Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places ☆

Toru Ishikawa [*,1], Daniel R. Montello

University of California, Santa Barbara, USA

Abstract

Existing frameworks for explaining spatial knowledge acquisition in a new environment propose either stage-like or continuous development. To examine the spatial microgenesis of individuals, a longitudinal study was conducted. Twenty-four college students were individually driven along two routes in a previously unfamiliar neighborhood over 10 weekly sessions. Starting Session 4, they were also driven along a short connecting route. After each session, participants estimated spatial properties of the routes. Some participants' knowledge improved fairly continuously over the sessions, but most participants either manifested accurate metric knowledge from the first session or never manifested accurate metric knowledge. Results are discussed in light of these large individual differences, particularly with respect to the accuracy and development of integrated configurational knowledge.
© 2005 Elsevier Inc. All rights reserved.

1. Introduction

As people act in the environment, they perceive surrounding space and acquire knowledge about it. Downs and Stea (1973) called this fundamental process *cognitive mapping*. Knowledge acquired during cognitive mapping includes the identities of places and landmarks, the patterns of path connections between places, distances, and directions between places, and so on. People use spatial knowledge of the environment to get to destinations such as home and work, to give and interpret navigational instructions, to interpret maps, to plan efficient trips, and more. These kinds of knowledge help guide people's actions in adaptive ways, in other words, so that their behavior is coordinated not only to the environment as perceived but also to the environment as conceived and remembered. Many researchers from various disciplines have investigated human spatial knowledge of environments, particularly since the late 1960s (see Cox & Golledge, 1969; Evans, 1980; Liben, Patterson, & Newcombe, 1981; Lynch, 1960).

A major research question that has attracted much theoretical interest concerns the structure of spatial knowledge about environments and the process of spatial knowledge acquisition in new environments. As Ittelson (1973) discussed, the *environment* is larger than and surrounds the human body, so that a person cannot grasp the layout of the environment in its entirety from a single viewpoint. Instead, one must locomote about the environment, integrating knowledge acquired from separate viewpoints and travel experiences. Due to this unique characteristic, learning the spatial layout of environments from direct experience is not a straightforward task, especially for people with a poor "sense-of-direction" (e.g., Kozlowski & Bryant, 1977). In fact, researchers have shown that people's knowledge about environments tends to be distorted, fragmented, and schematized (e.g., Golledge & Spector, 1978; Lynch, 1960; Stevens & Coupe, 1978).

1.1. Spatial microgenesis: Theoretical frameworks

Siegel and White (1975) proposed a theoretical framework for describing and explaining the process of knowledge development over time in new environments (called *spatial cognitive microgenesis*). In their framework, internal representations of spatial knowledge of a new place progress over time from an initial stage of landmark knowledge to a stage of route knowledge to an ultimate stage of survey knowledge. *Landmark knowledge* is knowledge about the identities of discrete objects or scenes that are salient and recognizable in the environment. *Route knowledge* consists of sequences of landmarks and associated decisions (e.g., "turn left at the gas station and go straight for three blocks"). According to Siegel and White, the space between landmarks is at first "empty" and receives "scaling" with accumulated experience (p. 29); in other words, route knowledge is initially nonmetric. The final and most sophisticated stage of knowledge in their framework is *survey knowledge.* This is a two-dimensional and "map-like," quantitatively scaled representation of the layout of the environment. Survey knowledge represents distance and directional relationships among landmarks, including those between which direct travel has never occurred. For survey maps to emerge, routes need to be metrically scaled and interrelated into a global allocentric reference system.

Siegel and White's framework was so influential in the scientific literature (including psychology, geography, and robotics) that Montello (1998) called it the *dominant framework*. This was certainly true for the time period up until the early 1990s, about which Montello was writing (see, e.g., Chase & Chi, 1981; Clayton & Woodyard, 1981; Golledge, Smith, Pellegrino, Doherty, & Marshall, 1985; Hirtle & Hudson, 1991; Kuipers & Levitt, 1988; Lloyd, 1989; McDermott & Davis, 1984; Moeser, 1988; Thorndyke & Hayes-Roth, 1982). Even though writers no longer cite Siegel and White's framework so commonly, for the most part they have not replaced it with another coherent theoretical framework, so we continue to refer to Siegel and White's proposal as the dominant framework. And even though Siegel and White may not be cited by name so often anymore, their ideas are still influential. For instance, many researchers continue to refer to the construct of route knowledge with terms such as *procedural knowledge* (Golledge, 1991) or *topological knowledge* (Allen, Kirasic, Dobson, Long, & Beck, 1996), terms that emphasize the sequential and nonmetric nature of the knowledge. There are computational process models of spatial knowledge that are based on discrete landmarks and sequential routes as conceived by the dominant framework (Yeap & Jefferies, 2000, described such models as "object-based" in contrast to "space-based" approaches). Aguirre and D'Esposito (1997) provide a recent example from cognitive neuroscience, a field in its childhood when Siegel and White first published their framework.

A comment is in order about landmark recognition. People recognize a landmark for its salience in the surroundings in terms of size, shape, or color, and for its functionality as a navigation clue. So at least approximate metric information may exist at or in the vicinity of landmarks that allows people to attend to the size and shape of objects. But the important property of landmark and route knowledge as conceived by the dominant framework is that individual landmarks are encoded as individual entities, and related to other landmarks in terms of only connectivity and order. This "skeletal" nature of people's spatial knowledge after early exposure to environments is similarly reflected in Golledge's (1978) conceptual framework, called the *anchor point theory*.

Its popularity aside, however, Siegel and White's (1975) dominant framework has not received convincing empirical support as a model of the structure and microgenetic course of spatial knowledge. Particularly troubling, as pointed out by Montello (1998), are the ideas that metric knowledge takes so long to begin developing, and that landmark knowledge is a necessary prerequisite for route knowledge, which in turn is a necessary prerequisite for survey knowledge. Research has shown, on the contrary, that with minimal exposure to a new environment (on the order of seconds or minutes), people can perform tasks that require some metric configurational knowledge[2] at a level at least better than chance—tasks such as taking shortcuts, returning directly back to starting locations, and estimating distances and directions directly between places (e.g., Klatzky et al., 1990; Landau, Spelke, & Gleitman, 1984; Loomis et al., 1993). Such tasks require that people

---

[2] It has repeatedly been shown that estimates of distances made independently (e.g., by pairwise estimates rather than sketch maps), even by spatially competent long-term residents, violate metric axioms (e.g., Burroughs & Sadalla, 1979); Montello (1992) cites such findings and considers possible interpretations of this ubiquitous phenomenon. The distance estimates cannot therefore be considered evidence for a unitary metric cognitive map. We refer to spatial knowledge as *metric* in the less strict sense that distances and directions are known as quantities, on at least an approximate interval or ratio scale. We do not mean, for example, that independent estimates of distances are necessarily symmetric. By *configurational*, we mean that spatial knowledge represents layouts as two-dimensional (or higher) spaces and is not restricted to (near) one-dimensional spatiality.

understand quantitative distances and directions between places (perhaps somewhat vaguely), which the dominant framework claims should not be possible to do by people without accumulated experience.

Besides being apparently contradicted by empirical evidence, the dominant framework provides only a general description of the developmental course of spatial microgenesis (hence we call it a *framework* rather than a *theory*). It does not tell us how much time or effort is necessary to proceed from one stage to another. Blades (1991) made a similar criticism of the descriptive, nonexplicit nature of the dominant framework.

In response to these problems with the dominant framework, Montello (1998) proposed an alternative framework for explaining people's spatial microgenesis. This framework posits continuous (or quantitative) development of metric knowledge, rather than discrete (or qualitative) development, as the dominant framework posits. This echoed suggestions by writers such as Evans (1980), Hirtle and Hudson (1991), and Thorndyke (1981). The idea that spatial knowledge, including metric knowledge, is acquired in new environments relatively continuously is at the core of Montello's framework, so we refer to it as the *continuous framework.* Other aspects of his framework are presented below.

1.2. Empirical studies of spatial microgenesis

Most of the empirical studies of the microgenetic development of spatial knowledge have used hand-drawn sketch maps as a measure of knowledge (e.g., Appleyard, 1970; Beck & Wood, 1976; Devlin, 1976) or have used cross-sectional designs that implicitly equate length of residence with amount of exposure to the environment (e.g., Herman, Kail, & Siegel, 1979; Stern & Leiser, 1988; Thorndyke & Hayes-Roth, 1982). These studies have failed to yield consistent results about the structure and development of spatial knowledge, due mainly to (a) differences in the structure and complexity of the study area, (b) difficulties in objectively classifying the style and degree of sophistication of sketch maps, and (c) lack of control over respondents' amount of experience and sources of knowledge (such as direct travel or maps).

Only a few studies have experimentally examined changes in the accuracy of people's spatial knowledge over time. Gärling, Böök, Lindberg, and Nilsson (1981) and Herman, Blomquist, and Klein (1987) had participants travel a route, and asked them to do spatial tasks such as distance and direction estimation. Participants repeated this set of route travel and estimation tasks a couple of times. Adult participants in these studies acquired knowledge of landmarks immediately, and their direction and distance estimates got better during early sessions and then leveled off, never attaining perfect accuracy.

Overall, the results from these empirical studies, which used a variety of different methods, seem to indicate that route knowledge and at least an approximate knowledge of distances and directions are acquired early on in a new environment. However, the results are not consistent enough to allow one to draw a solid conclusion about spatial microgenesis. Missing are studies that examine microgenetic development over an extended period of time with appropriate experimental control.

1.3. Integration of separately learned places

Both the dominant framework (Siegel & White, 1975) and Montello's (1998) continuous framework argue that the acquisition of coordinated and comprehensive survey

knowledge is an important process in microgenetic development. For survey maps to emerge, routes need to be metrically scaled and interrelated into a global allocentric reference system. In other words, places and routes learned during separate travel experiences are integrated[3] and interrelated with each other in a common frame of reference. This is a fairly sophisticated step in spatial microgenesis. For example, researchers (e.g., Lynch, 1960; Rand, 1969) have often observed that people who are familiar with two separate regions may not understand the spatial relationship between them.

Some researchers have examined how accurately people can integrate separately learned places into a common frame of reference (e.g., Hanley & Levine, 1983; Holding & Holding, 1988; Golledge, Ruggles, Pellegrino, & Gale, 1993; Moar & Carleton, 1982; Montello & Pick, 1993). These researchers presented two separate routes to participants, and then assessed their spatial knowledge about each route (within-route tasks) and about the spatial relation between the two routes (between-route, or integration, tasks).

The results of these studies on integration are somewhat inconsistent. Participants in some studies (Holding & Holding, 1988; Moar & Carleton, 1982) successfully integrated two separate routes, showing evidence of spatial knowledge that was just as accurate for the integrated routes as for the separate routes. Participants in other studies (Golledge et al., 1993; Hanley & Levine, 1983; Montello & Pick, 1993) failed to integrate the separate routes, or were less accurate for the integrated than for the separate routes. This inconsistency may have been due to differences in (a) the source of information (direct navigation vs. slides), (b) the spatial scale (environmental routes vs. tabletop paths), (c) the complexity of the routes, or (d) the information given to participants that allowed integration of the two routes (the two routes shared a common segment vs. the two routes were completely separate and information about the relationship between the two was given later as verbal descriptions or direct experience).

1.4. Individual differences

Montello's (1998) continuous framework states that individuals will vary in the extent and accuracy of the spatial knowledge that they acquire from direct experience. He posited that this variation, especially in the ability to integrate separately learned places, was so large that it deserved special note because it could confound general statements about the form of microgenesis. For example, individual differences could confound questions of how much metric information is acquired early, how precise is the metric knowledge, whether integration occurs, and if so, when. Such individual differences are ignored (relegated to the error term) in an aggregate analysis. There are studies of individual differences in spatial cognition, for example, with respect to underlying spatial abilities (Allen et al., 1996), sense-of-direction (Hegarty, Richardson, Montello, Lovelace, & Subbiah, 2002), and neural correlates of spatial thinking (Hartley, Maguire, Spiers, & Burgess, 2003). The study we report here specifically addresses in detail individual differences in the developmental

---

[3] This concept of integration is to be distinguished from the concept of *path integration* commonly used by navigation researchers, particularly those who study nonhuman animals (e.g., Mittelstaedt & Mittelstaedt, 1980), to mean the mechanism of updating position in space without reference to landmarks. The name comes from its equivalence to integrating velocity or acceleration over time (*dead reckoning* is a term often used more or less synonymously). Path integration in this sense almost certainly plays an important role in the acquisition of metric knowledge, so it is especially critical to microgenesis, according to the continuous framework.

pattern of people's spatial knowledge, which is quite unusual in the literature on spatial microgenesis.

1.5. Objectives of this research

Except for the few attempts discussed above, microgenetic development of spatial knowledge has not been investigated much in controlled longitudinal studies. As we discussed above, the dominant framework, explicitly proposed 30 years ago, has strongly influenced spatial cognition research in several disciplines; even if not cited by name, its concepts of nonspatial landmark knowledge and, especially, nonmetric and sequential route knowledge still echo through the writings of various cognitive-science fields (as we cited above). This is true even though it has never received clear empirical support. Likewise, critiques of the dominant framework and the ideas suggested as alternatives have not had much empirical research specifically directed at them. In our research, therefore, we aim to investigate how people's spatial knowledge of the environment, particularly the accuracy and precision of their knowledge, develops over time.

We are specifically motivated by three major research questions. Our first question concerns the development of metric knowledge (distances and directions) in a new environment learned directly via locomotion. In particular, we look at the level of people's configurational understanding of routes (i.e., beyond the sequential, nonmetric route knowledge of the dominant framework's terminology). Our second question concerns people's ability to integrate separately learned places into a common frame of reference (i.e., the acquisition of integrated survey knowledge). Related to this question, we also look at whether people acquire spatial knowledge differently when they travel a route in two opposite directions, compared to when they travel a route in only one direction. Our third major question concerns individual differences in people's spatial knowledge; specifically, how different individuals' developmental curves look with respect to accuracy and a developmental pattern, and whether the difference is especially large for integration.

To investigate these questions, we conducted a longitudinal experiment in a naturalistic setting, with as much experimental control as possible concerning the amount and method of exposure to the environment. We examined participants' performance on tasks that assessed their spatial knowledge, both metric and nonmetric, once a week for 10 consecutive weeks.

1.6. Predictions of the dominant and continuous frameworks

If spatial microgenesis progresses through qualitatively distinct stages as conceived by the dominant framework, participants' performance on metric tasks should be at or near chance level in early sessions. Only after repeated sessions, participants' performance should become better than chance and begin to improve. On the other hand, if spatial microgenesis is continuous, as the continuous framework argues, participants' performance should be at least better than chance from the first session and gradually improve. Both the dominant and continuous frameworks agree that the acquisition of integrated configurational knowledge (or survey knowledge) is a large, sophisticated step in microgenetic development. If so, people's performance should be poor on between-route, as opposed to within-route, tasks, at least in early sessions. Also, their performance should be affected by the complexity of routes. Fig. 1 schematically illustrates the predicted patterns of development by the two frameworks.
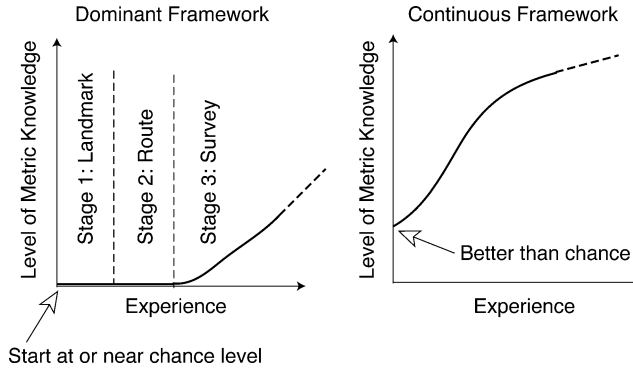
Fig. 1. Schematic illustrations of the predicted patterns of development by the dominant and continuous frameworks. According to the dominant framework, participants should do poorly on metric tasks during the stages of landmark knowledge and route knowledge. The continuous framework proposes continuous development of metric performance, which should start above chance level.

As well as comparing participants' performance to chance performance, which assumed complete lack of metric knowledge, we considered possible performance by "hypothetical" participants that we gave varying degrees of accuracy and precision of metric knowledge. We examined the performance of these hypothetical participants by conducting Monte Carlo simulations (Table 1). First, we simulated participants who would know only the identities of the landmarks. These participants can be thought of as being in the stage of landmark knowledge as conceived by the Siegel and White framework. When asked to draw a sketch map, such participants would simply locate points randomly on paper. Other such participants might locate points based on some heuristic rule, such as aligning the points in a straight line, or locating landmarks from one route on the right side of paper and landmarks from the other route on the left side. Researchers have reported that people tend to use these kinds of simplifying heuristics to judge or recall angles of turns (Byrne, 1979; Lynch, 1960; Moar & Bower, 1983).

Second, we simulated participants who would know the identities *and* the sequence of landmarks. These participants can be thought of as being in the stage of route knowledge as conceived by the Siegel and White framework. When asked to estimate the direction from one landmark to another, they would randomly choose a forward or backward direction, depending on whether the target landmark came after or before the current landmark in the sequence. Montello and Frank (1996) conducted similar simulations of qualitative direction judgments.

Table 1
Different types of knowledge possessed by hypothetical agents and simulation models of task performance

Type of knowledge	Task	Simulation model
Identities of landmarks	Sketch map	Random 6, random 12
		Straight-line, random LR
Identities/sequence of landmarks	Direction	Front-back, left-right
Some quantitative understanding	Direction, sketch map	90°-random, 90°-ordinal
	Distance	Two-category, ordinal

Note. Performance by models with these types of knowledge was examined by Monte Carlo simulations.

Third, we simulated participants who would know the identities and sequence of landmarks, but also possess some degree of quantitative knowledge. These participants can be thought of as being in transition from the stage of route knowledge to that of survey knowledge as conceived by the Siegel and White framework. Such participants would notice major changes in heading during travel and encode them as right-angled turns to the right or left. Other such participants might roughly conceive of the spatial relation between the two routes as one being to the right or left of the other. Other participants might possess minimal distance knowledge, in terms of two categories of relatively short or long, or in a little more detail, in terms of a rank order. These kinds of qualitative spatial reasoning have been discussed in the literature, particularly by computer scientists (Forbus, Nielsen, & Faltings, 1991; Frank, 1996).

2. Method

2.1. Participants

Twenty-four students (11 male and 13 female) at the University of California, Santa Barbara, participated in the experiment. Their ages ranged from 18 to 23 with a mean of 20.2 years. Among people who signed up for the experiment (announced in an introductory geography class), only those who indicated on a screening questionnaire that they had never been to the study area were selected as participants. The screening questionnaire asked how many times they had been to 10 places in the Santa Barbara area, one of which was the study area; in this way, participants were kept uninformed as to the identity of the study area until they were told after all the sessions were finished. They were paid $10 per session for participating and attended 10 sessions ($100 in total).

2.2. Materials

2.2.1. Study area and routes

We used as the study area the exclusive private residential area of Hope Ranch in Santa Barbara, California (Fig. 2). This area is full of hills and winding roads, and provides few distant landmarks that may be used consistently as orientation clues. Its road signs are small, inconspicuous, and few in number. Hence, accurate orientation on the first visit to the area, if not after several visits, is quite challenging. The area is relatively isolated from other areas and is rarely driven through by most people, so people who are not residents of the area rarely have much experience exploring or even passing through it. We selected two test routes in the area that were both about 2.2 km (1.4 miles) long and had no common segments (see Fig. 2). One route was "U-shaped" (we call it the *U-route*) and the other was roughly straight (we call it the *S-route*). The U-route had some big turns and caused great changes in a traveler's heading; the S-route had some turns but did not cause much change in a traveler's overall heading. We also selected a short route that connected the two (we call it the *connecting-route*). This connecting-route would be used starting in the fourth session of testing to directly expose participants to the spatial relationship of the two test routes, hopefully allowing the routes to be integrated into a common representation.

We chose four landmarks on each route that were visible in both directions of travel, and supported distance and direction estimates covering a wide range (0.5–2.2 km; 6–351° from straight ahead). Landmarks on the U-route were named *Trees, White Gate, Green*
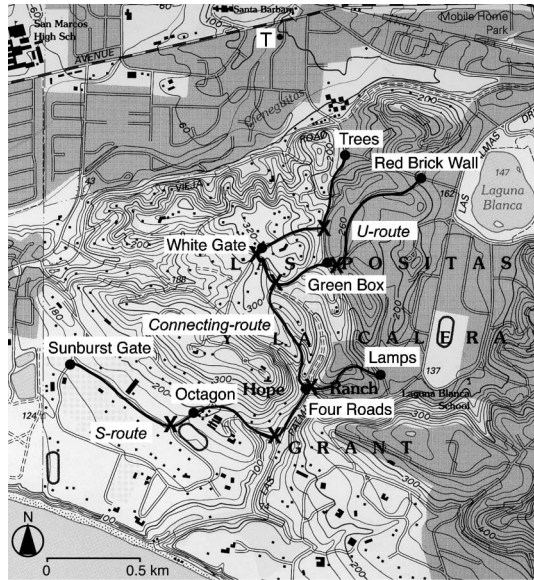
Fig. 2. Map of the study area (Hope Ranch, Santa Barbara, CA). Solid circles show landmarks on the two routes (both about 2.2 km). X's show major turns used in sketch-map analyses and simulations. T (at the top) shows the test location where participants conducted experimental tasks in a car. The connecting-route joins the two test routes. *Source*: 7.5-Minute Topographic Quadrangle, Goleta, CA, 1:24,000, US Geological Survey, 1995.

Box, and *Red Brick Wall*; landmarks on the S-route were named *Sunburst Gate, Octagon, Four Roads*, and *Lamps* (see Fig. 2). These landmarks were pointed out to participants by the experimenter, but they were clearly identifiable in the environment, could be mnemonically named, and posed no difficulty of recognition to participants. At a particular landmark, participants could not see the other landmarks, either on the same route or on the other route. This, together with the lack of distant views and the fact that participants wore a blindfold whenever they were not on a test route, ensured that participants learned the two routes separately, so they would not start to integrate the two routes before traveling the connecting-route during the fourth session.

2.2.2. Direction estimation task

At the center of an 8.5 × 11 in. (21.6 × 27.9 cm) sheet of paper, a circle with a radius of 5 cm was drawn, and the names of a pair of landmarks were shown above it. Participants imagined being at a particular landmark, seated in the car facing straight ahead, as when they had learned the landmark, and drew a line from the center of the circle to indicate the direction to a target landmark.

2.2.3. Distance estimation task

On an 8.5 × 11 in. sheet of paper, a standard line (2 cm) was drawn at the top. This represented a standard distance of 0.6 km, which was shown to participants at the beginning of every session (see Section 2.4). Below the standard line, response lines (23 cm) were drawn. For each pair of landmarks, the names of which were shown on the left, participants marked the response line so that its ratio to the standard line matched the ratio of the

distance between the two landmarks to the standard distance. Route distances (along the test routes) and straight-line distances (across the test routes) were estimated on separate sheets. The response lines were sufficiently longer than the standard line so that participants had plenty of room to overestimate as well as underestimate.

2.2.4. Map sketching task

During even-numbered sessions, participants were given an 8.5×11 in. blank sheet of paper and drew a sketch map of the routes, showing the four landmarks on each route and the shapes of the routes.

2.2.5. Self-report sense-of-direction

Participants filled out the *Santa Barbara Sense-of-Direction* (SBSOD) Scale, which was developed and validated by Hegarty et al. (2002). It consists of 15 statements to which participants express their degree of agreement on a 7-point Likert scale. Seven of the questions are stated positively (e.g., "I am very good at giving directions") and the other eight negatively (e.g., "I very easily get lost in a new city").

2.3. Design

Six participants were randomly assigned to each of four conditions with respect to two experimental factors: (a) whether they traveled each route twice in one direction (unidirectional) or once in both directions (bi-directional); and (b) whether they started from Trees on the U-route and Sunburst Gate on the S-route, or from Red Brick Wall on the U-route and Lamps on the S-route. In each condition, the order of presenting the two routes was counterbalanced: half the participants traveled the U-route first and the other half traveled the S-route first.

2.4. Procedure

On a separate day before the experimental sessions began, the experimenter (the first author) met participants individually in his office and explained the outline of the experiment. Participants practiced estimating directions and distances (both route and straight-line) using places on campus as examples. After practice, the experimenter ensured that participants completely understood what and how they would estimate. They next filled out the SBSOD questionnaire. Finally, participants signed a form promising they would neither personally visit, nor look at a map of, the study area during the period of the study, even if they happened to know where it was. This was important for controlling the amount and method of exposure to the environment.

Participants took part in an experimental session individually, once a week for 10 consecutive weeks. The experimenter drove a car with the participant seated in the passenger seat throughout the session. The experimenter kept the car's speed constant at 30 miles (48 km) per hour throughout the session, with only minimal deviations that were unavoidable. At the beginning of every session, before going to the study area, the experimenter showed participants the standard distance; participants were driven between two streetlights twice in opposite directions and told to memorize how far they were apart for later use in distance estimation. Then participants put on a blindfold and were driven to the study area.

In the first three sessions, participants learned the two routes separately and conducted experimental tasks that assessed their knowledge of each route. On arriving at the starting point of the first route, the experimenter asked participants to remove the blindfold and then drove them along the route without making conversation, so as not to distract their attention. At each landmark, the experimenter stopped the car briefly, stated the landmark's name, and told participants to memorize its name and location. Participants in the unidirectional condition traveled the route twice in the same direction (after the first travel, the experimenter asked them to wear the blindfold, drove them back to the starting point along a circuitous route, and then started the second travel); participants in the bi-directional condition traveled the route first in one direction and then in the opposite direction (i.e., the car turned around and returned). After traveling the route the second time, the experimenter asked participants to wear the blindfold and drove them circuitously to a test location outside the study area (shown in Fig. 2). At the test location, participants removed the blindfold and conducted the tasks designed to test their knowledge of the route they had just traveled. Participants (a) named the four landmarks in order of appearance; and (b) estimated directions, route distances, and straight-line distances between four landmarks (six pairs, presented in random order). In estimating directions, participants in the bi-directional condition were instructed to imagine the direction of the second travel (back to the starting point). When they completed these tasks, the experimenter asked them to wear the blindfold and drove them circuitously to the starting point of the second route. Then they traveled the second route and conducted experimental tasks in exactly the same way as they had for the first route. In the second session, the experimenter also asked participants to draw a sketch map of each route on separate sheets of paper. No feedback was given to participants about their performance on any task.

At the end of the third session, the experimenter also asked participants to guess straight-line distances and directions between landmarks on one route and landmarks on the other route (eight pairs of landmarks, presented in random order). Because participants had not yet been exposed to the connecting-route between the two routes, the experimenter explained that he merely wanted to know how well they could guess the spatial relationship between the two routes. Their performance on guessing in Session 3 provided baselines that we use to interpret participants' performance during subsequent sessions, after which they had been exposed to the connecting-route.

Starting from the fourth session, the experimenter drove participants along the connecting-route and prompted them to integrate the two test routes by having them conduct between-route tasks in addition to within-route tasks. After participants completed the usual tasks for the second route, the experimenter drove them blindfolded to an endpoint of the first route: Red Brick Wall (if they first traveled the U-route) or Lamps (if they first traveled the S-route). Then the experimenter asked participants to remove the blindfold, drove them along the first route and then onto the connecting-route, saying "We now go onto a route that connects the two routes." After driving the connecting-route, the experimenter turned onto the second route, saying "We now go onto the second route," and drove along it to an endpoint: Lamps (if they started from Red Brick Wall) or Red Brick Wall (if they started from Lamps). Then the experimenter drove them blindfolded to the test location along a circuitous route. At the test location, participants removed the blindfold and performed between-route tasks by estimating straight-line distances and directions between landmarks on one route and landmarks on the other route, for the same

Table 2
Timeline of the experiment over 10 sessions

Event/Task order	Session									
	1	2	3	4	5	6	7	8	9	10
Travel U (or S)	*	*	*	*	*	*	*	*	*	*
W/in direction/dist.	*	*	*	*	*	*	*	*	*	*
Map of U (or S)		*		*		*		*		*
Travel S (or U)	*	*	*	*	*	*	*	*	*	*
W/in direction/dist.	*	*	*	*	*	*	*	*	*	*
Map of S (or U)		*		*		*		*		*
Travel connecting				*	*	*	*	*	*	*
B/w direction/dist.			G	*	*	*	*	*	*	*
Integrated map				*		*		*		*

Note. In each session, participants did the tasks with an asterisk (*) from top to bottom. For example, everyone traveled the U- or S-route first in a session and sketched the integrated map last. Half the participants traveled the U-route first, and the other half traveled the S-route first. **G**, guess.

eight pairs of landmarks (in random order) that were used for guessing at the end of Session 3. At the end of each even-numbered session, the experimenter also asked participants to draw a sketch map of the two routes together on the same sheet of paper, showing their spatial relationship. As with the within-route tasks, no feedback was given to participants.

At the very end of the last session, participants were given three questionnaires. First, they filled out the SBSOD questionnaire, as they had before the experiment. Second, they rated the difficulty of each of the experimental tasks in the first and last sessions. Finally, the experimenter asked participants whether they had visited or looked at a map of the study area during the study period. The importance to the research of being forthcoming was stressed, as was the fact that no penalty would follow from an affirmative answer. No participants said they had done either. After these questionnaires were completed, a map of the study area was shown to participants. Sessions 1–3 took about 75 min each; Sessions 4–10 took about 90 min each. The timeline of the experiment is summarized in Table 2.

3. Results

3.1. Aggregate analyses

3.1.1. Names and sequence of landmarks

Participants named landmarks in order of appearance with perfect accuracy; all the participants correctly ordered the four landmarks on both routes in all sessions.

3.1.2. Direction estimates

To examine the accuracy of participants' direction estimates, we analyzed absolute errors. Absolute error is a good index of accuracy in this case, in that it reflects the probability that a particular participant's response falls within a particular range around the correct target (Spray, 1986). An α level of .05 was used for all the statistical tests below, unless otherwise noted.
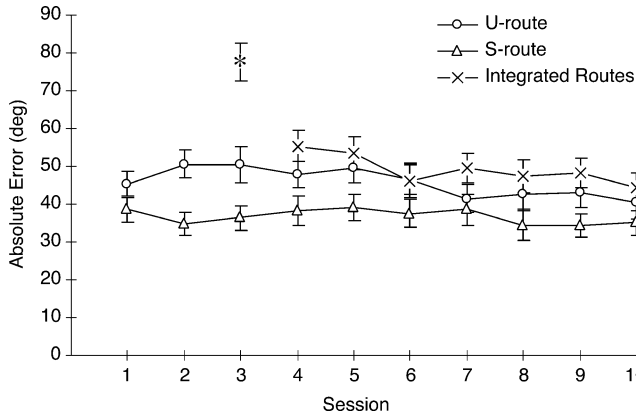
Fig. 3. Developmental curves for mean performance on direction estimates. Points represent the mean absolute error (in degrees) over 24 participants. Vertical lines depict standard errors of the means. ∗ represents the error of the guesses made in Session 3. The errors in the initial session were smaller than 90°. The developmental curve for the U-route showed significant linear and cubic trends; the developmental curve for the integrated routes showed a significant linear trend.

Fig. 3 shows developmental curves for mean performance on direction estimates, aggregated over participants. A repeated measures ANOVA[4] on the mean absolute error across all sessions (Sessions 1–10 for the U- and S-routes; Sessions 4–10 for the integrated routes) revealed a significant difference among errors for the U-route, S-route, and integrated routes, $F(2, 46) = 18.38$, $MSE = 55.18$, $p < .001$; $T^2 = 2.04$, $F(2, 22) = 22.45$, $p < .001$. Post hoc paired comparisons (Bonferroni, $\alpha = .05/3$) showed that the error for the S-route was smaller than that for the U-route and the integrated routes, t's$(23) = -3.84$ and -6.83, respectively, p's $< .001$; the latter two did not differ significantly from each other.

We conducted mixed ANOVAs on mean absolute errors for the U-route, S-route, and integrated routes separately, to examine the effects of (a) the method of learning a route (unidirectional vs. bi-directional); (b) the direction that participants imagined facing at each landmark at the time of estimation (two directions depending on which endpoint of a route they started from); and (c) session number (10 sessions, Sessions 1–10, for the U- and S-routes; 7 sessions, Sessions 4–10, for the integrated routes). The first two factors were between-subject and the third was within-subject. No significant main effects or interactions were found for the between-subject factors. The main effect of session number was statistically significant for the U-route and the integrated routes, $F(9, 180) = 3.08$ and $F(6, 120) = 3.04$, respectively, $MSEs = 112.49$ and 121.44, p's $< .01$. Trend analyses revealed significant linear and cubic trends for the U-route, and a significant linear trend for the integrated routes, Fs$(1, 20) = 4.82$, 7.69, and 7.82, respectively, $MSEs = 375.91$, 63.55, and 191.31, p's $< .05$. These significant trends reflect improvement over time for the U-route and

---

[4] Following a general recommendation (e.g., Girden, 1992), we conducted both the univariate and multivariate tests, each at the .025 level of significance. In each univariate test, use of unadjusted and adjusted degrees of freedom yielded the same result with respect to significance, so we report only the unadjusted univariate tests (this holds for all repeated measures ANOVA analyses below). For significant results in the univariate test, we report the F ratio and MSE. When the results are also significant in the multivariate test, we report Hotelling's T^2 and the F ratio as well.

integrated direction estimates, but no change for the S-route estimates. Also, performance on the U-route actually got a little worse for three sessions before it improved.

We next compared mean absolute errors in the initial session (Session 1 for the U- and S-routes; Session 4 for integration) to 90°, the absolute error expected from chance performance. The mean absolute errors for the initial session on the U-route, the S-route, and the integrated routes were all smaller than 90°, t's(23) = -12.97, -16.35, and -8.41, respectively, p's < .001. The mean absolute error of the guesses made in Session 3 about the directions between landmarks on the two routes was 77.6°, which was near chance but in fact smaller than 90°, $t(23) = -2.53$, $p < .05$. However, this guessing error was significantly larger than initial performance on the U-route, S-route, and integrated routes, t's(23) = 6.21, 6.89, and 4.28, respectively, p's < .001.

3.1.3. Distance estimates

To examine the accuracy of participants' distance estimates, we analyzed correlations between their estimates and the actual distances as a measure of relative accuracy (Montello, 1991, discusses the relative vs. absolute accuracy of distance estimates, including the ambiguity of inferring absolute accuracy from ratio estimation). In analyzing distance correlations, we applied Fisher's r-to-z transformation (but the numbers reported as correlations below are retransformed back into Pearson's r's). To help interpret patterns of estimates for the two routes, it is useful to note that actual route and straight-line distances were almost perfectly correlated (.99) for the S-route but weakly (.15) for the U-route.

Fig. 4 shows the developmental curves for mean performance on five sets of distance estimates, aggregated over participants. A repeated measures ANOVA on the mean correlation across all sessions (Sessions 1–10 for the U- and S-routes; Sessions 4–10 for the integrated routes) revealed a significant difference among the five sets, $F(4, 92) = 249.63$, $MSE = 0.07$, $p < .001$; $T^2 = 35.57$, $F(4, 20) = 177.84$, $p < .001$. Post hoc paired comparisons
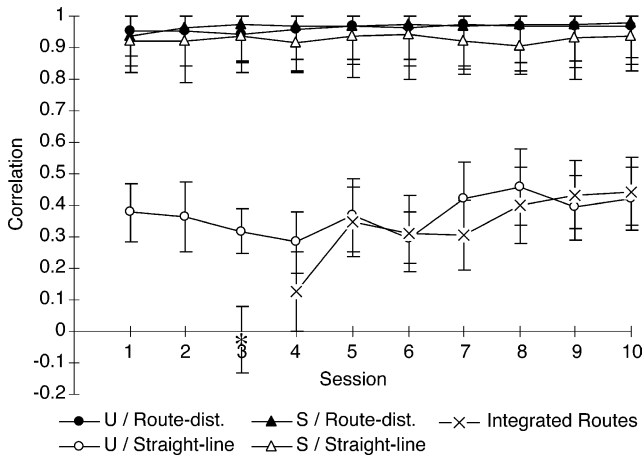

Fig. 4. Developmental curves for mean performance on distance estimates. Points represent the mean correlation (Pearson's r) over 24 participants. Vertical lines depict standard errors of the means. * represents the mean correlation obtained from the guesses made in Session 3. The correlations in the initial session were larger than 0, except for the integrated routes. The developmental curve for the integrated routes showed a significant linear trend.

(Bonferroni, $\alpha = .05/10$) showed that the five sets of correlations were divided into three groups that differed significantly from each other: (a) route distances for the U-route and the S-route; (b) straight-line distances for the S-route; and (c) straight-line distances for the U-route and the integrated routes. That is, distance estimates were extremely metrically accurate, in a relative sense, for route distances along both routes, only slightly less accurate for straight-line distances on the more direct S-route (not surprising considering their near perfect correlation with route distances on the S-route), and quite inaccurate for straight-line distances on the more indirect U-route and across the integrated routes.

For each of the five sets of distance estimates, we conducted a mixed ANOVA on distance correlations to examine the effects of the two between-subject factors (unidirectional vs. bi-directional; the imagined facing direction at each landmark) and one within-subject factor (session number). No significant main effects or interactions were found for any factor. Although the effect of session number did not reach significance for any of the sets of distance estimates, a trend analysis did reveal a significant linear trend of distance accuracy for the integrated routes, $F(1, 20) = 4.48$, $MSE = .35$, $p < .05$.

We next compared distance correlations after the initial session to 0, the correlation expected from chance performance. The correlations for route and straight-line distances for the S-route and U-route were larger than 0, t's$(23) = 22.80$, 4.23, 17.48, and 15.77, respectively, p's $< .001$. The correlation for the integrated routes after Session 4 was not significantly different from 0. The correlation obtained from the guesses made in Session 3 between landmarks on the two routes was $-.03$, which of course was not significantly different from 0.

3.1.4. Sketch maps

To examine the overall metric accuracy of sketch maps, we employed bidimensional regression (Tobler, 1965). We selected six "anchor" points on each route (the four landmarks and major turns[5] shown in Fig. 2), and digitally overlaid the actual map onto a sketch map so that the anchor points on the two maps matched as closely as possible, after translation, scaling, and rotation. Then we calculated a bidimensional correlation coefficient, which gave the degree of correspondence between the two maps ranging from 0 to 1 (the larger, the better correspondence). As in the analyses of distance correlations, Fisher's r-to-z transformation was applied. The focus here was on participants' abilities to depict the spatial layout in map form; this does not assume, however, that participants' internal representations were necessarily map-like (e.g., that their knowledge was stored in the same form as shown on their sketch maps). To help interpret the bidimensional correlations, four sketch maps drawn by four different participants in the last session are shown in Appendix A, with their bidimensional correlations.

Fig. 5 shows developmental curves for mean performance on sketch mapping, aggregated over participants. While the mean correlations for the three maps were high, a repeated measures ANOVA revealed a significant difference among the three, $F(2, 46) = 43.66$, $MSE = 0.08$, $p < .001$; $T^2 = 3.89$, $F(2, 22) = 42.75$, $p < .001$. Post hoc paired comparisons (Bonferroni, $\alpha = .05/3$) showed that the mean bidimensional correla-

---

[5] These major turns were selected because they were at an intersection and caused great changes in a traveler's heading. These turns were clearly identifiable on all the sketch maps drawn by participants. Three of the major turns were at or by a landmark (White Gate, Green Box, Four Roads).
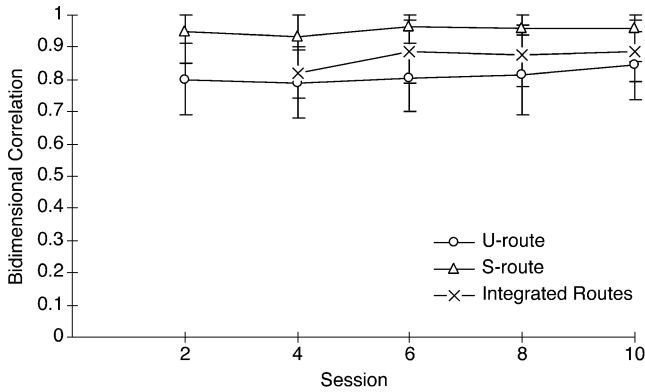
Fig. 5. Developmental curves for mean performance on sketch mapping. Points represent the mean bidimensional correlation over 24 participants. Vertical lines depict standard errors of the means. The bidimensional correlations in the initial session were larger than 0. The developmental curve for the integrated routes showed significant linear and quadratic trends.

tion for the S-route was larger than that for the U-route and the integrated routes, t's$(23) = 7.90$ and 8.34, respectively, p's $< .001$; the latter two did not differ significantly from each other.

For each of the three sets of maps, we conducted a mixed ANOVA on bidimensional correlations to examine the effects of the two between-subject factors (unidirectional vs. bi-directional; the imagined facing direction at each landmark) and one within-subject factor (session number). No significant main effects or interactions were found for the between-subject factors. The main effect of session number was statistically significant for the integrated routes, $F(3,60) = 5.05$, $MSE = 0.07$, $p < .01$; $T^2 = 0.99$, $F(3,18) = 5.96$, $p < .01$. A trend analysis revealed significant linear and quadratic trends, $Fs(1,20) = 6.40$ and 5.49, respectively, $MSEs = 0.09$ and 0.04, p's $< .05$. Thus, sketch maps did not change in accuracy over sessions for the separate routes, but got a little more accurate for the integrated routes, especially between Sessions 4 and 6.

Finally, we compared the accuracy of sketch maps after the initial session to chance performance of 0 on bidimensional correlation. The bidimensional correlations for the U-route, the S-route, and the integrated routes were all larger than 0, t's$(23) = 9.90$, 18.92, and 14.61, respectively, p's $< .001$.

3.1.5. Ratings of task difficulty

After the final session, participants rated how difficult each experimental task had been for each route and the integrated routes in both the first and last sessions, on a scale from 1 (*easy*) to 7 (*difficult*). All tasks were rated as being significantly easier in the last session than in the first session (an average drop in ratings of 2.1). Generally, participants rated recalling the landmarks as very easy (and the easiest of all tasks; an average rating of 2.7 in the first session, and 1.1 in the last session). They rated estimating straight-line distances for the separate routes, and both directions and distances for the integrated routes, as especially difficult (an average rating of 6.4 in the first session, and 4.4 in the last session).

3.2. Individual analyses

The aggregate analyses presented above provide a description of "average" performance for our group of research participants. However, theories of microgenesis are theories of development for individual people. Aggregate analyses may produce a misleading picture of performance, especially if the developmental trends of individuals are not just quantitatively distinct but qualitatively distinct. In addition, our motivation in this research explicitly includes the question of how individual developmental trends compare to each other. In this section, therefore, we report analyses of individual participants' developmental curves. We also examine the relationship of performance to individual differences on self-report sense-of-direction.

3.2.1. Individual differences in accuracy and the developmental pattern

Shifting our focus to individual performance, the existence of large individual differences in the developmental pattern and accuracy of spatial knowledge are revealed. Although individual performance did not change much over the 10 weeks (i.e., did not reveal much "development"), some participants' developmental curves did have a statistically significant trend indicating overall improvement (Table 3).

Although the lack of much improvement over the 10 weeks is consistent with the aggregate analyses, individual analyses revealed large differences among participants in accuracy; that is, individual data points were scattered widely, throughout the experiment, around the mean performance shown in Figs. 3–5. To get a sense of this, we report the range of performance on each task in the first session and in the last session (Table 4). These patterns show that individual differences in accuracy were especially large for tasks that required understanding the shapes of the curved U-route and the integrated routes.

From these analyses, participants' developmental curves can be classified into three groups. Participants in a first group ($n = 4$) acquired very accurate configurational knowledge early on in a new environment and kept doing well throughout the experiment. Participants in a second group ($n = 2$) failed to acquire accurate knowledge of spatial layout and did poorly from beginning to end. Participants in a third group ($n = 18$) were intermediate throughout the experiment. Some of these intermediate participants ($n = 12$) improved monotonically over the course of the experiment on at least one task that required configurational understanding, but the increase in performance was not so large—after the final session, only half of them ($n = 6$) estimated directions as well as participants in the first group had from the beginning (an absolute error below 30°).

Table 3
Number of participants (out of 24) with significant developmental trends indicating overall improvement

Task	U-route	S-route	Integrated routes
Direction	13 (−12.9°)	7 (−15.3°)	8 (−17.5°)
Route distance	7 (.05)	10 (.07)	
Straight-line distance	6 (.23)	5 (.12)	6 (.54)
Sketch map	1 (.04)	1 (.07)	1 (.12)

Note. We had few degrees of freedom for the analyses of sketch-map trends. Numbers in parentheses show the mean increase in performance (final minus first) in terms of the corresponding dependent measures.

Table 4
Ranges of performance in the initial and last sessions

Task	Initial session	Last session
Direction (°)		
U-route	18.3 to 83.8	14.2 to 100.7
S-route	14.3 to 74.0	11.0 to 71.5
Integrated routes	17.9 to 99.6	22.4 to 88.1
Route distance (r)		
U-route	.78 to .99	.56 to 1.00
S-route	.81 to .99	.81 to 1.00
Straight-line distance (r)		
U-route	−.12 to .94	−.17 to .90
S-route	−.62 to .99	−.55 to .99
Integrated routes	−.91 to .85	−.68 to .90
Sketch map (r)		
U-route	.17 to .98	.17 to .97
S-route	.63 to .99	.66 to .99
Integrated routes	.15 to .93	.32 to .98

To illustrate this point, we select one participant from each of the three groups described above, and show their developmental curves for direction estimates, distance estimates, and sketch maps (Fig. 6). Because our longitudinal study was carried out with many replications of task performance by each participant, even single participants provide a relatively reliable sample of evidence.

Participant P1 was very good throughout the experiment, even on the difficult integration tasks. His mean pointing error within the S-route averages less than 10° across all sessions. Even on the U-route, he estimates directions with only about 30° of error after the first session, improves to 20° after the second session, and maintains a mean error of only 17° across all sessions. These errors are remarkably small when one considers that our direction estimation procedure had an inherent measurement error of at least 5° (considering random error introduced by the orientation of the pointing circle, the direction the participant faced while pointing, and so on). In other words, we consider P1's direction estimates on the U-route to have been excellent, and on the S-route nearly as good as theoretically possible. Even his straight-line distance estimates on the curved U-route correlate over .60 in the first session, averaging .73 over all sessions. His sketch maps correlated with the actual layouts at over .90, starting from the first session (see his map in Appendix A). Moreover, P1's ability to integrate the two routes is remarkable. Although he estimates directions between the routes with nearly 90° of error when guessing in Session 3 (chance performance), he estimates them with only about 20° of error as soon as he is driven along the connecting-route in Session 4. His distance estimates across routes correlate .80 after first exposure to the connecting-route. It should be remembered that he performed so well even though he claimed he never visited the area before, and never looked at a map or visited the area during the 10 weeks of the study, other than while he was actually participating.

Participant P17 did poorly throughout the experiment. Her direction estimates were almost at chance level after the first session, and failed to improve or even got worse over the course of the experiment (absolute errors averaging 70° to 90°). As did many other participants, she estimated route distances for both routes and straight-line distances for the
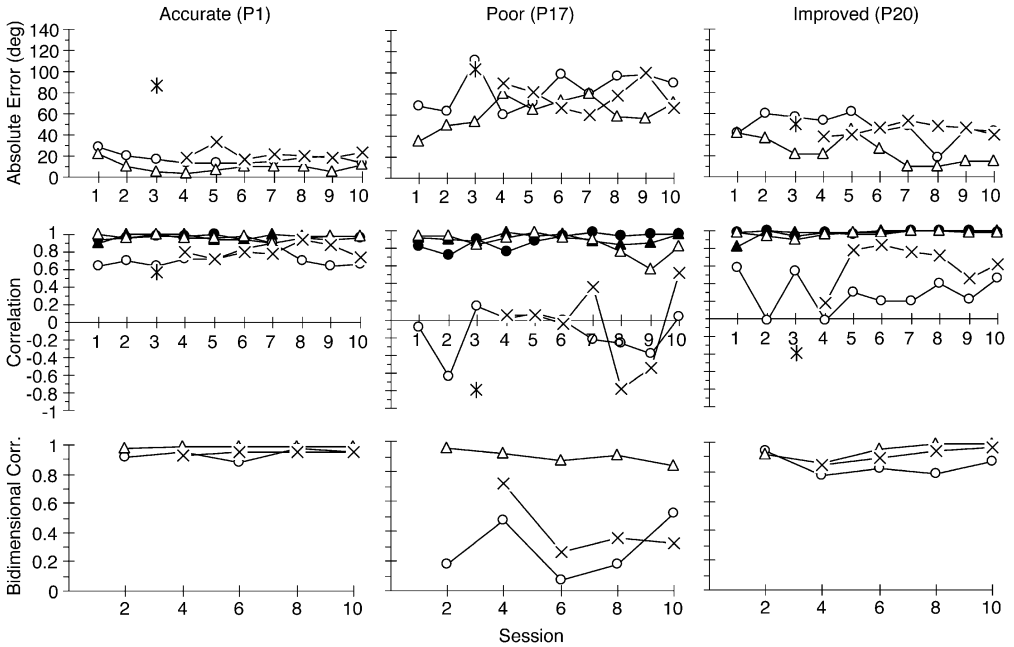
Fig. 6. Developmental curves for three selected participants from three groups (P1, accurate; P17, poor; P20, improved), for direction estimates (mean absolute errors in degrees), distance estimates (correlations between actual and estimated distances, Pearson's r), and sketch-map accuracy (bidimensional regression coefficients). Symbols are the same as in Figs. 3–5, respectively.

S-route quite accurately (correlations > .90). However, her estimates of straight-line distances for the U-route and the integrated routes are quite poor—correlations are near 0 or even negative. P17's sketch maps are also quite inaccurate, particularly her maps of the U-route and the integrated routes; her maps correlate with the actual layout around .30, which is chance performance on bidimensional correlation (see her map in Appendix A).

Finally, participant P20 was not good in early sessions, but improved toward the end of the experiment. His developmental curves for direction estimates and straight-line distance estimates for the U- and S-routes and the integrated routes, and those for sketch mapping for the S-route and the integrated routes, had a significant trend. In particular, his mean pointing error on the S-route was 41.2° in the first session, but decreased to 15.2° in the final session. His distance estimates across routes correlated only .17 in the first session, but correlated .62 in the final session. His sketch map of the integrated routes correlated with the actual layout .84 in the first session, and .96 in the final session.

3.2.2. Relationships with self-report sense-of-direction

Participants assessed their sense-of-direction by filling out the SBSOD questionnaire, before and after the 10 weeks of the experiment (we reversed answers to positively worded questions so that a higher score would indicate a better sense-of-direction). The mean scores on the two administrations did not significantly differ. The correlation between scores on the two administrations (test-retest reliability) was $r = .85$ ($p < .001$); Hegarty et al. (2002) reported a test-retest correlation of .91 over a period of 40 days. Because of this

Table 5
Correlations between self-report sense-of-direction and mean performance across all sessions

Task	U-route	S-route	Integrated routes
Direction	−.43*	−.36	−.45*
Route distance	−.09	.05	
Straight-line distance	.48*	.00	.29
Sketch map	.59**	.32	.32

* $p < .05$.
** $p < .01$.

high correlation, we used the mean of each participant's two administrations as our measure of self-report sense-of-direction.

Table 5 shows the correlations between self-report sense-of-direction and participants' mean performance across all sessions. Negative correlations with pointing errors and positive correlations with distance and bidimensional correlations mean that participants with a better sense-of-direction tended to perform better on the tasks. Participants' self-ratings of sense-of-direction correlate weakly, if at all, with direction estimates on the S-route, route distance estimates on both routes, and straight-line distance estimates and sketch mapping on the S-route and the integrated routes. In contrast, sense-of-direction ratings moderately and significantly correlated with direction estimates on the U-route and between the routes, straight-line distance estimates on the U-route, and sketch mapping on the U-route. We calculated a composite performance score based on these latter measures (after converted into *z*-scores), and found that it correlated .56 ($p < .01$) with sense-of-direction ratings. Fig. 7 shows a scatterplot of this relationship for the 24 participants.

3.3. Sex-related differences

This research was not specifically designed to address sex-related differences in large-scale spatial abilities, but the data showed a pattern (Fig. 8) similar to what has been reported in the literature on environmental spatial cognition (e.g., Montello, Lovelace, Golledge, & Self, 1999; for a discussion of sex-related differences in general spatial abilities, see McGee, 1979). Mean performance by men ($n = 11$) was significantly better than that by women ($n = 13$) on tasks that required accurate survey knowledge of the layout rather than route knowledge: direction estimates for both routes separately and the integrated routes; straight-line distance estimates for the U-route and the integrated routes; sketch mapping for both routes and the integrated routes. On average, men rated their own sense-of-direction as better than did women, but the difference did not reach significance.

4. Simulations

Our aggregate analyses included a comparison of mean performance in the first session to a pure chance level (i.e., an absolute direction error of 90°, a distance correlation of 0, a sketch-map correlation of 0). Performance was better than chance for all tasks, except for distance correlations on the integrated routes after guessing in Session 3 and after first exposure to the connecting-route in Session 4. Comparison to pure chance is not a very stringent test of performance, however; it only tells us that participants had at least some very minimal spatial knowledge. Pure chance performance implies an agent that neither
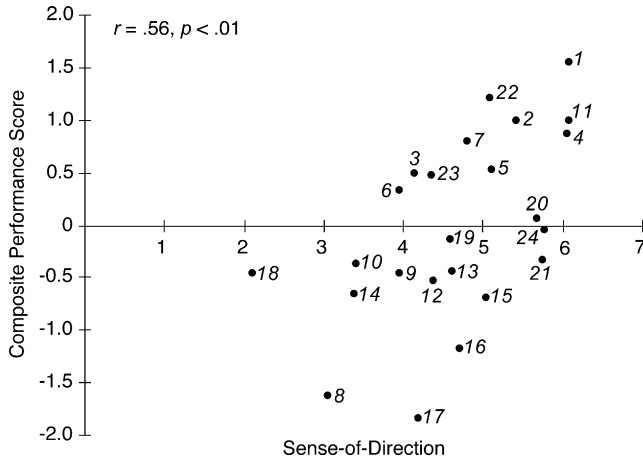
Fig. 7. Relationship between self-report sense-of-direction and a composite survey-task performance score. Sense-of-direction was rated on a 7-point scale. The composite score was based on z-scores for direction estimates on the U-route and the integrated routes, straight-line distance estimates on the U-route, and sketch mapping on the U-route. Both variables were transformed so that larger scores represented better sense-of-direction and task performance. Numbers in italics indicate participant numbers (1–24).
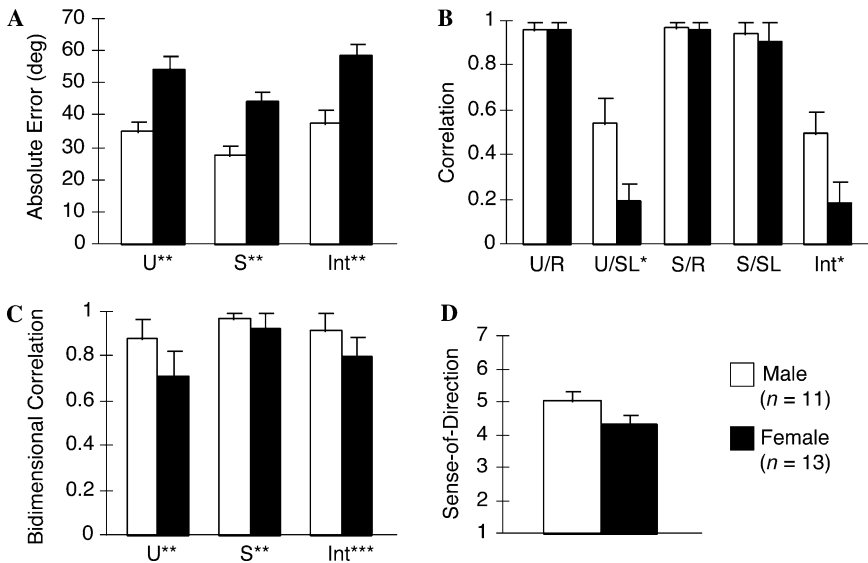

Fig. 8. Means (+SEs) for men and women: (A) direction estimates, (B) distance estimates, (C) sketch mapping, (D) self-report sense-of-direction. U, U-route; S, S-route; Int, integrated routes; R, route distance; SL, straight-line distance. Men did better than women on tasks that required configurational understanding of the routes ($^*p < .05$; $^{**}p < .01$; $^{***}p < .001$).

has any specific knowledge of the spatial layout of the environment nor makes inferences based on even minimal assumptions (or heuristic rules) about layout. But probabilistic models of minimal quantitative knowledge beyond pure chance are possible and
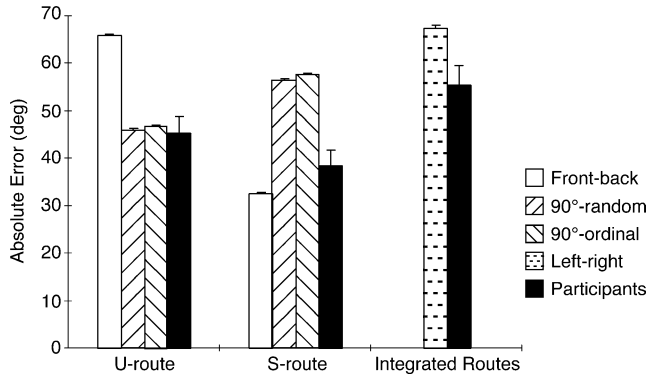
Fig. 9. Comparison of direction estimation by simulations to that by participants. Bars represent mean absolute errors (+*SE*s) in degrees. For the U-route (or the S-route), performance by the front-back model was worse (or better), and that by the two 90°-turn models was not different (or worse), than participants' mean performance. For the integrated routes, performance by the left-right model was worse than participants' performance.

informative. To learn more about the sophistication of spatial knowledge required to attain the particular levels of performance that we observe in our data, we compared performance by participants after the first session to different models of minimal spatial knowledge, by using a Monte Carlo simulation method.[6] See Table 1 for the type of knowledge that is possessed by an agent assumed by each simulation model.

4.1. Simulated direction estimates

Three models for within-route direction estimates and one model for between-route direction estimates (for the integrated routes) were evaluated. Results of the simulations (absolute errors) after 1000 iterations are shown in Fig. 9, along with performance by participants for each route separately and the integrated routes.

4.1.1. Front-back model

This model assumed an agent that knows the order of landmarks but not their distances apart, regards the routes as being roughly straight, and is able to judge correctly whether a target landmark is in front of or behind itself. If a target landmark was in front, the model randomly chose a direction within a 90° cone in front (between 315° and 45°, directions being measured clockwise from the facing direction); if a target landmark was in back, it randomly chose a direction within a 90° cone behind (between 135° and 225°).[7] For the U-route, simulated performance was worse (i.e., larger absolute error) than performance by participants, $z = 11.07$, $p < .001$; for the S-route, simulated performance was better (i.e., smaller absolute error), $z = -3.35$, $p < .001$.

---

[6] Comparison to data from the last session yielded similar results, with some improvements evident over the first session. For the sake of brevity, we report here only comparison with the first session.

[7] Because no significant main effects or interactions had been found for the facing direction at each landmark, we collapsed the two conditions with respect to the facing direction here (as we did with all simulations of direction estimates reported below).

4.1.2. 90°-turn random model

This model assumed an agent that knows the order of landmarks but not their distance apart, and regards the size of the major turns (the turns that we used in the analysis of sketch maps; see Fig. 2 and Footnote 5) as 90° left or right (whether left or right was judged correctly). Also, because direction estimates along a route that has turns requires distance as well as direction knowledge, this model randomly guessed (from a uniform distribution) the lengths of segments between the six points. For the U-route, simulated performance was not significantly different than performance by participants; for the S-route, simulated performance was worse than performance by participants, $z = 12.20$, $p < .001$.

4.1.3. 90°-turn ordinal model

This model was the same as the random model, except that it assumed that the agent additionally has an ordinal-level knowledge of distance. For the U-route, simulated performance was not significantly different than performance by participants; for the S-route, simulated performance was worse than performance by participants, $z = 16.18$, $p < .001$.

4.1.4. Left-right model for the integrated routes

This model simulated between-route direction estimates for the integrated routes in Session 4 (the first session after the connecting-route had been traveled). It assumed an agent that only knows whether one route is located to the left or right of the other route. If a target landmark was on a route to its right, the model randomly chose (from a uniform distribution) a direction between 0° and 180°; if a target landmark was on a route to its left, the model randomly chose a direction between 180° and 360°. Simulated performance was worse than performance by participants, $z = 3.60$, $p < .001$.

4.2. Simulated distance estimates

Two models for distance estimates were evaluated, against performance on both route and straight-line estimation tasks. Results of the simulations (actual–estimated correlations) after 1000 iterations are shown in Fig. 10 (for route distances) and Fig. 11 (for straight-line distances), along with performance by participants.

4.2.1. Two-category model

This model assumed an agent that classifies the three shorter distances into a *short* category and the three longer distances into a *long* category, classifications that were always made correctly. Given that our participants estimated distances by marking response lines on a sheet of paper, a distance in the *short* (or *long*) category means that the model placed a mark randomly on the left (or right) half of the response line. For route distances, simulated performance for both routes was worse than performance by participants, z's $= -11.96$ and -7.63, respectively, p's $< .001$. For straight-line distances, simulated performance for the U-route was better than performance by participants, while that for the S-route was worse, z's $= 7.36$ and -6.07, respectively, p's $< .001$. Applying this model to between-route distance estimates on the integrated routes, simulated performance was much better than performance by participants in the initial session, $z = 16.57$, $p < .001$.
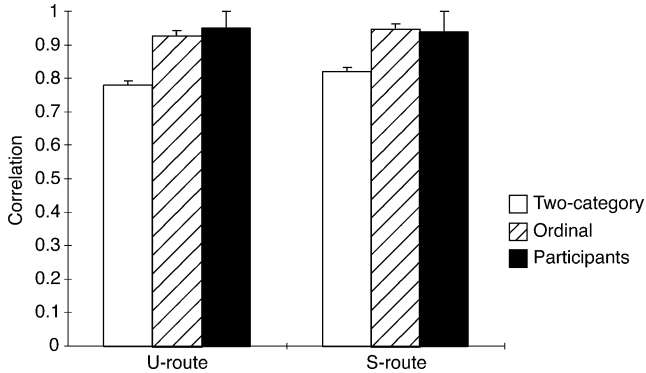
Fig. 10. Comparison of route-distance estimation by simulations to that by participants. Bars represent mean correlations (+*SE*s) in Pearson's *r*. Performance by all route-distance simulation models was worse than participants' mean performance, except that by the ordinal model for the S-route, which was not different from participants' performance.
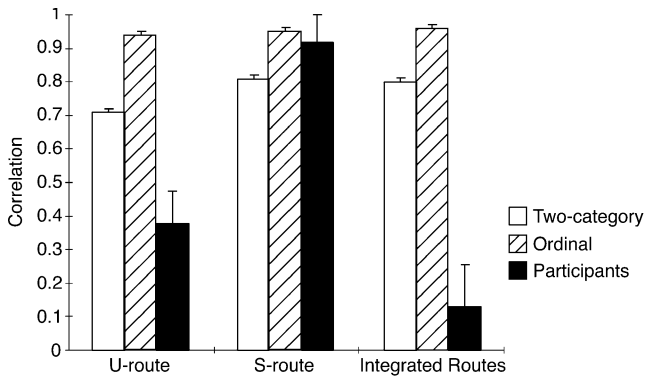

Fig. 11. Comparison of straight-line distance estimation by simulations to that by participants. Bars represent mean correlations (+*SE*s) in Pearson's *r*. Performance by all straight-line distance simulation models was better than participants' mean performance, except that by the two-category model for the S-route, which was worse than participants' performance.

4.2.2. Ordinal model

This model assumed an agent that has an ordinal-level knowledge of distance (i.e., the ratio of estimated distances was not correct but the rank order was correct). For route distances, simulated performance for the U-route was worse than performance by participants, $z = -2.17$, $p < .05$, but that for the S-route was not significantly different than performance by participants. For straight-line distances, simulated performance for both U- and S-routes (but especially the U-route) was better than performance by participants, z's $= 18.37$ and 3.28, respectively, p's $< .001$ and $.01$. Applying this model to between-route distance estimates on the integrated routes, simulated performance was much better than performance by participants in the initial session, $z = 25.25$, $p < .001$.

4.3. Simulated sketch maps

Three models for within-route sketch-mapping accuracy and two models for between-route sketch-mapping accuracy (for the integrated routes) were evaluated. Results of the simulations (bidimensional correlations) after 1000 iterations are shown in Fig. 12, along with performance by participants.

4.3.1. Random 6 model

This model assumed an agent that draws a sketch map of each route by placing six points randomly on an 8.5×11 in. (21.6×27.9 cm) sheet of paper. Simulated performance for both U- and S-routes was worse than performance by participants, z's $= -11.99$ and -28.91, respectively, p's $< .001$. It should be noted, however, that random sketch maps still correlate about .3 to .4 with the actual configuration. This suggests that bidimensional correlations, based as they are on rotated, translated, and scaled configurations, have an expected value much greater than 0 even under chance conditions. We comment on this further in Section 5 below.

4.3.2. Straight-line model

This model assumed an agent that places six points randomly on the paper but in a straight line. Simulated performance for both U- and S-routes (but especially the U-route) was worse than performance by participants, z's $= -49.39$ and -6.79, respectively, p's $< .001$.

4.3.3. 90°-turn ordinal model

In this model, the 90°-turn ordinal model used in the simulation of direction estimates was applied to sketch mapping. Simulated performance for the S-route was worse than performance by participants, $z = -57.79$, $p < .001$, but that for the U-route was not significantly different than performance by participants.
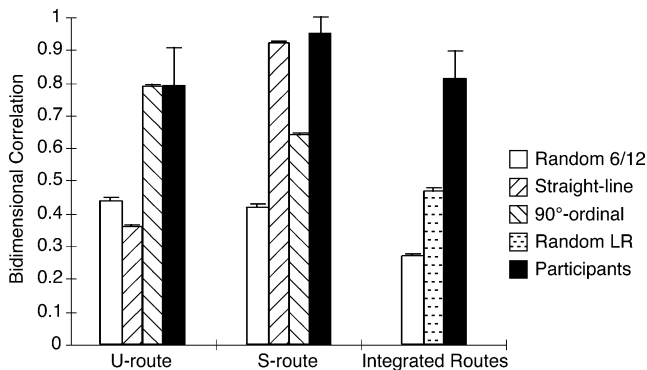

Fig. 12. Comparison of sketch mapping by simulations to that by participants. Bars represent mean bidimensional correlations (+*SE*s). (Open bar = random 6 model for the U- and S-routes; random 12 model for the integrated routes.) Performance by all sketch-mapping simulation models was worse than participants' mean performance, except that by the 90°-ordinal model for the U-route, which was not different from participants' performance.

4.3.4. Two models for the integrated routes

Sketch mapping of the two integrated routes drawn together in Session 4 (after first traveling the connecting-route) was simulated by two models: a *random 12 model* and a *random LR model*. The first model generated 12 points randomly within the space of the map drawing. The second model generated six points randomly on the right half, and another six points randomly on the left half. Simulated performance by the two models was quite a bit worse than performance by participants, z's $= -28.74$ and -17.68, respectively, p's $< .001$.

5. Discussion

Detailed analyses of the rich empirical data from the present study reveal various aspects of people's spatial knowledge acquired directly in a large-scale environment. In short, the process of spatial knowledge acquisition does not proceed in the way posited by Siegel and White's (1975) framework. This dominant framework does insightfully specify types of knowledge that people may have about spatial environments: knowledge of landmarks, knowledge of routes, and knowledge of configurations. What is questionable about the framework is not so much the content, but the structure and developmental course of spatial knowledge it proposes. Also, neither the dominant nor the continuous framework adequately takes into account the large individual differences among people in their accuracy and developmental pattern of spatial knowledge. We consider the implications of our results for theories of the microgenesis of spatial knowledge in environments.

5.1. Acquisition of metric spatial knowledge within the separate routes

5.1.1. First exposure

Participants' perfect performance on naming landmarks in order of appearance shows that they acquired minimal "landmark knowledge" and "route knowledge" (i.e., the identities and sequence of landmarks) right away, upon first exposure to a complex new environment. This empirical finding is evidence against the dominant framework's separation of the acquisition of landmark knowledge from route knowledge, putting the former before the latter in a sequence. Instead, the two types of knowledge are acquired at the same time, more or less immediately.

As discussed in Section 1, the dominant framework uses the term *route* to mean a sequence of landmarks, the space between the landmarks being relatively empty or unscaled (Siegel & White, 1975, p. 29). That is, at least at an early stage of exploration in a new environment, the dominant framework argues that people have no metric knowledge, only connectivity and sequence knowledge. On the contrary, our results indicate that participants quickly acquire some metric knowledge of environmental layout, as evidenced by the fact that they estimate directions and distances, and draw sketch maps, more accurately after first exposure to the routes than would be expected by pure chance alone.

5.1.2. Final exposure and developmental trends

Turning now to performance after the final of the 10 sessions, we find that participants on average show surprisingly little improvement in performance beyond that found after the first session. Changes in direction estimates on the U-route do show significant linear and cubic trends across the 10 sessions, though the total variation in mean performance is only about 10° across the sessions. Changes on the S-route do not reveal any significant trends.

Participants estimated route distances quite accurately after the final session. Given that estimation of route distances is so accurate even after the first session, there is no room for significant improvement over the course of the study, and we find no significant developmental trends on route distance estimation.

For straight-line distances, the correlation of estimated with actual distances is again very high for the S-route. This is not particularly revealing, given that route and straight-line distances are so highly correlated on the S-route. In contrast, straight-line distance estimates on the U-route, which are correlated .38 with actual distances after the first session, correlate only .42 after the final session.

For sketch mapping, mean bidimensional correlations with the actual layouts of the two routes are still quite large after the final session, but given how high they are after the first session, there is no average developmental trend over the 10-week course of the study.

5.1.3. Simulations

The fact that participants learn more than connectivity and sequence after first exposure to the routes is not to say that participants immediately acquire an extremely precise metric understanding of the layout of the environment. Results of the Monte Carlo simulations provide some insight into the nature of people's spatial knowledge. Results from the sketch-mapping simulations show that participants clearly know more than the identities of landmarks, and do not simply locate points on paper randomly or based on crude heuristic rules. Results from the direction simulations show that participants do not simplify the task by heuristic rules of front-back or 90° turns. They pay attention to the curved shape of the U-route with its big turns, and the nearly straight shape of the S-route.

Results from the distance simulations show that the precision of participants' knowledge of route distances is beyond the categorical short-or-long level; it is at the ordinal level after the first session and better than that after 10 sessions. The fact that the relative metric accuracy of participants' performance (as measured by a product-moment correlation) is equal to or better than performance by the ordinal simulation suggests that participants are sensitive to the ratios of route distances (i.e., metricity), not just mere rank order. As for straight-line distances, the average participant does not acquire the spatial knowledge of the U-route after first exposure, and even after 10 sessions, to support very accurate straight-line distance estimation, not even at the categorical level.

In sum, participants on average reveal sensitivity to metric properties of individual routes essentially right away, after the first session. But even after 10 sessions, performance shows surprisingly little improvement in the absence of specific feedback or instruction, even on tasks where initial performance is poor enough to leave a great deal of statistical room for improvement. However, we discuss below the fact that a disaggregate analysis of individual participants sheds important additional light on conclusions about the accuracy and precision of spatial knowledge acquired during microgenetic development.

5.2. Integration of the two routes

5.2.1. First exposure

As part of Session 4, participants were driven along a short route connecting the U- and S-routes. This provided sufficient information for participants to deduce the spatial relationship of the two test routes to each other; that is, it potentially allowed participants to integrate their knowledge of the two separate routes, an important achievement in theories

of spatial learning. The average participant can do that, to some degree, after first being exposed to the connecting-route. Mean pointing error between landmarks on the two routes is significantly better than chance performance and guessing performance. In contrast, mean distance correlation is not significantly better than chance or guessing.

The sketch maps drawn after Session 4 correlate a robust and significant .82 with the actual layout of the two routes. As discussed in Section 4.3 on simulated sketch mapping, it is important to note that bidimensional correlations based on random configurations of points are still well above 0 (as high as .4). Also, bidimensional correlations reflect direction as well as distance information; thus, the small distance correlation is not necessarily inconsistent with the high bidimensional correlation. Participant P18's map (shown in Appendix A), for instance, depicts the configuration of the two routes only roughly, although it correlates .79 with the actual layout.

5.2.2. Final exposure and developmental trends

Compared to within-route tasks, performance on between-route tasks shows more improvement over the course of the study, even though it is based on only seven sessions. After the final of the 10 sessions, we find that participants understood the layout of the integrated routes quite a bit better than after the first session of exposure to the connecting-route. Both mean pointing error and mean distance correlation between landmarks on the two routes are significantly better than chance performance, guessing performance, and performance after first exposure. These improvements over the course of the study are reflected in significant linear trends in direction and distance estimates for the integrated routes. Sketch maps drawn after the final session are bidimensionally correlated .88. The linear and quadratic trends in bidimensional correlations were statistically significant, showing that the sketch-map accuracy increased first and then leveled off.

5.2.3. Simulations

Results from the between-route direction and sketch-mapping simulations show that participants know more than the identities of landmarks and the general relationship of which route was on which side (left or right). In contrast, participants' knowledge of direct distances between landmarks on the two routes does not even reach the categorical level, in both initial and final sessions.

We can conclude that the average participant was somewhat able to integrate the two separately learned routes after a single drive along the connecting-route, although this was clearly a challenging task that was not performed very well. In particular, the integration experience that we gave to participants allowed them to approximately orient their representations of the two routes to each other, but it allowed almost no scaling of the straight-line distances between the two.

5.3. Individual differences: Aggregate versus individual data

The analyses of mean performance aggregated over the 24 participants enable us to discuss the nature of people's metric spatial knowledge. In particular, very little developmental change is observed across the 10 weeks of the study—mean performance is either very good right from the beginning (e.g., route-distance estimates) or mediocre without much improvement (e.g., direction estimates on the U-route). These results might lead us

to conclude that participants in this experiment failed to acquire accurate metric knowledge of the routes within the 10-week period, except perhaps for distance estimates along the routes.

However, when we perform disaggregate analyses on the 24 participants' individual developmental curves, a rather different picture emerges: Large individual differences in the accuracy and developmental pattern of spatial knowledge exist. Individuals differ strikingly in their levels of performance and their tendency to improve (or even deteriorate) across the course of the experiment. Some participants were quite accurate after the first session or two and maintained a high level of performance throughout the experiment. Others were quite poor early and never showed much improvement, even after 10 sessions, totaling something like 12–14 h of exposure to the routes. Still others were intermediate in their performance for most of the experiment's sessions. Half of the 24 participants showed enough improvement in performance across the 10 weeks to result in monotonic trends in their estimates, but these trends were not dramatic for the most part (which in fact agrees with the aggregate analyses that suggest stability more than developmental change). To a surprisingly large degree, good performers are good and poor performers are poor, from beginning to end. These individual variations were especially great on the tasks that required understanding the curved shape of the U-route and integrating the two routes. Had the present data been analyzed only at the aggregate level, these interesting and important differences would not have been brought to light.

It is also noteworthy that performance coheres rather strongly across tasks. Landmark sequence recall on both routes, all other estimates on the nearly straight S-route, and route-distance estimates on the U-route were performed quite well by almost all participants, even from early sessions. They might be called "route tasks" (as discussed, e.g., by Montello et al., 1999). Distinct from these route tasks, participants tended to perform consistently on estimates that required (at least approximate) quantitative comprehension of the shape of the curved U-route and the integrated routes—direction, straight-line distance, sketch mapping. These tasks, which might be called "survey tasks," apparently tap into a common set of internal representations and abilities, largely distinct from those involved in route tasks.

The validity of the distinction between route and survey tasks is further supported by the patterns of correlations that we find between participants' self-report sense-of-direction and their performance on the various spatial estimation tasks. SOD ratings correlate robustly with performance on survey tasks, but weakly, if at all, with that on route tasks, which is in line with Hegarty et al.'s (2002) findings. Furthermore, the validity of the route-survey distinction is echoed in the pattern of participants' ratings of task difficulties, and in the pattern of sex-related differences in spatial estimation. That is, men tend to outperform women on survey tasks but not route tasks (see Montello et al., 1999).

5.4. Theoretical implications

5.4.1. Spatial microgenesis: The dominant or the continuous framework?

The dominant framework of Siegel and White (1975) proposes that adults visiting a new place (such as a new city) will acquire spatial understanding of the place according to a microgenetic progression from landmark knowledge (knowledge of noticeable features without locational information) to route knowledge (knowledge of nonmetric networks of landmark nodes linked by travel connections) to survey knowledge (knowledge

of the metrically scaled, two-dimensional layout of landmarks and routes).[8] Although Siegel and White do not specify a time frame for this progression, it is apparently intended to be on the order of months or years. Our results clearly argue against this framework. In particular, some of our participants acquired route and even survey knowledge after the first session of exposure to the test environments; in the case of at least a couple participants, this early survey knowledge was even quite precise. Conversely, other participants never acquired particularly good survey knowledge, even after 10 sessions over more than two months of time, totaling some 12 or more hours of exposure. It appears that the developmental progression from route to survey knowledge proposed by Siegel and White, and promoted by many other researchers, is not a *progression* after all. We find instead that the distinction between route and survey knowledge is more correctly characterized as an individual difference *trait*, as evidenced by the large differences in the accuracy of configurational understanding of the routes and the generally stable performance across sessions.

Montello's (1998) continuous framework accounts better for our results, but it falls short to some extent too. The continuous framework is apparently correct in positing that metric knowledge of layout begins to be acquired more or less immediately, upon first exposure to a new place. However, this occurs only for some individuals. Other individuals do not acquire survey knowledge early; in fact, they may never achieve such an understanding of environmental layout. So even though the continuous framework stresses the importance of large individual variations, it still does not do full justice to the significant qualitative as well as quantitative nature of these variations. The continuous framework appears to be a developmental theory for people with a good sense-of-direction.

Just as some people can grasp the layout of a single curved route (like our U-route) quite quickly, and others cannot, some people can apparently grasp the integrated relationship of separately learned routes fairly quickly when they are given adequate information about the relationships of the routes. But even our best performing participants found this task somewhat more difficult than grasping the layout of a single route. We find this to be consistent with the idea, found in both the dominant and continuous frameworks, that the integration of separately learned places and routes is a difficult but possible achievement in the microgenesis of environmental knowledge. However, our data do not tell us whether this integrated knowledge exists as a stored internal representation of the two routes combined into a single representation (an "integrated cognitive map") or as two representations that remain essentially separate but can be integrated in working memory when a between-route task requires it. That is, we cannot tell whether between-route tasks depend on a stored representation of the two routes or an inference process that occurs at the time of task administration (this issue is discussed at length by Montello, Waller, Hegarty, & Richardson, 2004).

---

[8] As we discussed in Section 1, important characteristics of large-scale environments that are dealt with by Siegel and White (and other researchers, including Ittelson) are that a person cannot see the layout of the space in its entirety from a single viewpoint, and that the person can experience the space by moving through it, not necessarily that there are multiple routes and landmarks in the space. So, although the two routes selected in our study constitute a subset of the study area, we believe that it is a legitimate realization of the environment as described by the Siegel and White framework. If the space gets much larger, spatial knowledge is usually acquired from indirect sources, such as maps.

5.4.2. What is route and survey knowledge?

We believe that these results and the conceptual considerations they engage lead to fundamental questions about the concepts of route and survey knowledge. As defined by the dominant framework, route knowledge contains only sequences of landmarks along a route, something like a one-dimensional chain of landmarks and actions. However, as shown above, at least some people's knowledge of a route is partially metrically scaled (i.e., contains quantitative information), even after very little exposure to the environment. Even though routes that are curved like our U-route are topologically equivalent to straighter routes, people with a good sense-of-direction are sensitive to configurational differences between the two routes even after brief exposure. This suggests that metric knowledge can follow quickly after ordinal knowledge (i.e., sequential route learning) or even develop concurrently with it. Namely, people's spatial knowledge is more than an ordered list of landmarks; it contains "spatiality" from an early stage of microgenetic development. One need not undervalue what Tolman (1948) meant by the term *cognitive maps*, at least not for humans.

It appears therefore that the idea of ordinal route knowledge (in the sense that it contains only ordered sequences of landmarks) being a necessary precursor to metric survey knowledge is incorrect. In fact, we believe that it is reasonable to have metric knowledge (in the sense discussed in Footnote 2) without having accurate ordinal knowledge. The idea that ordinal knowledge is a necessary precursor to metric, as found in the dominant framework, was directly influenced by Piaget's theories of spatial cognitive ontogenesis. He in turn was inspired by the mathematical notion of a series of geometries, each preserving the properties possessed by less "sophisticated" geometries, from topological to projective to affine to metric (e.g., Van Fraassen, 1985). Similarly, the measurement levels of Stevens (1951) are progressive like this in that the information expressed by a nominal variable is also expressed by an ordinal variable, which in turn expresses information that is also expressed within a metric (interval or ratio) variable.

But we do not believe that this hierarchical progressiveness is necessarily a feature of the microgenesis of mentally represented spatial knowledge. Consider a set of distance estimates of routes of varying lengths, including some that are nearly the same length. A person's estimates could be rather impressive in a metric sense because they are accurate (whether relatively or absolutely) to within 5–10% of the quantitative length of the routes. On the other hand, the same person's estimates could be rather mediocre in an ordinal sense, because his or her knowledge is not precise enough to differentiate routes of nearly the same length (say, 245 m vs. 260 m). The person's metric accuracy clearly suggests that they acquired more than ordinal information, but that does not mean that the person has perfect ordinal knowledge. In fact, we find a pattern like this in our data. During the first session, only six participants correctly rank ordered all six route distances for the U-route, and only one participant did for the S-route. That is, the rank-order correlation between actual and estimated route distances was less than 1.0 (around .90). The ordinal simulation of route distances outperforms this in terms of a rank-order correlation, because it is designed to have perfect ordinal accuracy. However, performance by the ordinal simulation is worse than performance by our participants in terms of relative metric accuracy (product-moment correlation). We thus conclude that it is incorrect to posit that metric spatial knowledge has been acquired only if it can first be shown that accurate ordinal knowledge has been acquired. (White, 1983, discussed the tendency of developmental theories to assume "recapitulation" between different types of development and to force facts into an a priori idea of what development is, without much formal evidence.)

Likewise, the notion of metric survey knowledge needs refinement. Apparently some people can and do acquire surprisingly accurate metric knowledge, including knowledge of relations between places that were not directly traveled. This is not to say, however, that the metric knowledge that people (even "hawkeyes" with an excellent sense-of-direction) acquire is necessarily metric in the mathematical sense of obeying the metric axioms, as in Footnote 2, or very precise, as discussed above. We argue that the metric knowledge contained in people's internal representations of spatial environments is quantitative but *approximate*. A certain level of approximation is undoubtedly desirable in the face of the limited cognitive capacity of humans. Researchers in computer and information science, in particular, have been interested in modeling qualitative knowledge and reasoning for several years (e.g., Forbus et al., 1991). When qualitative knowledge includes approximate or vague quantitative information, it may be called *qualitative metric knowledge* (e.g., Frank, 1996). We believe that spatial knowledge of the environment is qualitatively metric, which is critically different than saying it is nonmetric (i.e., ordinal). Nonmetric knowledge would contain no knowledge of distances or directions as quantities. This is clearly not the case, even at very early stages of learning. An important goal for spatial-cognition researchers is to determine the precision of spatial knowledge learned in different ways after different amounts of exposure (e.g., the precision of directional knowledge is considered by Montello & Frank, 1996; Yeap & Jefferies, 2000, explicitly acknowledge this characteristic of qualitative metricity in their computational model).

5.4.3. Development, training, and motivation

It is quite notable that our participants showed so little improvement with repeated exposure alone. This is, on the surface, inconsistent with the oft-reported *power law of practice*, that people's performance on many tasks (ranging from typing to solving mathematics problems) improves rapidly at first and then gradually levels off. It has also been found that practice alone is not enough to ensure learning, however. For effective learning, timely and informative feedback, rather than mere repetition of an activity, is necessary (e.g., see an early study by Thorndike, 1931; see also Ericsson, Krampe, & Tesch-Römer, 1993). Thus, it is interesting to ask whether we can train people to have an accurate "cognitive map" of the environment, and if so, how. What are good strategies to be taught, when and in what form should feedback be provided, and how should the nature of strategies and feedback be adapted to people's traits (e.g., see research by Cornell, Heth, & Rowat, 1992, on route learning; and by Thorndyke & Stasz, 1980, on map learning)?

Concerning possible strategies, we asked participants, at the very end of the experiment, what kinds of environmental features they paid attention to while traveling in the study area (remember that the study area was hilly and had almost no distant views available). Four participants said they had used the position of the sun as an orientation clue, but their performance was not consistently good or poor. Thus, although it has been reported that skilled navigators use such celestial clues (e.g., Gladwin, 1970), not many people like our participants seem to voluntarily seek to locate the sun in the sky for orientation and navigation. The effectiveness of training culturally "modern" people from technologically developed societies to use the position of the sun is an intriguing area for research.

Another issue that needs further investigation concerns participants' motivation. Participants knew before the experiment what they would be asked to do, and so in that sense learning was intentional rather than incidental. But they might not have tried to fine-tune their estimates once they formed a certain "image" of the layout in their minds. Overall lack of improvement and the stable route-survey distinction might indicate that "poor surveyors" are not naturally motivated to learn environments in terms of configurations or layout (in fact, one of the SBSOD questions was, "It's not important to me to know where I am"). Ericsson et al. (1993) stressed the importance of being motivated to improve in deliberate practice. Thus it would be interesting to pose a task in a way that would increase participants' motivation. For example, one could have participants choose landmarks that they think would serve as good navigation clues, and examine their spatial knowledge using those landmarks. Or one could have participants actively
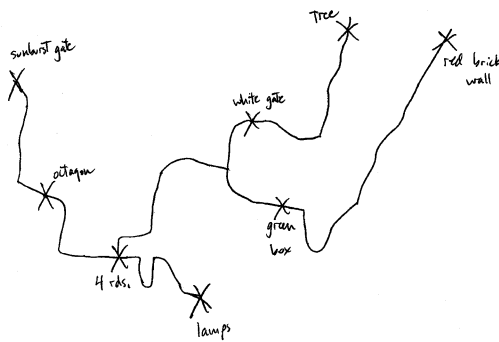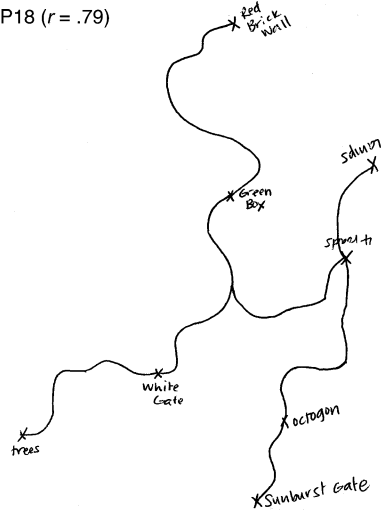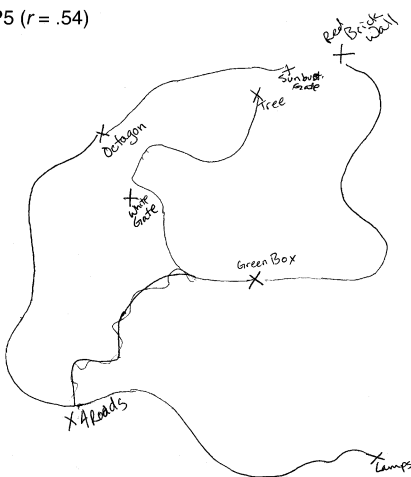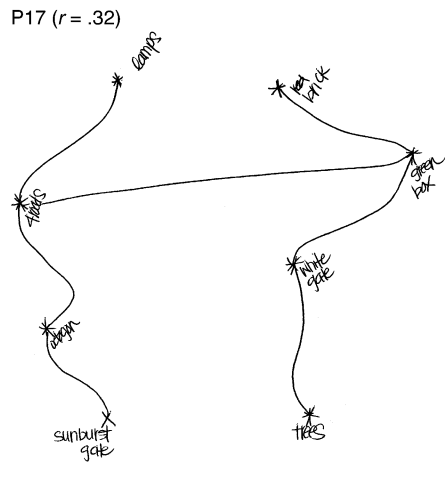

Fig. 13. Sample sketch maps drawn by four different participants in the last session, with their respective bidimensional correlations.

make navigational decisions, such as finding a new route in the study area. Of course, participants could be given monetary rewards for correct performance.

A final issue concerns the mode by which one experiences the environment. In our experiment, participants were driven in a car. The differences between active and passive learning, and between walking and driving, deserve further research. Waller, Loomis, and Steck (2003) and Waller, Loomis, and Haun (2004) found that vestibular sensing of accelerations in a car (inertial information) contributed minimally to spatial learning in large-scale environments, but that the additional proprioceptive information provided by walking (such as kinesthesis and efference copy) did facilitate environmental learning.

6. Conclusion

In conclusion, we believe that models of spatial learning, including explicit computational models (e.g., Chown, Kaplan, & Kortenkamp, 1995; Kuipers, 2000; Yeap & Jefferies, 1999), must be able to account for people who perform at high levels, sometimes astonishingly high levels, not just people who are average or even poor. In that respect, the most important implication of our research may be that some people can and do acquire surprisingly accurate metric knowledge, even relatively quickly, including relational knowledge about places between which they have not directly traveled.

Appendix A

These sample sketch maps (Fig. 13) were drawn by four different participants in the last session; their respective bidimensional correlations are shown. They exemplify how the degree of correspondence between sketch maps and the actual layout changes as the bidimensional regression becomes smaller. The remarkable accuracy of participant P1's sketch map ($r = .95$) is evident when it is compared to the actual map of the study area in Fig. 2. At the same time, it should be noted that all the maps accurately show qualitative properties of the routes such as the names and sequences of landmarks.

References

Aguirre, G. K., & D'Esposito, M. (1997). Environmental knowledge is subserved by separable dorsal/ventral neural areas. *Journal of Neuroscience, 17*, 2512–2518.

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.

Appleyard, D. (1970). Styles and methods of structuring a city. *Environment and Behavior, 2*, 100–117.

Beck, R., & Wood, D. (1976). Comparative developmental analysis of individual and aggregated cognitive maps of London. In G. T. Moore & R. G. Golledge (Eds.), *Environmental knowing: Theories, research, and methods* (pp. 173–184). Stroudsburg, PA: Dowden, Hutchinson & Ross.

Blades, M. (1991). Wayfinding theory and research: The need for a new approach. In D. M. Mark & A. U. Frank (Eds.), *Cognitive and linguistic aspects of geographic space* (pp. 137–165). Dordrecht, The Netherlands: Kluwer.

Burroughs, W. J., & Sadalla, E. K. (1979). Asymmetries in distance cognition. *Geographical Analysis, 11*, 414–421.

Byrne, R. W. (1979). Memory for urban geography. *Quarterly Journal of Experimental Psychology, 31*, 147–154.

Chase, W. G., & Chi, M. T. H. (1981). Cognitive skill: Implications for spatial skills in large-scale environments. In J. H. Harvey (Ed.), *Cognition, social behavior, and the environment* (pp. 111–136). Hillsdale, NJ: Erlbaum.

Chown, E., Kaplan, S., & Kortenkamp, D. (1995). Prototypes, location, and associative networks (PLAN): Toward a unified theory of cognitive mapping. *Cognitive Science, 19*, 1–51.

Clayton, K., & Woodyard, M. (1981). The acquisition and utilization of spatial knowledge. In J. H. Harvey (Ed.), *Cognition, social behavior, and the environment* (pp. 151–161). Hillsdale, NJ: Erlbaum.

Cornell, E. H., Heth, C. D., & Rowat, W. L. (1992). Wayfinding by children and adults: Response to instructions to use look-back and retrace strategies. *Developmental Psychology, 28*, 328–336.

Cox, K. R., & Golledge, R. G. (Eds.). (1969). *Behavioral problems in geography: A symposium*. Evanston, IL: Northwestern University Press.

Devlin, A. S. (1976). The "small town" cognitive map: Adjusting to a new environment. In G. T. Moore & R. G. Golledge (Eds.), *Environmental knowing: Theories, research, and methods* (pp. 58–66). Stroudsburg, PA: Dowden, Hutchinson & Ross.

Downs, R. M., & Stea, D. (1973). Cognitive maps and spatial behavior: Process and products. In R. M. Downs & D. Stea (Eds.), *Image and environment: Cognitive mapping and spatial behavior* (pp. 8–26). Chicago: Aldine.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*, 363–406.

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

Forbus, K. D., Nielsen, P., & Faltings, B. (1991). Qualitative spatial reasoning: The CLOCK project. *Artificial Intelligence, 51*, 417–471.

Frank, A. U. (1996). Qualitative spatial reasoning: Cardinal directions as an example. *International Journal of Geographical Information Systems, 10*, 269–290.

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

Girden, E. R. (1992). *ANOVA: Repeated measures* (Sage University Paper Series on Quantitative Applications in the Social Sciences, Series No. 07–084). Newbury Park, CA: Sage.

Gladwin, T. (1970). *East is a big bird*. Cambridge, MA: Harvard University Press.

Golledge, R. G. (1978). Learning about urban environments. In T. Carlstein, D. Parkes, & N. Thrift (Eds.), *Timing space and spacing time* (Vol. 1, pp. 76–98). London: Arnold.

Golledge, R. G. (1991). Cognition of physical and built environments. In T. Gärling & G. W. Evans (Eds.), *Environment, cognition, and action: An integrated approach* (pp. 35–62). New York: Oxford University Press.

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

Golledge, R. G., Smith, T. R., Pellegrino, J. W., Doherty, S. E., & Marshall, S. P. (1985). A conceptual model and empirical analysis of children's acquisition of spatial knowledge. *Journal of Environmental Psychology, 5*, 125–152.

Golledge, R. G., & Spector, A. N. (1978). Comprehending the urban environment: Theory and practice. *Geographical Analysis, 10*, 403–426.

Hanley, G. L., & Levine, M. (1983). Spatial problem solving: The integration of independently learned cognitive maps. *Memory & Cognition, 11*, 415–422.

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

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.

Herman, J. F., Blomquist, S. L., & Klein, C. A. (1987). Children's and adults' cognitive maps of very large unfamiliar environments. *British Journal of Developmental Psychology, 5*, 61–72.

Herman, J. F., Kail, R. V., & Siegel, A. W. (1979). Cognitive maps of a college campus: A new look at freshman orientation. *Bulletin of the Psychonomic Society, 13*, 183–186.

Hirtle, S. C., & Hudson, J. (1991). Acquisition of spatial knowledge for routes. *Journal of Environmental Psychology, 11*, 335–345.

Holding, C. S., & Holding, D. H. (1988). Acquisition of route network knowledge by males and females. *Journal of General Psychology, 116*, 29–41.

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

Klatzky, R. L., Loomis, J. M., Golledge, R. G., Cicinelli, J. G., Doherty, S., & Pellegrino, J. W. (1990). Acquisition of route and survey knowledge in the absence of vision. *Journal of Motor Behavior, 22*, 19–43.

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. (2000). The spatial semantic hierarchy. *Artificial Intelligence, 119*, 191–233.

Kuipers, B. J., & Levitt, T. S. (1988). Navigation and mapping in large-scale space. *AI Magazine, 9*(2), 25–43.

Landau, B., Spelke, E., & Gleitman, H. (1984). Spatial knowledge in a young blind child. *Cognition, 16*, 225–260.

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

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

Loomis, J. M., Klatzky, R. L., Golledge, R. G., Cicinelli, J. G., Pellegrino, J. W., & Fry, P. A. (1993). Nonvisual navigation of blind and sighted: Assessment of path integration ability. *Journal of Experimental Psychology: General, 122*, 73–91.

Lynch, K. (1960). *The image of the city*. Cambridge, MA: MIT Press.

McDermott, D., & Davis, E. (1984). Planning routes through uncertain territory. *Artificial Intelligence, 22*, 107–156.

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

Mittelstaedt, M. L., & Mittelstaedt, H. (1980). Homing by path integration in a mammal. *Naturwissenschaften, 67*, 566–567.

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

Moar, I., & Carleton, L. R. (1982). Memory for routes. *Quarterly Journal of Experimental Psychology A, 34*, 381–394.

Moeser, S. D. (1988). Cognitive mapping in a complex building. *Environment and Behavior, 20*, 21–49.

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

Montello, D. R. (1992). The geometry of environmental knowledge. In A. U. Frank, I. Campari, & U. Formentini (Eds.), *Theories and methods of spatio-temporal reasoning in geographic space* (pp. 136–152). Berlin: Springer.

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

Montello, D. R., & Frank, A. U. (1996). Modeling directional knowledge and reasoning in environmental space: Testing qualitative metrics. In J. Portugali (Ed.), *The construction of cognitive maps* (pp. 321–344). Dordrecht, The Netherlands: Kluwer.

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., Jr. (1993). Integrating knowledge of vertically aligned large-scale spaces. *Environment and Behavior, 25*, 457–484.

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

Rand, G. (1969). Pre-Copernican views of the city. *Architectural Forum, 132*(9), 77–81.

Siegel, A. W., & White, S. H. (1975). The development of spatial representations of large-scale environments. In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 10, pp. 9–55). New York: Academic Press.

Spray, J. A. (1986). Absolute error revisited: An accuracy indicator in disguise. *Journal of Motor Behavior, 18*, 225–238.

Stern, E., & Leiser, D. (1988). Levels of spatial knowledge and urban travel modeling. *Geographical Analysis, 20*, 140–155.

Stevens, A., & Coupe, P. (1978). Distortions in judged spatial relations. *Cognitive Psychology, 10*, 422–437.

Stevens, S. S. (1951). *Handbook of experimental psychology*. New York: Wiley.

Thorndike, E. L. (1931). *Human learning*. New York: Century.

Thorndyke, P. W. (1981). Spatial cognition and reasoning. In J. H. Harvey (Ed.), *Cognition, spatial behavior, and the environment* (pp. 137–149). Hillsdale, NJ: Erlbaum.

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

Thorndyke, P. W., & Stasz, C. (1980). Individual differences in procedures for knowledge acquisition from maps. *Cognitive Psychology, 12*, 137–175.

Tobler, W. R. (1965). Comparison of the correspondence of geographic patterns. *Papers and Proceedings of the Regional Science Association, 15*, 131–139.

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

Van Fraassen, B. C. (1985). *An introduction to the philosophy of time and space.* New York: Columbia University Press (First published 1970).

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

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

White, S. H. (1983). The idea of development in developmental psychology. In R. M. Lerner (Ed.), *Developmental psychology: Historical and philosophical perspectives* (pp. 55–77). Hillsdale, NJ: Erlbaum.

Yeap, W. K., & Jefferies, M. E. (1999). Computing a representation of the local environment. *Artificial Intelligence, 107*, 265–301.

Yeap, W. K., & Jefferies, M. E. (2000). On early cognitive mapping. *Spatial Cognition and Computation, 2*, 85–116.