A Direct Comparison of Real-World and Virtual Navigation Performance in Chronic Stroke Patients

Michiel H.G. Claessen,[1,2] Johanna M.A. Visser-Meily,[2] Nicolien K. de Rooij,[2] Albert Postma,[1] AND Ineke J.M. van der Ham[3]

[1]Experimental Psychology, Helmholtz Institute, Utrecht University, Utrecht, the Netherlands
[2]Brain Center Rudolf Magnus and Center of Excellence for Rehabilitation Medicine, University Medical Center Utrecht and De Hoogstraat Rehabilitation, Utrecht, the Netherlands
[3]Department of Health, Medical and Neuropsychology, Leiden University, Leiden, the Netherlands

Abstract

Objectives: An increasing number of studies have presented evidence that various patient groups with acquired brain injury suffer from navigation problems in daily life. This skill is, however, scarcely addressed in current clinical neuropsychological practice and suitable diagnostic instruments are lacking. Real-world navigation tests are limited by geographical location and associated with practical constraints. It was, therefore, investigated whether virtual navigation might serve as a useful alternative. **Methods:** To investigate the convergent validity of virtual navigation testing, performance on the Virtual Tübingen test was compared to that on an analogous real-world navigation test in 68 chronic stroke patients. The same eight subtasks, addressing route and survey knowledge aspects, were assessed in both tests. In addition, navigation performance of stroke patients was compared to that of 44 healthy controls. **Results:** A correlation analysis showed moderate overlap ($r = .535$) between composite scores of overall real-world and virtual navigation performance in stroke patients. Route knowledge composite scores correlated somewhat stronger ($r = .523$) than survey knowledge composite scores ($r = .442$). When comparing group performances, patients obtained lower scores than controls on seven subtasks. Whereas the real-world test was found to be easier than its virtual counterpart, no significant interaction-effects were found between group and environment. **Conclusions:** Given moderate overlap of the total scores between the two navigation tests, we conclude that virtual testing of navigation ability is a valid alternative to navigation tests that rely on real-world route exposure. (*JINS*, 2016, *22*, 467–477)

Keywords: Spatial navigation, Stroke, Virtual navigation test, Real-world navigation test, Ecological validity, Neuropsychological practice

INTRODUCTION

Spatial navigation is an ability that enables us to find our way from one location to another. Whether we walk, ride a bike, or drive a car, we rely on the ability to navigate to arrive at our planned destination. Navigation ability is thus crucial for everyday life, as it allows us to function independently in the community. Notwithstanding the notion of the cognitive complexity of navigation ability (Brunsdon, Nickels, & Coltheart, 2007; Wiener, Büchner, & Hölscher, 2009; Wolbers & Hegarty, 2010), researchers usually distinguish between two fundamentally different memory representations for navigation (Montello, 1998; Siegel & White, 1975). Route knowledge concerns information related to a specific route, such as distinctive features in the environment (landmarks), associations between landmarks and directional information (place-action associations) and the temporal order of landmarks or turns. Survey knowledge, on the other hand, refers to an integrated geometrical representation of the environment which also includes information about distances and angles.

Inherent to the cognitive complexity of navigation ability is its vulnerability to the effects of brain damage. Based on self-report data, nearly a third of stroke patients experience navigation difficulties after their stroke event (Van der Ham, Kant, Postma, & Visser-Meily, 2013). Other studies have provided evidence for this notion using objective navigation ability assessments in stroke patients (e.g., Van Asselen et al., 2006). Special attention to navigation ability should be paid in neglect patients, as deficits in this ability have shown to be associated with the neglect syndrome (De Nigris et al., 2013; Guariglia, Piccardi, Iaria, Nico, & Pizzamiglio, 2005; Nico et al., 2008). Recent studies have indicated that navigation impairment can also be found in other patient groups with acquired brain injury (ABI), including traumatic brain injury

Correspondence and reprint requests to: Michiel H. G. Claessen, Heidelberglaan 1, 3584 CS Utrecht, The Netherlands. E-mail: m.h.g.claessen@uu.nl

(e.g., Livingstone & Skelton, 2007), Korsakoff's syndrome (Oudman et al., in press) and Alzheimer's disease (e.g., Cushman, Stein, & Duffy, 2008). In general, these and many other studies clearly illustrate the importance of evaluating the status of navigation ability in ABI patients. Strikingly, this skill is scarcely addressed in an explicit manner in current clinical neurological and neuropsychological practice.

The lack of studies with a specific focus on navigation ability in ABI patient groups may partly be due to the fact that no valid objective navigation tests are currently generally available for use in neuropsychological practice. A further obstacle lies in the finding that common spatial neuropsychological tests, such as the Judgment of Line Orientation, the Rey-Osterrieth/Taylor Complex Figure and the Corsi Block-Tapping Task, are hardly able to reliably predict navigation behavior (e.g., Nadolne & Stringer, 2001; Van der Ham et al., 2013). It has been argued that this might result from neuropsychological tests falling short in ecological validity (Chaytor & Schmitter-Edgecombe, 2003), with regard to the ability to navigate. Ecological validity refers to the extent to which a neuropsychological test is representative of everyday situations and denotes the degree to which the test results are generalizable to and predictive of everyday life performance (Burgess et al., 2006).

A cognitive explanation for the inadequate ecological validity of common neuropsychological spatial tests lies in the fact that they are carried out within near or reaching space. Spatial navigation, in contrast, concerns interaction with large or navigational space. Behavioral and neuropsychological studies have drawn attention to this notion by showing that small-scale and large-scale spatial learning abilities can be dissociated (e.g., Piccardi et al., 2010; Piccardi, Iaria, Bianchini, Zompanti, & Guariglia, 2011) and rely on partly independent neural circuits (Nemmi, Boccia, Piccardi, Galati, & Guariglia, 2013). That is, patients suffering from navigation impairment do not necessarily fail on the small-scale spatial tests currently used in neuropsychological practice. These findings thus clearly indicate that assessment of navigation behavior should be based on large-scale tasks that closely resemble everyday navigation situations rather than using existing small-scale spatial neuropsychological tests.

For scientific purposes, researchers have generally adopted two different approaches to measure navigation ability in an objective manner: real-world and virtual reality (VR) navigation tests. In a typical real-world navigation test, the researcher takes the participant along a specific route in a building (for example, a hospital) or on the streets. After this learning phase, participants are asked to retrace the studied route (e.g., Barrash, Damasio, Adolphs, & Tranel, 2000) or tested on their knowledge of it (e.g., Van Asselen et al., 2006). As the participant has to physically follow the route, real-world navigation tests are likely to be closely related to actual navigation performance. Nonetheless, real-world navigation tests are also characterized by several serious limitations.

First, real-world navigation tests are, by definition, bound to a specific indoor or outdoor environment, for instance a particular hospital building (e.g., Barrash et al., 2000). This is an essential problem, as a navigation test validated in a specific environment is of limited use to clinicians at other locations. A second limitation of real-world navigation testing lies in the fact that identical exposure to the test environment during the learning phase of the route cannot be guaranteed across participants, for example due to differences in exposure time. Moreover, it is hard to control many other potential disturbing factors such as weather conditions, traffic and noise (Van der Ham, Faber, Venselaar, Van Krefeld, & Löffler, 2015). Another potential confounding factor is the participant's familiarity with the test environment. Some recent studies have shown that the degree of familiarity is an important factor to address, as higher familiarity generally leads to better performance on navigation tests (De Goede & Postma, 2015; Iachini, Ruotolo, & Ruggiero, 2009; Prestopnik & Roskos-Ewoldsen, 2000). More specifically, higher familiarity is associated with higher sense of direction and greater reliance on a survey/allocentric navigation strategy (Iachini et al., 2009). Apart from the above limitations, real-world navigation test procedures also have some practical drawbacks. These tests can be rather time-consuming and require the participant to be physically able to traverse the route. These disadvantages make it nearly impossible to develop a well-validated real-world navigation test that is widely applicable in neuropsychological practice around the world.

Virtual navigation tests have been proposed as a potential alternative to real-world navigation tests, because VR testing is not restricted by the above limitations. VR does not only allow for developing novel environments (to avoid issues with the participant's familiarity with the test environment), but also offers the researcher the ability to generate realistic and highly controllable real-world simulations (Rose, Brooks, & Rizzo, 2005). Most importantly, assessment of a well-validated virtual navigation test is not bound to a specific location.

It should, however, also be mentioned that virtual navigation is associated with an important drawback; the absence of locomotion. When passively studying a virtual route, participants can only rely on visual cues or external landmarks. That is, passive exposure to a virtual route does not provide the participant with vestibular cues or the possibility to internally perceive the body in space. Yet, the sensory input of moving through the environment has been implicated in the creation of an environmental mental map (e.g., Chrastil & Warren, 2013; Van der Ham et al., 2015), which contains information about the relative positions of landmarks in an environment. It might thus be possible that the validity of virtual navigation is compromised when it comes to testing the survey knowledge aspects of a route.

The validity of virtual navigation tests has been studied several times in healthy participants. Studies have not only suggested that transfer from real-world to virtual environments is possible (Péruch, Belingard, & Thinus-Blanc, 2000; Wilson, Foreman, & Tlauka, 1997; Witmer, Bailey, Knerr, & Parsons, 1996), but have also shown equivalent navigation performance across real-world and virtual navigation tests

(Lloyd, Persaud, & Powell, 2009; Richardson, Montello, & Hegarty, 1999). Three studies have addressed the equivalence of real-world and virtual navigation tests in ABI patient groups. Cushman and colleagues (2008) compared performance on a real-world navigation test to that on a virtual version. They found a strong correlation ($r = .73$) across all participants, including MCI and early Alzheimer's disease patients. Sorita and colleagues (2013) compared a real-world and a virtual navigation test by testing traumatic brain injury patients in a between-participants design. Whereas route retracing performance was comparable in the real-world and virtual conditions, patients in the real-world condition were better in scene ordering and a trend existed for better sketch-mapping performance in this condition. The authors, therefore, concluded that the spatial representations probably differed between the real-world and virtual conditions. These two studies share the use of identical environments in their real-world and virtual navigation tests. Busigny and colleagues (2014), in contrast, applied different navigation tasks in their real-world and computerized tests. Nonetheless, they still reported a strong correlation ($r = .80$) between performances on the two test procedures in the patient group. They, however, also argued that their real-world navigation tests were more sensitive in revealing navigation impairment in their patients with posterior cerebral artery infarctions.

In the current study, a group of 68 chronic stroke patients completed both a virtual navigation test, that is, the Virtual Tübingen test (Claessen, Van der Ham, Jagersma, & Visser-Meily, in press; Claessen, Visser-Meily, Jagersma, Braspenning, & Van der Ham, 2016; Van der Ham et al., 2010), and a real-world navigation test. This was done to verify the convergent validity of virtual navigation testing. The study focused on this patient group, as they frequently complain about navigation problems after their stroke event (Van der Ham et al., 2013). The approach taken here is unique in two respects. First, the study relies on a large and representative sample of chronic stroke patients, which is uncommon in the clinical literature on navigation ability. And, second, the within-participants design allows a direct investigation of the relationship between virtual and real-world navigation performance for which significant correlations are expected. Stroke patients' performances on the two navigation tests were compared to that of a group of healthy control participants. It is expected that stroke patients have more difficulties with the navigation tasks than controls and that performance is comparable for the real-world and virtual environments. In contrast to Cushman and colleagues (2008), different rather than identical environments were used in the real-world and virtual tests to prevent unwanted learning effects.

METHODS

Participants

Sixty-eight chronic stroke patients (time post-stroke varied between 14 and 86 months, $M = 38.4$; $SD = 15.3$) were recruited from the rehabilitation clinic of De Hoogstraat Revalidatie and the rehabilitation department of the University Medical Center Utrecht (Utrecht, the Netherlands). Inclusion criteria were the ability to walk independently and the absence of severe aphasia. In addition, 44 healthy participants served as controls. Most of them were directly recruited by the experimenters (relatives or acquaintances) or were partners of patients. None of them reported a history of visual, neurological, psychiatric, or mobility problems, or substance abuse. Demographic data (gender, age, and educational level) of all participants and stroke characteristics (type and location) of the patients are provided in Table 1.

All participants provided written consent after being informed about the study's purpose. Participants received a small monetary compensation for engaging in the study and their travelling costs were reimbursed. The procedures reported here satisfied the regulations as set by the Declaration of Helsinki and were approved by the medical ethical review board of the University Medical Center Utrecht (protocol no. 12-198). This study's dataset results from a larger project on navigation ability in stroke patients. A small proportion of these data are also presented in Claessen, Visser-Meily, and colleagues (2016).

Materials and Procedure

Each assessment started with participants completing a brief neuropsychological screening comprising four common neuropsychological tests. Next, the virtual and real-world navigation tests were administered in fixed order. The virtual test was always presented first, as we aimed to assess the virtual navigation test in as many stroke patients as possible for the larger project. Participants were required to take a break after the virtual navigation test to prevent fatigue. Additional breaks were given on request in between the neuropsychological tests. It took participants two and a half hours on average to complete the full assessment procedure. Six experimenters were involved in data collection. All experimenters were trained and supervised by the same researcher to minimize differences in the assessment of the tests (M.B.; see Acknowledgements).

Neuropsychological screening

The screening consisted of four neuropsychological tests administered in the order as listed below. These commonly used tests were included to obtain a general indication of the participants' neuropsychological profile. The screening contained only tests assessing the most relevant cognitive functions and was kept brief to ensure feasibility of the entire assessment procedure for the stroke patients.

The Dutch version of the Adult Reading Test (DART; in Dutch: NLV, Nederlandse Leestest voor Volwassenen) was applied as a measure of premorbid intelligence (Schmand, Lindeboom, & Van Harskamp, 1992). Raw scores were converted to an estimated premorbid intelligence quotient adjusted for age, gender and level of education.

Table 1. Demographic data for patients and controls, and patients' stroke characteristics

	Patients ($n = 68$)	Controls ($n = 44$)	test value	p	Effect size
Age in years	59.5 (12.5)	60.3 (10.2)	$t < 1$	.708	—
Male/female (%)	57.4%/42.6%	45.5%/54.5%	$\chi^2 = 1.52$	.218	*Phi* $= -0.12$
Education	5.2 (1.4)	5.7 (0.9)	$U = 1211$	.077	$r = -0.17$
Stroke type					
Ischemic stroke	54 (79.4%)				
Hemorrhagic stroke					
Intracerebral	10 (14.7%)				
Subarachnoid	3 (4.4%)				
Unspecified/unavailable	1 (1.5%)				
Stroke location					
Supratentorial region					
Left	27 (39.7%)				
Right	29 (42.7%)				
Bilateral	1 (1.5%)				
Infratentorial region					
Left	2 (2.9%)				
Right	2 (2.9%)				
Bilateral	6 (8.8%)				
Unspecified/unavailable	1 (1.5%)				

Note. The upper part of the table displays demographic data (age, gender and educational level based on Verhage (1964, possible range: 1–7)) for patients and healthy controls. Differences in demographics were assessed using an independent *t* test (age), a chi-square test (gender), and a Mann-Whitney test (educational level). Standard deviations are displayed in parentheses for age and educational level. The bottom part of the table provides descriptive information on the stroke characteristics of the patient group.

The Corsi Block-Tapping Task was used as a representative of visuospatial attention span (forward condition: Kessels, Van Zandvoort, Postma, Kappelle, & De Haan, 2000) and visuospatial working memory capacity (backward condition: Kessels, Van den Berg, Ruis, & Brands, 2008). Raw data were converted to percentiles correcting for age.

The *Trail Making Test* (TMT; Reitan, 1992) was administered to obtain measures of mental processing speed (part A) and divided attention (part B). Raw scores were converted to percentiles based on the norms provided by the Neuropsychology section of the Dutch Association of Psychologists (Schmand, Houx, & De Koning, 2012). These norms correct for the effects of age, gender and educational level and provide three scores: part A, part B, and part B corrected for performance on part A.

The *Digit Span subtest of the WAIS-III* (Wechsler, 1997) was used to measure verbal working memory span. Norms correcting for age from the Dutch manual were used to convert raw scores to a scaled score and an accompanying percentile score.

Navigation test batteries

Participants completed a virtual navigation test (Virtual Tübingen test; see Claessen, Van der Ham et al., in press; Claessen, Visser-Meily et al., 2016; Van der Ham et al., 2010) and a real-world navigation test. Knowledge of the studied route was assessed by way of eight subtasks in both navigation tests.

Virtual environment

In the learning phase, participants were shown one of two routes (see Figure 1a) through a photorealistic virtual rendition of the German city Tübingen (Van Veen, Distler, Braun, & Bülthoff, 1998), twice in immediate succession. The two movies were nearly comparable in duration (A: 210 s and B: 253 s), similar in distance (400 m) and in movement speed (somewhat above walking speed). Each route contained 11 decision points. An actual left or right turn was taken at seven of these decision points, whereas the route continued in straight-ahead direction at the other four decision points. Eight subtasks were used to assess the participants' knowledge of the studied route in the testing phase (see below).

Real-world environment

A route (426 m) through the immediate vicinity of the rehabilitation clinic of De Hoogstraat Revalidatie was used for the real-world navigation test (see Figure 1b). This environment is located in an urban area (Utrecht, the Netherlands). No exceptionally salient landmarks or route signs were present along the test route. This environment was used for the real-world navigation test for practical reasons. It would have been impossible to take all participants to another test location that would be unfamiliar to all of them.

The participant followed the experimenter throughout the route, which lasted 324.9 s ($SD = 78.5$ s) on average. Experimenters were instructed to take the walking speed of the participant into account. The configuration of the real-world
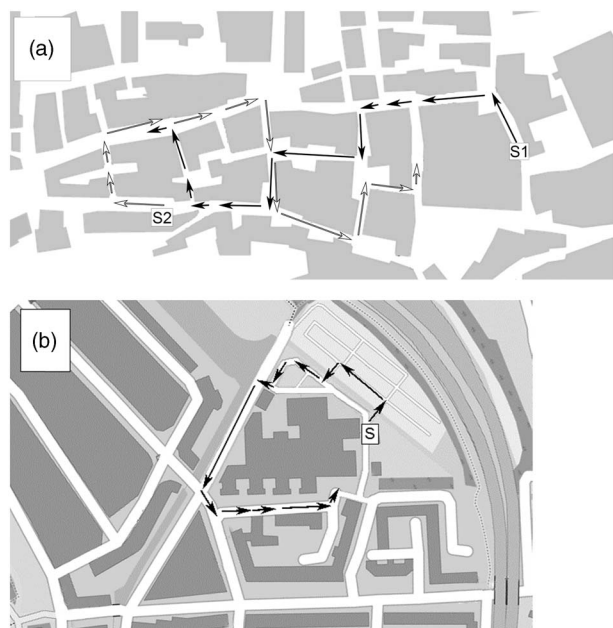
Figure 1. Maps of the two Virtual Tübingen routes and the route applied in the real-world test. (a) The first map displays the two Virtual Tübingen routes (black and white arrows). Each route segment is represented as an arrow. Starting locations of the routes are marked with an S and the corresponding route number. (b) The second map shows the route as used in the real-world navigation test (direct vicinity of rehabilitation clinic "De Hoogstraat Revalidatie" in Utrecht, the Netherlands). Route segments are displayed as arrows and the starting position of the route is marked with an S. (Figure 1A adapted from Claessen, M.H.G., Visser-Meily, J.M.A., Jagersma, E., Braspenning, M.E., & van der Ham, I.J.M. (2016). Dissociating spatial and spatiotemporal aspects of navigation ability in chronic stroke patients. *Neuropsychology*, doi: 10.1037/neu0000260 [Epub].)

route was matched as closely as possible to the virtual route: it also contained 11 decision points including seven actual left or right turns. The route continued in straight-ahead direction at the remaining four decision points. The participant was requested to perform the eight subtasks as described below for the real-world route upon return in the test room. Participants were asked to indicate their familiarity with the real-world environment at the end of the test procedure (1 = "not familiar at all" to 7 = "highly familiar"). We asked for this information, as nearly half of the patients had completed their rehabilitation in the rehabilitation clinic of the De Hoogstraat Revalidatie. They might thus have been more familiar with the test environment than the patients recruited through the University Medical Center Utrecht and the healthy control participants.

Navigational subtasks

The navigation tests contained eight subtasks assessed in the order of appearance below. The first four subtasks address route knowledge aspects, while the latter four subtasks rely on integration of the geometrical aspects of the environment, which is considered survey knowledge.

Scene Recognition. Twenty-two images of decision points taken from the studied route were presented to the participant. Eleven of these images[1] were targets (i.e., encountered during the route), whereas the other 11 scenes were distractors. Scoring: Number of correct responses, range: 0–22.

Route Continuation. The 11 decision points taken from the route were presented one-by-one in random order. Participants were asked to indicate the direction in which the route continued at each decision point. Scoring: Number of correct responses, range: 0–11.

Route Sequence. Participants were requested to indicate the sequence of turns as taken during the route. They responded by arranging a set of arrow cards. Only actual turns (i.e., left and right turns) were considered. Accuracy: Number of correctly indicated turns in the sequence, range 0–7.

Route Order. Participants were instructed to reconstruct the order in which 11 images of decision points occurred during the route. Scoring: Three points for each image assigned to its correct position in the sequence; two points for images assigned one position too late or too early; one point for images assigned two positions away from correct placement (range: 0–33).

Distance Estimation. Participants were requested to provide a distance estimate of the studied route. Scoring: Absolute deviation from the correct response in meters.

Duration Estimation. Participants were required to provide a duration estimate of the studied route. Scoring: Absolute deviation from the correct response in seconds.

Route Drawing. Participants were asked to draw the studied route on a map of the test environment. Only the starting point and the correct starting direction were already provided. Scoring: One point for each correctly indicated direction (left turn, straight forward or right turn) at relevant decision points, range: 0–11.

Map Recognition. Participants had to select the correct map of the route out of four options. Scoring: Correct or incorrect.

Statistical Analysis

Differences in demographics were assessed using an independent *t* test (age), a chi-square test (gender), and a Mann-Whitney test (educational level). Group differences on neuropsychological measures were investigated using independent *t* tests. Self-rated familiarity with the real-world environment between the groups was tested using an independent *t* test. Relationships between familiarity and real-world subtask performance were investigated by way of a Pearson correlation analysis. A semi-partial correlation analysis was performed to assess relationships between subtask scores on the real-world and virtual navigation tests while controlling for the effect of familiarity on real-world navigation performance. A repeated measures analysis of covariance (ANCOVA) was then performed for each subtask, with environment (real-world *vs.* virtual) as within-subject

---

[1] These 11 decision point images were also used for the route continuation and route order subtasks. Images were taken right in front of the decision point depicting all possible directions.

factor and group (healthy controls *vs.* stroke patients) as between-subject factor. ANCOVAs were corrected for educational level and familiarity with the real-world environment, due to the (trend-level) differences between controls and patients on these variables (see Tables 1 and 3). Due to its ordinal scale, educational level was recoded into low and high levels (1–4 *vs.* 5–7; Verhage, 1964) and included as a between-subject factor rather than as a covariate. Familiarity with the real-world environment was taken into account as a covariate. In case the initial analysis indicated a significant contribution of educational level and/or familiarity ($p < .05$), these variables were maintained in the ANCOVA.

The real-world Scene Recognition score of one patient was missing due to a technical problem. Moreover, one patient did not provide distance and duration estimates for the real-world route. Alpha level was set to .05 for all statistical tests. Analyses were performed using IBM SPSS Statistics version 22.0.

RESULTS

Demographics and Neuropsychological Screening

Patients and controls were comparable in terms of age and gender ($p = .708$ and $p = .218$, respectively, see Table 1). The comparison of educational level between the groups was also nonsignificant, but a trend ($p = .077$) existed for patients being slightly lower educated than controls. Patients obtained significantly lower scores on the majority of the neuropsychological screening tasks compared to controls (see Table 2).

Self-Rated Familiarity with the Real-World Environment

Patients were significantly more familiar ($M = 4.88$; $SD = 1.88$) with the real-world environment than controls ($M = 1.66$;

Table 3. Correlations between self-rated familiarity with the real-world environment and performance on the real-world navigation subtasks for patients and controls

	Patients		Controls	
	r	*p*	*r*	*p*
Scene Recognition	.352	.003*	.330	.031*
Route Continuation	.206	.092	.213	.166
Route Sequence	.237	.052	.171	.266
Route Order	.414	<.001*	.211	.114
Distance Estimation	.114	.356	.021	.891
Duration Estimation	−.074	.551	−.163	.291
Route Drawing	.192	.116	−.039	.802
Map Recognition	.130	.292	−.001	.992

Note. Displayed correlations are based on Pearson correlation coefficients, only the correlations of the Map Recognition subtask concern point-biserial correlation coefficients.
*$p < .05$.

$SD = 1.40$), $t(107.80) = -10.39$, $p < .001$, $r = .71$. Hence, a Pearson correlation analysis was conducted to verify the relationship between self-rated familiarity with the environment and performance on the real-world navigation subtasks (see Table 3). Only one significant correlation was found in the control group (Scene Recognition). In the patient group, two correlations were found to be significant (Scene Recognition and Route Order) and two other correlations reached trend level (Route Continuation and Route Sequence).

Relationship between the Real-World and Virtual Tübingen Navigation Tests

Semi-partial correlations reached significance for three subtasks in controls, together with four significant correlations and one trend-level ($p = .077$) correlation in the patient group

Table 2. Neuropsychological screening results for patients and controls

	Patients ($n = 68$)	Controls ($n = 44$)	*t*	*p*	Effect size *r*
Dutch Adult Reading Test (IQ)	97.5 (17.0)	110.9 (12.9)	4.70	<.001*	0.41
Corsi Block-Tapping Task					
forward (span × score)	37.8 (15.8)	42.6 (12.9)	1.67	.097	0.16
backward (span × score)	38.8 (20.5)	46.6 (16.3)	2.25	.027*	0.21
Trail Making Test					
Part A (seconds)	57.3 (39.0)	35.6 (12.3)	−4.26	<.001*	0.42
Part B (seconds)	124.8 (85.6)	74.0 (24.9)	−4.54	<.001*	0.45
Part B (B/A)	2.4 (1.0)	2.2 (0.6)	−1.59	.116	0.15
Digit Span (WAIS-III)					
forward (score)	7.7 (1.8)	8.7 (1.6)	3.00	.003*	0.27
backward (score)	5.1 (1.9)	6.2 (2.0)	2.91	.004*	0.27

Note. Group differences were tested by way of independent *t* tests. Effect size *r* is reported for significant results. Standard deviations are displayed in parentheses.
*$p < .05$.

Table 4. Performance on the eight subtasks of the virtual and real-world navigation tests and their correlations, displayed for patients and controls separately

	Virtual environment	Real-world environment	Correlation	
	Mean (SD)	*Mean* (SD)	VT-Eco, *r*	*p*
Stroke patients				
Scene Recognition	16.8 (2.4)	18.6 (2.2)	.216	.077
Route Continuation	7.0 (2.0)	8.0 (1.6)	.269	.027*
Route Sequence	3.5 (1.9)	4.7 (1.7)	.266	.029*
Route Order	15.0 (6.6)	25.6 (6.8)	.350	.003*
Distance Estimation	1468.7 (1342.3)	535.7 (1320.5)	.306	.012*
Duration Estimation	427.7 (796.0)	319.5 (274.4)	.035	.776
Route Drawing	4.0 (3.0)	8.0 (3.0)	−.028	.823
Map Recognition	44.1% correct	85.3% correct	−.141	.253
Healthy controls				
Scene Recognition	18.0 (2.0)	18.5 (1.8)	.071	.649
Route Continuation	8.1 (1.9)	8.4 (1.3)	.371	.013*
Route Sequence	3.7 (1.9)	5.3 (1.4)	.039	.801
Route Order	18.3 (7.2)	25.2 (7.2)	.306	.043*
Distance Estimation	1294.3 (1209.9)	174.7 (373.9)	.558	<.001*
Duration Estimation	103.2 (471.1)	117.2 (194.9)	.206	.170
Route Drawing	5.3 (3.2)	8.8 (2.3)	.210	.170
Map Recognition	63.6% correct	93.2% correct	.170	.269

Note. Relationships between virtual and real-world navigation performance were investigated by semi-partial correlation coefficients to correct for the effect of self-reported familiarity on real-world navigation performance. The (uncorrected) point-biserial correlation was applied for the Map Recognition subtask. Possible range of scores: Scene Recognition = 0–22, Route Continuation = 0–11, Route Sequence = 0–7, Route Order = 0–33, Distance Estimation = Absolute deviation from correct response in meters, Duration Estimation = Absolute deviation from correct response in seconds, Route Drawing = 0–11, and Map Recognition = correct or incorrect.
*$p < .05$.

(see Table 4). A composite score of overall performance was calculated for the real-world and virtual navigation tests in the patient group (based on the means and standard deviations of controls). The semi-partial correlation between the two composite scores was moderate in degree, $r = .535, p < .001$, indicating moderate overall overlap between the two navigation tests in patients. Two further analyses were performed using separate composite scores for the route and survey knowledge subtasks (see Methods section). Moderate overlap was found between the two route knowledge

composite scores, $r = .523, p < .001$, whereas the correlation between the two survey knowledge composite scores was weak to moderate, $r = .442, p < .001$.

Effects of Group and Environment on Navigation Performance

Results of the repeated measures ANCOVAs for each of the eight subtasks are presented in Table 5. A significant main-effect of group was found for seven out of the eight subtasks

Table 5. Main effects of group and environment on the navigation tests, together with the interaction effect between group and environment

	Group	Environment	Group × Environment
Scene Recognition	$F(1,108) = 8.89, p < .01, \eta p^2 = .08$*	$F < 1$	$F < 1$
Route Continuation	$F(1,110) = 8.15, p < .01, \eta p^2 = .07$*	$F(1,110) = 10.32, p < .01, \eta p^2 = .09$*	$F(1,110) = 3.83, p = .053, \eta p^2 = .03$
Route Sequence	$F(1,109) = 6.46, p = .01, \eta p^2 = .06$*	$F(1,109) = 5.46, p = .02, \eta p^2 = .05$*	$F(1,109) = 1.65, p = .20, \eta p^2 = .02$
Route Order	$F(1,109) = 8.55, p < .01, \eta p^2 = .07$*	$F(1,109) = 10.95, p < .01, \eta p^2 = .09$*	$F < 1$
Distance Estimation	$F(1,109) = 1.33, p = .25$	$F(1,109) = 51.86, p < .01, \eta p^2 = .32$*	$F < 1$
Duration Estimation	$F(1,109) = 7.98, p < .01, \eta p^2 = .07$*	$F < 1$	$F < 1$
Route Drawing	$F(1,109) = 10.28, p < .01, \eta p^2 = .09$*	$F(1,109) = 23.26, p < .01, \eta p^2 = .18$*	$F < 1$
Map Recognition	$F(1,110) = 5.98, p = .02, \eta p^2 = .05$*	$F(1,110) = 37.09, p < .01, \eta p^2 = .25$*	$F(1,110) = 1.00, p = .32, \eta p^2 = .01$

Note. ANCOVAs were corrected for educational level and familiarity with the real-world environment, in case a significant contribution of these variables to performance on that subtask existed ($p < .05$).
*$p < .05$.

showing that patients had more difficulties with the navigation tasks than controls. The main-effect of environment was significant for six of the eight subtasks indicating higher performance based on the real-world environment in comparison to the virtual environment. More importantly, the interaction-effect between group and environment was nonsignificant for all subtasks (except for one trend-level interaction-effect, $p = .053$, on the Route Continuation task), meaning that the differences in performance between patients and controls were similar in the real-world and virtual environment.

DISCUSSION

The primary objective of this study was to establish the relationship between performance on a real-world and a virtual navigation test in chronic stroke patients. This was done to investigate whether virtual navigation testing might be a valid alternative to real-world navigation testing, as the latter type of testing is usually associated with many practical limitations.

In line with expectations, there were significant correlations between four subtasks as assessed in both navigation tests in the group of stroke patients. More specifically, real-world and virtual performance on subtasks addressing place-action associations (Route Continuation), order of turns (Route Sequence), order of scenes (Route Order), and Distance Estimation was significantly correlated. These findings seem to suggest that virtual navigation testing is only valid for the administration of route knowledge aspects. That is, three of the four route knowledge subtasks correlated across the environments, whereas this was only the case for one of the four survey knowledge subtasks. Further analyses based on separate composite scores for route and survey knowledge subtasks, however, indicate that this initial conclusion is not correct. Route knowledge composite scores were moderately correlated across the real-world and virtual environments, whereas this correlation was lower but still weak to moderate in degree for the survey knowledge composite scores.

Furthermore, the composite scores of overall performance were found to be moderately related indicating moderate overlap between performance on the real-world and virtual navigation tests in patients. These correlation analyses were based on semi-partial correlation coefficients to correct for the effect of self-rated familiarity on real-world performance. With regard to the administration of route knowledge, the current findings thus provide evidence in favor of the convergent validity of virtual navigation testing as an alternative to real-world navigation tests. In addition, when performance on the survey knowledge subtasks is combined into a single composite score, virtual navigation testing might also be suitable for measuring survey knowledge.

A different series of analyses was performed to compare navigation performance of stroke patients to that of healthy controls. The hypothesis that patients would experience more difficulties with the navigation tasks than controls was

supported by this analysis. Patients indeed scored significantly lower than controls on seven subtasks with the exception of the Distance Estimation subtask. Furthermore, it was found that the real-world and virtual navigation tests were not equal in their level of difficulty. Regardless of group, performance on the real-world test was significantly better on six out of the eight subtasks as compared to performance on the virtual navigation task. Nevertheless, none of the interaction-effects between group and environment reached significance. This finding indicates that the difference in real-world and virtual navigation performance was thus similar for patients and controls. Importantly, these results were obtained after statistical corrections for the (trend-level) differences between patients and controls on educational level and self-reported familiarity with the real-world environment were applied.

The correlational analysis as described above has indicated moderate overlap between scores on the virtual and real-world navigation tests. Although this result corroborates findings of earlier studies showing overlap between real-world and virtual navigation performance (Busigny et al., 2014; Cushman et al., 2008; Sorita et al., 2013), the correlation between the composite scores was somewhat weaker than reported by two of these studies (Busigny et al., 2014; Cushman et al., 2008). This might in part be a result of methodological differences between these studies and ours. Cushman and colleagues (2008) used exactly the same environment and subtasks in both test procedures, whereas Busigny and colleagues (2014) used rather different navigation tasks in the real-world and virtual conditions. Both studies relied on a within-subject design. In contrast, we administered the same eight subtasks in the real-world and virtual tests, but used different environments. As a consequence, learning effects with regard to the environment cannot have occurred in our study between the real-world and virtual navigation tests.

When comparing navigation performance based on the two different environments, results showed that the virtual navigation test was consistently more difficult than the real-world navigation test in both groups. Several factors could be responsible for this difference in performance, for example differences in the scenery of the environments or in the configuration of the routes. In our view, however, the higher performance level in the real-world test is the primary result of the fact that the exposure to the real-world environment allowed for a more complete navigation experience. More specifically, previous studies have argued that information from multiple sensory systems contributes to navigation behavior: visual, vestibular, and proprioceptive information (Berthoz & Viaud-Delmon, 1999). Whereas exposure to the virtual environment provided participants only with visual information, exploring the real-world environment allowed for integration of visual and physical information (i.e., vestibular and proprioceptive cues). We pose that elevated performance in the real-world test relative to the virtual test follows from the fact that multisensory integration is only possible in the former.

Recent studies have speculated that locomotion contributes to the acquisition of survey knowledge, while visual information alone might be sufficient for acquiring route knowledge (e.g., Chrastil & Warren, 2013; Van der Ham et al., 2015). In our study, three of the four subtasks that correlated significantly across the real-world and virtual tests in the patient group concerned route knowledge aspects (i.e., place-action associations, the order of turns and scenes). On the other hand, most of the subtasks relying on survey knowledge aspects showed no significant correlations between real-world and virtual performance. When performance on individual survey knowledge subtasks was, however, combined into a composite score, a weak to moderate correlation was found between the real-world and virtual tests. Although it might thus be necessary to combine performances due to the single-trial nature of three survey knowledge subtasks, these findings suggest that the acquisition of survey knowledge can be measured in a virtual navigation test.

In the current study, self-reported familiarity with the real-world environment was taken into account, as the patient group was more familiar with the real-world environment than controls. This was due to the fact that half of the patients had stayed in the rehabilitation center that is situated in the environment that was used for the real-world navigation test. A correlation analysis showed that familiarity was positively correlated to performance on tasks assessing route knowledge (i.e., recognition of scenes and their order; trends for place-action associations and order of turns) but not to the survey knowledge subtasks in patients. We hypothesize that previous exposure or exposures to the real-world environment might have helped them to infer what landmarks could or could not be present or logically follow each other in the studied route.

The current study has several notable strengths. An important strength is that, in comparison to earlier work, this study incorporates a relatively large sample of chronic stroke patients. The fact that patients with various stroke types and locations are included in our sample allows the current results to be broadly generalized to the stroke patient population. A further strength of our study lies in the fact that the same eight subtasks were assessed for the real-world and virtual navigation tests, while each test was based on a different environment. This enabled us, due to the within-subject design, to directly compare real-world and virtual navigation performance within each participant.

Some limitations should be discussed. First, information with regard to the neuropsychological status of the patients was relatively limited. For example, no information was available on the presence of visuospatial neglect, a syndrome that might affect navigation performance (De Nigris et al., 2013; Guariglia et al., 2005; Nico et al., 2008). Furthermore, for practical considerations, the virtual navigation test was administered first in all participants. Performance in the real-world test might thus be elevated because the participants were already familiar with the content of the eight subtasks. However, it remains unlikely that this fixed order influenced the relationship between real-world and virtual navigation performance itself. Furthermore, the fact that the patient group was more familiar with the environment as used in the real-world navigation test, might be regarded as a limitation of the study. However, statistical corrections for this difference were applied by taking self-rated familiarity with the real-world environment into account. We also state that this group difference in familiarity with the real-world environment clearly illustrates an important practical limitation associated with any real-world navigation test. In contrast, a virtual navigation test can be assessed in a highly standardized manner, guaranteeing equal exposure and familiarity across participants.

In summary, this study compared performance on a real-world and a virtual navigation test in 68 chronic stroke patients. Results demonstrated a moderate correlation between composite scores on the two navigation tests. Additional analyses indicated moderate overlap between real-world and virtual performance on route knowledge subtasks, whereas this relationship was weak to moderate for subtasks addressing survey knowledge aspects. These findings suggest that virtual navigation testing could serve as a valid alternative to real-world navigation tests. As a next step in this line of research, the Virtual Tübingen test should be administered in a large, heterogeneous group of healthy participants. This is necessary to generate normative data which would allow implementation of the test in clinical neuropsychological practice.

REFERENCES

Barrash, J., Damasio, H., Adolphs, R., & Tranel, D. (2000). The neuroanatomical correlates of route learning impairment. *Neuropsychologia*, *38*(6), 820–836. doi:10.1016/S0028-3932(99)00131-1

Berthoz, A., & Viaud-Delmon, I. (1999). Multisensory integration in spatial orientation. *Current Opinion in Neurobiology*, *9*(6), 708–712. doi:10.1016/S0959-4388(99)00041-0

Brunsdon, R., Nickels, L., & Coltheart, M. (2007). Topographical disorientation: Towards an integrated framework for assessment. *Neuropsychological Rehabilitation*, *17*(1), 34–52. doi:10.1080/09602010500505021

Burgess, P.W., Alderman, N., Forbes, C., Costello, A., Coates, L. M.-A., Dawson, D.R., … Channon, S. (2006). The case for the development and use of "ecologically valid" measures of executive function in experimental and clinical neuropsychology. *Journal of the International Neuropsychological Society*, *12*, 194–209. doi:10.1017/S1355617706060310

Busigny, T., Pagès, B., Barbeau, E.J., Bled, C., Montaut, E., Raposo, N., … Pariente, J. (2014). A systematic study of topographical memory and posterior cerebral artery infarctions. *Neurology*, *83*(11), 996–1003. doi:10.1212/WNL.0000000000000780

Chaytor, N., & Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. *Neuropsychology Review*, *13*(4), 181–197. doi:10.1023/B:NERV.0000009483.91468.fb

Chrastil, E.R., & Warren, W.H. (2013). Active and passive spatial learning in human navigation: Acquisition of survey knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(5), 1520–1537. doi:10.1037/a0032382

Claessen, M.H.G., Van der Ham, I.J.M., Jagersma, E., & Visser-Meily, J.M.A. (in press). Navigation strategy training using virtual reality in six chronic stroke patients: A novel and explorative approach to the rehabilitation of navigation impairment. *Neuropsychological Rehabilitation*. doi:10.1080/09602011.2015.1045910

Claessen, M.H.G., Visser-Meily, J.M.A., Jagersma, E., Braspenning, M.E., & Van der Ham, I.J.M. (2016). Dissociating spatial and spatiotemporal aspects of navigation ability in chronic stroke patients. *Neuropsychology*, doi:10.1037/neu0000260 [Epub].

Cushman, L.A., Stein, K., & Duffy, C.J. (2008). Detecting navigational deficits in cognitive aging and Alzheimer disease using virtual reality. *Neurology*, *71*, 888–895. doi:10.1212/01.wnl.0000326262.67613.fe

De Goede, M., & Postma, A. (2015). Learning your way in a city: Experience and gender differences in configurational knowledge of one's environment. *Frontiers in Psychology*, *6*, 402. doi:10.3389/fpsyg.2015.00402

De Nigris, A., Piccardi, L., Bianchini, F., Palermo, L., Incoccia, C., & Guariglia, C. (2013). Role of visuo-spatial working memory in path integration disorders in neglect. *Cortex*, *49*, 920–930. doi:10.1016/j.cortex.2012.03.009

Guariglia, C., Piccardi, L., Iaria, G., Nico, D., & Pizzamiglio, L. (2005). Representational neglect and navigation in real space. *Neuropsychologia*, *43*(8), 1138–1143. doi:10.1016/j.neuropsychologia.2004.11.021

Iachini, T., Ruotolo, F., & Ruggiero, G. (2009). The effects of familiarity and gender on spatial representation. *Journal of Environmental Psychology*, *29*, 227–234. doi:10.1016/j.jenvp.2008.07.001

Kessels, R.P.C., Van den Berg, E., Ruis, C., & Brands, A.M.A. (2008). The backward span of the Corsi Block-Tapping Task and its association with the WAIS-III Digit Span. *Assessment*, *15*(4), 426–434. doi:10.1177/1073191108315611

Kessels, R.P.C., Van Zandvoort, M.J.E., Postma, A., Kappelle, L.J., & De Haan, E.H.F. (2000). The Corsi Block-Tapping Task: Standardization and normative data. *Applied Neuropsychology*, *7*(4), 252–258. doi:10.1207/S15324826AN0704_8

Livingstone, S.A., & Skelton, R.W. (2007). Virtual environment navigation tasks and the assessment of cognitive deficits in individuals with brain injury. *Behavioural Brain Research*, *185*(1), 21–31. doi:10.1016/j.bbr.2007.07.015

Lloyd, J., Persaud, N.V., & Powell, T.E. (2009). Equivalence of real-world and virtual-reality route learning: A pilot study. *Cyberpsychology & Behavior*, *12*(4), 423–427. doi:10.1089/cpb.2008.0326

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

Nadolne, M.J., & Stringer, A.Y. (2001). Ecologic validity in neuropsychological assessment: Prediction of wayfinding. *Journal of the International Neuropsychological Society*, *7*, 675–682.

Nemmi, F., Boccia, M., Piccardi, L., Galati, G., & Guariglia, C. (2013). Segregation of neural circuits involved in spatial learning in reaching and navigational space. *Neuropsychologia*, *51*, 1561–1570. doi:10.1016/j.neuropsychologia.2013.03.031

Nico, D., Piccardi, L., Iaria, G., Bianchini, F., Zompanti, L., & Guariglia, C. (2008). Landmark based navigation in brain-damaged patients with neglect. *Neuropsychologia*, *46*(7), 1898–1907. doi:10.1016/j.neuropsychologia.2008.01.013

Oudman, E., Van der Stigchel, S., Nijboer, T.C.W., Wijnia, J.W., Seekles, M.L., & Postma, A. (in press). Route learning in Korsakoff's syndrome: Residual acquisition of spatial memory despite profound amnesia. *Journal of Neuropsychology*. Online early view version. doi:10.1111/jnp.12058

Péruch, P., Belingard, L., & Thinus-Blanc, C. (2000). Transfer of spatial knowledge from virtual to real environments. In Freksa, C., Bauer, W., Habel, C., & Wender K. (Eds.), *Spatial cognition II, Lecture notes in artificial intelligence*, *Vol. 184*, pp. 253–264). Berlin: Springer. doi:10.1007/3-540-45460-8_19

Piccardi, L., Berthoz, A., Baulac, M., Denos, M., Dupont, S., Samson, S., & Guariglia, C. (2010). Different spatial memory systems are involved in small- and large-scale environments: Evidence from patients with temporal lobe epilepsy. *Experimental Brain Research*, *206*, 171–177. doi:10.1007/s00221-010-2234-2

Piccardi, L., Iaria, G., Bianchini, F., Zompanti, L., & Guariglia, C. (2011). Dissociated deficits of visuo-spatial memory in near space and navigational space: Evidence from brain-damaged patients and healthy older participants. *Aging, Neuropsychology, and Cognition*, *18*(3), 362–384. doi:10.1080/13825585.2011.560243

Prestopnik, J.L., & Roskos-Ewoldsen, B. (2000). The relations among wayfinding strategy use, sense of direction, sex, familiarity, and wayfinding ability. *Journal of Environmental Psychology*, *20*, 177–191. doi:10.1006/jevp.1999.0160

Reitan, R.M. (1992). *Trail Making Test. Manual for administration and scoring*. Tucson, AZ: Reitan Neuropsychological Laboratory.

Richardson, A.E., Montello, D.R., & Hegarty, M. (1999). Spatial knowledge acquisition from maps and from navigation in real and virtual environments. *Memory & Cognition*, *27*(4), 741–750. doi:10.3758/BF03211566

Rose, F.D., Brooks, B.M., & Rizzo, A.A. (2005). Virtual reality in brain damage rehabilitation: Review. *Cyberpsychology & Behavior*, *8*(3), 241–262. doi:10.1089/cpb.2005.8.241

Schmand, B., Houx, P., & De Koning, I. (2012). *Norms for neuropsychological tasks (Verbal Fluency, Stroop Color Word test, Trail Making Test, Story recall of the Rivermead Behavioural Memory Test and the Dutch version of the Rey Auditory Verbal Learning Test)*. Amsterdam: The Neuropsychology section of the Dutch Association of Psychologists.

Schmand, B., Lindeboom, J., & Van Harskamp, F. (1992). *Dutch Adult Reading Test*. Lisse, The Netherlands: Swets and Zeitlinger.

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

Sorita, E., N'Kaoua, B., Larrue, F., Criquillon, J., Simion, A., Sauzéon, H., … Mazaux, J.-M. (2013). Do patients with traumatic brain injury learn a route in the same way in real and

virtual environments? *Disability & Rehabilitation*, *35*(16), 1371–1379. doi:10.3109/09638288.2012.738761

Van Asselen, M., Kessels, R.P.C., Kappelle, L.J., Neggers, S.F.W., Frijns, C.J.M., & Postma, A. (2006). Neural correlates of human wayfinding in stroke patients. *Brain Research*, *1067*(1), 229–238. doi:10.1016/j.brainres.2005.10.048

Van der Ham, I.J.M., Faber, A.M.E., Venselaar, M., Van Krefeld, M.J., & Löffler, M. (2015). Ecological validity of virtual environments to assess human navigation ability. *Frontiers in Psychology*, *6*, 637. doi:10.3389/fpsyg.2015.00637

Van der Ham, I.J.M., Kant, N., Postma, A., & Visser-Meily, J.M.A. (2013). Is navigation ability a problem in mild stroke patients? Insights from self-reported navigation measures. *Journal of Rehabilitation Medicine*, *45*, 429–433. doi:10.2340/16501977-1139

Van der Ham, I.J.M., Van Zandvoort, M.J.E., Meilinger, T., Bosch, S.E., Kant, N., & Postma, A. (2010). Spatial and temporal aspects of navigation in two neurological patients. *Neuroreport*, *21*, 685–689. doi:10.1097/WNR.0b013e32

Van Veen, H.J., Distler, H.K., Braun, S., & Bülthoff, H.H. (1998). Navigating through a virtual city: Using virtual reality technology to study human action and perception. *Future Generation Computer Systems*, *14*, 231–242. doi:10.1016/S0167-739X(98)00027-2833aea78

Verhage, F. (1964). *Intelligence and age: Study with Dutch people from age 12 to 77*. Dissertation. Assen: Van Gorcum. (In Dutch).

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale—Third edition (WAIS-III)*. San Antonio, TX: Psychological Corporation.

Wiener, J.M., Büchner, S.J., & Hölscher, C. (2009). Taxonomy of human wayfinding tasks: A knowledge-based approach. *Spatial Cognition & Computation*, *9*, 152–165. doi:10.1080/13875860902906496

Wilson, P.N., Foreman, N., & Tlauka, M. (1997). Transfer of spatial information from a virtual to a real environment. *Human Factors*, *39*, 526–531. doi:10.1518/001872097778667988

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*(4), 413–428. doi:doi:10.1006/ijhc.1996.0060

Wolbers, T., & Hegarty, M. (2010). What determines our navigational abilities? *Trends in Cognitive Sciences*, *14*(3), 138–146. doi:10.1016/j.tics.2010.01.001