials. We examined differences between participants based on their performance on the two different types of pointing trials, Between-route or Within-route. We hypothesized that Between-route trials would be more difficult than Within-route trials and thus separated trials based on whether the building being pointed to was on the same route as the one the participant was currently standing (Within trials) or on the other main route (Between trials). Dividing the trials in this manner resulted in 24 Within trials and 32 Between trials per participant. For the VE, a paired-sample *t*-test on Within versus Between trial types revealed a significant difference, $t(48) = 12.28$, $p < .001$, $r = .55$ such that error on Within trials ($M = 24.06$, $SD = 12.13$) was significantly lower than error on Between trials ($M = 46.47$, $SD = 14.54$).

Trials were also examined based on whether the building being pointed to was Seen or Unseen from the participant's location. Splitting the trials in this manner resulted in 14 Seen trials and 42 Unseen trials (10 of which were Within-route). Performance was significantly better for Seen-Within trials ($M = 19.81$, $SD = 14.32$), than Unseen-Within trials ($M = 30.64$, $SD = 13.77$), $t(48) = 5.05$, $p < .001$, $r = .43$, and Between trials, $t(48) = 11.94$, $p < .001$, $r = .41$. Unseen-Within trials were also significantly easier than Between trials, $t(48) = 7.88$, $p < .001$ $r = .51$. Note that, while they were the easiest trials, substantial errors on Seen-Within trials are most likely attributed to the participant failing to learn the name of the building.

Figure 3 plots each participant’s average error for Within-route trials (on the Y-axis) by the average error for Between-route trials (on the X-axis). Although the linear correlation between the two trial types was high and significant, $r = .55$, $p < .001$, the pattern of results was different for the Within versus the Between pointing trials. Nearly all participants with good performance on the Between trials had good performance on the Within trials. But, participants with bad performance on the Between trials exhibited a range of values on the Within trials. Only two participants exhibited average accuracy on the Between trials and low accuracy on the Within trials – observe that the upper left quadrant of the scatter plot in Figure 3 is nearly vacant.

In order to analyze whether performance on the psychometric and self-report measures was related to performance on the two types of pointing trials, we used Within and Between pointing error in a cluster analysis to assemble the participants into groups. Using SPSS 18 statistical software's two-step cluster analysis algorithm with log-likelihood as the distance measure, we clustered participants based on their scores on Between and Within trials. The two-step algorithm first assigns individual values into pre-clusters, which in turn are clustered together to maximize the log-likelihood of a case belonging to that cluster. Nearly identical results were obtained after clustering in several different ways[2].

First, participants were clustered on each of the variables (Between and Within pointing trial average) separately. The analysis clustered participants into two groups – nominally good and bad performance - for each variable. Based on their group for each variable, participants were then assigned into one of four possible combinations: Good Between / Good Within, Good Between / Bad Within, Bad Between / Good Within, and Bad Between / Bad Within. The Good Between / Bad Within group only had one participant, while the other three groups had 12, 17, and 19 participants, respectively.

Based on this analysis and the distribution of participants along these two dimensions, a second two-step cluster analysis was conducted with both variables entered at the same time, and the number of clusters constrained to three. Conducting the cluster analysis in this manner resulted in the same clusters as when the variables were entered separately. This grouping is displayed by the three different colors in Figure 3 and used in all subsequent analyses. Using the alternative groupings did not substantially alter the results for the subsequent analyses.

3.1.3 Individual differences on pointing performance. We wanted to know how Pointing Groups, distinguished based on their performance (Good or Bad) on the two different trial types (Within or Between) differed on the psychometric and self-report tests we

[2] The same analyses were performed a) using the *k*-means algorithm and b) using Seen and Unseen trial divisions, resulting in clusters that differed by only one or two cases. In the interest of space, only the two-step results are reported here.

administered. First, the psychometric and self-report measures were Z-scored to normalize the scales and allow easier comparison between measures. Results from one-way ANOVAs with Pointing Group as a between-subject variable, and Z-scores of each of the other measures as dependent variables, are displayed in Figure 4. We observed significant differences between groups for the mental rotation test, $F(2, 46) = 3.35$, $p = .04$, $\eta^2 = .13$, and Santa Barbara sense of direction scale, $F(2, 46) = 5.64$, $p = .006$, $\eta^2 = .20$. The groups were not significantly different on the perspective taking test, $F(2, 46) = 2.78$, $p = .07$, $\eta^2 = .12$, Philadelphia spatial abilities scale, $F(2, 46) = 1.47$, $p = .24$, $\eta^2 = .06$, or the Philadelphia verbal abilities scale, $F(2, 46) = 0.21$, $p = .81$, $\eta^2 = .01$. Follow-up post-hoc pairwise contrasts using $\alpha = .016$ revealed that the Good Between / Good Within group outperformed the Bad Between / Bad Within group on the MRT, and SBSOD. The Bad Between / Good Within and Bad Between / Bad Within groups were not significantly different for any post-hoc contrasts.

3.1.4 Model-building task. The model-building task was analyzed using bidimensional regression (Friedman & Kohler, 2003; Tobler, 1994). Bidimensional regression is the correlation between a set of independent X-Y points (in this case, the correct locations of all eight buildings) and a set of dependent A-B points (the participant's placement of the eight buildings). The set of eight dependent points are optimally rotated, scaled, and translated to match the fixed independent points on these dimensions. The correlation coefficient indicates the remaining correlation between the correct answer and the participant's response. Squared, the correlation coefficient describes the proportion of variance explained in the actual layout of buildings by the participant's arrangement of buildings. Like the pointing task, participants exhibited a range of performance on the model building task ($M = .48$, $SD = .27$, $Range = .93$). There were no gender differences on the model building task, $t(47) = 1.05$, $p = .30$, and thus the analyses were collapsed across this variable. A one-factor ANOVA using the above cluster analysis groups, revealed significant differences in performance on the model building task, $F(2, 46) = 7.99$, $p < .001$, $\eta^2 = .33$. Post-hoc follow-up contrasts indicated the Good Between / Good Within group also outperformed the Bad Between / Good Within group on the SBSOD and the Model building task.

To assess whether average VE performance was above chance, a Monte Carlo simulation was conducted. Random X and Y coordinates were independently generated for each of the buildings and entered into the bidimensional regression formula as the dependent variables. This process was repeated 1000 times, with each set of eight points representing a randomly chosen set of positions for the eight buildings. Participant's performance on the model building task was significantly better than the Monte Carlo simulation average ($M = .35$, $SD = .17$), one-sample $t(48) = 3.58$, $p = .001$, $d = 1.03$.

3.1.5 Correlations and regression analyses. The correlations displayed in Table 1 show that the SBSOD, MRT, and PTT (error) correlate significantly with both Within and Between pointing error, as well as the model building task. Dividing the pointing trials further into Seen Within trials, Unseen but Within route trials and Between route trials (all Unseen), reveals significant relationships between PTT and Unseen Within, $r = .29$, $p = .04$, and Between, $r = .39$, $p = .01$, but not between PTT and Seen trials, $r = .22$, $p = .12$. The MRT, on the other hand was significantly correlated with all three trial types (Seen Within, $r = -.30$, $p = .04$; Unseen Within, $r = -.32$, $p = .02$; and Between, $r = -.35$, $p = .01$).

To assess the amount of variability the psychometric and self-report measures were able to account for in the dependent variables, multiple two-step regressions were conducted with the navigation tasks as dependent variables (See Table 2). In the first step, MRT, PTT, PSAS, and

PVAS were first included in the model. These measures did not explain a significant proportion of the variance for any of the dependent variables. In the second step, SBSOD was added to the model. SBSOD was a significant predictor of the Between pointing trials and the model-building task. The models from the second step explained a significant proportion of the variance for both Between pointing trials and the model-building task, but not Within pointing trials. Adding SBSOD also explained a significantly higher proportion of variance over the variables from the first step.

3.2 Comparison of Virtual Environment and Real World

As explained above, the real world environment methodology differed slightly from the virtual environment. Participants in the real world first learned both main routes and were then tested (session 1). Participants returned one week later, reviewed both main routes, learned one connecting route, and were then tested again (session 2). Participants returned one week later again, reviewed both main routes, learned the second connecting routes, and were tested a third time (session 3). Both the pointing and model-building tasks can be compared across all three real-world sessions to determine how performance differed between the two environments.

For the first session of the real world experiment, participants' performance on the pointing task was comparable in terms of distribution (M = 36.55, SD = 12.85, *Range* = 37.81), and was not significantly different from performance in the VE, $t(63) = 0.04$, $p = .97$, $d = .01$. After participants in the real world learned the connecting paths, however, they surpassed performance on the pointing task in the VE. Participants in the real world performed significantly better in both session 2 (M = 23.06, SD = 5.39, *Range* = 18.08), $t(63) = 4.33$, $p < .001$, $d = 1.09$, and session 3(M = 20.65, SD = 2.77, *Range* = 8.12), $t(63) = 5.22$, $p < .001$, $d = 1.32$.

The final article is available, upon publication, at: https://doi.org/10.1037/a0035261

Comparing real and virtual environments on the basis of performance on Between and Within pointing trials revealed that, for Within pointing trials, participants did not differ between environment types for session 1, 2, or 3 (all *p*'s > .07). For Between pointing trials, participants did not differ at session 1, $t(63) = 1.48$, $p = .14$, $d = 0.37$. However, participants in the real world performed significantly better than participants in the VE for session 2, $t(63) = 5.17$, $p < .001$, $d = 1.30$, and session 3, $t(63) = 6.35$, $p < .001$, $d = 1.60$.

However, unlike the pointing task, participants' overall score on the model building task was substantially lower in the VE than in even the first session of the real world study ($M = .70$, $SD = .05$, $Range = .79$), $t(63) = 2.90$, $p = .005$, $d = .73$. This difference continued for both session 2, $t(63) = 5.05$, $p < .001$, $d = 1.27$, and session 3, $t(63) = 5.78$, $p < .001$, $d = 1.46$.

4. Discussion

The present study expanded upon current understanding of individual differences in navigation performance. Data from the pointing task revealed that participants exhibit individual differences not only in integrating two large areas together, but in forming accurate representations of spatial relations between buildings along individual routes. Moreover, the full pattern of individual differences was not fully predicted by self-report data. The SBSOD, while correlated with both Between and Within trial errors, explained a significant portion of the variance in the regression analyses for the Between trials but not Within trials. The SBSOD was also unable to distinguish between two of the three groups from the cluster analysis: Bad Between / Good Within and Bad Between / Bad Within.

Our results support the idea that *some* people can form accurate cognitive maps of a novel environment quickly, as demonstrated by their ability to point between two places they had never directly travelled between, while others have great difficulty with such tasks. While comparing results from the VE to real world data suggests that learning is more accurate overall in the real world, particularly for Within trials, the pattern of individual differences with regard to cognitive map formation is similar (Schinazi et al., in press). This finding reinforces hypothesis that the formation of integrated spatial representations of data gathered during navigation is a challenging task, one that can be accomplished in principle but that not everyone finds easy.

4.1 Individual Differences

Our results strongly suggest that forming an accurate spatial representation of buildings within one route is a prerequisite for integrating the spatial positions of the two routes of buildings together. As the scatter plot in Figure 3 shows, the different trial types exhibited separate patterns of performance. Notably, no participants performing relatively well on the Between trials performed poorly on the Within trials. However, participants performing poorly on Between trials varied widely on Within trials.

The methodology used for the VE study discouraged direct integration between the two main routes, allowing a meaningful distinction to be drawn for Within-route trials which the participant had learned as part of the same route and Between-route trials which were explicitly part of a separate route. Participants were therefore tasked to learn the positions of buildings along one route and then integrate sets of buildings across different routes. Interestingly, these abilities, though correlated, may be dissociable and supported by different neural structures. (Burgess, 2006).

Brain imaging data from the real world study (Schinazi et al., in press) demonstrated that the neural correlates are different for these different navigational processes. Participants with the largest hippocampi performed best on an offsite pointing task which required taking an imagined

The final article is available, upon publication, at: https://doi.org/10.1037/a0035261

orientation and pointing to buildings around the environment. In addition, participants performing better on Within-route trials tended to have smaller caudates. One plausible explanation for this finding lies in the dissociation found between the hippocampus and the caudate based on the navigation strategy their activation predicted (Hartley et al., 2003; Iaria et al., 2003; Marchette, Bakker, & Shelton, 2011). In one virtual environment study, for instance, caudate activation was shown to be a strong predictor of response-based strategies for navigation while hippocampal activation predicted a place-based strategy (Marchette & Shelton, 2011). A place-based strategy may support a more accurate representation of the spatial properties of buildings, while a response-based strategy may support more accurate route knowledge. This hypothesis awaits further investigation.

These individual differences in navigation ability go beyond what self-report measures are able to predict. In particular, SBSOD was highly predictive of performance on Between-route pointing trials, and on the model building task, but was unable to predict a significant proportion of variance on Within-route pointing trials (see Table 2). Additionally, despite correlating with error rates for both Between and Within trials, the SBSOD could not distinguish between participants performing poorly on both types of trials from participants performing poorly just on Between-route pointing judgments. The correlations (shown in Table 1) between SBSOD and both types of pointing tasks are driven primarily by high-SBSOD participants scoring well on both trial types, while poor-SBSOD participants scored poorly on both trial types. This correlation obscures the fact that participants scoring poorly only on Between trials rated themselves no higher on SBSOD than participants scoring poorly on both trial types. Indeed, post-hoc contrasts indicated that the Good Between / Good Within group scored higher on SBSOD than the other two groups but no difference between the other two groups.

As Hegarty et al. (2002) observed, the SBSOD is best at predicting performance for locating buildings that are not mutually visible. It is particularly noteworthy that dividing trials by whether the building was mutually visible (seen) or not (unseen) yielded identical results from the group-based and regression analyses when trials were divided by Within and Between-route. That is, SBSOD scores were not significantly different between Bad Unseen / Good Seen participants and Bad Unseen / Bad Seen participants. The lack of discrimination by the SBSOD suggests that there is a factor of navigation not being measured that explains a significant portion of behavioral results. Whether that factor relates to a general resources detriment (i.e., attention, working memory, etc.) or a more specific navigation-related impairment (i.e., lack of binding the scene information of the building to the identity or name of that location) cannot be determined without further research. The SBSOD, however, is not the only measure of spatial ability that does not distinguish these groups of performers.

In the real world study, PTT score mediated the relationship between hippocampal volume and performance on the offsite pointing task, suggesting that perspective taking, instantiated in the hippocampus, is a significant component of establishing an orientation in an environment and representing the surrounding spatial properties of the environment. Yet in the VE study, the MRT and PTT, psychometric measures of figural spatial ability, while significantly correlated with both Within and Between pointing trials, were also not significantly different between the Bad Between groups. In the current study, MRT, but not PTT, was significantly correlated with pointing judgments on Seen trials alone. These results suggest that the individual differences in representing the spatial properties of separate routes of an environment have distinct self-report and psychometric correlates from individual differences in representing the spatial properties within one route.

While it is unclear what, precisely, underlies the difference between the two lower-performing groups, one way to address this issue for the future of self-report scale development would be to add items which measure how well an individual encodes scenes, attends to locations of places in space, and binds the identity to places. This may violate the one-factor structure of the SBSOD, but it would allow researchers to measure more precisely individual differences among navigators.

4.2 Cognitive Map Formation

In the virtual environment study reported here and previous real world studies (Ishikawa & Montello, 2006; Schinazi, in press), results show that some, but not all, participants form an accurate cognitive map. Although participants in the VE were tested using the onsite pointing task, and the correlation found in the real world study was between offsite pointing and hippocampal volume, there is a key similarity between these tasks that allowed us to investigate the presence or absence of cognitive map-like representations. Specifically, in both tasks, participants are required to reestablish their orientation at each pointing site without the aid of idiothetic cues to update as one moves from site to site. Thus, in both tasks, accuracy on the Between-route trials of the onsite pointing task still indicates the hallmark of cognitive maps – the ability to take novel shortcuts. Both pointing tasks require hippocampal involvement, insofar as formation and use of a cognitive map, but the offsite pointing task has the additional component of requiring participants to imagine a position in an environment and establish an orientation. The onsite pointing task allows participants to use cues that are directly visible in the environment, but still requires them to indicate a direct line to buildings which they cannot see, and to which they have never travelled directly.

In the real world study, which had much longer learning (within sessions and overall) and full sensory cues, all 16 participants were eventually able to learn the spatial properties of the Within-route buildings after just the first session of learning. This finding can likely be attributed to the increase in accuracy concomitant with improved sensory cues or due to increased attention and motivation in the real world as opposed to a virtual environment (Chrastil & Warren, 2011). While the range of learning was not observed for Within-route trials in the real world, all participants who performed well on Between-route trials were also accurate on Within-route trials. In the VE, however, not all participants who formed accurate Within-route representations were able to form accurate Between-route representations.

This study was not designed to test the two competing frameworks of cognitive map development, Siegel and White's (1975) stage framework and Montello's (1998) continuous framework, but our results allow for some speculation. In Siegel and White's (1975) framework, landmark and sequential knowledge may make up the Within-route pointing ability, while survey knowledge comprises Between-route pointing ability. A recent review of neuroimaging studies also adds graph knowledge as an intermediary between sequential and survey knowledge (Chrastil, 2012). Graph knowledge is beyond route knowledge in the sense that multiple (even novel) routes can be taken between points in an environment by understanding the nodes and paths, without understanding the metric properties (i.e., distance and direction) that characterize survey knowledge. Although graph knowledge certainly describes spatial behavior in other tasks, participants in the present study did not learn a network of routes, and thus the ability to point between two places in the environment never travelled between must be taken as indicative of survey knowledge.

We did not find evidence of any participants with survey knowledge (Between trial success) who did not evince strong sequential knowledge (Within trial success). This suggests two possibilities. First, some individuals may quickly proceed through Siegel and White's stages, learning different types of spatial knowledge concurrently, and thus improve their representation of a space continually, while others never proceed past the sequential stage to the survey stage. Second, the individuals who perform well on the Within-route but poor on the Between-route trials could be gleaning metric information about the two routes, but after a short amount of learning, their survey knowledge is weak. Specific tasks that are more sensitive to metric knowledge (e.g., judging route vs. straight-line distance) might be able to determine the plausibility of this second possibility.

There is a separate possible property of the representation that merits some consideration: good navigators may have used a hierarchically-structured representation of the two main routes that contains both coarse information (the route a particular building is on, and the general spatial position of buildings in that area) and more fine-grained information. This property of integrating qualitative (or categorical) information with quantitative (or fine-grained) information has been called the category adjustment model (Huttenlocher, Hedges, & Duncan, 1991). At least one study provides evidence that when a familiar environment contains natural categories, navigators bias their estimates of locations toward the center of that category (Uttal, Friedman, Hand, & Warren, 2010). Interestingly, the bias observed by Uttal and colleagues increased with familiarity, but self-reported navigation ability was not collected. Unfortunately, while this seems like an excellent heuristic, the locations used in the environment for this study are too close together to measure any bias toward the center of the buildings for Between route trials. Future research should investigate whether this representation structure underlies learning of new environments, and whether it differs for good and bad navigators.

4.3 Implications

While self-report measures have important uses, there are strong reasons to consider moving beyond them. Mounting evidence suggests that self-report measures are unable to predict the variance observed in more nuanced investigations in individual differences in navigation tasks (Shelton, Furman, Marchette, & Brockman, in prep). As a rough measure of individual differences, self-reports are an excellent starting point, but to understand the nuanced differences in navigation ability that characterize individuals, more ecologically valid measures are required.

For instance, the difference discussed above between the Bad Between / Good Within and Bad Between / Bad Within implies that different training strategies would be recommended for people in different categories of performance (i.e., performing well on Within-route trials, or not). To reiterate, these groups are not distinguishable using extant self-report measures alone. For participants performing well on the Within-route trials, strategies focusing their learning on globally available cues (e.g., buildings visible along both routes, the sun, etc.) may boost their Between-route scores. This strategy would be less effective for participants who struggle with creating accurate Within-route representations because they are integrating two routes that are themselves inaccurate. Training these participants to focus on path integration, the angle of turns, and distances between buildings – in short, a place-based strategy – should be more effective.

A VE, such as the one reported here, which has the capability to measure navigation ability has great potential for experimentation in cognitive psychology. While a great deal is known about the components of navigation (e.g., Wolbers & Hegarty, 2010), not enough is known about other important research questions, for which existing self-report data are insufficient. Spatial ability in general, and navigation in particular, decline with normal aging

(Bohbot et al., 2012), and this trend is exacerbated in cases of Parkinson's and Alzheimer's disease (Gazova et al., 2012). VEs can be used to understand how and when navigation ability declines among normal aging and other populations, as well as the brain mechanisms that underlie this decline.

Perhaps the largest gap in current understanding is how little is known about the relationship between navigation ability and other cognitive abilities. There is a large amount of evidence that spatial ability is a significant and unique predictor of entrance into STEM disciplines (Wai, Lubinski, & Benbow, 2009), but spatial ability has been exclusively measured in these studies with paper-and-pencil tests. Despite factor analyses demonstrating that large-scale navigation may rely upon, but is certainly distinct from figural, or small-scale, spatial abilities (Hegarty et al., 2006), no studies have shown a link between navigation ability and educational outcomes. Given what we know about the correlation between hippocampal volume and navigation ability, exercising one's navigation ability in daily life may be a way to boost other cognitive abilities known to be based in the hippocampus; in other words, do abilities supporting navigation transfer (e.g., binding, episodic memory, memory consolidation)? Understanding this relationship has implications on the increasing reliance on global positioning systems (GPS) to supplement or replace "non-caudal" navigation. The VE used for this study is an excellent candidate to begin to address these possibilities by testing the relationship between behavioral or neural differences in various cognitive abilities along with navigation ability.

References

Blajenkova, O., Motes, M. & Kozhevnikov, M. (2005). Individual differences in the representations of novel environments. *Journal of Environmental Psychology*, *25*, 97-109.

Bohbot, V. D., McKenzie, S., Konishi, K., Fouquet, C., Kurdi, V., Schachar, R., ... & Robaey, P. (2012). Virtual navigation strategies from childhood to senescence: evidence for changes across the life span. *Frontiers in Aging Neuroscience*, *4*, 28. doi: 10.3389/fnagi.2012.00028

Burgess, N. (2006). Spatial memory: how egocentric and allocentric combine. *Trends in Cognitive Science*, *10*(12), 551-557.

Chrastil, E. R. (2012). Neural evidence supports a novel framework for spatial navigation. *Psychonomic Bulletin & Review*. 10.3758/s13423-012-0351-6

Chrastil, E. R., & Warren, W. H. (2012). Active and passive contributions to spatial learning. Psychonomic Bulletin & Review, 19, 1–23. doi:10.3758/s13423-011-0182-x

Chai, X. J., & Jacobs, L. F. (2010). Effects of cue types on sex differences in human spatial memory. *Behavioural Brain Research*, 208(2), 336-342. doi: 10.1016/j.bbr.2009.11.039

Epstein, R. A., Higgins, J.S., & Thompson-Schill, S.L. (2005). Learning places from views: Variation in scene processing as a function of experience and navigational ability. *Journal of Cognitive Neuroscience*, 17, 73-83.

Foo, P., Warren, W. H., Duchon A., & Tarr, M. J. (2005). Do humans integrate routes into a cognitive map? Map- versus landmark-based navigation of novel shortcuts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(2), 195-215.

Friedman, A. & Kohler, B. (2003). Bidimensional regression: Assessing the configural similarity and accuracy of cognitive maps and other two-dimensional data sets. *Psychological Methods*, *8*, 468-491.

Gazova, I., Vlcek, K., Laczo, J., Nedelska, Z., Hyncicova, E., Mokrisova, I., Sheardova, K., Hort, J. (2012). Spatial navigation-a unique window into physiological and pathological aging. *Frontiers in Aging Neuroscience*, 4, 1-6.

Hartley, T., Maguire, E. A., Spiers, H. J., & Burgess, N. (2003). The well-worn route and the path less traveled: Distinct neural bases of route following and wayfinding in humans. *Neuron*, *37*(5), 877–888.

Hegarty. M. & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, *32*, 175-191.

Hegarty, M., Crookes, R. D., Dara-Abrams, D., & Shipley, T. F. (2010). Do all science disciplines rely on spatial abilities? Preliminary evidence from self-report questionnaires. In C. Hölscher, T. F. Shipley, M. Olivetti Belardinelli, J. Bateman, & N. S. Newcombe (Eds.), *Spatial Cognition VII: Lecture Notes in Computer Science 6222* (pp. 85-94). Springer.

Hegarty, M., Montello, D. R., Richardson, A. E., Ishikawa, T., & Lovelace, K. (2006). Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial-layout learning. *Intelligence, 34*(2), 151-176. doi: 10.1016/j.intell.2005.09.005

Hegarty, M., Richardson, A. E., Montello, D. R., Lovelace, K., & Subbiah, I. (2002). Development of a self-report measure of environmental spatial ability. *Intelligence, 30*, 425-447.

Heth, C. D., Cornell, E. H., & Flood, T. L. (2002). Self-ratings of sense of direction and route reversal performance. *Applied Cognitive Psychology, 16*(3), 309-324.

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

Iaria, G., & Barton, J. J., (2010). Developmental topographical disorientation: A newly discovered cognitive disorder. *Experimental Brain Research* (206), 189-196.

Iaria, G., Petrides, M., Dagher, A., Pike, B., & Bohbot, V. D. (2003). Cognitive strategies dependent on the hippocampus and caudate nucleus in human navigation: Variability and change with practice. *Journal of Neuroscience, 23*(13), 5945–5952.

Ishikawa, T., & Montello, D.R. (2006). Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places. *Cognitive Psychology*, *52*, 93-129.

Ishikawa, T., & Nakamura, U. (2012). Landmark selection in the environment: Relationships with object characteristics and sense of direction. *Spatial Cognition and Computation*, 12, 1-22.

Kozhevnikov, M. & Hegarty, M. (2001). A dissociation between object-manipulation and perspective taking spatial abilities. *Memory & 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*(4), 590.

Maguire, E. A., Woollett, K., & Spiers, H. J. (2006). London taxi drivers and bus drivers: A structural MRI and neuropsychological analysis. *Hippocampus*, *16*(12), 1091-1101.

Marchette, S. A., Bakker, A., & Shelton, A. L. (2011). Cognitive mappers to creatures of habit: Different engagement of place and response learning mechanisms predicts human navigational behavior. *Journal of Neuroscience, 31*(43), 15264–15268.

Mittelstaedt, M. L., & Mittelstaedt, H. (2001). Idiothetic navigation in humans: estimation of path length. *Experimental Brain Research*, *139*(3), 318-332.

Montello, D.R. (1998) A new framework for understanding the acquisition of spatial knowledge in large-scale environments. In R.G. Golledge and M.J. Egenhofer (Eds.*), Spatial and temporal reasoning in geographic information systems*. New York: Oxford University Press, pp. 143–154.

Montello, D. R., Waller, D., Hegarty, M & Richardson, A. E. (2004). Spatial memory of real environments, virtual environments, and maps. In G. Allen (Ed.). *Human Spatial Memory: Remembering Where*, 251-285.

Pazzaglia, F., & De Beni, R. (2001). Strategies of processing spatial information in survey and landmark-centred individuals. *European Journal of Cognitive Psychology*, *13*(4), 493-508

Pazzaglia, F., & Taylor, H. A. (2007). Perspective, instruction, and cognitive style in spatial representation of a virtual environment. *Spatial Cognition and Computation*, *7*(4), 349-364.

Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., & Richardson, C. (1995). A redrawn Vandenberg and Kuse mental rotations test - different versions and factors that affect performance. *Brain and Cognition*, *28*(1), 39-58. doi: 10.1006/brcg.1995.1032

Schinazi,V.R., Nardi, D., Newcombe, N.S., Shipley, T.F., & Epstein, R.A. (in press). Hippocampal size predicts rapid learning of a cognitive map in humans. *Hippocampus*.

Shelton, A., Furman, A.J., Marchette,S., A., & Brockman, M.A. (in preparation). Distinguishing preference from ability in selecting navigational solutions.

Sholl, J.M. (1988). The relation between sense of direction and mental geographic updating. *Intelligence*, *12*(3), 299-314.

Siegel, A. W. & White, S. H. (1975). The development of spatial representations of large-scale environments. *Advances in Child Development and Behavior*, *10*, 9-55.

Tobler, W. R. (1994). Bidemensional regression. *Geographical Analysis*, *26*(3). 187-212.

Tolman, E. C., (1948). Cognitive maps in rats and men. *The Psychological Review*, *55*(4), 189-208.

Uttal, D. H., Friedman, A., Hand, L. L., & Warren, C. (2010). Learning fine-grained and category information in navigable real-world space. *Memory & Cognition*, *38*(8), 1026-1040.

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations: A group test of three-dimensional spatial visualization. *Perceptual and Motor Skills*, 47, 599-604.

Wai, J., Lubinski, D., & Benbow, C., P. (2009). Spatial ability for stem domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, *101*(4) 817-835.

Wen, W., Ishikawa, T., & Sato, T. (2011). Working memory in spatial knowledge acquisition: Differences in encoding processes and sense of direction. *Applied Cognitive Psychology*, 25, 654-662.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Teleoperators and Virtual Environments*, *7*(3), 225-240.

Wolbers, T. & Hegarty, M. (2010). What determine our navigational abilities? *Trends in Cognitive Science*, *14*, 138-146

Woollet, K., Maguire, E. A., (2011). Acquiring “the knowledge” of london’s layout drives structural brain changes. *Current Biology*, *21*(24), 2109-2114.

Figure and Table Titles and Captions

Figure 1. Title: Aerial view map of main routes, connecting routes, and buildings.

Caption: An aerial view map depicts the layout of buildings, main routes, and connecting routes for the virtual environment. Note that the spatial arrangement of buildings was identical to the real world study from Schinazi et al., in press. Participants never saw this representation. The letter-number combinations indicate starting and ending points along each of the routes learned. All participants began each route at 1, travelled the entire route to 2, and walked back to 1. Participants always learned the main routes (in red) first, with route A and route B counter-balanced between participants. Then participants learned both connecting routes (in blue), with route C and route D similarly counter-balanced.

Figure 2. Title: Screenshots from virtual environment

Caption: Screenshots from main route B (left) and connecting route D (right). The screenshot at right displays an arrow which marked the path participants were instructed to follow in the virtual environment, a blue diamond which indicated a building participants had to remember, and a sign with the name of the building (Golledge Hall).

Figure 3. Title: Scatterplot of individuals' accuracy on within and between pointing trials

Caption: Participants' error was calculated for each trial type and plotted. Each dot represents one participant's error rates on Within (y-axis) and Between (x-axis) trials. Note that no participants who scored below 40 degrees on the between trials scored above 20 degrees on the within trials. This suggests that accurate representation of local spatial position is important for forming an accurate global spatial configuration.

Figure 4. Title: Mean Z-scores for participants grouped by performance on both types of pointing trial

The final article is available, upon publication, at: https://doi.org/10.1037/a0035261

Caption: Participants were divided into three groups by means of a cluster analysis based on their average scores on Between pointing trials and Within pointing trials. The graph presents the average z-score for participants within each Pointing group on each of the psychometric and self-report measures, as well as the model building task. Omnibus ANOVAs demonstrate significant differences between the groups on the MRT test, the SBSOD, and the model building task. [Perspective Taking Test is reverse Z-scored]. $*p < .05$, $**p < .01$, $***\ p < .001$. MRT = Mental Rotation Test; PTT = Perspective Taking Test; SBSOD = Santa Barbara Sense of Direction; PSAS = Philadelphia Spatial Abilities Scale; PVAS = Philadelphia Verbal Abilities Scale.

Table 1. Intercorrelations, means, and standard deviations for psychometric and navigation tasks

Caption: Bivariate correlations, means, and standard deviations for all participants (N = 49) are presented above. Rows labeled [Error] should be interpreted as higher values indicate worse performance. Rows not labeled [Error] should be interpreted as accuracy, where higher values indicate better performance. MRT = Mental Rotation Test, PTT = Perspective Taking Test, SBSOD = Santa Barbara Sense of Direction, PSAS = Philadelphia Spatial Ability Scale, PVAS = Philadelphia Verbal Ability Scale. $*p < .05$, $**p < .01$

Table 2.Title: Regression models for model-building and pointing tasks.

Caption: Regression analyses. Controlling for gender, psychometric tests, and other self-report measures, SBSOD explains a significant portion of the variance, and improves the fit of the model for the model building task, and Between pointing trials, but is not a significant predictor and does not significantly improve the model for Within pointing trials.

The final article is available, upon publication, at: https://doi.org/10.1037/a0035261