Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
INTELLIGENCE
Intelligence 34 (2006) 151–176

Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial-layout learning

Mary Hegarty [a,*], Daniel R. Montello [b], Anthony E. Richardson [c,1], Toru Ishikawa [d], Kristin Lovelace [a]

[a] Department of Psychology, University of California, Santa Barbara, Santa Barbara, CA 93106, United States
[b] Department of Geography, University of California, Santa Barbara, Santa Barbara, CA 93106, United States
[c] Department of Psychology, Hamilton College, United States
[d] Lamont–Doherty Earth Observatory, Columbia University, United States

Received 4 November 2004; received in revised form 22 June 2005; accepted 26 September 2005
Available online 14 November 2005

Abstract

Most psychometric tests of spatial ability are paper-and-pencil tasks at the "figural" scale of space, in that they involve inspecting, imagining or mentally transforming small shapes or manipulable objects. Environmental spatial tasks, such as wayfinding or learning the layout of a building or city, are carried out in larger spaces that surround the body and involve integration of the sequence of views that change with one's movement in the environment. In a correlational study, 221 participants were tested on psychometric measures of spatial abilities, spatial updating, verbal abilities and working memory. They also learned the layout of large environments from direct experience walking through a real environment, and via two different media: a desktop virtual environment (VE) and a videotape of a walk through an environment. In an exploratory factor analysis, measures of environmental learning from direct experience defined a separate factor from measures of learning based on VE and video media. In structural-equation models, small-scale spatial abilities predicted performance on the environmental-learning tasks, but were more predictive of learning from media than from direct experience. The results indicate that spatial abilities at different scales of space are partially but not totally dissociated. They specify the degree of overlap between small-scale and large-scale spatial abilities, inform theories of sex differences in these abilities, and provide new insights about what these abilities have in common and how they differ.

1. Introduction

There has been a long tradition of research on the measurement and classification of individual differences in spatial abilities (e.g., Carroll, 1993; Eliot & Smith, 1983; Lohman, 1988; McGee, 1979). In this literature, measures of spatial abilities have included tasks such as mental rotation of shapes, solving mazes, imagining the folding and unfolding of sheets of paper, and finding hidden figures. These "small scale" psychometric tests are similar insofar as they are all paper-and-pencil tests, and almost all involve perceptually examining, imagining, or mentally transforming representations of small shapes or manipulable objects, such as blocks or sheets of paper.

In contrast to small-scale spatial abilities, there have been relatively few attempts to assess individual differences in larger-scale or "environmental" spatial abili-

* Corresponding authors. Hegarty is to be contacted at Department of Psychology, University of California, Santa Barbara, Santa Barbara, CA 93106, United States. Montello, Department of Geography, University of California, Santa Barbara, CA 93106, United States.
E-mail addresses: hegarty@psych.ucsb.edu (M. Hegarty), montello@geog.ucsb.edu (D.R. Montello).
[1] Current address: Department of Psychology, St. Michael's College, VT, United States.

0160-2896/$ - see front matter
doi:10.1016/j.intell.2005.09.005

ties. Environmental spatial tasks include learning the layout of new environments, such as buildings or cities, navigation in known environments, and giving and interpreting verbal navigation directions (see reviews by Evans, 1980; Gärling & Golledge, 1987; Liben, Patterson, & Newcombe, 1981; Spencer, Blades, & Morsley, 1989). One goal of the research project reported below is to characterize the sources of individual differences in environmental spatial tasks, assessing whether performance of different environmental spatial tasks reflects a single underlying ability or a disparate set of abilities. The second goal is to examine the extent to which processing of spatial information at different scales of space reflects the same or different underlying abilities, as articulated by Cooper and Mumaw (1985): "To what extent [do] common processes and representations underlie skill in dealing with large-scale space and spatial ability as measured by standard aptitude tests?" (pp. 91–92).

Fig. 1 diagrams different possible models of the relationship between small-scale and large-scale spatial abilities. The Unitary model assumes that spatial abilities at the two scales of space are completely overlapping, as would be the case if they depended on exactly the same cognitive processes. The Total Dissociation model proposes that the two sets of abilities depend on completely distinct cognitive processes. The Partial Dissociation model proposes that the two sets of abilities rely on some common processes, but that ability at each scale of space depends on some unique processes not shared by abilities at the other scale of space. If this is the preferred model, we also need to specify the amount of common variance and the cognitive processes that are shared by the two sets of abilities. The Mediation model proposes that the two sets of abilities are completely dissociated, but that both are related to a third ability that mediates the relationship between large- and small-scale spatial ability.

1.1. Background literature

Although previous studies of spatial abilities have not made strong claims that abilities at different scales of space are either completely overlapping (Unitary model) or completely dissociated (Total Dissociation model), most studies emphasize either the commonalities or the dissociations between these abilities. First, by using "pictorial" scale stimuli, such as images on paper or computer monitors, to study large-scale spatial tasks such as navigation, the majority of studies in the literature have implicitly adopted the Unitary model (Montello, 1993). The fact that almost all tests of spatial abilities involve paper-and-pencil depictions of small objects (Eliot & Smith, 1983) also reflects an implicit assumption that all spatial cognition can be studied with small-scale stimuli. Other research has addressed the unitary model more explicitly. For example, many theories of the evolution of sex differences propose that differences in small-scale spatial abilities (e.g., mental rotation) reflect different selection pressures for navigational (large-scale) abilities between males and females in evolutionary history (Gaulin, 1995; Kimura, 2000). One recent review (Jones, Braithwaite, & Healy, 2003) considers seven theories of the evolution of sex

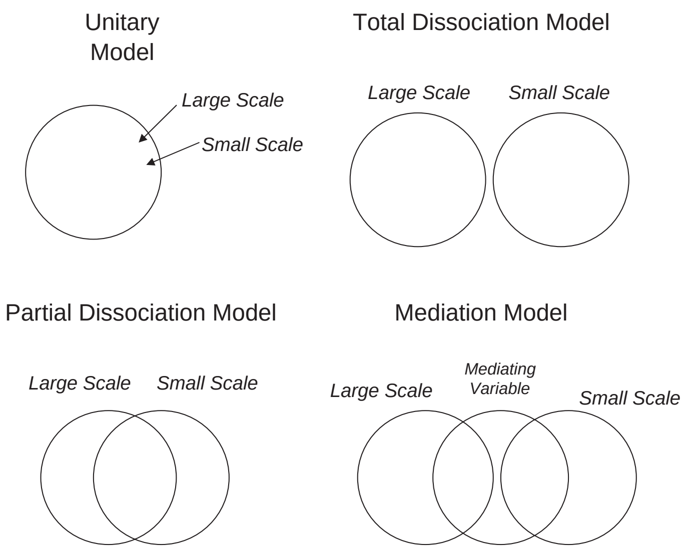

Fig. 1. Four models of possible relations between small-scale and large-scale spatial abilities.

differences in spatial abilities, all of which are based on demands for the sexes to have different amounts of mobility in the environment (e.g., for foraging, warfare, or finding mates). These theories assume that large-scale and small-scale spatial skills reflect the same abilities, and are causally related.

In contrast, some other theorists have emphasized the dissociations between the mental systems for processing spatial information at different scales of space. In comparing environmental space with the object space of traditional space perception, Ittelson (1973) was the first to point out that in contrast with objects, environments are larger than the individual, allow viewing from multiple vantage points, and require locomotion and information integration over time for their apprehension. Several other researchers in cognition and perception (e.g., Acredolo, 1981; Cutting & Vishton, 1995; Freundschuh & Egenhofer, 1997; Gärling & Golledge, 1987; Kuipers, 1982; Mandler, 1983; Montello, 1993; Montello & Golledge, 1999; Tversky, Morrison, Franklin, & Bryant, 1999; Zacks, Mires, Tversky, & Hazeltine, 2000) and some evolutionary psychologists (e.g. Silverman & Eals, 1992) have echoed this distinction in theorizing that spatial entities at different scales involve distinct cognitive structures and processes. These proposals are inconsistent with the Unitary model in Fig. 1, and consistent with the dissociation models, but do not specify whether the dissociation is partial or complete.

Of particular importance to our work are the distinctions between *figural*, *vista,* and *environmental* space (Montello, 1993; Montello & Golledge, 1999). Figural space is small in scale relative to the body and external to the individual, and can be apprehended from a single viewpoint. It includes both flat pictorial space and volumetric object space (e.g., small, manipulable objects) associated with psychometric tests of spatial ability. Vista space is projectively as large or larger than the body, but can be visually apprehended from a single place without appreciable locomotion. It is the space of single rooms, town squares, small valleys and horizons. Environmental space is large in scale relative to the body and "contains" the individual. It includes the spaces of buildings, neighborhoods, and cities, and it typically requires locomotion for its apprehension (see Montello, 1993, for a discussion of other scales of space).

The neuroscience literature provides considerable evidence that processing spatial information at different scales of space involves different brain structures and mechanisms. Whereas small-scale spatial tasks such as mental rotation are associated primarily with activation of the parietal lobes (see Kosslyn & Thompson, 2003, for a recent review), learning and remembering the layout of large-scale spaces is associated with processing in the hippocampus and surrounding regions in the medial temporal lobes (e.g., Morris & Parslow, 2004). Patients with various forms of topographical disorientation, who are impaired in wayfinding and environmental layout learning following brain lesions, do not typically show impairments in small-scale spatial abilities, such as mental rotation (Aguirre & D'Esposito, 1999). Parietal patients, showing typical impairments in small-scale mental spatial manipulations, in turn, show intact spatial updating during locomotion in large-scale space (Philbeck, Behrmann, Black, & Ebert, 2000). In an extensive review of the neuroscience literature, Previc (1998) proposed four "behavioral realms" for spatial behaviors, based on the sensorimotor systems involved in various actions, such as reaching and locomotion. These realms are quite tightly correlated with spatial scale. They are: peripersonal (near-body space), focal extrapersonal (the space of visual search and object recognition), action extrapersonal (orienting in topographically defined space), and ambient extrapersonal (orienting in earth-fixed space). Previc provides evidence that that the four realms are largely associated with distinct cortical networks: dorsolateral for peripersonal, ventrolateral for focal extrapersonal, ventromedial for action extrapersonal, and dorsomedial for ambient extrapersonal.

Studies of individual differences can provide important insights into the relation between large-scale and small-scale spatial cognition. If large- and small-scale spatial tasks depend on common basic capacities and processes, and there are individual differences in these capacities and processes, then measures of spatial cognition at different scales of space should be highly correlated.[2] In fact, previous individual differences studies have generally reported only weak, if any, relations between large-scale and small-scale spatial abilities. In a review of 12 studies of the relation between large- and small-scale spatial abilities (Allen, Kirasic, Dobson, Long, & Beck, 1996; Bryant, 1982; Goldin & Thorndyke, 1982; Juan-Espinosa, Abad, Colom, & Fernandez-Truchaud, 2000; Kirasic, 2000; Lorenz, 1988; Meld, 1985; Pearson & Ialongo, 1986; Rovine & Weisman, 1989; Sholl, 1988; Waller, 2000; Walsh, Krauss, & Regnier, 1981), Hegarty and Waller (2005)

[2] Note that the converse is not necessarily true. That is, if two measures are highly correlated, it does not necessarily imply that they are based on common capacities and processes. Two tasks that depend on different cognitive processes could be good predictors of each other, for example if they are influenced by similar genetic and/or environmental factors.

reported that there were only two studies in which the median correlation exceeded .3. The majority of these correlations were not statistically significant.

Previous factor-analytic studies also support the separability of small-scale and large-scale spatial abilities. Lorenz (1988; Lorenz & Neisser, 1986) administered several different measures of environmental spatial ability, including giving directions, recalling a route, and pointing to local and distant locations. In exploratory factor analyses, environmental abilities loaded on different factors than pencil-and-paper measures of spatial abilities. A similar dissociation between large- and small-scale spatial abilities was found by Pearson and Ialongo (1986). In the most extensive individual differences study to date, Allen et al. (1996) measured performance on six psychometric spatial tests (Surface Development, Cube Comparison, Hidden Figures, Gestalt Completion, Map Memory and Map Planning) and seven measures of learning of an environment from a walk through a small city. Five of the measures of learning (scene recognition, sequencing scenes along the route, placement of landmarks on a map of the route, intra-route distance estimates and route reversal) defined a large-scale learning factor that Allen et al. called “topological knowledge.” This factor was unrelated to the spatial ability factor defined by the psychometric tests. However another factor called “spatial sequential memory,” defined by measures of maze learning and maze reversal measured in a laboratory task, was related to both the psychometric test factor and the topological knowledge factor and therefore mediated the relationship between these two factors. Other measures of large-scale learning, ability to make straight-line distance estimates and direction estimates, did not load on the topological knowledge factor or on any other factor. There was again no direct relationship between these measures and small-scale spatial ability. However, in one experiment Allen et al. found that a measure of perspective taking mediated the relationship between paper-and-pencil measures of spatial ability and the measure of direction estimation (ability to point to non-visible locations in a learned environment).

In summary, previous individual differences research is inconsistent with the Unitary model in Fig. 1. Some individual differences studies are consistent with the Total Dissociation model while others are consistent with the Partial Dissociation model but suggest that the amount of overlapping variance between small- and large-scale spatial tasks is relatively small. Finally, the results of Allen et al. (1996) are consistent with the Mediation model depicted in Fig. 1, suggesting that other abilities mediate the relationship between small-scale spatial abilities and different aspects of large-scale learning.

1.2. *Information processing analysis of large-scale spatial learning*

An analysis of the perceptual and cognitive processes involved in large-scale spatial tasks can help us identify the possible sources of variance in those tasks, and inform questions of the amount and nature of the dissociation between large- and small-scale spatial abilities. The information processing approach has been used productively in the analysis of small-scale spatial abilities, to reveal that these abilities rely on differences in speed of encoding and transforming spatial information, spatial working memory capacity, and strategies (Hegarty & Waller, 2005; Just & Carpenter, 1985; Lohman, 1988; Pellegrino & Kail, 1982; Shah & Miyake, 1996). Although large-scale spatial cognition involves many complex activities, including learning the layout of new spaces, using existing knowledge of an environment to plan routes and navigate, and communicating about space, it is necessary to restrict our focus in an information-processing analysis. Consistent with previous individual-difference studies (Allen et al., 1996; Bryant, 1982; Goldin & Thorndyke, 1982; Kirasic, 2000; Pearson & Ialongo, 1986), we focus on ability to learn the layout of novel environments.

In studies of large scale learning, participants are given a controlled amount of exposure to a large scale environment, in our study by being led on a route through the environment, being shown a video, or interacting with a desktop virtual environment. They are tested on various “outcome measures” of learning, such as route retracing, pointing to non-visible locations, straight-line distance estimates and map sketching. Fig. 2 shows a preliminary analysis of the representations and processes involved in learning the layout of an environment and performing outcome measures of that learning. First, the layout of the environment must be encoded from the various sensory inputs available. This leads to an internal representation of the environment, which might be a route representation (i.e., a representation of the sequence of landmarks encountered and movements made in locomoting through the environment) or a survey representation (i.e., a two-dimensional representation of the configuration of the environment). The internal representation cannot be measured directly, but must be inferred from performance on the outcome measures. Furthermore the performance of a particular outcome measure (e.g.,

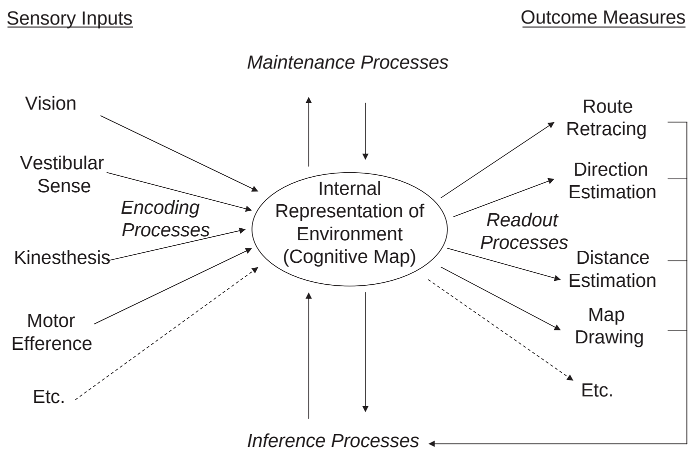

Fig. 2. Schematic depiction of the perceptual and cognitive processes involved in constructing a cognitive map of a large-scale space and using that cognitive map to perform different outcome measures.

direction estimation) may involve transformations or inferences from a person's initial representation of the environment that occur either when the person is exposed to the learning medium, or when his or her knowledge of the environment is being tested (Montello, Waller, Hegarty, & Richardson, 2004). A task analysis of large-scale spatial learning must therefore specify the cognitive processes involved both in constructing an internal representation of an environment from exposure to that environment and in transforming that internal representation to perform different outcome measures (see Fig. 2).

Based on the task analysis, we hypothesize three main sources of individual differences in large-scale spatial cognition: ability to encode spatial information from sensory experience, ability to maintain a high-quality internal representation of that information in memory, and ability to perform spatial transformations in order to make inferences from this spatial information. First, learning the spatial layout of an environment depends on the ability to encode spatial information from the various sources of sensory information provided in the learning experience. Vision is probably the main sensory modality in humans for sensing the spatial layout of an environment. However, when one directly navigates through a real environment, one also senses one's own movement through non-visual senses. Movement is sensed by the vestibular system, which provides information about linear and angular accelerations, kinesthesis, which senses movement of the limbs, and efference copy, which is based on signals from the central nervous system to the muscles. Recent research has indicated that these senses contribute to spatial updating and learning of spatial layout independently of vision (Chance, Gaunet, Beall, & Loomis, 1998; Klatzky, Loomis, Beall, Chance, & Golledge, 1998; Waller, Loomis, & Haun, 2004; but see Waller, Loomis, & Steck, 2003), so that it is likely that individual differences in learning from these senses also make an independent contribution to variability in learning of spatial layout. (Spatial layout can also be sensed to some extent by audition, but this will not be addressed in the current study).

Second, various aspects of memory affect the quality of the internal representations or cognitive maps constructed from a given amount of exposure to an environment. Working memory is a key factor underlying environmental learning, because environmental spaces cannot be apprehended in a single view (e.g., Ittelson, 1973; Montello, 1993), and these spaces must therefore be learned by maintaining information over time. The format of the memory constructed from exposure to an environment is another possible source of variance. Information might be retained in verbal working memory, as a sequence of route directions (Allen et al., 1996) or in spatial working memory, as a configuration. Sex differences in large-scale spatial cognition are often characterized as differences in the format of the internal representation, with females depending more on route representations and males depending more on configural representations of space (Lawton, 1994; Lawton, Charleston, & Zieles, 1996; Montello, Lovelace et al., 1999). Finally metric precision is an important way in which spatial memories differ, and can have a large influence on the types of inferences made from memories of environmental spaces.

Third, ability to infer new information from spatial memories is clearly essential to many large-scale cognitive tasks. When one learns spatial layout from direct experience and visual media, one encounters the environment from a sequence of viewpoints as one moves through the environment. A memory of this sequence of viewpoints and movements alone is not sufficient for performing such tasks as estimation of straight-line distances and directions between landmarks that were encountered at different stages of the route. Therefore performance on these tasks requires an inference from the information that was directly perceived.

Encoding, memory, and inference processes are highly related. For example, the ability to infer the straight-line direction or distance between landmarks from a route representation depends on the ability to encode metric information about distances and turns between landmarks and to maintain this information in working memory. Integration over a section of route, in turn, changes the nature of the spatial memory from a sequential route representation to a configural representation. Finally, integration over sections of a route is less effortful when one physically moves in the environment (e.g., Klatzky et al., 1998; Loomis, Klatzky, Golledge, & Philbeck, 1999; Waller, Montello, Richardson, & Hegarty, 2002) so that encoding of self-motion on the basis of body-based senses may facilitate the process of inferring spatial configuration from the sequence of viewpoints encountered as one moves through an environment.

1.3. The present study

1.3.1. Sources of variance in large-scale learning

In the present study, we measured people's ability to learn the layout of three different novel environments, one from direct experience walking in the environment, another from watching a video of a route through the environment, and a third by interacting with a desktop virtual environment. After each of these learning experiences, we administered three different measures of learning: estimation of the direction between landmarks in the environment, estimation of the straight-line distance to non-visible landmarks, and map sketching. Compared to other common measures of spatial learning (e.g., scene recognition, landmark sequencing, and route retracing), these tasks cannot be performed on the basis of a route representation alone, and depend more on metric and configural knowledge of the environment (see Kitchen, 1996; Montello et al., 2004; Newcombe, 1985).

We can isolate the different sources of variance identified in our information-processing analysis by comparing patterns of performance across different learning experiences and different outcome measures. Learning spatial layout from direct experience, video, and desktop virtual environments (VEs) is similar in the sense that in all cases learning depends on perception of a sequence of viewpoints as one moves through the environment, memorization of that sequence of viewpoints in memory and updating of one's position in the environment on the basis of visual cues including self-to-object relations and optical flow. In all of these learning experiences, the ability to compute straight-line distances and directions relies on the ability to make inferences from the memorized sequence of viewpoints encountered on the route.

The different learning experiences also vary in important ways. First, although all include visual information, the field of view is greater in real-world navigation than in a video or virtual environment, so that optic flow information is less available in the visual media. Second, when learning (and being tested) in a real environment, as opposed to a video or desktop virtual environment, updating of body position in the environment can be sensed by vestibular input and kinesthesis in addition to vision. Finally, motion is more active and self-directed in a real environment (and to some extent in a desktop virtual environment), whereas in watching a video, movement through the environment is more passive. When motion is self-directed, efference copy of the muscle commands is an additional source of information for updating of self-motion. Because of these differences, learning from different media makes greater or lesser demands on different basic cognitive processes, such as ability to update one's position in an environment from visual information alone. On the basis of our task analysis, therefore, we hypothesize that ability to learn from real world experience and from visual media will be partially dissociated.

Because a correlational study of this type requires that all participants perform all of the same measures, it was necessary to have participants learn three different environments from the three different learning experiences. These environments necessarily had different layouts. To minimize confounds due to the shape of the environments we used three environments that were made up of mostly straight segments and right-angled turns. The environments learned from direct experience and from a video were of similar complexity. The virtual environment was simpler, because previous research (Richardson, Montello, & Hegarty, 1999) and our pilot studies indicated that learning a complex layout from a desktop virtual environment was more difficult and might lead to a floor effect unless the environment was simplified.

Performing different outcome measures based on an internal representation of an environment can also be a source of variance in measures of large-scale learning (see Fig. 2). Whereas all outcome measures depend on how well the learner has memorized the sequence of views that he or she encountered while learning the environment (so that more accuracy and metric precision in that memory will lead to better performance), outcome measures depend on different "readout processes" and can require individuals to make inferences that were not made at the time of learning (Kitchen, 1996; Montello et al., 2004; Newcombe, 1985). For example, ability to point accurately to non-visible locations or to estimate straight-line distances to these locations requires ability to infer a spatial configuration from information that is encoded as one traverses an environment, as well as metric knowledge of the segments and turns along the route. Map sketching may be less dependent on the internal integration of information in memory, in that a relatively accurate map can be sketched from an internal route representation, as long as metric distances and directions are represented, such that the integration over segments and turns of the route occurs when the route representation is externalized in the sketching process. Furthermore, it is possible to draw an accurate map from a verbally encoded memory, as long as suitably precise metric information is included in the verbal memory. Finally ability to represent one's current heading in a memorized environment during testing is necessary for pointing performance, but not for straight-line distance estimates or map drawing.

Because of the cognitive demands made by different outcome measures (Kitchen, 1996; Montello et al., 2004; Newcombe, 1985), we also consider the possibility that different outcome measures may not reflect the same spatial abilities. If all three measures reflect the same underlying ability to acquire or infer configurations from spatial information learned over time, they should load on the same factor. In contrast, if estimating distances, estimating directions, and map drawing rely on abilities associated with different task demands, they should define different factors.

1.3.2. Relations between large- and small-scale abilities

The information processing analysis in Fig. 2 suggests that large-scale spatial learning depends on processes of encoding spatial information from visual input, and maintaining and mentally transforming spatial representations, which may be shared with small-scale spatial tasks (e.g., mental rotation and finding hidden figures), while also depending on processes such as spatial updating that are not shared with small scale tasks. To test this information-processing model, we examined the relation between large-scale learning, and measures of basic capacities and processes in visual–spatial information processing, including psychometric tests of spatial ability and a test of spatial working memory. Although they come from different research traditions, tests of spatial ability are highly related to tests of visual–spatial working memory (Miyake, Rettinger, Friedman, Shah, & Hegarty, 2001; Shah & Miyake, 1996). In the current study, we are concerned with the variance that they share, i.e., ability to encode, maintain and mentally transform visual spatial information. In addition to psychometric tests and working memory, we include measures of perspective taking ability to examine Allen's alternative model in which perspective taking mediates the relation between small-scale spatial ability and ability to acquire configural knowledge of an environment.

We also included a measure of self-reported sense-of-direction. Such measures are highly correlated with tasks that involve reorienting oneself in the environment (Bryant, 1982, 1991; Kozlowski & Bryant, 1977; Sholl, 1988) and measures of spatial knowledge acquired from direct experience (Hegarty, Richardson, Montello, Lovelace, & Subbiah, 2002; Lorenz & Neisser, 1986; Montello & Pick, 1993). In contrast, they are typically less highly correlated with measures of spatial knowledge acquired from maps and very weakly correlated with paper-and-pencil measures of spatial ability (Bryant, 1982; Hegarty et al., 2002; Sholl, 1988). Self-report sense-of-direction is therefore largely independent of small-scale spatial ability and appears to reflect an ability to update one's location in space as a result of self-motion.

Finally, we measured verbal ability, verbal working memory and abstract reasoning ability. To the extent that environmental information is coded verbally rather than spatially in memory, verbal abilities, including verbal working memory (cf. Shah and Miyake, 1996) should be predictive of performance of large-scale spatial tasks. To the extent that environmental learning reflects general intelligence rather than spatial ability, abstract reasoning ability should be predictive of large-scale spatial tasks.

We hypothesize that all forms of spatial layout learning will be predicted by visual encoding abilities, spatial working memory, and spatial transformation abilities as measured by small-scale spatial ability measures. We expect that learning from direct experience will also be predicted by measures of spatial updating and encoding based on body-based senses. Because spatial updating involves relatively little effort when body-based senses and optic flow information are available (Klatzky et al., 1998; Loomis et al., 1999; Waller et al., 2002) learning from direct experience should be less dependent on individual differences in visual encoding, working memory and effortful spatial transformation abilities than learning from visual media. Therefore we predict that small-scale spatial abilities will be less predictive of performance when learning is based on direct experience. Finally, although the focus of this study is on individual differences in spatial cognition, we also examine sex differences specifically, to inform theoretical debate on their nature and causes.

2. Method

2.1. Participants

Two hundred and eighty-six participants took part in the study. They were recruited by announcements in a local newspaper and on flyers that were posted around the UCSB campus. They were paid $40 for their participation, which took approximately 3.5 h, spread over two sessions. Fifty-eight participants had missing data on more than one of the variables of interest in the study (most of these failed to return for the second data-collection session) and 7 reported that they were already very familiar with one of the environments studied. The data from these participants were not included in the analyses.

Of the 221 participants included in the final sample, 135 were female, 83 were male, and information about the sex of 3 participants was not recorded. Their mean age was 22.0 years (SD = 7.1) and they had been on the UCSB campus 4.1 (SD = 4.7) quarters on average. Although the ages of participants ranged from 17 to 59 years, most were students at area colleges and 80.6% were between 17 and 22. They rated their prior familiarity with the environment used as a measure of real world learning (Ellison Hall) as 1.9 on average (SD = 1.1, range = 1–5) and with the environment used in our measure of learning from a video (the Santa Barbara courthouse) as 1.6 (SD = 1.1, range 1–5), on a scale of 1 (“very unfamiliar”) to 7 (“very familiar”).

2.2. Tasks and measures

2.2.1. Demographics questionnaire

Participants completed a demographics questionnaire, which asked them to indicate their age, sex, SAT Verbal and Mathematics scores, and the number of quarters they had been at UCSB. The questionnaire also asked them to estimate the number of times they had visited the Santa Barbara courthouse (depicted in the video described below) and how often they had visited a number of different locations in Ellison Hall on the UCSB campus. Finally they were asked to indicate their overall familiarity with Ellison Hall and the Santa Barbara courthouse.

2.2.2. Measures of visual–spatial abilities

Four measures of visual–spatial abilities were administered, as follows.

1. The *Group Embedded Figures Test* (Oltman, Raskin, & Witkin, 1971) measures ability to encode a spatial pattern and recognize it in a complex figure. In this test, participants are given a sheet showing several simple 2-D geometric figures. On each trial they are shown a complex 2-D figure, and their task is to locate the simple figure within the complex figure and to trace it in pencil. There are three sections of the test: an initial practice section with 7 items, lasting 2 min, and two sections with 9 items each, for which participants are allotted 5 min apiece. The internal reliability of the test is .82.
2. The *Vandenberg Mental Rotations Test* (Vandenberg & Kuse, 1978) is a measure of spatial visualization ability. In this test, participants view a three-dimensional target figure and four test figures. Their task is to determine which of the test figures are rotations of the target figure as quickly and accuracy as possible. It consists of two sections of 10 items, for which participants are allotted 3 min each. Its internal reliability is .88.
3. The *Arrow Span Test*, adapted from Shah and Miyake (1996), measures ability to maintain spatial information in working memory. This test was presented on a PC-compatible 486 computer running MEL software (Schneider, 1990). On each trial, a set of arrows was presented on the computer screen, one at a time, in one of eight possible directions (upright and increments of 45° from upright). The sequence of directions was randomized for each trial set and each participant. After all the arrows of a set were presented, the participant’s task was to indicate the directions of the arrows in their order of presentation. Participants indicated their answers by typing the keys on a standard numeric keypad. The numerals on the keys were covered with arrows indicating directions; ‘7’ was 45° to the left from upright, ‘8’ was upright, ‘9’ was 45° to the right from upright, and so on (‘5,’ in the middle, was not

used). The test consisted of 15 sets of arrows, three at each level ranging from 2 to 6 arrows. Participants completed all of the items on the test. The dependent measure was the number of arrows identified in the correct order across the 15 trials (maximum possible score = 60). The internal reliability of the this measure is .78 (Shah and Miyake, 1996).

4. A test of *perspective-taking ability*, based on the stimulus materials of Huttenlocher and Presson (1973, 1979), measures ability to encode, maintain and transform spatial representations at the vista scale of space. Four objects (a lamp, a plant, a trash can, and a box) were placed at the center of each wall in an 8-m square room. Participants stood in front of one object looking into the center of the room and were instructed to learn the locations of the 4 objects. They were allowed as much time as they needed, which was usually less than 3 min. After viewing the objects, participants sat at an IBM-compatible 486 computer, out of view of the objects. On each trial they were asked to imagine standing in front of one of the objects facing the middle of the room, and to point to another object as quickly and accurately as possible. They pointed by pressing arrow keys (for forward, right, and left) on the computer keyboard. There were 12 trials, three for each object, which were self paced and presented in a random order. We computed two measures of performance, the number of items answered correctly and the average reaction time. Measures of internal reliability (Chronbach's alpha) were .74 for the accuracy measure and .61 for the reaction time measure.

2.2.3. Measure of self-reported sense of direction

Participants were administered the Santa Barbara Sense-of-Direction Scale, which consists of 15 Likert-type items adapted from previous self-report scales of environmental spatial abilities (Hegarty et al., 2002). Each item was a self-referential statement about some aspect of environmental spatial cognition; participants responded by circling a number from 1 ("strongly agree") to 7 ("strongly disagree"). The items are phrased such that approximately half are stated positively, and half negatively. An example of a positively stated item is "I am very good at judging distances"; an example of a negatively stated item is "I very easily get lost in a new city." In scoring, positively stated items were reversed so that a higher score indicates a better sense-of-direction. Sums of the 15 items were used for the analyses. The internal reliability of the scale is .88.

2.2.4. Measures of verbal and reasoning abilities

We included two measures of verbal intelligence (a vocabulary test and a test of verbal working memory).

1. *Extended Range Vocabulary Test–V3* (Ekstrom, French, Harman, & Derman, 1976) is a test of verbal ability. On each trial, participants are given a target word, and their task is to select from five other words the one that is closest in meaning to the target word. One section of the test was administered, with 24 items, for which participants were allotted 6 min. Reliability estimates for this test range from .76 to .89 for different samples (Ekstrom et al, 1976).
2. *Reading Span Test* (Daneman & Carpenter, 1980) measures verbal working memory. In this task, participants read a set of unrelated sentences (each printed on an index card) aloud one at a time and recall the final word of each sentence at the end of the set. Participants are presented sentences in sets of increasing size, starting from two sentences per set, finishing at five sentences per set. There are 20 sentence sets in this test, five sets for each set size. A participant is scored as passing a level (set size) if he or she gives correct answers to at least four of the five items at that level and is given half credit for producing only three correct answers at that level. The final score is the highest level at which the participant passed the test or received half credit. For example, if a participant answered all the items correctly for a set size of two, four of the items correctly for a set size of three, and three of the items correctly for a set size of four, he or she would be assigned a reading span of 3.5. The test is ended when a participant gives correct answers to no more than two sets at a given level. Its internal reliability is .85 (Shah & Miyake, 1996).
3. *Abstract Reasoning Test, Form S* (Bennett, Seashore, & Wesman, 1972) is a measure of non-verbal reasoning, which is highly related to general intelligence. Each has four figures that form a series. Participants select from another set of five figures the one which would be the next item in the series. Participants are allowed 25 min to complete 50 items. Reliability estimates for this test range from .85 to .95 for different samples (Bennett et al, 1972).

2.2.5. Spatial learning from navigation directly in a real environment

Participants learned the layout of a route through two floors of a building on the UCSB campus (Ellison Hall), along with the locations of 8 landmarks on that

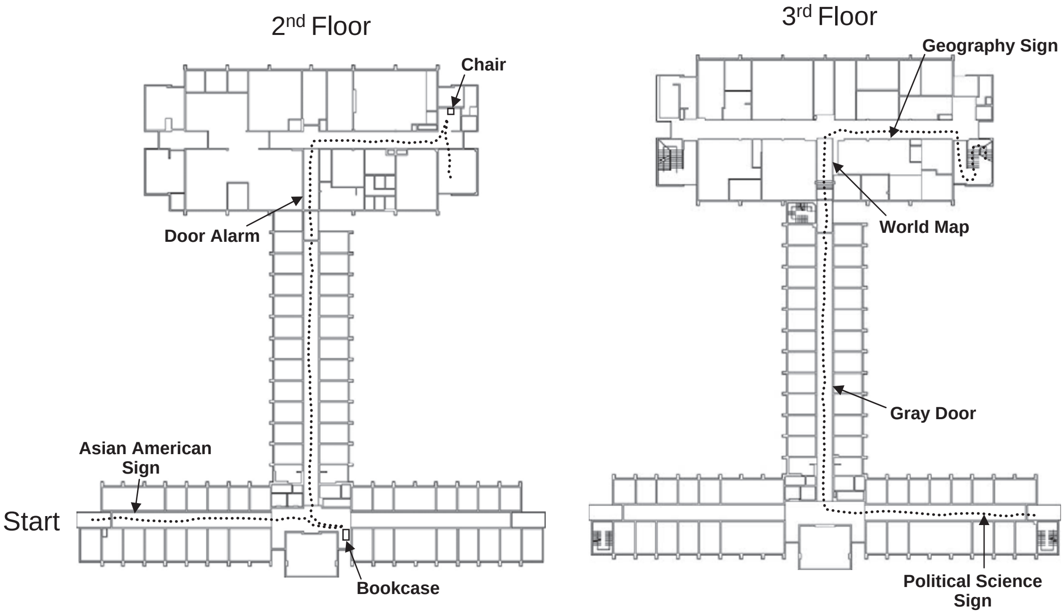

Fig. 3. Map of the environment learned from direct navigation, showing the locations of the landmarks.

route. Shown in Fig. 3, the route covered approximately 300 ft and took about 6 min to traverse. The experimenter first led each participant through the building, stopping at each landmark to point it out and name it. After traversing the route once, the participant was taken outside the building and back to the start of the route for testing. The experimenter again led the participant along the route and at each of the 8 landmarks, instructed him or her to make straight-line distance and direction judgments to two other landmarks that were not visible from that location, for a total of 16 distance and 16 direction estimates. Distance estimates were made directly in feet. Direction estimates were made by means of a circular pointing dial (see Montello, Richardson, Hegarty & Provenza, 1999), described above. In making distance and direction judgments between landmarks on different floors of the building, participants were instructed to ignore the vertical dimension. After completing all estimates on the route, participants drew a sketch map of the route including the locations of the 8 landmarks.

2.2.6. Spatial learning from navigation in a desktop virtual environment (VE)

Participants learned the layout of a single-floor route depicted in a desktop VE, along with the locations of 4 landmarks on that route. The VE was constructed using the Duke Nukem 3D game engine, a desktop virtual-reality type adventure game created by 3D Realms Entertainment. The environment was a long hallway with six segments, connected by 90° turns (see Fig. 4). Participants traversed the route using the forward, right, and left arrow keys on the

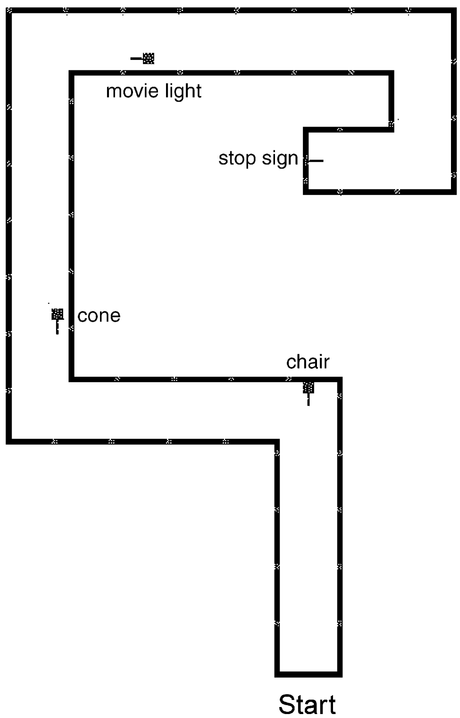

Fig. 4. Map of the desktop virtual environment, showing the locations of the landmarks.

keyboard. They first spent time in a practice environment to become familiar with the interface. In this practice environment, participants moved along a square-shaped route and were informed that one complete circuit was 100 ft (to give them a sense of scale for the environmental learning task). They then traversed the test route twice, in response to verbal instructions by the experimenter, who pointed out the 4 landmarks in the environment as they were encountered. Each traversal of the route took approximately 2 min (the maximum walking speed in the VE was 4.8 km/h). On the third traversal, participants were stopped at each landmark and instructed to make distance and direction judgments to two other landmarks, for a total of 8 distance and 8 direction estimates. Estimates were collected as for the real-environmental task, and participants also drew a sketch map of the virtual route with its 4 landmarks.

2.2.7. Spatial learning from viewing a videotaped environment

Participants learned the layout and locations of 7 landmarks in a local building (the Santa Barbara Courthouse) by viewing a videotape. The videotape lasted 5 1/2 min and was made by walking a route through two floors of the building using a handheld Hi8 camera. Each landmark was indicated by stopping, pointing the camera at the landmark, and naming it aloud. The videotaped route is shown in Fig. 5. Participants viewed the video twice on a monitor mounted in the corner of the room, and faced the video during both the learning and testing phases of this task. After the second viewing, the video was replayed for testing purposes. This time, the video was paused at each landmark and participants made direction and distance judgments to two of the other landmarks (a total of 14 distance and 14 direction judgments). Estimates were collected as for the real- and VE tasks, and participants also drew a sketch map of the video route with its 7 landmarks. For each direction judgment, participants were given a circle with a straight line pointing downwards from the center of the circle. They were instructed to point the line towards themselves and to draw an arrow on the circle indicating the direction to the landmark in question.

2.2.8. Scoring of the environmental learning measures

Three dependent measures were scored for each of the environmental-learning tasks: distance estimates, direction estimates, and sketch maps. The score for the distance estimates was the correlation of a participant's estimates across trials with the correct distances for these trials. A participant who accurately estimates distances proportionately across trials would therefore receive a high correlation. We advocate this measure of relative distance accuracy over a measure of absolute distance accuracy because individuals have varying conceptions of standard distances, such as a foot (Mon-

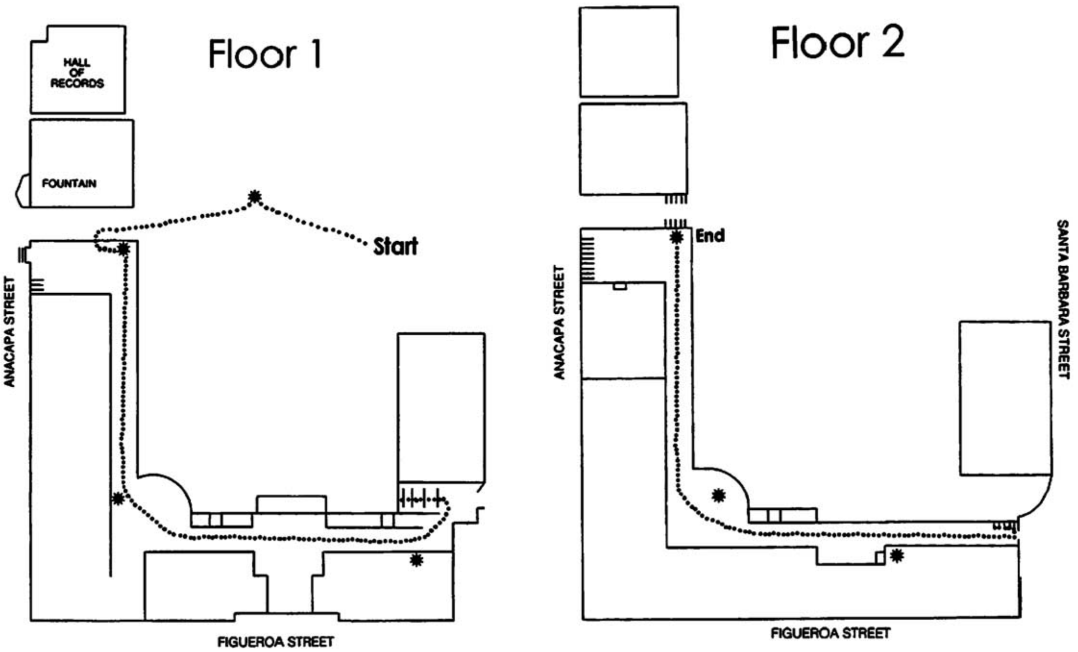

Fig. 5. Map of the environment learned from a videotape, showing the locations of the landmarks.

tello, 1991).[3] Correlations of estimated with correct distances were transformed to reduce deviations from normality using Fisher's r-to-z transform.

Direction estimates were scored as the mean absolute error in degrees (absolute deviation of pointing direction from the correct direction) across trials. Separation of absolute error into constant and variable errors has been advocated for experimental studies (Montello, Richardson et al., 1999; Schutz & Roy, 1973). However we have argued that absolute error provides the best summary measure of individual differences in direction estimation (Hegarty et al., 2002) because it best predicts whether a person would be closest to a correct answer on a given trial, an aspect of performance that requires both low constant error and low variable error, on average. For example, one would not consider a person to have high spatial ability if he or she had a large bias (constant error) in pointing to targets but little variability, or if his or her estimates always centered on the correct direction but were highly variable. Reliability estimates (coefficient alphas) for the direction estimation measures were .85 for learning from direct experience, .84 for learning from videotape, and .66 for learning from the VE.

Sketch maps were scored by counting the number of qualitative errors made in drawing the maps. Qualitative errors included the number of landmarks omitted (based on the landmarks pointed out to participants during the learning phase), the number of missing or additional segments of the route, and the number of turns that were qualitatively in the wrong direction (i.e., to the left instead of the right). For maps of the Ellison Hall and Courthouse routes, each of which traversed two floors of a building, we counted an additional error if the two floors were not aligned approximately correctly on participants' maps. Although this scoring system did not measure precise metric deviations of the sketch maps from the correct spatial layout, good sketch map performance in this scoring scheme requires sophisticated knowledge of spatial layout, especially to correctly align the two floors in the real and videotaped environments. The errors were summed to produce a measure of sketch-map performance for each participant. Inter-rater reliability, based on two independent codings of 30 sketch maps, was .87.

2.3. Procedure

Participants were tested in two sessions. In the first session, they were tested in groups of up to 12 at a time; in the second session they were tested individually. Tasks were included in one or the other session based on whether they could be administered in groups or required individual administration. The tasks were administered to all participants in the same order, as is customary in individual-differences studies, to minimize any measurement error due to participant-by-order interactions.

The first session took place in a large laboratory room, approximately 8 m square. Participants sat at tables with wooden barriers dividing the tables into four partitions, which prevented them from seeing others' responses. They first completed the demographics questionnaire, followed by the SBSOD. They were then administered the Group Embedded Figures test followed by the Extended Range Vocabulary test, according to the standard instructions for these tests. Then all participants were shown the videotape of the Courthouse and performed the spatial learning tests based on it. Finally, participants were administered the Vandenberg Mental Rotation Test in the group session, followed by the Abstract Reasoning test.

For the second session, participants arrived at the same laboratory alone. They first performed the perspective-taking test, followed by the test of spatial learning from the desktop VE. Then they were led out of the laboratory to the entrance of Ellison Hall and performed the test of spatial learning from the real environment. Next they performed two trials of a blindfold updating task which will not be reported on in this paper, because we concluded in retrospect that two trials were not sufficient to provide an adequate measure of this ability. Finally, participants were brought back to the laboratory and administered the Arrow Span test, followed by the Reading Span test.

[3] With the correlation metric being used here, what is computed is a single correlation across several trials (path segments). Measures of internal consistency are problematic in that the "trials" in this case are *necessarily* dependent on each other, so that if the estimate of one segment is large, this affects how the estimate of another segment is treated by the correlation metric. We computed split-half reliabilities for the correlation measures. This is a less than ideal as a way of estimating reliability, however, in that each "half" that makes up the split-half measure has half as many items as the whole scale. The split-half reliabilities of the three distance scales are .43 (direct-experience condition), .44. (video condition), and .16 (VE condition). The fact that the distance measures correlated more strongly with other measures in our data set, such as direction estimates for the same environments (see Table 2), is evidence that they are more reliable measures than suggested by these low reliability coefficients.

3. Results

3.1. Descriptive statistics

Because the multivariate techniques used in this study assume normal distributions and are sensitive to extreme outliers, we screened our data as follows. For each variable, any observations with values that exceeded 3 standard deviations from the mean were set to be equal to 3 standard deviations from the mean (winsorized). This conservative procedure allowed us to retain extreme observations, important in an individual differences study, while minimizing the effects of these observations. This procedure affected less than .007 of the observations across the 18 variables of interest in the study.

Descriptive statistics for the measures are included in Table 1. Note that the standard deviations of the learning measures did not differ appreciably for learning from different experiences, indicating that there were substantial individual differences in all measures of environmental learning. Several of the measures (e.g., direction estimation and map sketching) are based on errors; therefore higher scores on these items indicate less ability. For ease of interpretation, these measures were transformed (by changing the sign of the variables) so that higher scores indicate more ability, and the transformed scores were used for the correlational analyses, factor analysis and structural equation models. About half of the subjects (48.9%) made no errors on the perspective-taking test (number of errors ranged from 0 to 10), so that the distribution of observations on this measure departed considerably from normality. A new dichotomous variable was computed and used in all further analyses, such that participants were assigned a value of 1 if they made no perspective-taking errors, and a value of 0 if they made at least one error (for the resulting variable Mean = .49, SD = .50, Skewness = .05, Kurtosis = −2.02). The measure of reaction time on the perspective-taking test was subjected to a log transformation, as is customary to reduce the influence of very long times (for the transformed variable, Mean = 1.65, SD = .40, Skewness = .56, Kurtosis = 0.00).

Table 1
Descriptive statistics for variables in the study

Variable	Range	Mean	SD	Skewness	Kurtosis
Visual–spatial abilities					
Embedded Figures Test	0–18	12.4	4.7	−0.9	0.0
Vandenberg MRT	5–39	19.7	6.9	0.5	0.1
Arrow Span Test	29–60	48.6	6.6	−0.7	0.1
Perspective taking					
Accuracy	2–12	10.8	1.8	−2.2	6.0
Reaction time (s)	2.3–16.5	5.7	2.6	1.7	3.5
Self-reported sense of direction					
SBSOD	1.6–6.0	3.6	1.0	0.4	−0.4
Verbal/general abilities					
Vocabulary	2.4–23.0	9.8	4.7	0.3	−0.4
Reading Span Test	1.0–4.5	2.4	0.7	0.6	0.7
Abstract Reasoning Test	23 – 50	40.9	5.9	−0.8	0.3
Learning from direct experience:					
Direction estimation	10–101	38.7	20.8	0.6	−0.4
Distance estimation[a]	−.49–.97	0.38	0.31	−0.2	0.7
Map sketching	0–7	1.5	2.0	1.3	0.8
Learning from video					
Direction estimation	9 – 137	56.3	26.9	0.2	−0.5
Distance estimation[a]	−.57–.87	0.33	0.28	−0.3	−0.6
Map sketching	0 – 7	1.6	1.9	1.0	0.3
Learning from VE					
Direction estimation	5–110	45.5	21.6	0.5	0.1
Distance estimation[a]	−.45–.91	0.32	0.30	−0.2	−0.7
Map sketching	0–4	0.9	1.2	1.2	0.6

[a] Measures of distance estimation are correlations of a participant's estimates across trials with the correct distances for those trials. Correlations were subjected to an r-to-z transform for all subsequent analyses.

3.2. Correlations

Correlations between the measures are given in Table 2. All but one correlation are positive, indicating that, in general, the ability measures are positively related. Consistent with previous studies (Allen et al., 1996; Pearson and Ialongo, 1986), correlations of paper-and-pencil spatial abilities (the Embedded Figures Test and the Vandenberg Mental Rotation Test) with measures of learning from direct experience are relatively low (ranging from .11 to .28); correlations of self-report sense-of-direction with these measures are higher (ranging from .34 to .45) (cf. Bryant, 1982, 1991; Kozlowski and Bryant, 1977; Montello and Pick, 1993). In contrast, measures of paper-and-pencil spatial ability are correlated more highly with learning from the video (range = .16 to .46) and the VE (range = .08 to .45). The self-report measure is correlated less with learning from the video (range = .23 to .34) or the VE (range = .20 to .24) than with learning from direct experience (range = .36 to .46).

3.3. Statistical procedure for the multivariate analyses

The goals of the multivariate analyses were to examine whether performance of different environmental tasks reflect a single underlying ability or a dispa-

Table 2
Univariate correlations between variables in the study

	1	2	3	4a	4b	5	6	7	8	9	10	11	12	13	14	15	16
1. Embedded figure																	
2. Mental rotation	.31																
3 Arrow Span	.36	.31															
4a. Perspective accuracy	.20	.20	.13														
4b. Perspective RT	.34	.24	.30	.24													
5. Sense-of-direction	.15	.09	.23	.12	.07												
6. Vocabulary	.28	.05	.06	.21	.13	.07											
7. Reading span	.22	−.01	.10	.18	.23	.06	.39										
8. Abstract reasoning	.54	.23	.35	.22	.48	.21	.35	.35									
9. Direct experience distance	.27	.14	.32	.18	.13	.37	.12	.17	.27								
10. Direct experience direction	.28	.21	.39	.18	.24	.46	.15	.17	.34	.67							
11. Direct experience map	.19	.11	.36	.18	.22	.36	.11	.14	.31	.52	.67						
12. Video distance	.29	.16	.14	.19	.20	.26	.21	.14	.39	.24	.28	.28					
13. Video direction	.46	.33	.25	.25	.29	.34	.22	.18	.51	.31	.43	.35	.50				
14. Video map	.30	.22	.26	.13	.28	.23	.23	.15	.42	.30	.38	.40	.30	.49			
15. VE distance	.19	.08	.14	.09	.10	.10	.07	.20	.19	.20	.27	.15	.20	.20	.12		
16. VE direction	.45	.40	.34	.22	.27	.24	.12	.11	.50	.29	.41	.25	.27	.53	.25	.37	
17. VE map	.42	.25	.39	.14	.28	.22	.11	.14	.43	.25	.41	.33	.26	.46	.36	.23	.48

Note. All variables have been transformed so that higher scores indicate more ability.

rate set of abilities, and to examine the extent to which processing of spatial information at different scales of space reflects the same or different underlying abilities. To achieve these goals, we first conducted an exploratory analysis of the environmental spatial tasks, and then used structural equation modeling to examine the extent to which latent variables (or factors) underlying the predictor variables accounted for the variance in latent variables underlying the environmental spatial measures.

3.3.1. Exploratory factor analyses of large-scale spatial abilities

Our data included nine measures of environmental learning, that is, three measures of environmental learning (direction estimation, distance estimation, map sketching) for each of three different learning media (direct experience, video, VE). These were subjected to an exploratory factor analysis, using the Maximum Likelihood extraction method with direct oblimin rotation. Although our task analysis predicted that learning from direct experience and from visual media would be partially dissociated, an exploratory factor analysis was used because we also allowed for the possibility that the nine measures of environmental learning might reduce to factors based on how spatial knowledge is measured (i.e., by distance estimation, direction estimation and map sketching), rather than how it is learned; they might also be based on the complexity of the environments (e.g., the need to integrate over two floors of a building in the real and video environments but not in the VE).

Two factors had eigenvalues greater than 1.0. The scree plot also suggested a two-factor solution. The rotated pattern matrix is shown in Table 3. Loadings higher than .4 are used to interpret the factors. Factor 1 can be interpreted as measuring ability to learn spatial layout from visual media. Measures of learning from both a video and a virtual environment loaded on this factor. The only measure of learning from visual media that did not have a loading of greater than .4 was the measure of distance estimation after learning a virtual environment. In contrast none of the measures of learning from direct experience had loadings of greater than .4 on this factor. Factor 2 can be interpreted as a measure of learning from direct experience. All of the measures of learning from direct experience loaded

Table 3
Rotated factor matrix for the factor analysis of environmental learning measures

Measure	Factor 1	Factor 2
Learning from direct experience		
Direction estimation	.03	.91
Distance estimation	.03	.73
Map sketching	.04	.70
Learning from video		
Direction estimation	.96	.14
Distance estimation	.58	.05
Map sketching	.48	.13
Learning from VE		
Direction estimation	.57	.07
Distance estimation	.19	.16
Map sketching	.49	.13

highly on this factor, whereas the measures of learning from visual media had loadings of less than .3 on this factor. The correlation between the two factors was .61.

In summary, as predicted by our task analysis the nine variables could be reduced to two factors indicating ability to learn spatial layout from visual media and from direct experience, respectively. In contrast, there was no evidence that direction estimation, distance estimation and map drawing defined different factors, or that the complexity of the environment learned in different media affected the factor structure.

3.3.2. Relating environmental spatial abilities to small-scale spatial and other abilities

We now turn to the question of the relation of environmental learning to small-scale spatial abilities and other predictors. To answer this question, we tested structural-equation models that included latent variables measuring spatial ability in small-scale space, verbal ability, and self-reported sense-of-direction.[4]

The structural equation modeling was carried out using the AMOS program (Arbuckle, 1999), which uses maximum-likelihood estimation to derive the specified parameters based on the covariance matrix. We use several indices to evaluate the fit of the models, as recommended by Hu and Bentler (1998). The most common fit index is the χ^2 statistic, with a significant χ^2 indicating a poor fit to the data. However, the χ^2 statistic is correlated with sample size and is consequently significant with large samples even when differences between the model and data are small (Kline, 1998). For this reason, many researchers have advocated the χ^2/df statistic, with a value less than 2.0 indicating a good fit. Another index of fit is the Standardized Root Mean Square Error of Approximation (RMSEA); a value of .08 or below indicates a fair fit, and a value no higher than .05 indicates a good fit (Hu and Bentler, 1998). The Comparative Fit Index (CFI) measures the extent to which the examined model fits better than a baseline model, with a CFI of at least .9 indicating a fair fit and a value of at least .95 indicating a good fit to the data. Another important principle underlying the structural equation models is that of parsimony. Although models of arbitrary complexity can be evaluated using structural equation models, the reliability of a model is inversely proportional to its complexity, or the number of estimated parameters in the model.[5] Therefore in all of the analyses, we sought to find the most parsimonious model that accounted for the data. The RMSEA index also takes the parsimony of a model into account, (i.e., it penalizes complex models).

The full structural-equation model relating the predictor variables to the measures of environmental learning is shown in Fig. 6. The fit parameters for this model, shown in Table 4, indicate that it is a fair to good fit to the data. However, the path coefficients from verbal ability to the two measures of environmental learning were not significant. Path coefficients can be interpreted as standardized regression weights indicating the degree of relation between the predictor and predicted variables after controlling for the effects of the other variables. The non-significant path coefficients indicate that verbal ability is not a significant predictor of environmental-learning ability after controlling for the effects of small-scale spatial ability and sense-of-direction. Inclusion of this construct also reduces the parsimony of the model (number of estimated parameters = 38). We therefore tested a reduced model, shown in Fig. 7, which omits the verbal factor. All of the path coefficients in this model were statistically significant and did not vary significantly from the corresponding path coefficients in the full model. Fit indices for the reduced model, shown in Table 4, indicate that it also fits the data reasonably well and is more parsimonious (number of estimated parameters = 30); therefore this model is preferred.

Although verbal ability does not make an independent contribution to explaining large-scale spatial learning, it is possible that general intelligence does, and it is important to test whether the significant paths from spatial ability to the measures of large-scale learning reflect variance specific to spatial as opposed to general intelligence. We tested another model in which we added abstract reasoning ability as a measured variable to the model depicted in Fig. 7. Fit parameters for this model indicated that it is a good

[4] As recommended by Kline (1998), a latent variable measuring Self-Report Sense-of-Direction was formed by fixing the error variance of the SBSOD using the formula

$$X = (1 - r_{xx})(\mathrm{Sx}^2)$$

where r_{xx} is the reliability of the SBSOD scale (.88) and Sx^2 is the variance in this measure (1.0). This method was used because there was only one manifest variable measuring sense-of-direction. The alternative method of using individual items from the SBSOD in the structural equation models was not possible because it would lead to a much more complex model, which could not be evaluated reliably with the sample size in this study.

[5] Although there are no absolute standards in the structural equation modeling literature regarding sample size, a sample size of 10–20 participants per parameter is desirable, and a sample size of less than 5 subjects per parameter produces an unreliable model (Kline, 1998).

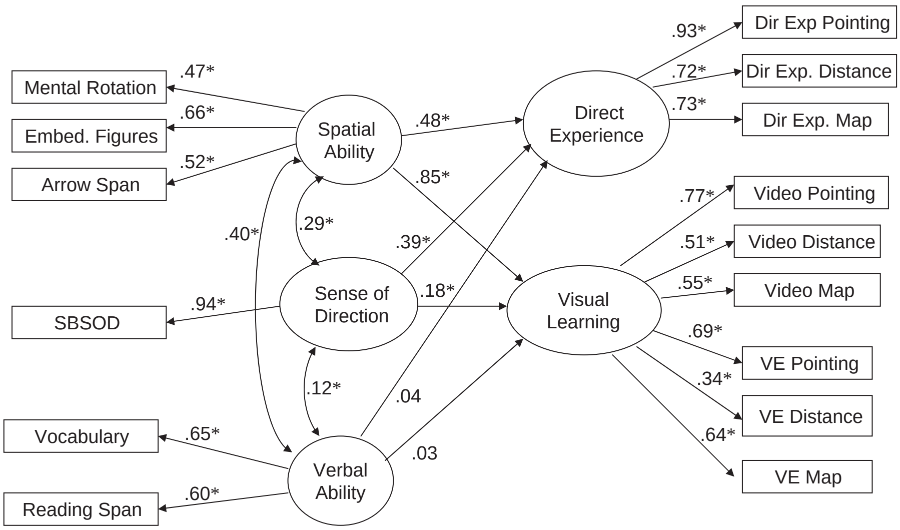

Fig. 6. Results of the "full" structural-equation model relating the environmental learning factors to the predictor factors. Asterisks indicate significant path coefficients.

fit to the data (see Table 4). However coefficients of the paths from abstract reasoning to learning from direct experience (.16) and learning from visual media (.03) were not statistically significant. Furthermore, addition of abstract reasoning to the model did not change the significance of any of the other paths; it increased the path from spatial ability to direct experience learning from .50 to .61, and did not change any of the other path parameters by more than .04. Since abstract reasoning does not add to the prediction of spatial layout learning and this model is less parsimonious (number of estimated parameters = 37), the reduced model in Fig. 7 is preferred.[6]

Table 4
Fit indices for structural-equation models of abilities and performance

Model	χ^2 (*df*)	χ^2/df	RMSEA	CFI
Full model (Fig. 6)	142.48 (82)	1.74	.06	.94
Reduced model without verbal ability (Fig. 7)	117.64 (61)	1.93	.07	.94
Reduced model with abstract reasoning added as a predictor	111.86 (68)	1.65	.05	.96
Model with path coefficients from spatial ability to environmental learning abilities constrained to be equal	133.66 (62)	2.15	.07	.92
Model with path coefficients from SOD to environmental learning abilities constrained to be equal	122.88 (62)	1.66	.07	.93
Reduced model with sex added as a predictor	139.80 (70)	1.99	.06	.92

The reduced structural-equation model in Fig. 7 indicates that small-scale spatial ability is a predictor of both environmental spatial factors. It is a strong predictor of learning from visual media (path coefficient = .87 with a 95% confidence interval of .74 to .99)[7] and a moderate predictor of learning from direct experience (path coefficient = 50 with a 95% confidence interval of .36 to .63). These data support the Partial Dissociation model depicted in Fig. 1. They are inconsistent with the Unitary model in that the 95% confidence intervals do not include 1. They are inconsistent with the Total Dissociation model in that the 95% confidence intervals do not include 0. They also indicate that, as predicted, small-scale spatial ability is more predictive of learning from visual media than from direct experience in an environment. As a more formal test of this prediction, we used a procedure in which the model shown in Fig. 7 was compared to a nested restricted model in which these path coefficients are constrained to be equal. The two models can be com-

[6] As another method of controlling for the effects of general ability, we performed a factor analysis of all of the variables in this study. We extracted a general ability factor, on which abstract reasoning, verbal ability and verbal working memory but none of the other variables had loadings greater than .4. All other variables were regressed on this factor, and the residuals were used to repeat the analyses reported in Figs. 7 and 8. None of the results of the analyses (e.g., significance of the paths) changed appreciably from those reported.

[7] To calculate the confidence intervals, the standard error of the path coefficients were estimated using the bootstrapping method (Klein, 1998).

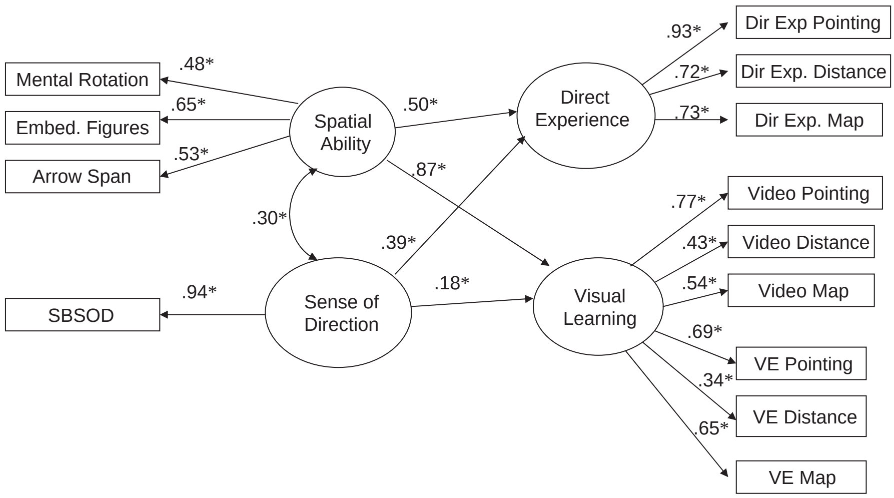

Fig. 7. Results of the reduced structural-equation model relating the environmental learning factors to the predictor variables. Paths in the full model with non-significant path coefficients are omitted from this model. All path coefficients are significant.

pared by subtracting the χ^2 for the unconstrained model from the χ^2 for the constrained model (degrees of freedom for this comparison are calculated with an analogous subtraction). The restricted model (constraining the paths from spatial ability to the two environmental learning factors to be equal) provided a significantly poorer fit to the data than the unrestricted model shown in Fig. 7 (see the fit parameters for these models in Table 4, χ^2 [1, N=221]=16.02, p<.001). This result indicates that as we hypothesized, small-scale spatial ability is a significantly stronger predictor of learning from visual information alone than of learning from direct experience.

The significant path coefficients from self-report sense-of-direction to the latent variables measuring learning from direct experience (path coefficient=.39 with a 95% confidence interval of .24 to .54) and learning from visual media (path coefficient=.18 with a 95% confidence interval of .01 to .34) indicate that sense-of-direction makes independent contributions to predicting environmental learning after controlling for the effects of small-scale spatial ability. To examine whether there is a significant difference between the strengths of these paths, we compared a restricted model in which these paths were constrained to be equal to the full model. The restricted model (see Table 4) provided a significantly poorer fit to the data than the unrestricted model (χ^2 (1, N=221)=5.25, p<.05). Thus, we can conclude that self-report sense-of-direction is a significantly stronger predictor of learning from direct experience than of learning from visual information alone.

Is there any evidence for the Mediation model suggested by Allen et al. (1996), in which perspective taking mediates the relationship between small-scale ability and measures of large-scale learning? To test this model, we performed analyses in which we assumed both a direct path from the small-scale spatial abilities factor to the environmental learning factors, and an indirect path assuming that perspective taking mediates the relation between small-scale and large-scale spatial ability. These analyses are summarized in Fig. 8. The spatial-abilities and environmental learning factors in these analyses are de-

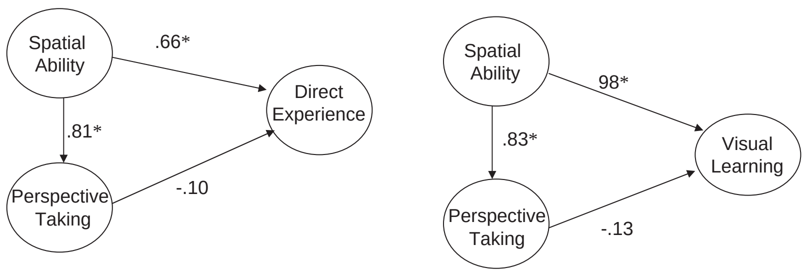

Fig. 8. Results of the analyses testing the hypothesis that perspective taking mediates the relationship between small-scale spatial ability and learning from direct experience and from visual media. Asterisks indicate significant path coefficients.

fined by the same variables as in previous analyses, but for simplicity of presentation, we have omitted the loadings of the different measures on these factors. Note that path coefficients from spatial ability to the environmental learning factors are higher, because shared variance with other measures (e.g., sense-of-direction) is not controlled in this analysis. The most important result of these analyses is that the direct path between spatial ability and environmental learning is significant in both cases, even when we allow for the possibility of a mediated path through perspective taking. Furthermore, the direct path from perspective taking is not significant and has a negative coefficient in both cases, indicating that if anything, perspective taking acts as a suppressor variable. It is clear therefore that although small-scale spatial ability and perspective taking ability are highly related, there is no evidence for the Mediation model in this study.

3.4. Sex differences

Table 5 presents data on sex differences, including a measure of effect size, Cohen's *d* (Cohen, 1988), which is equal to the difference in means between males and females divided by the pooled standard deviation. Males outperformed females in some but not all of the spatial measures. Consistent with previous research (Linn & Petersen, 1985; Voyer, Voyer, & Bryden, 1995) the largest sex difference in small-scale spatial ability was in mental rotation. There was also a significant sex difference in favor of males on the Arrow Span Test, but not on the Embedded Figures test or the Perspective Taking measures. Males significantly outperformed females in some of the measures of environmental learning from both direct experience and visual media. These differences were most pronounced when environmental learning was assessed with measures of direction estimation. Males also reported a better sense-

Table 5
Sex differences for variables in the study

Variable	Males (N=83)		Females (N=135)		t	*df*	Cohen's *d*
	Mean	SD	Mean	SD			
Visual–spatial abilities							
Embedded Figures Test	13.0	4.9	12.0	4.6	1.43	166	.21
Vandenberg MRT	22.6	7.5	17.8	5.8	5.00***	142	.70
Arrow Span Test	50.2	6.3	47.6	6.7	2.86**	181	.39
Perspective taking							
Accuracy	11.0	1.6	10.7	1.9	1.46	193	.17
Reaction time (s)	5.9	3.0	5.5	0.9	0.90	140	.15
Self-report sense of direction							
SBSOD	3.9	0.9	3.4	1.0	3.33**	197	.50
Verbal/general abilities							
Vocabulary	9.4	5.2	10.1	4.3	1.03	148	.15
Reading Span Test	2.2	0.7	2.5	0.7	2.44*	182	.42
Abstract Reasoning Test	41.2	5.8	40.6	5.9	0.70	175	.09
Learning from direct experience							
Direction estimation	33.4	19.2	42.4	21.1	3.20**	187	.43
Distance estimation	.45	.29	.33	.31	2.60*	171	.39
Map sketching	1.2	1.9	1.7	2.1	1.56	188	.33
Learning from video							
Direction estimation	47.9	24.6	62.0	26.7	3.99***	185	.52
Distance estimation	.36	.29	.31	.27	1.26	155	.18
Map sketching	1.4	1.9	1.8	1.8	1.37	166	.21
Learning from VE							
Direction estimation	37.6	19.8	50.8	21.1	4.64***	181	.61
Distance estimation	.39	.30	.29	.29	2.20*	171	.33
Map sketching	0.7	1.0	1.1	1.3	2.60*	204	.33

Note. Significance tests are based on independent samples t-tests (two-tailed) with equal variance not assumed. Data from 3 participants who did not report their sex are not included in these analyses.
*$p<.05$. **$p<.01$. ***$p<.001$.

of-direction. Finally, there was a significant sex difference in favor of females on the Reading Span test.

To examine whether sex predicts spatial-layout learning independently of small-scale spatial ability and self-reported sense-of-direction, we conducted an analysis in which we included sex (measured variable) as an additional predictor variable to the model presented in Fig. 7. Although the correlations of sex with small-scale spatial ability (.34) and sense-of-direction (.23) were statistically significant, the path coefficients from sex to learning from direct experience (.04) and visual learning (.02) were very small and non-significant, indicating that sex does not make an independent contribution to predicting spatial-layout learning, above and beyond the contributions of small-scale ability and sense-of-direction. Furthermore, adding sex as a variable did not change the significance of any of the paths in the model presented in Fig. 7, and did not change any of the path coefficients in that model by more than .06.

4. Discussion

This study had two major goals. The first was to characterize the sources of individual differences in environmental spatial abilities, examining whether they reflect a single underlying ability or a disparate set of abilities. We found that measures of environmental learning defined separable factors that were characterized by whether the environment was experienced directly, or from visual media (a video and a desktop VE). The second goal was to assess the degree to which large-scale spatial abilities can be accounted for by measures of small-scale spatial ability, spatial updating ability (including self-reported sense-of-direction) and verbal ability. Our results indicate that small-scale spatial ability has considerable shared variance with large-scale spatial learning, although it is a much stronger predictor of learning assessed using visual media than of learning assessed in a real environment. Self-report sense-of-direction independently predicts learning assessed in a real environment and, to a lesser degree, from visual media. In contrast, verbal ability and general ability are not independent predictors of large-scale spatial abilities.

We considered four different models of the relation between small-scale spatial abilities and learning of spatial layout at the environmental scale (Fig. 1): the Unitary model, the Total Dissociation model, the Partial Dissociation model and the Mediation model. Our results support the Partial Dissociation model and are inconsistent with the other three models. Moreover, they specify the degree of overlap between small-scale and large-scale abilities, and provide new insights about what these abilities have in common and how they differ. First, it is clear that the variance they share is specific to spatial information processing, and does not reflect verbal or general intelligence. Second, our results suggest that the shared variance is specific to spatial representations constructed from visual inputs. Thus, measures of small-scale spatial ability are very highly predictive of environmental learning from visual media, but less predictive of learning from direct experience, which occurs in environmental spaces and involves other sensory inputs such as kinesthesis. Ability to process spatial information based on these other sensory inputs, in turn, accounts for the dissociation that we observed between small-scale spatial abilities measures and measures of learning from direct experience. The elaborated Partial Dissociation model suggested by our research is depicted in Fig. 9. Note that this model does not specify which, if any, processes are specific to small-scale spatial tasks, and this question is beyond the scope of the present study.

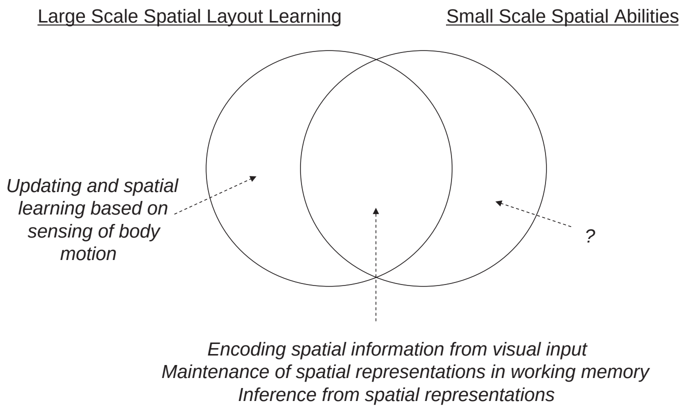

Fig. 9. Model of the relation between small-scale and large-scale spatial abilities proposed on the basis of this research.

The model proposed in Fig. 9 is consistent with the theories of behavioral scientists (Acredolo, 1981; Cutting & Vishton, 1995; Freundschuh & Egenhofer, 1997; Gärling & Golledge, 1987; Kuipers, 1982; Mandler, 1983; Montello, 1993; Montello & Golledge, 1999; Tversky et al., 1999) and neuroscientists (e.g., Aguirre & D'Esposito, 1999; Previc, 1998) who have emphasized the dissociation between the mental systems for processing spatial information at different scales of space, and provides new information regarding the nature of this dissociation. On the other hand, it suggests that theorists need to pay more attention to the ways in which these systems work together to accomplish complex spatial cognitive tasks, and the extent to which spatial tasks at different scales of space rely on common processes for encoding, maintaining and transforming spatial representations. This study has shown that the study of individual differences, especially when analyzed using structural equation models, is a promising method for investigating these questions.

4.1. Learning from direct experience and visual media

An important new finding of this study is that measures of spatial learning from a real environment and measures of learning from visual media are partially dissociated in that they load on different factors. We can interpret this dissociation in terms of the task analysis presented in Fig. 2. First, real navigation by walking and simulated navigation by watching a video or a desktop VE differ in that the former includes sensory inputs from body-based senses, whereas the others involve visual input alone. In regard to learning, recent research has suggested that vestibular information does not contribute to learning of spatial layout in spaces at the scale studied in our video environment (Waller et al., 2003), although other body-based senses involved in active motion (i.e., kinesthesis and efference copy of the muscle commands) do affect learning of spatial layout at this scale (Waller et al., 2004). Vestibular information does contribute to spatial updating at smaller scales, similar to that of our VE simulation (Chance et al., 1998; Klatzky et al., 1998).

The quality of visual information sensed by participants also differed for the real environment and visual media. The building space that our participants walked through surrounded them and was larger than their bodies, both distally and proximally on the retina. In contrast, both the video and computer monitors that were used to display the video and virtual environments show flat pictures of less than a meter square distally, and one half or less of the visual field proximally. This field of view in the visual media eliminates most peripheral vision, which means that a given visual feature disappears from the field of view more quickly in a visual medium than in the real world, placing more demand on working memory to maintain an internal representation of the feature location (Sholl, 1996). Thus, the dissociation between learning from direct experience and learning from visual might also reflect differential demands on internal maintenance of information involved in processing different displays.

The dissociation of ability to learn from direct experience and ability to learn from visual-media might also reflect individual differences in ability to update one's position in an environment while performing outcome measures such as direction and distance estimation, which rely on an accurate representation of one's location. Participants in our study were tested "in situ," such that they walked through the real environment when their knowledge of that environment was being tested, but they sensed their motion by vision alone during the testing phase for the video and virtual environments. Walking in a real environment enables one to automatically update one's location, whereas this updating is more effortful when movement is sensed by vision alone (Klatzky et al., 1998; Loomis et al., 1999). Testing in a visual medium therefore relies more on effortful updating.

An issue that arises in interpreting the dissociation between learning from real environments and from visual media is that the environments learned from these experiences were different; that is, had different shapes and varied in complexity. Although it is important to replicate our results in other environments, a few considerations suggest that differences between the environments are unlikely to be a major cause of the dissociation observed in this study. First, while the environments had different shapes, all three were composed of mostly straight segments and right-angled turns. Second, measures of learning the video and virtual environment loaded on the same factor, although the environments learned had different shapes. Third, in terms of complexity, the environments learned from direct experience and from video were similar in that they both involved integrating two floors of a building and learning a relatively large number of landmarks (8 and 7, respectively), in contrast with the virtual environment that had just one floor with 4 landmarks. However, measures of learning these complex environments loaded on different factors in the exploratory factor analysis.

The partial dissociation between learning from real environments and from visual media has important methodological implications for spatial cognition research, because visual media such as video and virtual environments are increasingly being used to assess environmental skills (e.g., Goldin & Thorndyke, 1982; Juan-Espinosa et al., 2000; Moffat, Hampson, & Hatzipantelis, 1998; Pearson & Ialongo, 1986; Waller, 2000). Virtual environment technology is increasingly used in behavioral studies because of the experimental control that it affords (Blascovich, Loomis, & Beall, 2002), and it is necessary to use visual media in brain imaging studies of spatial-layout learning because participants cannot physically move while being scanned. Our study suggests that researchers should be cautious in drawing conclusions about real-world spatial learning from measures of environmental learning in simulations that provide visual information alone. Specifically, measures based on visual simulations may lead one to overestimate the relation between small- and large-scale spatial abilities. On the other hand, learning from navigation in a real world and from visual media also share considerable variance ($r = .61$ in this study; see also Richardson et al., 1999). Therefore, depending on the research goal, in some situations a measure of learning from a video or virtual environment may provide an adequate approximation of real-world learning ability.

In contrast to different learning media, different outcome measures did not define separate factors in our analyses, indicating that they measure a common ability. In terms of the task analysis presented in Fig. 2 this suggests that there are not large individual differences in the “readout processes” involved in using the internal representation of the environment to perform a given outcome measure, or that abilities associated with these readout processes are highly correlated. The fact that these different measures were not separable in the factor analysis is somewhat surprising, because there has been much discussion of the differential task demands potentially imposed by various measures of environmental knowledge (e.g., Kitchen, 1996; Newcombe, 1985). In particular, map drawing, which depends less on ability to internally integrate spatial information (and depends to some extent on drawing ability) was not separable from distance and straight-line direction measures, suggesting that all three measures reflect ability to acquire and infer knowledge of the configuration of an environment from information sensed over time as one navigates the environment. Although we have stressed the importance of taking task demands into account in making inferences from measures of environmental knowledge (Montello et al., 2004), the present study suggests that their contribution to individual differences in performance is not decisive, at least for the set of measures used here.

4.2. Relations between large and small-scale spatial abilities

In providing evidence for a partial rather than total dissociation between large- and small-scale spatial abilities, our study argues for a stronger relation between these abilities than previous individual differences studies. There are several possible reasons for this stronger association. First, in contrast to previous studies, which selected small-scale spatial abilities on the basis of their prevalence in the psychometric literature, we selected our measures of small-scale spatial ability on the basis of a cognitive task analysis. As a result, we sampled a broader set of measures of small-scale spatial ability, including a measure of spatial working memory, which reflects ability to maintain spatial representations in memory that probably underlies both small- and large-scale spatial tasks. Second we used structural equation modeling, which enabled us to factor out extraneous sources of variance in our analyses. Consistent with previous studies of environmental spatial abilities (Allen et al., 1996; Bryant, 1982; Goldin & Thorndyke, 1982; Lorenz, 1988; Lorenz & Neisser, 1986; Pearson & Ialongo, 1986), we observed relatively low zero-order correlations between small-scale and environmental spatial abilities. However, when we partialed out the variance associated with specific tests and controlled for other variables such as verbal intelligence in structural equation models, we observed a much stronger relationship (a similar relationship was reported by Kirasic, 2000, who also used structural equation modeling). Thus, task-specific variance and error variance may have masked a stronger relationship in previous studies. A third possible reason for the stronger relationship observed in this study is that our measures of large-scale spatial ability were based on learning of new environments, rather than existing knowledge of familiar environments (as studied by Bryant, 1982; Goldin & Thorndyke, 1982), which can be confounded by factors such as amount of exposure to the environment.

Finally, and perhaps most importantly, our outcome measures primarily assessed the ability to infer configural properties of environments in contrast with the large-scale spatial ability factors identified in previous studies (Allen et al., 1996; Lorenz, 1988; Pearson & Ialongo, 1986), which were defined by tasks that can be accomplished from a route representation and do not necessarily involve metric configural knowledge. This is most clearly seen in a comparison of our study with that of Allen et al. (1996). They identified a large-scale spatial factor called topological knowledge, which was totally dissociated from their small-scale spatial factor. The topological knowledge factor is very different from the environmental learning factor that we measured, in that it was based on measures of scene recognition, scene sequencing, map placement, intra-route distance error and route reversal error. This factor can be characterized as a measure of route memory. The measures making up this factor require little or no configurational knowledge or inference from the memory of the route encountered. Taking our study together with that of Allen et al., it is likely that the ability to remember the sequence of landmarks along a route does not share common processes with small-scale spatial cognition, but the ability to infer the configuration of an environment from route experience does. Thus, the two studies complement each other.

Although Allen et al. (1996) did not identify a configural knowledge factor (in their study, ability to point to unseen landmarks in a newly learned environment and ability to estimate straight-line distances between landmarks did not load on a common factor), they found in one experiment that perspective taking ability mediated the relationship between small-scale spatial ability and pointing to unseen landmarks. However, this result did not replicate in a second experiment. In our study, we observed a strong correlation between perspective taking ability and small-scale spatial ability factors (see also Kozhevnikov & Hegarty, 2001; Hegarty & Waller, 2004), and there was no evidence that perspective taking made a unique contribution to predicting spatial-layout learning when its common variance with small-scale spatial ability was partialed out. The discrepancy between our study and theirs might be explained by the fact that both perspective taking and large-scale spatial abilities were measured differently in the two studies. In particular, the perspective taking test used in our study was simpler in that it involved remembering the locations of 4 objects and imagining perspectives that were orthogonal to the studied view. In contrast, the task of Allen et al. involved taking perspectives that were both orthogonal and oblique to the viewed perspective of a model town. Although the perspective taking abilities measured in these two studies might not be comparable, the idea that tasks at the vista scale of space (such as perspective taking) might mediate the relation between large and small-scale spatial cognition is intriguing and worthy of future study.

4.3. Sex differences

Although there has been much speculation about the nature of sex differences in large-scale spatial cognition and its relation to sex differences in small-scale spatial abilities, few studies have actually measured sex differences in larger-scale tasks, especially in real (as opposed to virtual) environments. Our study informs theories of the nature and causes of sex differences in spatial cognition. First, the sizes of the sex differences in small-scale spatial abilities (mental rotation and embedded figures) replicate previous studies (reviewed by Linn & Petersen, 1985; Voyer et al., 1995), so that our sample is comparable to others in the literature. Second it indicates that there are sex differences in measures of large-scale spatial learning, although these differences are not evident in all measures and are smaller than the sex difference in mental rotation (the largest sex difference documented in the literature). Specifically, sex differences are especially evident in direction-estimation tasks, which rely most on the ability to infer configural and metric knowledge from spatial information learned sequentially, on a route. In contrast, they are less pronounced in map-drawing tasks, which can be accomplished with route knowledge. Although certainly not conclusive, our results are consistent with the view that there may be a stylistic difference in how males and females represent space, with males depending more on configural representations and females depending more on route representations (Lawton, 1994; Lawton et al., 1996; Montello, Lovelace, et al., 1999).

Our study also informs debate about the evolution of sex differences. This debate has revolved around questions of whether sex differences in tasks such as mental rotation can be attributable to selection pressures for males to have superior navigation ability, suggesting that mental rotation and navigation reflect the same abilities (Gaulin 1995; Jones et al., 2003; Kimura, 2000), or whether different selection pressures (e.g., in hunting vs. gathering) produce dissociable sets of spatial abilities that were subject to independent selection pressures (Silverman & Eals, 1992). Our study adds plausibility to the view that sex differences in small-scale spatial abilities may partially result from evolved sex differences in navigation, because navigation abilities shared some common variance with small-scale tasks, and sex was not a significant predictor of environmental learning when its common variance with small-scale tasks and sense-of-direction was partialed out (see also Waller, 2000). (A post hoc analysis indicated that this was true, even when environmental learning was not controlled for variance shared with sense-of-direction). On the other hand, it would be wrong to conclude from these data that sex differences in mental rotation specifically can be attributed primarily to differential selection pressures for navigation ability between the sexes, because the simple correlations between mental rotation and learning from direct experience are low (ranging from .11 to .21 for different outcome measures, see Table 2). Given evidence that spatial ability can be modified by experience and training (e.g., Baenninger & Newcombe, 1989), it is likely that evolutionary pressures alone do not fully explain sex differences in spatial ability.

4.4. Limitations

The present study has some limitations. Most notably, our measure of spatial updating ability, i.e. SBSOD, was an indirect measure based on self reports. In fact, we attempted to measure blindfolded updating in this study, but our measure was based on only two trials which were not sufficient to yield a reliable measure. Use of self-reported sense of direction as a measure of updating ability is justified because Kozlowski and Bryant (1977) found a large difference between good- and poor-sense-of-direction individuals in ability to update their position with respect to the start of a route through an underground maze, in which visual information was very impoverished. Furthermore, in a previous study (Hegarty et al., 2002), the SBSOD was correlated .40 with a more extensive measure of blindfold updating, and was correlated .44 with a task that included the ability to update one's position with respect to landmarks on a college campus as one walked from an outside location to an experimental room to be tested. The contribution of non-visual updating to environmental learning ability is an important issue that needs to be addressed more directly in future studies.

Another limitation of our study is that although we documented individual differences in large-scale spatial learning and can speculate about the nature of these differences based on their correlations with other abilities, we did not directly measure the specific strategies, representations or processes used by individuals with different spatial abilities. Much more research is needed to develop precise and comprehensive cognitive models of the performance of high- and low-ability individuals in environmental spatial cognition. Finally, like most other studies of large-scale spatial abilities, our study focused on learning of spatial layout, which is just one task that people accomplish in large-scale space. The approach used here also needs to be applied to study individual differences in other large-scale tasks, such as wayfinding in familiar environments, planning trips, and giving and understanding verbal directions. Finally, there was probably a restricted range in some of our measures, because most of our participants were college students, so it is important to replicate these findings with samples that are more representative of the general population.

4.5. Conclusion

In conclusion, perhaps one of the most striking findings of this study is that there are very large individual differences in environmental spatial abilities. Inspection of the range of values in Table 1 indicates that after a brief exposure to a new environment, the most able participants could point to unseen landmarks in the environment with less than 10° of absolute error, showed an almost perfect correlation between their estimates of distances among landmarks and the true distances, and made no qualitative errors in drawing sketch maps of the environments. In contrast, the least able participants had pointing performance that was not significantly different from chance, showed no correlation or even a negative correlation between estimated and true distances, and made numerous errors in drawing sketch maps. Individual differences in large-scale spatial cognition deserve more study, not just because they are so large and pervasive, but also because they can inform theories of the nature of spatial cognition.

Acknowledgement

We acknowledge the financial assistance of The Army Research Institute for Behavioral and Social Sciences (award DASW01-K-0014). We thank John Anderson, Stacey Caron, Chris Lentz, Chia-Lin Liang, Juan Navarro, Kimberly Pardinez, Erika Probst, Michael Provenza, Landon Romano, Lance Rushing, Chris Schwandt, and Lara Snyder for help with data collection and processing. We are also grateful to Nancy Collins, Mike Drillings, Sehee Hong, Jack Loomis, Joe Psotka, Ilavenil Subbiah, and David Waller who provided valuable comments on this research.

References

Acredolo, L. P. (1981). Small- and large-scale spatial concepts in infancy and childhood. In L. S. Liben, A. H. Patterson, & N. Newcombe (Eds.), *Spatial representation and behavior across the life span* (pp. 63–81). New York: Academic Press.

Aguirre, G. K., & D'Esposito, M. (1999). Topographical disorientation: A synthesis and taxonomy. *Brain*, *122*, 1613–1628.

Allen, G. L., Kirasic, K. C., Dobson, S. H., Long, R. G., & Beck, S. (1996). Predicting environmental learning from spatial abilities: An indirect route. *Intelligence*, *22*, 327–355.

Arbuckle, J. L. (1999). *Amos for Windows: Analysis of moment structures* (Version 4.0). Chicago, IL: Small Waters.

Baenninger, M., & Newcombe, N. (1989). The role of experience in spatial test performance: A meta-analysis. *Sex Roles*, *20*, 327–344.

Bennett, G. K., Seashore, H. G., & Wesman, A. G. (1972). *Differential aptitudes tests: Abstract reasoning, Form S*. New York, NY: The Psychological Corporation.

Blascovich, J. J., Loomis, J. M., & Beall, A. C. (2002). Immersive virtual environment technology as a basic research tool in psychology. *Behavior Research Methods, Instruments and Computers*, *314*, 557–564.

Bryant, K. J. (1982). Personality correlates of sense-of-direction and geographic orientation. *Journal of Personality and Social Psychology*, *43*, 1318–1324.

Bryant, K. J. (1991). Geographical/spatial orientation ability within real-world and simulated large-scale environments. *Multivariate Behavioral Research*, *26*, 109–136.

Carroll, J. (1993). *Human cognitive abilities: A survey of factor-analytical studies*. New York, NY: Cambridge University Press.

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

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Cooper, L. A., & Mumaw, R. J. (1985). Spatial aptitude. In R. F. Dillon (Ed.), *Individual differences in cognition* (pp. 67–94). Orlando, FL: Academic Press.

Cutting, J. E., & Vishton, P. M. (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In W. Epstein, & S. Rogers (Eds.), *Handbook of perception and cognition*, *vol. 5* (pp. 69–117). San Diego, CA: Academic Press.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450–466.

Ekstrom, R. B., French, J. W., Harman, H. H., & Derman, D. (1976). *Kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.

Eliot, J., & Smith, I. M. (1983). *An international directory of spatial tests*. Windsor, Berkshire: NFER-Nelson.

Evans, G. W. (1980). Environmental cognition. *Psychological Bulletin*, *88*, 259–287.

Freundschuh, S. M., & Egenhofer, M. J. (1997). Human conceptions of spaces: Implications for geographic information systems. *Transactions in GIS*, *2*, 361–375.

Gärling, T., & Golledge, R. G. (1987). Environmental perception and cognition. In E. H. Zube, & G. T. Moore (Eds.), *Advances in environment, behavior, and design* (pp. 203–236). New York, NY: Plenum Press.

Gaulin, S. J. C. (1995). Does evolutionary theory predict sex differences in the brain? In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 1211–1225). Cambridge, MA: MIT Press.

Goldin, S. E., & Thorndyke, P. W. (1982). Simulating navigation for spatial knowledge acquisition. *Human Factors*, *24*, 457–471.

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.

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

Hegarty, M., & Waller, D. (2005). Individual differences in spatial abilities. In P. Shah, & A. Miyake (Eds.), *The Cambridge handbook of visuospatial thinking* (pp. 121–169). Cambridge University Press.

Hu, L. -T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*, 424–453.

Huttenlocher, J., & Presson, C. C. (1973). Mental rotation and the perspective problem. *Cognitive Psychology*, *4*, 277–299.

Huttenlocher, J., & Presson, C. C. (1979). The coding and transformation of spatial information. *Cognitive Psychology*, *11*, 375–394.

Ittelson, W. H. (1973). Environment perception and contemporary perceptual theory. In W. H. Ittelson (Ed.), *Environment and cognition* (pp. 1–19). New York, NY: Seminar.

Jones, C. M., Braithwaite, V. A., & Healy, S. D. (2003). The evolution of sex differences in spatial ability. *Behavioral Neuroscience*, *117*, 403–411.

Juan-Espinosa, M., Abad, F. J., Colom, R., & Fernandez-Truchaud, M. (2000). Individual differences in large-spaces orientation: G and beyond? *Personality and Individual Differences*, *29*, 85–98.

Just, M. A., & Carpenter, P. A. (1985). Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability. *Psychological Review*, *92*, 137–171.

Kimura, D. (2000). *Sex and cognition*. Cambridge, MA: MIT Press.

Kirasic, K. C. (2000). Age differences in adults' spatial abilities, learning environmental layout, and wayfinding behavior. *Spatial Cognition and Computation*, *2*, 117–134.

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

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

Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York, NY: Guilford.

Kosslyn, S. M., & Thompson, W. L. (2003). When is early visual cortex activated during visual mental imagery? *Psychological Bulletin*, *129*, 723–746.

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

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

Kuipers, B. (1982). The "map in the head" metaphor. *Environment and Behavior*, *14*, 202–220.

Lawton, C. A. (1994). Gender differences in wayfinding strategies: Relationship to spatial ability and spatial anxiety. *Sex Roles*, *30*, 765–779.

Lawton, C. A., Charleston, S. I., & Zieles, A. S. (1996). Individual- and gender-related differences in indoor wayfinding. *Environment and Behavior*, *28*, 204–219.

Liben, L. S., Patterson, A., & Newcombe, N. (Eds.) (1981). *Spatial representation and behavior across the life span*. New York, NY: Academic Press.

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development*, *56*, 1479–1498.

Lohman, D. F. (1988). Spatial abilities as traits, processes, and knowledge. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 181–248). Hillsdale, NJ: Lawrence Erlbaum Associates.

Loomis, J. M., Klatzky, R. L., Golledge, R. G., & Philbeck, J. W. (1999). Human navigation by path integration. In R. G. Golledge (Ed.), *Wayfinding behavior: Cognitive mapping and other spatial processes* (pp. 125–151). Baltimore, MD: Johns Hopkins Press.

Lorenz, C. A. (1988).*The structure of the spatial domain. An analysis of individual differences*. Unpublished doctoral dissertation, Cornell University.

Lorenz, C. A., & Neisser, U. (1986). *Ecological and psychometric dimensions of spatial ability* (Rep. No. 10). Emory Cognition Project, Atlanta, GA: Emory University.

Mandler, J. M. (1983). Representation. In P. Mussen (Ed.), *Handbook of child psychology*, *vol. III* (4th ed.) (pp. 420–494). New York, NY: John Wiley and Sons.

McGee, M. G. (1979). Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences. *Psychological Bulletin*, *86*, 889–918.

Meld, A. (1985). Map- and route-learning strategies in spatial cognition: A comparison of survey and procedural knowledge. (Doctoral dissertation, University of Washington, 1986). *Dissertation Abstracts International, 46(7-A)*, 1880.

Miyake, A., Rettinger, D. A., Friedman, N. P., Shah, P., & Hegarty, M. (2001). Visuospatial working memory, executive functioning and spatial abilities. How are they related? *Journal of Experimental Psychology: General*, *130*, 621–640.

Moffat, S. D., Hampson, E., & Hatzipantelis, M. (1998). Navigation in a "virtual" maze: Sex differences and correlation with psychometric measures of spatial ability in humans. *Evolution and Human Behavior*, *19*, 73–87.

Montello, D. R. (1991). The measurement of cognitive distance: Methods and construct validity. *Journal of Environmental Psychology*, *11*, 101–122.

Montello, D. R. (1993). Scale and multiple psychologies of space. In A. U. Frank, & I. Campari (Eds.), *Spatial information theory: A theoretical basis for GIS. Proceedings of COSIT '93*. *Lecture notes in Computer Science*, *vol. 716* (pp. 312–321). Berlin: Springer-Verlag.

Montello, D. R., & Golledge, R. G. (1999). Scale and detail in the cognition of geographic information. *Report of the Specialist Meeting of Project Varenius, Santa Barbara, CA, May 14–16, 1998*. Santa Barbara: University of California.

Montello, D. R., Lovelace, K. L., Golledge, R. G., & Self, C. M. (1999). Sex-related differences and similarities in geographic and environmental spatial abilities. *Annals of the Association of American Geographers*, *89*, 515–534.

Montello, D. R., & Pick, H. L. (1993). Integrating knowledge of vertically aligned large-scale spaces. *Environment and Behavior*, *25*, 457–484.

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

Montello, D. R., Waller, D., Hegarty, M., & Richardson, A. E. (2004). Spatial memory of real environments, virtual environments, and maps. In G. L. Allen, & D. Haun (Eds.), *Remembering where: Advances in understanding spatial memory* (pp. 251–285). Mahwah, NJ: Lawrence Erlbaum Associates.

Morris, R. G., & Parslow, D. (2004). Neurocognitive components of spatial memory. In G. L. Allen, & D. Haun (Eds.), *Remembering where: Advances in understanding spatial memory* (pp. 217–247). Mahwah, NJ: Lawrence Erlbaum Associates.

Newcombe, N. (1985). Methods for the study of spatial cognition. In R. Cohen (Ed.), *The development of spatial cognition* (pp. 277–300). Hillsdale, NJ: Lawrence Erlbaum Associates.

Oltman, P. K., Raskin, E., & Witkin, H. A. (1971). *Group embedded figures test*. Palo Alto, CA: Consulting Psychologists Press.

Pearson, J. L., & Ialongo, N. S. (1986). The relationship between spatial ability and environmental knowledge. *Journal of Environmental Psychology*, *6*, 299–304.

Pellegrino, J. W., & Kail, R. V. (1982). Process analyses of spatial aptitude. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence*, *vol. 1* (pp. 311–365). Hillsdale, NJ: Erlbaum.

Philbeck, J. W., Behrmann, M., Black, S. E., & Ebert, P. (2000). Intact spatial updating during locomotion after right posterior parietal lesions. *Neuropsychologia*, *38*, 950–963.

Previc, F. H. (1998). The neuropsychology of 3-D space. *Psychological Bulletin*, *124*, 123–164.

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

Rovine, M. J., & Weisman, G. D. (1989). Sketch-map variables as predictors of way-finding performance. *Journal of Environmental Psychology*, *9*, 217–232.

Schneider, W. (1990). *MEL User's Guide: Computer techniques for real time experimentation*. Pittsburgh, PA: Psychology Software Tools.

Schutz, R. W., & Roy, E. A. (1973). Absolute error: The devil in disguise. *Journal of Motor Behavior*, *5*, 141–153.

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General*, *125*, 4–27.

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

Sholl, M. J. (1996). From visual information to cognitive maps. In J. Portugali (Ed.), *The construction of cognitive maps* (pp. 157–186). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Silverman, I., & Eals, M. (1992). Sex differences in spatial ability: Evolutionary theory and data. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 533–549). New York: Oxford University Press.

Spencer, C., Blades, M., & Morsley, K. (1989). *The child in the physical environment: The development of spatial knowledge and cognition*. Chichester: John Wiley and Sons.

Tversky, B., Morrison, J. B., Franklin, N., & Bryant, D. J. (1999). Three spaces of spatial cognition. *The Professional Geographer*, *51*, 516–524.

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

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin*, *117*, 250–270.

Waller, D. (2000). Individual differences in spatial learning from computer-simulated environments. *Journal of Experimental Psychology: Applied*, *6*, 307–321.

Waller, D., Loomis, J. M., & Haun, D. B. M. (2004). Body-based senses enhance knowledge of directions in large-scale environments. *Psychonomic Bulletin and Review*, *11*, 157–163.

Waller, D., Loomis, J. M., & Steck, S. B. (2003). Inertial cues do not enhance knowledge of environmental layout. *Psychonomic Bulletin and Review*, *10*, 987–993.

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

Walsh, D. A., Krauss, I. K., & Regnier, V. A. (1981). Spatial ability, environmental knowledge, and environmental use: The elderly. In L. S. Liben, A. Patterson, & N. Newcombe (Eds.), *Spatial representation and behavior across the life span* (pp. 321–357). New York, NY: Academic Press.

Zacks, J. M., Mires, J., Tversky, B., & Hazeltine, E. (2000). Mental spatial transformations of objects and perspective. *Spatial Cognition and Computation*, *2*, 315–332.