© 1999, ALL RIGHTS RESERVED

Running head: ORIENTATION AND WAYFINDING

ORIENTATION AND WAYFINDING: A REVIEW

EARL HUNT & DAVID WALLER

UNIVERSITY OF WASHINGTON

Abstract

Spatial orientation can take place in three separate scales: scenes within an individual's visual field, surrounds including information to the front, side, and rear, and neighborhoods, that contain points not visible from the current location. When asked to orient in a surround people are especially sensitive to information to their fronts and backs. However if the surround has been experienced by viewing a map time to access information about a point increases with the angle between the forward direction and the queried point. As people become familiar with neighborhoods they first notice landmarks, then paths between landmarks, and finally develop configurational knowledge of the key locations. The last stage is not always reached, even after years of experience. On the average, people can orient themselves toward an unseen point in a neighborhood with an accuracy of about twenty degrees. However there are very large individual differences in orienting ability. People can acquire orienting information from viewing a map, listening to a description, or experience in a computer-generated virtual environment. The characteristics of the representation may depend upon the medium through which the information is presented.

ORIENTATION AND WAYFINDING: A REVIEW

TABLE OF CONTENTS

ORIENTATION AND WAYFINDING: A REVIEW

Orientation is our awareness of the space around us, including the location of important objects in the environment. Orientation in space is crucial for finding one's way (or *wayfinding*) from one location to another. Most of us have a talent for wayfinding. We know our neighborhoods. We find our way to and from work every day. These skills both develop and deteriorate over the course of life. Children under seven have to be watched or they will become lost. Ten-year olds can walk about their neighborhood almost as well as an adult. Loss of orientation is a characteristic of senility. Patients suffering from Alzheimer’s disease may lose the ability to navigate in areas where they have lived for more than thirty years. On the other hand, some healthy elderly people retain considerable orienting ability. Deer hunters over 65 are, if anything, less likely to become lost in the woods than younger hunters (Hill, 1992).

Wayfinding is not exclusively a human trait. Certain species of birds migrate over thousands of miles. Based on laboratory studies, Tolman (1948) concluded that rats formed cognitive maps of their environment. Dogs calculate shortcuts, bees locate and return to food sources, and the lowly desert ant keeps track of her bearing back to the home burrow. Some animals use species-specific techniques of navigation, such as the salmon’s sensitivity to the chemical composition of outflow from its home stream. To complicate the issue, there may not be any one way of wayfinding. Rats, for example, are remarkable opportunistic. Sometimes they form the cognitive maps as Tolman said they do; at other times they learn to associate responses with local stimuli (Restle, 1957). As we shall see, humans are similarly flexible.

In this review we consider the psychological basis of human orientation. We begin with a discussion of some of the issues concerning the establishment and measurement of mental maps. Next, psychological findings on orientation are discussed. Separate subsections consider orientation in the space surrounding an observer, learning about a region by moving through it (wayfinding), and the development of a knowledge about the configuration of objects within a region (survey knowledge). We next discuss wide inter-individual differences in the ability to maintain orientation. The review concludes with an analysis of three artifacts that people use to understand space: maps, texts, and computer-generated virtual environments (VE's).

A psychological classification of environments and environmental information.

A person is oriented when he knows his own location relative to other important objects in the environment, and can locate those objects relative to each other. The word ‘relative’ is central. For instance, in the modern latitude-longitude system, geographic coordinates define locations with respect to the center of the Earth. There are a number of advantages to using a polar-Equator system, most of which are related to the fact that the poles are the endpoints of the Earth’s axis of spin. For purposes of location only, though, we could equally well define the polar axis by any two points at the opposite ends of a line passing through the Earth’s center.

The latitude-longitude system is an example of an *absolute* frame of reference because all objects on the surface being explored are positioned with respect to a fixed origin and direction. It provides unambiguous directions for movement on the surface -- the familiar East-West, North-South system for defining bearings. Other cultures have used other absolute systems. Prior to the European settlement of Hawaii, the Polynesians expressed direction as *mauka* (toward the mountains) and *makai* (toward the sea). Directions orthogonal to a beach-mountain line were expressed by reference to a prominent landmark near the end of an island. This system works quite well on an island with a well-defined central peak. Tzeltal, a Central American Mayan language makes similar use of “up the mountain” and “down the mountain” to express direction. The Tzeltal system can be applied to movement away from the mountain by adopting the convention that “toward the top of the mountain” always means a direction approximating North, even when the speaker is on a flat plain (Levinson, 1996b).

Relative frames of reference express direction and location of one object relative to another, but vary the location and sometimes the identity of the origin from one situation to another. *Egocentric* frames of reference are centered on the speaker, as in the familiar English above-below, front-back, right-left system. (Linguists sometime refer to a *dietetic* frame of reference, but egocentric seems a more direct term.) *Intrinsic* frames of reference are centered on an object in the environment. The distinction is captured by the ambiguity of the statement “The cat is behind the truck,” which may have a different meaning depending upon whether the speaker is using an egocentric frame of reference or an intrinsic one centered on the truck.

In an abstract sense, reference systems are adequate if they specify bearing and position accurately. Every object on Earth has a unique latitude and longitudinal position, and these positions imply bearings. Psychologically, a reference frame is useful if it is easy to establish a correspondence between concepts in the frame of reference and cues in the environment. However, ease of use can vary from situation to situation. Egocentric right and left are easy to determine, but hard to communicate to a second party. Intrinsic frames of reference force the wayfinder to consider a situation from a perspective other than his or her own. Absolute frames of reference can be extremely abstract, as our latitude or longitude system is, or may be applicable only within a certain region, as is the Hawaiian island-centered frame of reference.

In many situations people find it convenient to mix frames of reference. This is especially true in giving directions (Taylor and Tversky, 1996). Here is an example.

> “To get to my office, park in the underground parking structure and leave by the staircase marked ‘Administration building.’ On reaching the outside, face south, toward the fountain. Walk down the steps and enter the building on your right.”

The underground parking structure, the staircase, the fountain, the steps, and the building are features of the environment. South is a term in an abstract system of exocentric bearings. Right is an egocentric directional term.

Wayfinders use their cognitive representations of a space in order to guide behavior in it. According to Poucet (1993) the key behaviors are recognizing places, locating places with respect to each other, and planning routes from one's present location to a target location. In one of the seminal references in the field, Siegel and White (1975) claimed that when adults learn a new space they acquire these abilities in order. First they become familiar with landmarks, then routes, and finally configurations. Children do the same, except that children younger than age 12 do not seem to acquire configurational representations. Siegel and White's argument has received a great deal of support from subsequent research findings. Siegel and White's ideas have been generally supported as descriptive statements. The require expansion in order to present a theory of orientation.

People certainly use landmarks to describe familiar regions, but landmarks may not be good guides to wayfinding. Therefore it is useful to distinguish between landmarks, in the sense of locations whose (possibly non-geographical) characteristics define an area, and *control points* used to determine directions. For instance, in Washington D.C. the White House is an important landmark, but it is not a control point for motor vehicle travel, for it is difficult to turn near it. The 14th Street Bridge is a less memorable landmark but it is a control point for city traffic. The Washington Monument is both a landmark, for its architectural and cultural significance, and a control point, as it dominates an important traffic circle.

As knowledge of landmarks accumulates a wayfinder may begin to rely more heavily on route knowledge. Wayfinders have to know what action to take when they reach control points. Referring to the well-established distinction between types of information in memory, McDonald and Pellegrino (1993) distinguish between *procedural* and *declarative* knowledge about routes. Procedural knowledge is defined in terms of reactions to external stimuli, the sight of a location is associated with the appropriate response. Route knowledge refers to explicit knowledge of a route, in the sense that the stimulus-response sequences required at each control point are held in memory as part of a higher order entity: the route itself. McDonald and Pellegrino's distinction has an important psychological implication. Procedural knowledge is, by definition, not available for conscious inspection and discussion, while declarative knowledge is. Also, movements directed by procedural knowledge are typically carried out much more rapidly than movements directed by declarative knowledge. Unfortunately, very little research on wayfinding has been guided by the distinction between procedural and declarative knowledge.

Tolman used the appealing metaphor of a 'cognitive map' to explain how rats represented a maze environment. Kuipers (1982), and several authors since, caution that the metaphor must not be taken too literally. Personal experience and empirical studies (Foley & Cohen, 1984) show that under some circumstances people will create a visual image of a map of a familiar environment. However the image will be distorted from an actual map in a number of ways. Some of these are described below. Rather than thinking of an actual map in the head it is safer, although wordier, to speak of mental representations containing configural knowledge of distance and direction.

The title of Siegel and White's article contained the words 'representations of large-scale environments.' How large is large? Unfortunately, there is little consensus in the experimental literature.. The same terms and references have been applied to spatial reasoning at the scales of finger-mazes and the continental United States. This seems unwise, for different modes of reasoning may apply to spaces of such vastly different scale. We feel that it is important to distinguish between four classes of space.

> Spatial scenes. A spatial *scene* is an area that can be sensed at a glance, without moving. In normal human terms, this is essentially the space in front of us. In order to avoid circumlocutions, we will speak of it this way unless, when dealing with the blind or non-human wayfarers, there is reason to do otherwise. A spatial scene is a visual stimulus, it contains objects, can be imaged, and can be described by stating perceivable features.
>
> Spatial surround. A *surround* is the space consisting of all scenes that a stationary wayfarer could sense, if that observer were to make a 360^o sweep without moving from his or her position. A mathematician would say 'rotating without translating.' A somewhat obvious, but important, point is that surrounds cannot be perceived in their entirety. Accordingly the mental representation of a surround must be developed from memories of scenes.
>
> Neighborhoods. A *neighborhood* is a set of spatial surrounds through which a wayfarer moves. Therefore it is an ordered set of personally experienced surrounds. Just as the representation of the surround must be developed from representations of scenes, the representation of a neighborhood must be developed from the representations of surrounds.
>
> Geographic regions. A *region* is a geographically defined space that a person knows only indirectly, through exposure to cultural representations of it, such as maps or verbal instructions. The term 'person' rather than 'wayfarer' is now appropriate, for knowledge of geographic regions does not depend upon personal exploration. Because we acquire information about geographic regions from secondary sources, our representation of a region depends upon our ability to represent the secondary sources that have defined it. It follows that any distortion in memory that applies to the secondary sources will influence our representation of the region. If a region has been defined by a map, any principles of perception and memory that apply to visual diagrams will apply to memory of the map and, perforce, the representation of the region. Similarly, if a region is described to us our understanding of that region will depend upon our ability to comprehend discourse.

The distinctions made above are based on the information available to a wayfinder and the behaviors required to move in the relevant space. It is of interest that similar distinctions have been made on the bases of neurophysiological and neuropsychological data. In particular, there is a good deal of evidence that performance in what we have called scenes, surrounds, and neighborhoods is mediated by different neural structures (Previc, 1998).

Methodological issues

We now move to the issue of measuring a person’s mental representation of a space. A mental representation is, by definition, an internal summary of information a person has acquired, either by direct or vicarious exploration. The psychologist’s first task is to determine how faithfully the representation depicts the space. We want to know what information has been retained, what has been lost, and what has been distorted. The second task is to determine how that information is held. There are two answers to this question: a functional one and a neuroanatomic one. We will concentrate on the functional one, although neuroanatomical findings will be discussed where appropriate.

The usual way of determining that a person’s mental representation contains (or implies) a particular piece of information is to determine that the person in question can solve problems whose solution depends upon the relevant information. Lynch’s (1960) study illustrates perhaps the ‘most obvious’ way to find out what a person knows about an area. He asked his respondents to describe their home city and to draw a map of it. While the resulting answers have a great deal of face validity they are both indirect. Verbal recall is suspect because it is hard to state fine-grained metric information. While we have words to describe, say, an angle of 10^{o} 37” and a distance of 144.73 meters, saying “A bit to the right and about a block away” sounds more natural. Measures derived from maps can be similarly problematic, primarily because drawing skills vary tremendously. A good map is always evidence of a good representation, but a bad map may simply be a sign of a poor artist.

Three techniques have been used to deal with this problem. The simplest is to ask a person to choose between correct and incorrect maps or between views of an area from a particular perspective. Such recognition tasks avoid the skill problem, but correct recognition does not show that a respondent could generate the relevant information on his or her own. Cueing techniques have sometimes been used as an intermediate between drawing and recognition. The respondent is shown an incomplete map and asked to fill in blank regions (Kitchin, 1996). In this case, the accuracy of a person’s responses depends upon the cues provided, and in particular whether the cues are landmarks for the respondent. Construction techniques are intermediate between recognition and recall. In a construction task the respondent is given blocks or pieces of paper representing locations (e.g. toy houses) and asked to place them on a larger outline map (Siegel, 1981). Construction can be combined either with cueing or recognition, and has the advantage of avoiding problems associated with drawing skill.

Since relative locations are determined by direction and distance, many investigators have attempted to determine these directly. A frequently used paradigm for determining bearings is the directional pointing task, in which participants imagine themselves standing at point A, facing B. They are asked to point toward C. This provides a face-valid estimate of knowledge of relative direction.

Distance estimates are psychophysical functions of actual distance, and are not equivalent to actual distances even when a person looking at the distance to be estimated. As is often the case for psychophysical judgments, distance estimates are approximated by Stevens' power law for scaling:

$$E_{(x,\, y)} = a\, D(x, y)^{b} \qquad (1)$$

where $E_{(x,\, y)}$ is the estimated distance from x to y, D(x, y) is the actual distance, and parameters a and b are estimated from the data. Depending on the value of these parameters one may obtain accurate estimates, consistent overestimates, consistent underestimates, there may be underestimation up to a point, at which time the respondent switches to overestimation, or there may be overestimation at short distances followed by underestimation at longer ones. Estimation is also affected by other factors, including experience with a path, the number of distinct points intervening between x and y, and whether or not the two points belong to the same subregion of a space and scale of environment (for a review, see Wiest & Bell, 1985).

Locations in a mental representation can be inferred from directional pointing alone, or by using *mental triangulation*. Suppose that we wish to determine an observer's belief about the location of point C, with respect to points A and B. The respondent is asked to imagine standing at point A, and asked to point to B and C. This determines the angle < BAC. If we repeat this procedure at B, pointing to A and C triangle ABC has been defined, and the location of C has been determined relative to line AB. The procedure is then repeated using a new pair of base locations, A' and B', to determine the triangle A'B'C. The technique is shown in Figure 1. If the person has a Euclidean representation of space, point C should be at the same position in triangles ABC and A'B'C. In practice, it seldom is. The discrepancy between the two locations of C provides a measure of the consistency of a person's mental representation.

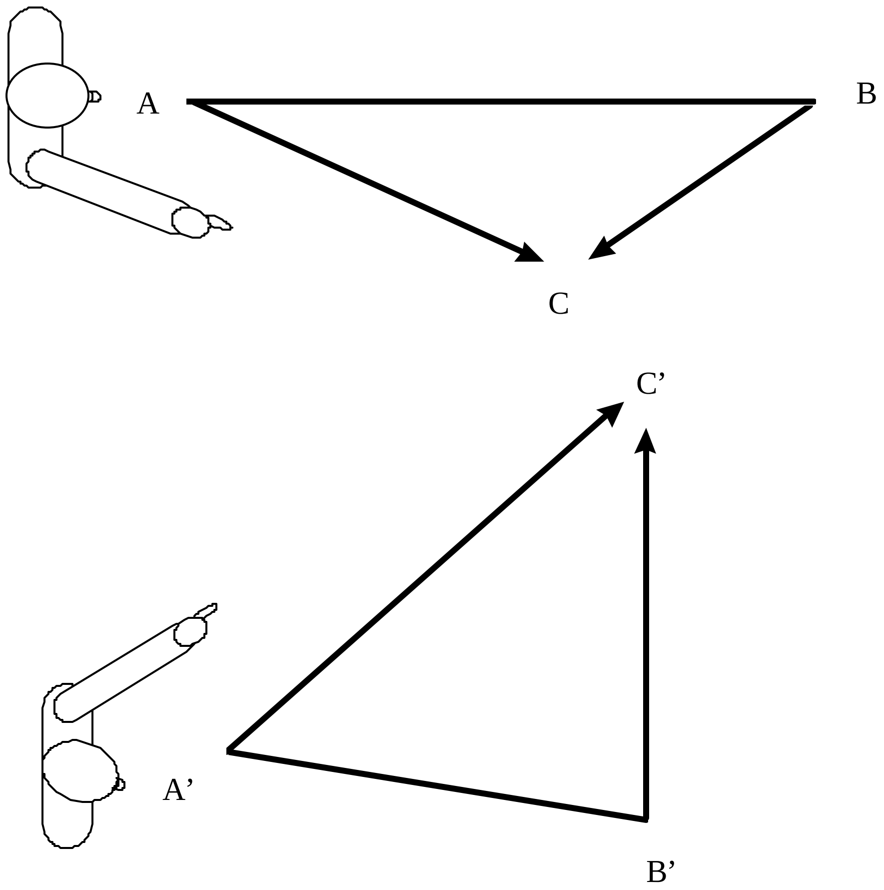


Figure 1. A mental triangulation experiment. A respondent is asked to indicate the direction to reference point C when standing at A and looking toward B, and vice versa. The resulting location, C, is the mental location of C. The procedure can be repeated for points A’ and B’, producing location C’. Inconsistency in the representation is indicated by the length of the line C-C’.

In evaluating a person’s mental representation, three properties are important: accuracy, consistency, and reliability. A straightforward way of measuring the accuracy of a representation is to use the discrepancy between a measurement made in the representation and a measurement made in the space being represented. For instance, if a mental triangulation study shows that a person thinks that $<ABC = 47^{o}$ and it is actually 35^{o} , then the signed error is $+12^{o}$, and the unsigned error is 12^{o}. While this is a useful and frequently applied procedure it does have one major drawback. When an experimenter requires a response he or she has to choose a frame of reference to use in analyzing the responses. If the observer’s mental representation uses a frame of reference that is misaligned with the experimenter’s, all responses may be systematically distorted in some way. Bovet (1994) reported an illustrative example. He drove students to sites several kilometers from their campus and asked them to point back to it. One site was beside a freeway that was labeled as running East-West, but in fact ran Southeast-Northwest. The students’ estimations at

that site were systematically biased in a counter-clockwise direction. The 'independent' measurements were affected by a common error.

Mental representations of bearing and distance can be estimated by having a wayfarer follow an indirect path from point A to point C, through intermediate points B, B', etc., and then return directly to A from C. In a variant of this method the wayfinder explores routes (e.g. route A-B-C-D and route B-R-Q-C). One segment is then blocked, forcing the wayfinder to find a detour. Continuing the example, if the task is to go from A to C and segment B-C is blocked, the wayfinder should take the route A-B-R-Q-C. This is the technique that Tolman (1948) used in his famous studies of cognitive maps in rats. The paradigm is still in use. For instance, Maguire and colleagues asked London taxidrivers to indicate the routes they would take if certain streets were blocked (Maguire, Frackowiak, & Frith, 1997).

Pointing, drawing, and locating objects on a map require that a person indicate some property of the mental representation that translates directly into distance or direction on a map. An alternative approach is to ask the respondent to make a series of simple geographic judgments, and then to apply mathematical techniques to determine the configuration of locations implied by these judgments. A favorite technique is to ask people to make relative distance judgments, and then to apply a conceptually elegant analysis known as *Multi-dimensional scaling* (MDS) to the results. First the respondent makes judgments of the relative distances between pairs of locations. A typical question might be "Is Paris closer to Madrid or London?" (In some applications the respondent is asked to estimate the distances.) The multidimensional scaling algorithm is then applied to determine a configuration of points in Euclidean space that maximally satisfy the constraints implied in the judgments.

Although geographic examples make good illustrations of the concepts behind MDS there is no requirement that "closer to" mean "closer to in a geographic sense." In practice MDS has been used in situations in a mental space is to be mapped out based on a general concept of similarity. For instance, spaces for animal names have been constructed by asking people questions like "Is a bear more like a dog or a horse?" Numerous other applications have also been made. (See Schiffman, Reynolds, and Young, 1981, for a discussion.) In such situations the right answer is not defined, for the experimenter's task is to find out how – not how accurately – objects are located in the appropriate mental space.

When the purpose is to investigate mental representation of physical space there is a right answer and, in spite of the popularity of the method, MDS may not produce it. Kitchin (1996) compared mental maps inferred using MDS to those obtained using mental triangulation, drawing, and other methods. The maps produced using MDS were substantially less accurate than maps produced by other methods.

Hirtle and Jonides (1985) described an alternative technique that is also based on judgments of closeness, but that is somewhat less demanding in its assumptions about the metrics of mental space than MDS is. First they asked people to recall key locations in a familiar space. They took as a measure of 'closeness' the distance between two items their list. For example, suppose a person were asked to recall the names of major U.S. cities and the responses were: New York, Philadelphia, Washington D. C., Chicago, St. Louis, and Los Angeles. New York would be assumed to be psychologically closer to Philadelphia than Los Angeles, because the first two cities are recalled close together. The clustering algorithm then grouped cities together based on relative closeness. Suppose that measures of closeness were obtained for the set of U.S. cities { New York, Philadelphia, Boston, Chicago, St. Louis, Cleveland, Los Angeles, San Diego, San Francisco}. If mental closeness resembled geographic closeness, which it surely would in this example, the Hirtle and Jonides procedure would group the items as { New York, Philadelphia } {Boston} {Chicago, Cleveland} {St. Louis} {Los Angeles, San Diego} {San Francisco}, and then aggregate these groupings into Eastern cities { New York, Philadelphia, Boston}, Mid-Western cities, { Chicago, Cleveland, St. Louis} and California Cities { Los Angeles, San Diego, San Francisco } and finally the California cities versus the rest. The clustering procedure can be used to identify psychologically real neighborhoods, but it does not provide any configurational information about locations other than relative proximity. Note that the clustering procedure is independent of Hirtle and Jonides' particular measure of closeness.

A third way of defining mental closeness utilizes semantic priming, a technique that is widely used in memory research as a way of showing how ideas in memory are tied to each other (McNamara; 1986, 1992). In one study, college students responded to a series of geographic questions about locations on their campus. An example might be " Is the Administration building located in the main quadrangle?" Consider two questions, about locations A and B, that are asked in the order A, then B. The time to answer the second question, about location B, will be influenced by the physical distance between A and B. McNamara concluded that, *other things being equal,* priming is influenced by recalled geographic distance. Therefore priming can be used as a measure of the psychological distance between two locations.

Priming also occurs when the questions are about non-geographic information, such as "What is the name of the building (location A) with a swimming pool?" followed by "Name the building (location B) that has a Diego Rivera mural." This finding is important because it shows that the mental representation of geographic space is one of the dimensions of a more general semantic space that ties meaningful objects together in memory.

Suppose that one or more of the above methods has been used to a representation of a person's mental map. The result is a statement of where *the experimenter* thinks *the person* thinks objects are in two (or occasionally, three) dimensional Euclidean space. In order to evaluate the respondent's knowledge, it is necessary to compare the mental map to a physical map. One way to do this is to ask informed judges to compare (the experimenter's representation of) the respondent's mental map to an actual map. This can be done with a reasonable degree of inter-judge agreement (Kitchin, 1996). Another method is to determine discrepancies between actual and mental locations, e.g. by calculating the mean squared distance between points on the actual map and points on the representation's map. Although this method has face validity, it is problematic because it only works if some way is found to transform the scale and orientation of one map into another, and because the location of points in the mental map are dependent on one another. Therefore if the participant misplaces a single key point, and judges all other points relative to it, the entire map may appear to be inaccurate even though there is only one error in it. Kitchin (1996) reports a way to handle this problem, by computing a two-dimensional analogue of the correlation coefficient. Waller (1999) has offered a critique and extension of the method.

In many studies it is assumed that our goal should be to derive a Euclidean representation of respondents' mental maps. McNamara (1992) has raised a question about this goal. Sketch maps, mental triangulation, and MDS force the representation of a respondent's cognitive map to obey the metric axioms. That is, for all points { x, y, z } and distances { d(x,y), d(x,z), d(y,z) } the following statements are assumed to be true:

(2) Positivity and an established zero: $d(x,x) = 0$; $d(x,y) > 0$ if $x <> y$.

(3) Symmetry. $d(x,y) = d(y,x)$.

(4) Triangle inequality. $d(x,z) \leq d(x,y) + d(y,z)$.

(5) Segmental additivity. Define a path going from x to y and then to z, with each segment of the path being a straight line. Then if p(x,y,z) is the length of the path, $p(x,y,z) = d(x,y) + d(y,z)$

McNamara cites several studies indicating that mental representations may not satisfy these axioms. The methodological point is that when an experimenter uses the form of a conventional map to represent a person's mental map, then the experimenter has forced the data to conform to the metric axioms. There is no guarantee that the cognitive map also conformed to them. Therefore before we impute metric properties to a cognitive representation of space we have to ask whether or not our method of data analysis forced the answer to obey metric axioms.

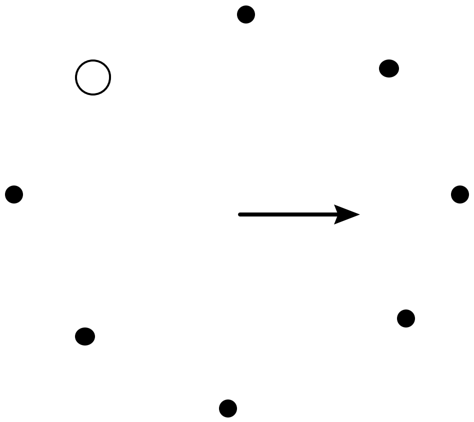
With these methodological comments in mind, let us look at how orientation is maintained in different types of spaces.

Scenes and surrounds

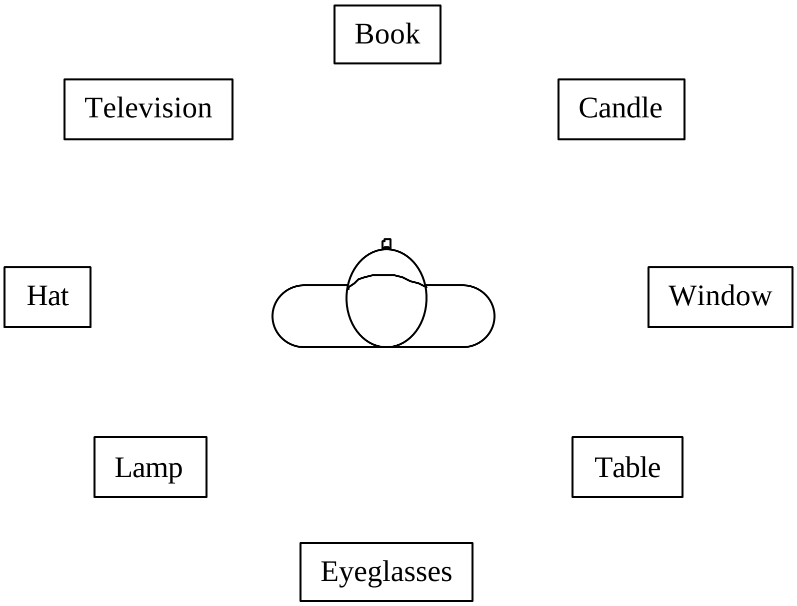
Assuming normal vision, it is easy to locate an object in a scene. You just look at it. But how do we locate objects in surrounds, where the object may be out of sight to our side or even behind us? How is this managed?

Baddeley and Lieberman (1980) showed that keeping track of objects is a visual function, even when visual input is not being used. Blindfolded participants pointed at a moving sound-source while performing either an auditory or a visual short-term memory task. Pointing at the sound source interfered with the visual task but not the auditory one. In terms of Baddeley's (1987) model of working memory, this showed that keeping track of locations in the immediate surround requires resources assigned to the visual-spatial scratchpad. We want to go beyond this to determine more precisely how visual memory represents a surround.

Hintzman, O'Dell, and Arndt (1981) conducted a series of experiments that addressed this question in considerable detail. Hintzman et al. had college students learn the locations of objects in a room. Learning took place under two different conditions. In the *map* condition, shown in the upper panel of Figure 2, participants memorized an upwards oriented map with object positions marked on it. In the *explore* condition (the lower panel of Figure 2) the participant explored the room itself. In the test, the map was removed and the objects taken from sight. The participant was told to imagine facing an object (The *orienting object*) and then to point to a second (*target*) object. Each trial involved two independent variables of interest: the identity of the orienting object and the angle between the orienting and the target object. People were able to point fairly accurately, but the time required depended upon the condition and the location of the objects. In the map condition, reaction time increased as a function of the angle between the upright and the orienting object. In the explore condition reaction times did not vary across orienting objects. This result has come to be called *orientation independence* of knowledge about spatial layouts. It contrasts with the *orientation dependence* displayed in the map condition.



The MAP condition of the Hintzman et al. experiment. The participant is to imagine themselves at the center of the circle of dots, facing in the direction shown. The task is to point to a target, here shown as a white dot. (After Hintzman et al.,1981, Fig. 1).



The EXPLORE condition of the Hintzman et al. studies. The participant sat in a chair surrounded by pictures of the named objects. Subsequently the participant imagined facing one object and pointing to another

Figure 2. The conditions of the Hintzman, O'Dell & Arndt (1981) study of how a surround is represented.

In both conditions, reaction times varied with the relative position of the target object. People were quick to point to objects directly forward in their imagined orientation, then to objects immediately to the back, and slowest to objects off to the side. Hintzman et al. refer to this as an M-shaped response profile, where the base points of the M represent facing directly forward (0^o or 360^o), and the low mid-point is the 180^o. The M shaped pattern would not occur if people were doing something analogous to a mental rotation, for in that case the longest reaction times should have been found for targets at 180^o, immediately behind the participant.

Two subsidiary experiments extended the results of the map condition. In one, people were asked to imagine themselves standing in an imaginary U.S. city, located in southern Illinois. The map experiment was then repeated, with cities replacing objects (e.g. look to Detroit, directly to the North, and point to Denver, which would now be to the left). A strong orientation effect was observed. People made pointing decisions much more rapidly when they were asked to imagine facing North (the conventional upward orientation of a map) than when facing in other directions.

In most of Hintzman et al.'s work participants were asked to orient themselves facing an object. Therefore 'forward' and 'direction of orientation object' are confounded, leaving open the question of whether the participants were using an egocentric frame of reference or an absolute one defined by room co-ordinates. (The reader is invited to stop and guess what reference system was used.)

In Experiment 13 of the Hintzman et al. series the participant was told to imagine standing with the orienting object to the right or left. In this case the egocentric and room co-ordinates conflict; 'forward' defined by body position is different from 'forward' defined by the direction to the orienting object. Figure 3 shows an abstraction of the results, across selected data points. The M shaped profile shifted to the right or left, with the shortest response times being those in the direction of the orienting object, rather than the participant's (imagined) forward direction. More abstractly, the preferred direction was established by the observer-orienting object axis, an exocentric frame of reference, instead of the egocentric axis determined by the observer's own direction. The result strongly suggests that Hintzman et al.'s participants developed a configurational representation of their immediate surround.

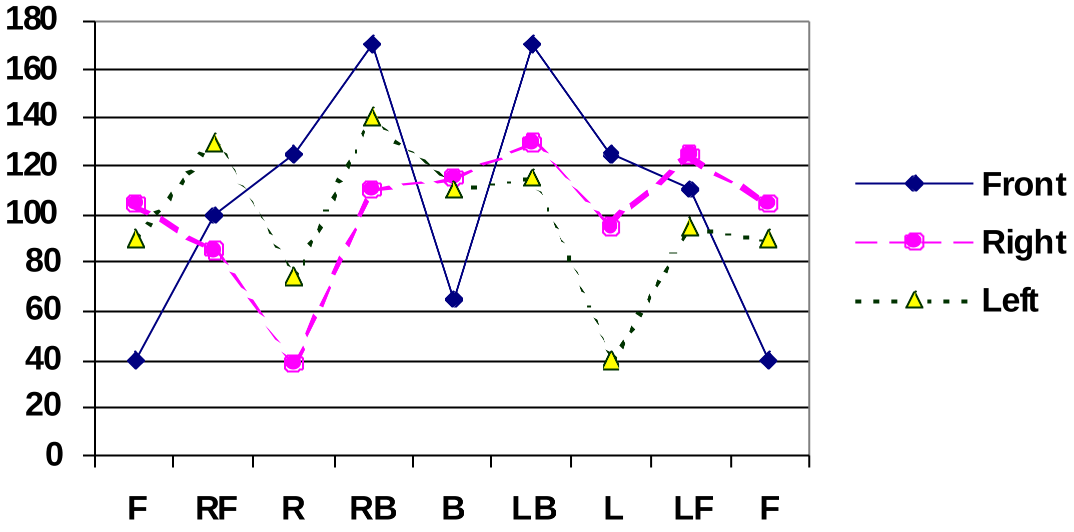


Figure 3. The M shaped pattern of pointing effects observed by Hintzman, O'Dell, & Arndt (1981). The data shown are the reaction times, as a percent of mean reaction times over conditions, plotted as a function of direction of pointing relative to the observer; Front (F), Right Front (RF), etc. F is shown twice for symmetry. When the participant is facing the orienting object an M shaped figure is obtained with the most rapid responding when the target is directly in front of the observer, and the next fastest when the target is immediately behind the observer (F and B conditions respectively). This is shown by the solid lines, FRONT condition. When the orienting object is to the right (RIGHT condition, dashed lines) the most rapid responding is to the right and next most rapid to the left of the observer (R and L conditions). This pattern is reversed when the orienting object is to the left (LEFT condition, triangles). The M shaped pattern is centered on exocentric alignment, not egocentric bearing. Data extrapolated from inspection of Hintzman et al., Figures 5 (first day responses) and 18.

Finally, Hintzman et al. made an important observation about individual differences. The M shaped profile was highly reliable across subjects. Mean correlations between pairs of subjects, across target objects, ranged from .84 to .95 in various experiments. The between-subjects effects of orientation were much less reliable. The interpair correlations of reaction times as a function of orientation objects ranged from .36 to .83 in the same experiments (Hintzman et al., Tables 3 and 4). The low correlations for orientation effects could be because orientation made no difference, and hence the low correlations arose because of variation due to chance. Alternatively, different participants might be systematically using different orientation strategies. Hintzman et al. (pg. 164-165) favor this interpretation. The fact that there were wide individual differences in orientation was probably due to systematic individual differences in the way in which orientation was established. On the other hand, once orientation was established, the calculation of bearing to a target seems to have been a much more systematic procedure across individuals.

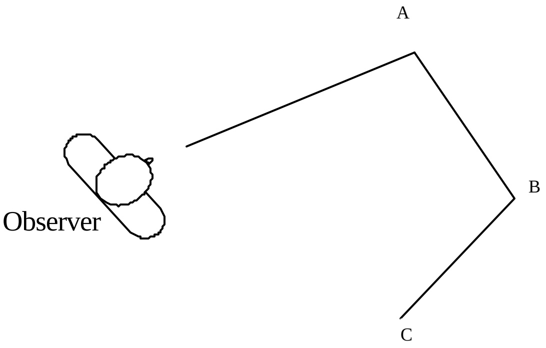
The phenomena observed by Hintzman et al. appear to be ubiquitous across studies of how people understand their surround. They can even be observed when the surround is described, instead of actually being seen (Franklin & Tversky, 1990). The privileged status of the space along to front-back axis is particularly striking. Suppose that an observer is shown an object, the object is removed, and the observer is asked to point to where the object was. Accuracy is greater for objects to the observer's front (Franklin, Henkel, & Zangas, 1995). Our language reflects this asymmetric degree of precision. When asked to describe objects as being in the front, on the right, etc. *front* is used with more precision, and is more likely to be a primary term (as in 'left front') than the other position terms. Clearly the space around us is not seen as equal; front is psychologically most important, back next most important, and the sides least important.

The Hintzman et al. studies were conducted in an impoverished environment in which the participant was at the center of a circle with objects on the circumference. "Exploration" consisted of shifting one's orientation, while remaining in the center of the circle. Geometrically, exploration was done by rotation without translation. Normal surrounds contain objects that vary in distance and location, and we typically explore a surround both by translation and rotation. Furthermore, exploration is seldom complete; people normally do not move to every possible location. Nevertheless, non-experienced perspectives can be imagined. For instance, we invite the reader to imagine what a scene would look like, standing on their kitchen sink, looking out the kitchen door.

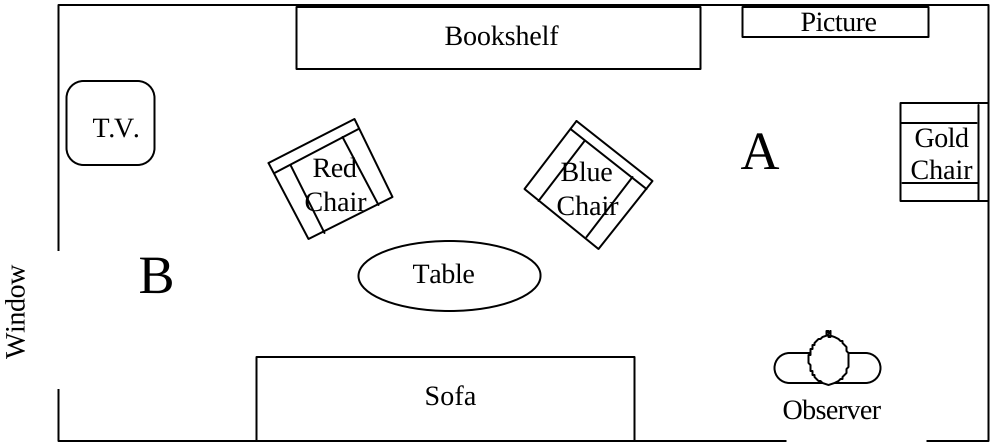
There are two ways by which a person might hold information about the scenes that would be experienced when viewing a surround from different perspectives. In *early computation* the requisite information is incorporated into the observer's mental model during exploration. Subsequently, when asked a question about a particular view, such as "When looking from your desk, is the chair to the right or left of the closet? " the observer can answer by retrieving information from memory. The time required to answer a question about a particular view should be independent of whether or not the observer had actually experienced it. In *late computation* the observer does not incorporate the information directly into a mental model, but does incorporate enough information so that the answer can be computed when the question is asked. Therefore questions about experienced views should be answered more rapidly than questions about unexperienced ones. While the early and late models imply different reaction times, the implications for accuracy are not so clear. Errors might occur either at the original time of viewing (early error in computation) or at the time of retrieval and inference (late computational model).

The distinction between early and late computation is useful in thinking about a number of psychological processes. For instance, it is used in models of text comprehension regardless of whether or not spatial information is involved. While the distinction is useful, it is not a sharp one. There seem to be relatively few situations in which people only memorize precisely what they see or hear, and there certainly are very few, if any, situations in which an observer draws all possible inferences from an observation. What psychologists can do is characterize those situations in which a particular type of question is likely to be answered by early or late computation.

Examinations of the orientation dependence-independence issue provide a good case in point. Suppose that an experimenter marks out a multi-segment path on a room, and asks an observer to stand at one end, as shown in Figure 4 (upper panel). (Although the term ‘path’ has been used, the observer is still in a surround, for the entire path can be seen at one time.) The observer then moves to a second room and performs pointing tasks (stand at X, look at Y, point to Z), as described above. Pointing tasks can be chosen so that the direction of pointing is either the same as that experienced during observation (*aligned*) or opposite (*contraligned*) with the directions that would have been experienced during the observation period. This is explained in the upper panel of Figure 4. The experiment can be conducted either with a simplified diagram, as shown in the upper panel of Figure 4, or in a room-like environment, as is shown in the lower panel. Orientation dependency is found if aligned pointing is more rapid than contraligned pointing. When found, the effect is evidence for late computation of directions from perspectives other than those experienced during exploration.



Orientation dependency effects. An observer stands at the point indicated and memorizes the path. Subsequently the observer is removed from the room and asked to indicate the direction from one point, either from the original point of observation or from some point on the path that has not been observed. For instance, the observer might be asked to point to B either when standing at the observation point facing C (experienced view) or when standing at C facing the observation point (inexperienced view). The diagram is based on procedures used by Presson & Hazelrigg, 1984.



An observer enters a room and examines the layout, using only designated positions (path). Subsequently the observer is asked to describe the room, either from the experienced point A or the unexperienced point B. Orientation dependence is found. The diagram is based on procedures used by Sholl & Nolin, 1997.

Figure 4. Two different techniques for contrasting the psychological effects of describing a surround from an experienced or unexperienced perspective.

What happens? The facts are in dispute. Presson and his colleagues claimed that aligned and contra-aligned views are equivalent, providing that (a) the information is not acquired from the map and (b) the path being viewed is relatively large, covering a space on the order of 4 square meters (Presson, DeLange, & Hazelrigg, 1989; Presson & Hazelrigg, 1984). Others have found strong alignment effects, both in viewing paths and in viewing room-like situations (Diwadkar & McNamara, 1997; Roskos-Ewoldson, McNamara, Shelton, and Carr, 1998; Shelton & McNamara, 1997; Sholl & Nolin, 1997). Perhaps the strongest evidence for late computation comes from studies by McNamara and his colleagues (Diwadkar & McNamara, 1997; Shelton & McNamara, 1997). They showed that if a person is permitted several, but not all possible, views of the same scene then the time taken to answer a question from a new perspective is related to the discrepancy between the new perspective to that of the closest experienced scene. This suggests that people use a late computation procedure somewhat similar to the computer-graphics technique of morphing, in which a new picture is created by blending two old pictures. Interestingly, Rhodes et al., (1998) reached a similar conclusion in studies in which people recognized faces and objects that were studied from one perspective and, and test, viewed from old or new perspectives.

This leaves us with a puzzling issue, why were Presson's obtained at all? To answer this question we have to look carefully at the paradigm used. Participants were briefly shown a path marked on the floor of a room, from a standing perspective, blindfolded, and then either asked to do a pointing task while blindfolded and

(a) in the original learning position,

(b) after sitting in a wheelchair and being wheeled to a new room and told to imagine being at the position required by the pointing task, or

(c) after being wheeled circuitously to the position that they were to imagine being at in the other conditions.

Sholl and Nolin (1997) repeated these procedures and found that orientation dependence or independence depends exactly on how the test is done. Orientation dependence is found in all conditions if the experiment is done as just described, providing that the participant stands up at the test site. Orientation independence occurs if the observer remains seated in the wheelchair and testing is on the path (Sholl & Nolin; contrasting their experiments 1 and 2 to the results of experiment 3). However orientation dependence returns if the person is tested at the site of original learning or in a room rather than on the path! Effects like these strongly suggest the orientation independence effect observed by Presson and his colleagues is associated with particulars of a rather unusual testing situation, in which a person views a path laid out on the floor of a room.

Sholl and Nolin (Experiment 5) then conducted a study that was logically similar to, but psychologically different from, Presson's path paradigm. A simplification of the procedure used is illustrated in the bottom panel of Figure 4. Observers entered a room and examined it from a fixed observation point. Subsequently they were removed from the room and told to imagine themselves either back at their original point ('Observer' in Figure 4) or at another point in the room ('A' in the figure), facing the actual observation point. In the first case directional pointing is aligned with experience, in the second case it is counteraligned. Orientation dependency was found. This is consistent with McNamara's results, and suggests that orientation dependency (and late computation) is the case following investigation of a ‘normal’ surround. Paths marked on the floor may be a special case.

Knowing where things are about you is certainly an interesting psychological ability. Nevertheless, it is not exactly what we mean when we speak of wayfinding. We look next at some of the problems associated with finding one’s way about in a neighborhood.

Learning about neighborhoods: Route knowledge.

According to the Random House dictionary a neighborhood is

> 1) The region surrounding or near some place or thing; a vicinity.
> 2) A district or locality, often with reference to its character or inhabitants.
>
> *Random House Dictionary, 1980.*

The first definition is geographic, the second social and semantic. For our purposes a psychological definition is needed. A geographic region is a (psychological) neighborhood if a wayfinder learns about it by traversing routes in it. This definition applies to conventional neighborhoods, multi-room buildings, and large ships. On the other hand, it rules out surrounds, which can be examined without translation, and geographic regions, such as the state of California, where spatial information must be acquired from secondary sources, such as maps and descriptions. Because the distinction depends upon how a wayfarer interacts with the environment, intermediate sized areas (and perhaps the state of Rhode Island!) might be geographic regions to some people and neighborhoods to others.

The *route* is a central concept in the discussion of neighborhoods. The term was used previously in its common sense meaning. It is now necessary to define *route* more precisely. A (decision) *node* is a location at which a wayfinder selects a new bearing. A *route* is a sequence of such nodes, together with the segment traveled from one node to another. Tourists visiting Washington D.C. are offered a classic tourist trip that includes the White House, the Capitol, and the Washington, Lincoln, Roosevelt and Jefferson memorials. There are multiple routes that can be used to visit these control points.

We will consider a wayfinder who moves along (*traverses*) routes without access to maps or aerial views. Therefore all knowledge of the neighborhood must be built up from perspective views of points on the ground. A complete theory of spatial orientation should describe the wayfinder’s mental representation of a neighborhood and explain how that representation is built up from experience with routes. In order to address these issues we shall examine several progressively more complicated techniques of wayfinding. Each time we ask what information processing characteristics are required to use the wayfinding technique and what the wayfinder can learn from applying them.

Tracking relies on local cues that identify a route. For instance, a businessperson flying to Seattle to visit the Microsoft Corporation could be told to

> "Follow the EXIT TO FREEWAY signs from the airport. Then take I-405 North from the Airport to the intersection with State Route 520."

In these instructions "North" is a tracking instruction rather than a reference to geographic alignment. The traveler is to take the highway marked *North*. I-405 actually goes East at the point at which the traveler enters.

Tracking is a useful method of navigation. People who follow signs in a large, complicated building arrive at their destination more rapidly than those who use You-Are-Here maps, presumably because the time saved going to the destination is less than the time required to read the map (Butler et al., 1993). Signs and markings "on the ground" have also shown to be an effective way of guiding search in virtual buildings (Satalich, 1995) and large virtual spaces (Darken & Sibert, 1996). A wayfinder can lay down cues during a traverse, and then use those cues to *backtrack* to a previously visited location.

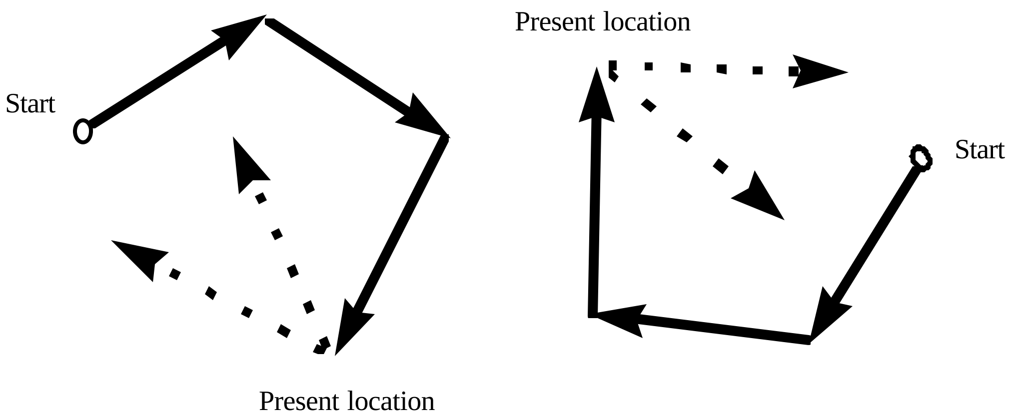
While tracking is often an efficient way of getting to a location it has several disadvantages. The wayfinder learns nothing about the environment except the route itself. The memory of the route lasts only so long as the cues last. (Cookie crumb trails are of little use after a rainstorm.) Still worse, at least for the development of a spatial representation, tracking may divert attention from the acquisition of geographic information, as the following examples show.

Manuel de Juan-Espinosa studied the Fang, a hunting and farming group living in the tropical rain forest of Equatorial Guinea. Fang hunters told Juan-Espinosa (personal communication) that they become lost if they are so intent on following the animal's sign that they forgot to update their location. This anecdotal report was documented, in a very different environment, in a Master's Thesis completed at the University of Washington's Human Interface Technology laboratory (Satalich, 1995). The environment to be explored was a virtual building, i.e. a building that existed as interior design plans in a computer. Participants in the study saw computer-controlled projections of the views that they would have seen had they been at a particular point and bearing inside the (non-existent) building. Several different methods of exploration were permitted: movement under computer control, free exploration, or directed exploration, in which the participants tracked arrows indicating the path that was to be followed. Wayfinders who had tracked had more trouble composing new routes when old ones were blocked than any of the other wayfinders did.

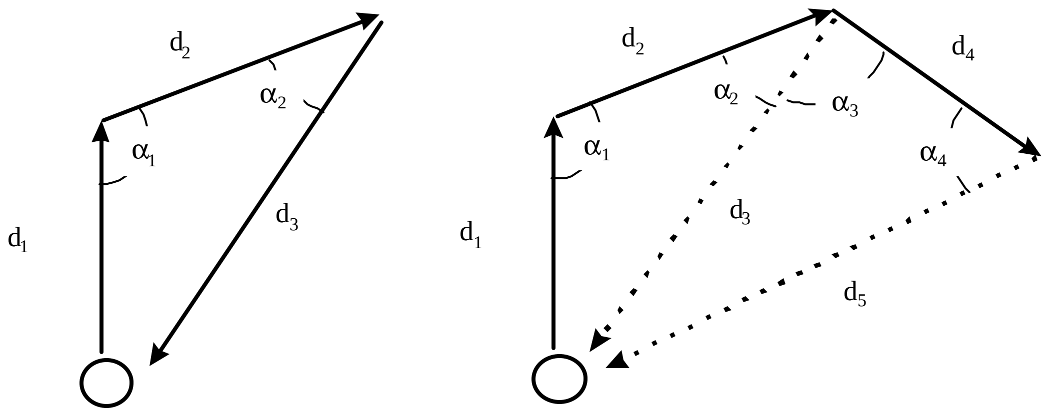
The fact that tracking is associated with fast performance but low memory is consistent with current theories of memory. Virtually every cognitive psychologist who has studied the topic has stressed that if you wish to have specific memory of an experience, think about it. Butler et al. point out that during tracking, temporary memory loads are minimized. This will maximize performance, but minimize the amount of information consolidated into declarative long-term memory.

Tracking relies entirely on exocentric, local cues. *Dead reckoning* is a method of navigation that relies on egocentric cues. In dead reckoning a route is defined by the wayfinder's records of the turns and distances traveled during a traverse. If these can be recorded perfectly the dead-reckoning wayfinder can learn quite a lot about a neighborhood.

The geometry of dead reckoning is shown in Figure 5. The two three-segment routes in panel A of Figure 5 show how a dead reckoner with perfect memory for turns and distance could move directly from one on-route node to another, following an off-route track (the dotted lines in the figure). This will be called *egocentric bearing* information, and is defined as the turn required in order to point the wayfinder toward a target location. Once this is done the wayfinder can return to the target without knowing the distance to it. This represents a primitive method of route planning. Of course, the wayfinder would not know whether the proposed route was possible, because by definition an egocentric wayfinder knows nothing of the environment off the route.



A. The information available to an egocentric navigator who records distance and turn information. The navigator can determine bearing from the present location to any other point on the route, including the starting point. The navigator cannot discriminate between the route on the right and the route on the left, because the initial bearing has not been defined. This is an exocentric term.



B. How bearing to home can be maintained. On the first segment (left hand side) of the route the navigator travels distance d_1. Bearing to the start point is achieved by turning 180^o and distance is simply d_1. The navigator then makes turn $\alpha_1$and travels distance d_2. By the law of cosines, bearing and distance to the start (α_2 and d_3) are established since d_1, α_1, and d_2 are known. At the next stage (right hand side) the navigator "forgets" all information not relevant to the current node, but retains bearing and distance (α_2 and d_3) at that point. The navigator makes the turn α_3 and travels distance d_4. At this point d_3, $\alpha_2 + \alpha_3$, and d_4 are known. By the law of cosines, α_4 and d_5, the new bearing and distance to the start, are defined.

Figure 5. A geometric analysis of the information available to navigator using egocentric cues only ("dead reckoning").

The distinction between early computation and late computation applies here. Dead reckoners who use early computation have to update their bearings to target locations during a traverse. The necessary computations are illustrated in the bottom panel of Figure 5. The early computation wayfinder must keep track of the node to target bearing computed at the last node, the turn taken at that node, and the distance traveled along the segment from the last node to the current location. Because bearings are updated continuously an early-computation dead reckoner can proceed directly to a target location (e.g. a home burrow) if the situation demands it. Animal wayfinders behave in this way. For instance, the desert ant follows a circuitous route when foraging, When it encounters food the ant carries it directly back to the home burrow. If a dog is forced to travel a particular route and then released it will calculate shortcuts from its present location to locations where food has been observed (Chapius & Varlet, 1987).

A wayfinder who uses late computation records the route itself as a set of distances and turns. Bearings from one arbitrary point on the route to another can be computed on demand. Therefore, the late-computation wayfinder does not rely on the advance specification of targets. On the other hand, the amount of computation required to determine the bearing from origin point A to target point B increases linearly with the number of segments that intervene between the origin and the target. Because the computations are inherently serial, the time required to determine bearing should also increase. If there are errors in the recorded segment and turn information they will accumulate over segments, so accuracy should decrease as the number of intervening segments increases. These effects will be referred to as *segment effects* for response time and accuracy, respectively. Both effects are found in human wayfinding situations (Loomis et al., 1993; Sholl, 1996).

A purely egocentric wayfinder cannot reproduce a route after learning it, because the starting alignment cannot be defined egocentrically. Therefore, the wayfinder cannot discriminate between a route and a rigid translation or rotation of it (Figure 5, top). Such errors can be corrected without referring to an absolute frame of reference. All that is required is that the initial segment of a route be aligned with the local geography of the region. This is often established by physical constraints. If there is only one door to your house (or one entrance/exit to your burrow) the initial alignment problem is solved.

The geometric analysis in Figure 5 shows that a dead reckoner can solve interesting spatial orientation problems, but it does not show how to solve them. Determining bearing and distance are exercises in trigonometry, a mathematical procedure that many people, and presumably all dogs and ants, do not consciously understand. However there are a number ways to approximate the mathematical solution to bearing problems. Early seafarers drew lines on a chart, then measured the lines and angles. Connectionist networks could be constructed to approximate trigonometric computations. For all we know, the brain may contain them. The important point is dead reckoners must have access to reasonably accurate information about turns and segment lengths. What ‘reasonably accurate’ means depends somewhat on the route, although in general, errors in measuring or remembering turns create more problems than errors in estimating distances (Rieser, 1989).

In theory, a dead reckoner can rely solely on egocentric cues, providing that the reckoner has a way of sensing turns and distances. For instance, a motorist could act as a dead reckoner by recording turns and mileage readings, both of which are egocentric cues with respect to the car. In more natural examples, several experimenters have shown that the blind and blindfolded people with normal vision can learn short routes (3 to 4 segments up to 4 meters long) quite rapidly (Levine, Janovic, & Palij, 1982; Loomis et al., 1993).The Loomis et al. study was particularly informative. Blind people and blindfolded people with normal vision walked a short route within a room and then attempted to walk directly back to the starting point. All participants performed well above chance, but there were systematic errors. The wayfinders behaved as if (a) turns were moved toward the nearest right angle turn and (b) distances were underestimated. Figure 6 shows how these errors influenced performance on the final test. When Loomis et al.'s participants attempted to return to the starting point they tended not to walk far enough, and to have their bearing displaced in a predictable way.

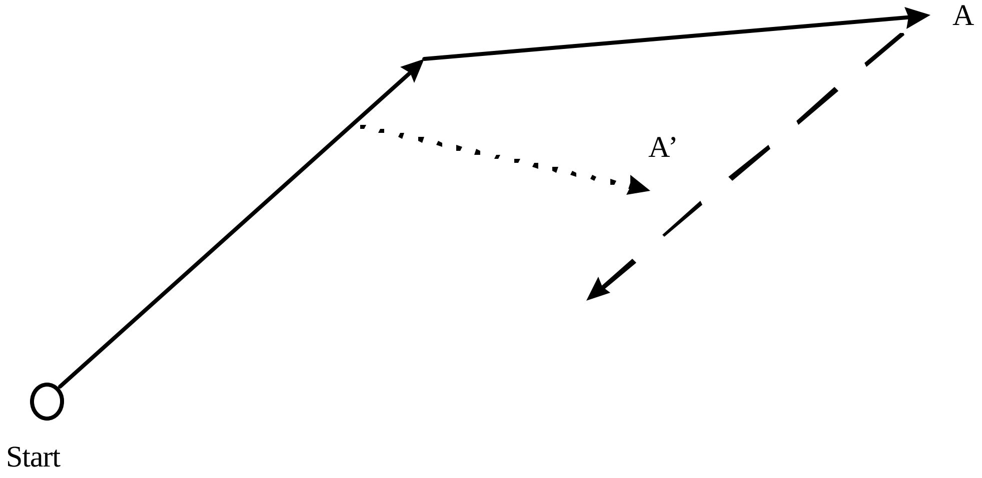
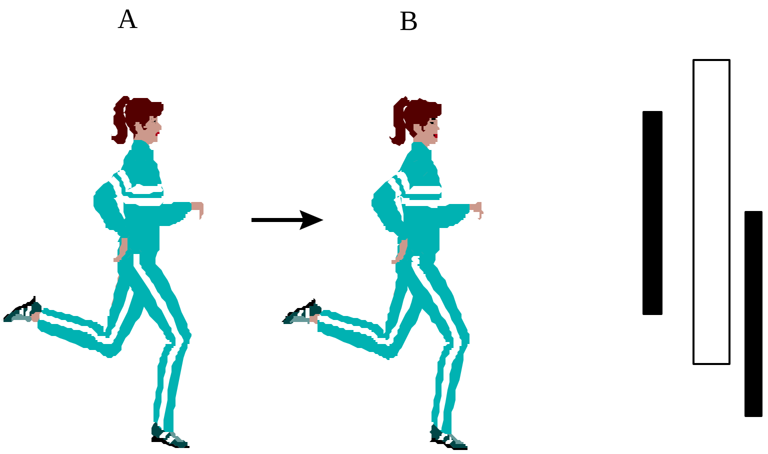


Figure 6. The effect of systematic errors upon maintenance of bearing. Navigators follow the route indicated by the solid arrows, and then attempt to return to the starting point. The route shown by dotted lines indicates the route as represented by a wayfinder who underestimates distances and "squares off" angles. Therefore when the wayfinder is at point A he or she thinks they are at point A'. When the wayfinder attempts to return to the start from point A, thinking that they are at A', they miss the return in a systematic way (dashed line).

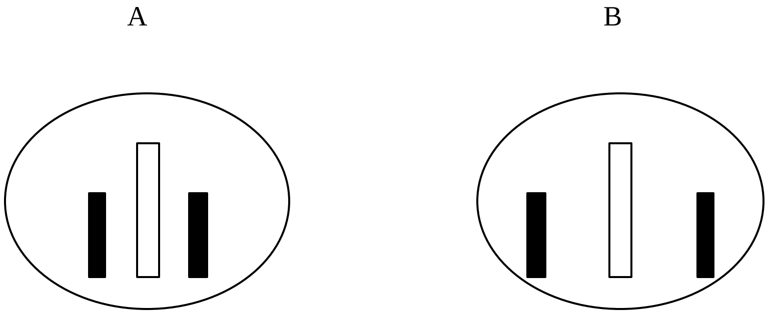
Dead-reckoning routes can also be defined by external cues, such as highway mileage indicators. In addition, and perhaps more naturally, people can estimate distances from visual cues that are less direct than a mileage sign. For instance, Golledge et al. (1993) found that adult wayfinders had a good sense of the relative length of segments in routes that they had 'traversed' by watching a slide show. In this case the wayfinders had no kinesthetic or proprioceptive cues.

What sort of cues are used in egocentric dead reckoning? Changes in bearing can be sensed directly by kinesthetic and vestibular cues from turns in the head and body. Determining the distance one has traveled is more problematical, for muscular effort has to be co-ordinated with distance on the ground.

Velocity and time determine distance. Let us assume some primitive capacity to judge time. The problem is to relate velocity to egocentric kinesthetic cues. Figure 7 shows a way that this can be done. The figure depicts a wayfinder moving directly toward a target. If the wayfinder fixes on the target it will remain in the center of the visual field, while objects to the side of the target move through the peripheral visual field and eventually out of sight. This is called *optic flow*. The rate of optic flow is determined by the wayfinder’s speed, so optic flow is a potential link between egocentrically defined muscle movements and exocentrically defined distance, and hence can be used to calibrate a muscular effort in terms of distance traveled. Wayfinders evidently use optic flow as a direct cue to velocity, and hence distance, for their ability to maintain bearing is reduced if their peripheral vision is occluded (Sholl, 1996). This may partially explain the fact that learning a route from watching video displays is less effective than learning by traversing (Gale, Golledge, Pellegrino, & Doherty, 1990).



A wayfinder moves rapidly directly toward a target, the white post, moving from point A to point B. The wayfinder looks directly at the target, so that the white post is in the center of the visual field and the two flanking black posts are to either side.



The runner’s view of the three posts as she approaches. The two black posts move to the periphery of the runner’s visual field at a rate that is determined by her speed. This provides a cue to exocentric speed that can be co-ordinated with egocentric kinesthetic cues.

Figure 7. Changes in the optic flow in the periphery of the visual field can be used to indicate exocentric velocity. These cues can be co-ordinated with egocentric cues from the runnerÕs musculature to establish a correspondence between the egocentric cues and distance traveled in a fixed time.

The importance of optic flow as a cue to distance leads us to expect that people with severe visual disabilities, especially in the peripheral visual field, would have difficulty maintaining bearing. In fact, such people do surprisingly well (Loomis et al., 1993; Rieser, Hill, Taylor, & Bradfield, 1992). How do they calibrate their motions to measurements on the ground? The answer turns out to be that we have to specify when the visual disability occurred. If motion-distance calibration occurs early in life individuals who have lost their sight through accident or disease should be better wayfinders than those who were born with impaired vision. In fact, they are. Rieser et al. (1992) asked sighted individuals and people who had suffered different degrees of visual loss to indicate bearings from one point to another in familiar neighborhoods. The average unsigned error for participants with full vision was 17°. This is typical of the value obtained in many studies. The mean unsigned error for people who had lost peripheral vision after they were 3 years old was 23°, only slightly more than for people with full vision. The mean unsigned error for people who had lost peripheral vision before they were three years old was 74°, not far from the 90° [1] that would be expected on the basis of chance. These results suggest that people learn to calibrate muscular effort in terms of velocity relatively early in life. Once this is done muscular effort alone can be used to estimate distance traveled.

A third method of wayfinding, *piloting*, associates egocentric bearing information with exocentric location information, as in this play on the plight of the French Foreign Legionnaires in the novel *Beau Geste*:

> "Leave the gate of Fort Zinderneuf, turn 75° to the right, and march until you reach the oasis."

Although it is natural to think of a pilot's route as a set of ordered instructions, ordering is not necessary. All that is necessary is that the decision points on a route be recognizable and that the actions to be taken at each decision point be clear. Here are the piloting directions for going to our laboratory, as an unordered set of productions.

> When at the parking lot and facing the fountain -> walk forward until the bottom of the steps is reached.
>
> When in front of the Johnson Hall door -> enter the building.
>
> When standing at the exit to the parking lot -> turn until facing the fountain.
>
> When at the bottom of the steps and facing the fountain -> turn right 90°.

This method of storing route information will not work if a node appears on more than one route. One way to avoid this restriction is to associate each rule with a contextual cue, such as the goal of the route:

> If the goal is to go to the office and you are at the parking facility -> Turn until you are facing the fountain and proceed forward.
>
> If the goal is go to the library and you are at the parking facility -> Turn until you are facing the library.

Pilots have to be able to recognize nodes when they reach them and to make sufficiently accurate changes of bearing when a change is required. A "pure" pilot thinks of a route only in terms of bearings and targets, and therefore has no need of distance estimation. Such a pilot would not be able to maintain bearings from a present location to any previously visited point except the last one. (Even this capability would be lost if the pilot were tracking a twisting trail between nodes.) Since people (along with dogs and the desert ant) do show an ability to maintain bearing to prominent previously visited points, such as the starting location, wayfinders must be doing some distance estimation as they explore a route.

On the other hand, the extent to which distances and bearings are computed should not be exaggerated. When people are asked to explain how they traverse routes a substantial number of them will give piloting instructions. These self-described pilots seem to be pure pilots, for they are rather poor at tasks that require maintenance of bearing, such as pointing to a distant, unseen object (Lawton, 1994, 1996).

People can continue as pure pilots, without acquiring bearing information, for a long time. In a widely cited study Thorndyke and Hayes-Roth (1982) found that people might work in a building for several months before they were able to perform a pointing task reliably. This estimate may have been optimistic. When Moeser (1988) asked nurses who had worked in a building for up to two years to describe it, they generally gave what amounted to piloting instructions for going from one workstation to another. Although they could find their way along familiar routes the nurses made pointing errors on the order of 40 to 50 degrees.

We have some understanding of the brain mechanisms that are used to acquire piloting information. Single unit recording techniques have located cells in the rat's hippocampus that become sensitive to a single place. Other cells, in the thalamic region, are sensitive to the orientation of the animal's head. (McNaughton et al., 1996). Connectionist modeling can be used to account for the way in which the orientation and location records become associated, resulting in piloting instructions of the sort described above. Further modeling offers an explanation of how piloting instructions at one decision point can be linked to the instructions associated with the next point on a path (Samsonovich & McNaughton, 1997). This provides a way of discriminating between paths containing common nodes, something both rats and people can do.

To what extent can information about brain activity in the rat be generalized to humans? Addressing this question poses a significant technical challenge. Single cell recording requires invasive surgery. Brain scanning studies are not feasible either, for the scanning apparatus is not portable and, in the case of Positron Emission Tomography (PET) scanning, the time available for study is limited because of the use of radioactive isotopes. Even if these technical challenges could be overcome neuroscientists would have to find some way of distinguishing between brain activity associated with wayfinding and the brain activity required to control walking and gazing at non-geographical information during a traverse. Examining EEG waveforms obtained from shoppers in a mall would not be a good way to study orientation!

Fortunately there is an interesting alternative: taking brain scans while people explore virtual environments. With some exceptions that at present are not clearly understood (Klatzky, Loomis, Beall, Chance, & Golledge, 1998), behavioral data obtained from exploration of virtual environments is fairly close to that obtained from exploration in analogous real environments (Ruddle, Payne, & Jones, 1997; Waller, Hunt, & Knapp, 1998, see the discussion of virtual environments, given below). Therefore generalizing results from virtual to real environments is probably not unwarranted.

The resulting studies, while not entirely consistent, paint a picture that agrees with the data obtained from the animal studies. Maguire and her colleagues (Maguire, Burgess et al., 1998; Maguire, Frackowiak, & Frith, 1997) asked participants to recall routes that they had either learned by experience with a virtual environment or, in an especially interesting case, learned by their experiences as taxi drivers. Route recall initiated activity in the right hippocampal area and the parahippocampal gyrus (a region near the hippocampus). Maguire and others (Aguirre & D'Esposito, 1997) also observed activity in the dorsal and ventral visual processing streams during wayfinding tasks, including a pointing task. Landmark recognition without a wayfinding task produces activity in the ventral stream alone, which is consistent with the distinction between visual processing to determine 'what' something is and to determine 'where' an item is relative to other items. The neuroimaging results complement neuropsychological observations that associate spatial disorientation with lesions in the higher order visual processing system. (Anderson, 1988; Kritchevsky, 1988; Levine, Warach, & Farah, 1985). However, as reviewers have noted, some neuropsychological cases present exceptions to the general picture.

The various studies we have cited show quite impressive abilities. Obviously we learn routes as we traverse them. The ability appears to be fully developed by age 11 or so. Younger children display difficulty in learning routes, at least in part because they are weak in identifying good landmarks (Cornell, Heth, and Alberts, 1994; Gale et al., 1990). What is perhaps most impressive is the number of ways in which people can learn routes through a neighborhood. Although walking is our natural mode of locomotion, we can learn routes without normal proprioceptive cues. We do this everyday when we learn a route by being driven through it. Vestibular as well as proprioceptive cues are removed when people learn by viewing videotapes (Gale et al, 1990) or exploring routes in desktop virtual environments (Klatzky et al., 1998; Ruddle et al., 1997; Waller et al., 1998). To top things off, it is possible to learn routes by viewing randomly ordered sequences pictures of locations along the route (Allen, Siegel, & Rosinski, 1978). Conversely, route learning can occur without vision, both in the blind and in normal adults who have been blindfolded (Loomis et al., 1993).

The variety of conditions under which learning occurs indicates that the route-learning mechanism must be independent of input from a particular sensory system. The brain regions involved include the thalamus, which is involved in detection of turns, the hippocampus and related structures, which are required for the establishment and recall of route information, and the dorsal and ventral paths of the higher order visual system. The latter may be primarily involved in maintaining orientation as a wayfarer moves through the surrounds within a route.

Impressive as route learning is, there is a wayfinder more powerful than the tracker, dead reckoner, or pilot. We next look at the psychological properties of the navigator.

Cognitive maps and configural representations.

The route learner thinks of space as a set of objects connected by paths. The navigator thinks of space as a system of positions from which distance and bearing (*configural information*) may be computed. The distinction is captured by two equally true statements:

(Route representation) The city of Washington D.C. is located at one end of the straight line known as the Boswash Axis, which extends from Boston through New York to Washington. Washington is 300 km. from New York.

(Configural representation) The city of Washington D.C. is located at 77^o w. longitude and 39^o n. latitude.

The navigator's advantage over the route finder rests on the use of exocentric cues to locate navigation points without having to link them by a traverse. This is illustrated in Figure 8. The figure shows the case of a navigator who has traversed two distinct, unlinked routes: ABCD and XY. The bearing to location M can be determined from at least two points on each route. This locates the routes in a common space. The resulting representation can be used to establish the bearing between any two positions.

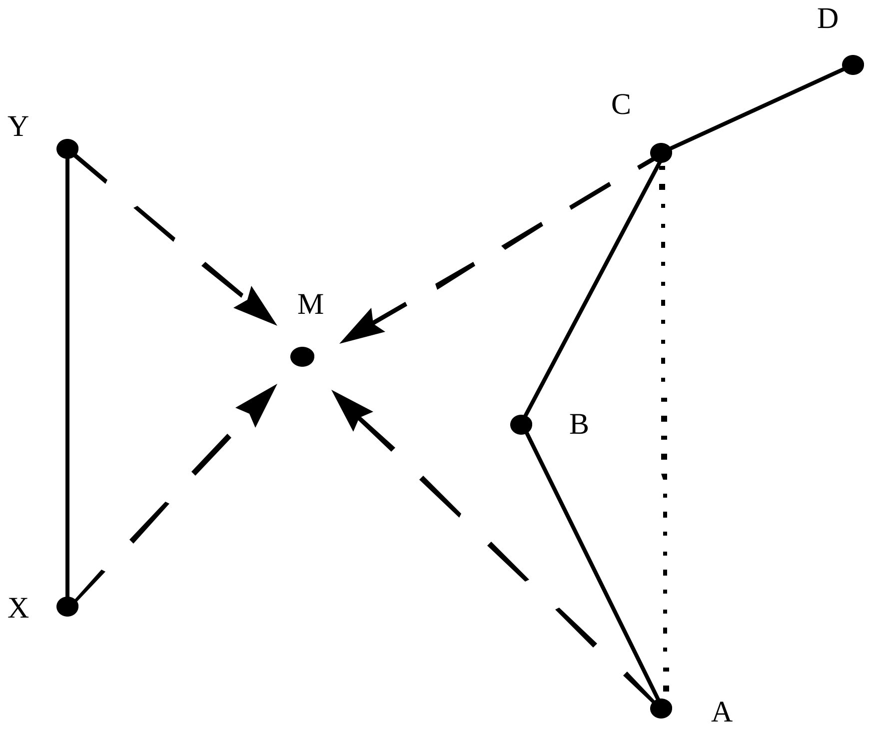


Figure 8. A navigator traverses routes XY and ABCD, and establishes bearings to M (relative to North) from points X, Y, A, and C (dashed arrows). Since the navigator has kept track of distance and bearing information the distance and bearing from A to C can be calculated (dotted line). Using triangulation, M can be located and used to position the two routes in the space. The navigator can now determine the bearing from any point on the route ABCD to any point on route XY, even though the points have never been linked by a traverse.

Navigators depend on both early and late computation. Early computation is used to calculate configurations of key objects. (The word ‘map’ is intentionally avoided.) Late computations are used to calculate routes by determining bearings from one point to another. The time required for computation will be independent of the route used during exploration. For example, a navigator using the chart shown in Figure 8 should determine the bearing D-A as rapidly as B-A. Therefore a navigator should not show an increase in the time required to calculate a bearing to a previously visited point as a function of the number of segments between the point and the present location.

Errors are another issue. If the representation of the route is less than perfect, as is almost certainly the case, wayfinders should also show functional dependency for errors in bearing as well, because errors will accumulate over segments. However the magnitude of the errors will depend upon the accuracy with which the wayfinder records and recalls turn and distance information. A very accurate wayfinder would make small errors. Similarly, a navigator might or might not display functional dependency for errors, depending upon how the navigator determined position. Consider a navigator who starts out from a known location and establishes positions along a route. If the navigator uses dead reckoning, as the early

European navigators did, then errors will accumulate and the later locations on the route will be improperly positioned. This would produce functional dependency for errors but not for latencies. If the navigator makes an independent determination of co-ordinates for each point, and as a modern sailor does using a satellite-based global positioning system does, errors will not accumulate.

In order to use a configural representation, navigators have to solve three problems. The first two are *alignment* of a direction in the representation with directions on the ground and *positioning* the starting point before exploration begins. Erroneous solutions to the alignment and positioning problems will result in systematic discrepancies between the navigator's configural representation and locations on the ground. For instance, in the Pacific Northwest region of the United States a navigator who confuses magnetic north (the direction to the Earth's north magnetic pole) with true north (the direction to the North Pole) will be misaligned by about 20^o.

The navigator's third problem is to determine distance. Suppose, for a moment, that a mental map corresponded exactly to a real map. We estimate distance in the world from distances on a map by measuring the distance between two points on a map, and then making a scale conversion to estimate distance in the world. This only works when two conditions are satisfied. The same equation for distance must be valid for measurements on the map and in the world that the map represents, and the same scale relationship must hold between map distance and real-world distance for all pairs of points. The first condition fails with conventional maps whenever two points are sufficiently far apart so that the curvature of the Earth has to be considered. If you use a conventional flat map the straight-line distance from New York to London is roughly 7800 km. The airline flight distance, which approximates the length of the shortest New York-London route on the earth's surface, is only 5600 km. However the navigational problems most people face, most of the time, cover much less space, so 'flat earth' maps are sufficiently accurate. The issue of scale relationships, which is subtler, will be dealt with below.

Let us begin with the alignment problem. If a person is properly educated, alignment can be determined by astronomical observations. However the most prominent astronomical marker, the Sun, indicates East and West in a very general way, unless the navigator has access to tables showing where the Sun 'rises' and 'sets' as a function of latitude and time of year. Using stars other than the Sun requires formal training. Compasses can be used to indicate magnetic north, but they are only used in a few settings. Most people solve the alignment problems in their everyday life by relying on cultural signals. People who live on 20th Street South know, by an act of faith in the city surveyor's office, that they live south of Main Street.

Unfortunately, cultural cues to alignments are not always correct. Recall Bovet's (1993) report that university students became disoriented when standing by a highway that was marked East-West but actually ran Northeast to Southwest. In the case of the students we know what cue they were (mis) using. McNaughton et al. (1996) report an analogous phenomena in rats. The rats' task was to enter one of several arms extending from a circular arena. When the animals took a "wrong turn" electrical recordings in the hippocampus appeared to be shifted 90°, i.e. the animals' brains were responding as if they had a correct map of the maze, but had misidentified their initial orientation. One wonders what would have been observed if Bovet had had pictures of activity in his students' hippocampi, or what the rats would have said if the experimenters knew how to ask them what they were doing.

Position and bearing are closely related, for one implies the other. Therefore pointing techniques are often used to establish, simultaneously, a person's representation of position and bearing. An example would be "Imagine that you are at the Washington Monument, facing the Lincoln Memorial. Point to the White House." After a number of such bearings have been established the experimenter should have a good idea of the participant's configural representation of the relevant space. Several studies have found that people who are familiar with a neighborhood can point to within about 20° to 30° of the true location (Moar & Bower, 1983, Rieser et al., 1992, Sholl, 1987). While an error of 20° may seem small it can have substantial consequences. If, for example, a target is 5 km. away, a navigator who proceeded along a route 20° from the true route would miss the target by 1.8 km. There are also large, systematic individual differences in bearing accuracy. These will be discussed in the next section.

Errors in alignment are systematic rather than random. We have previously seen that wayfinders tend to estimate turns as being closer to right angles than they are. Navigators make similar errors. Moar and Bower (1983) had people imagine standing at location A in a familiar urban neighborhood, and then asked them to draw the angle <BAC between two other familiar landmarks, points B and C. Reports were distorted toward the nearest right angle. For example, a true angle of 100° was drawn as 88°, and a true angle of 67° as 84°. Montello (1991) extended these results by asking passersby to point to distant but familiar objects. The study was conducted in an urban neighborhood in which most but not all roads were laid out in the grid pattern typical of U.S. cities. People were more accurate when they were on a street that was orthogonal to the long axis of the grid than when they were standing on a street that was oblique to it. We seem to believe that the world we live in is more regular than it is.

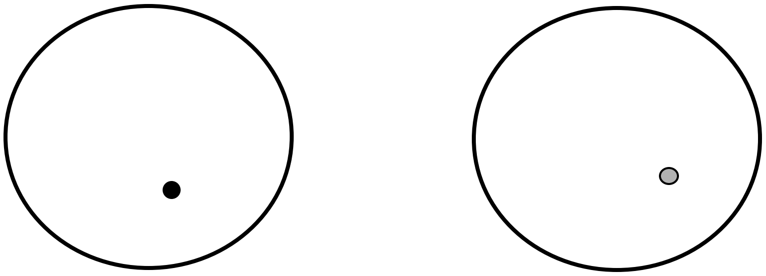
In addition to being useful as a way of establishing a person's configural representation of space, pointing tasks raise questions about the way that the configuration is used. The term 'navigator' suggests that when a wayfinder attempts a pointing task he or she examines a mental map. This is not a psychological theory, it is a metaphor and not an entirely accurate one. When people answer questions about familiar neighborhoods they do not behave in exactly the same way as when they answer questions about spaces they know only from maps. Pointing tasks based on map knowledge show an *alignment effect*. The time and accuracy of the response depends upon the alignment between egocentrically defined directions on the map, as inspected in its normal orientation, and egocentrically defined directions on the ground. To illustrate, imagine standing in St. Louis, looking towards either Chicago (to the north) or New Orleans (to the south). You are asked to point to Denver. Suppose, further, that you did this while looking at an actual map of the United States, in its normal North-up orientation. If you were 'looking' at Chicago, Denver would be to your left both on the ground and on the map. That is, the map and ground directions would be aligned. If you were to imagine looking toward New Orleans, Denver would be to your left on the map and to your right on the ground. In this case map and ground directions are said to be contra-aligned. Laboratory studies have established that when a person is actually looking at a map aligned responses are made more rapidly than contra-aligned responses (Levine, Janovic, & Palij, 1982; Shepard & Hurwitz, 1984). What happens when a cognitive map is used?

This depends upon whether the map is based on inspection of a real map or exploration of a neighborhood. In a widely cited study, whose basic findings have been replicated several times, Evans and Pezdek (1980) asked students to judge the spatial relationship between geographic regions and between locations in their familiar college campus. Alignment effects were found for geographic regions, but not for the neighborhood. [2]When people who were unfamiliar with the campus studied maps of it they showed an alignment effect. Sholl (1987) replicated the alignment effect finding, and in addition found a viewpoint-dependent orientation effect. In the neighborhood condition people were more rapid at pointing to objects in front of them than in back of them. This is consistent with Hintzman et al.'s finding of an M shaped pattern of responding when people are asked to demonstrate their knowledge of a surround (see above). The M shaped pattern was not found when people answered questions based on map knowledge.

There is a simple intuitive model that might account for these results. It could be that when people are asked to respond to pointing questions about a familiar neighborhood, they imagine themselves in that neighborhood, and answer the question by inspecting an image of the appropriate ground-level view. In such a case alignment problems go away, and the relation between orientation and pointing is reduced to the analogous case for a surround. We would expect some alignment effects if people are more familiar with one perspective at a point than others, and indeed these are found. We would also expect people to have more trouble pointing to a target that was visible from the imaged point than one that was not visible. The evidence on this point is equivocal (Sholl, 1987).

Distance estimation turns out to be a complex problem, for cognitive space has some unusual properties. In the external world, units of distance are constant over translation, a mile is a mile no matter where it is. This is in conflict with Lynch's (1960) observation that residents of a city think of it as being organized into neighborhoods, and that places within the same neighborhood are psychologically closer to each other than they are to geographically equidistant locations outside the neighborhood. Lynch's finding has been repeatedly confirmed in laboratory research. Studies using techniques as diverse as overt distance estimates, paired comparisons, and priming all come to the same conclusion: interpoint distances are magnified if the interval between two points crosses a conceptual boundary, such as a river, major street, or social subdivision. (Hirtle & Jonides, 1985; McNamara, 1986, 1992). A similar phenomenon has been noted in the perception of routes. Many routes are broken up into natural subroutes. For instance, in Washington D.C., a tour of the government buildings and monuments naturally breaks into three segments: from the Capitol to the end of the mall, from the mall to the Washington monument, and from the Washington monument to the Lincoln Memorial. Two points on the same segment of a subroute are judged to be closer to each other than are two equally dispersed points in different segments (Allen & Kirasic, 1985).

The Washington Mall, downtown neighborhoods, and other cultural groupings provide good illustrations of the hierarchical organizations of large-scale space. Huttenlocher, Hedges, and Duncan (1991) found hierarchical organization in a uniform, uninterpreted smaller-scale space. Adults were shown a piece of paper containing a circle and a dot. They then attempted to reproduced the dot position from memory. The displacement of the dots from their true position was consistent with the assumption that the space was divided into horizontal-vertical and radial regions, and that points were recalled as being displaced toward the center of the appropriate regions. This is illustrated in Figure 9. Subsequent experiments showed that displacements toward a regional center occurred when very young children attempted to retrieve toys after watching an experimenter bury a toy in a sandbox (Huttenlocher, Newcombe, & Sandberg, 1994).



The procedure used in Huttenlocher, Hedges, & DuncanÕs (1991) experiments on spatial coding. The observer is first shown a circle with a dot in it (left hand circle). Subsequently the observer attempts to reproduce the dot location from memory (gray dot in the right hand circle).

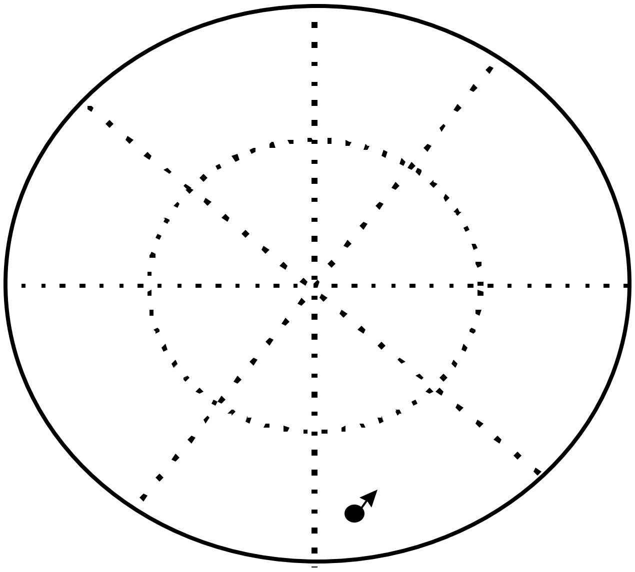


Figure 9. The theoretical model proposed by Huttenlocher et al. (1991). The circle is psychologically divided into quadrants by vertical and horizontal lines, and into radial sections based one or two smaller circles concentric with the large one. The division is shown by dotted lines. The quadrants are further divided by a diagonal indicating the prototypical angular position within a quadrant (dashed lines). The remembered dot is displaced toward the center of the region into which it falls.

Generalizing Huttenlocher’s model to configural representations, people’s memory of the location of landmarks within their own neighborhood should be distorted toward the center of the neighborhood. Presumably this could be tested by asking people to conduct mental triangulation studies, as described above. The locations inferred from the mental triangulation data could be compared to actual locations to see if the predicted distortion occurred.

Under certain circumstances, distances in cognitive space display an extremely counter-intuitive property: asymmetry. The distance from A to B is not always equal to the distance from B to A. Asymmetries are found when one of the locations is a landmark and the other is a less memorable object in the vicinity. People estimate the distance from the landmark to the non-memorable location to be less than the distance from the non-memorable location to the landmark (Holyoak & Mah, 1982; Sadalla, Burroughs, & Staplin, 1980). Returning once again to Washington D.C. the Lincoln Memorial is a dramatic landmark. Not far from the Lincoln Memorial there is a far smaller statue of Albert Einstein. Generalizing from the laboratory results, we would expect people to judge the distance from the Einstein Statue to the Lincoln Memorial to be further than the distance from the Lincoln Memorial to the Einstein statue.

Why should this be so? Sadalla et al. argue that when people are asked to estimate a distance from an initial point X to a target point Y they “cognitively locate” themselves at the initial point, and then (mentally) transit to the target point. The effort required to do this is used as a cue to distance. If the initial point is a landmark used to organize spatial knowledge, the task is easy because the initial location is easy. Conversely, if the initial point is not a reference point in a person’s mental representation of a neighborhood the task is difficult, and a longer distance in inferred. This explanation, which has not been tested extensively, offers a mechanism that might produce the observed asymmetry.

Sadalla’s explanation is restricted to memory for location. Holyoak and Mah (1982) have suggested that spatial assymetries are a special case of the general finding that under certain circumstances any similarity judgments may be assymetrical (A. Tversky, 1977). For instance, university students in the 1970’s regarded North Korea as more similar to the People’s Republic of China than China was similar to North Korea. Tversky pointed out that linguistically, a similarity comparison is not symmetrical. When we say “How similar is X to Y” we implicitly assign X a target role and Y a base role, and the two are not interchangeable. He then developed a model of the comparison process in which similarity judgments are made by considering what fraction of the features known to be applicable to the target are also applicable to the base. Symbolically, let X ={x} be the set of features applicable to object X, Y = {y} the set of features applicable to Y, and let N(X) be the number of features in the set X (or Y).[3] The similarity of X to Y, S(X,Y), is defined as

(6)
$$S(X,Y) = \frac{N(X \cap Y)}{N(X)}$$

Now imagine a person who knows more about China than North Korea, so that *N(China) > N(North Korea)*. It would follow from (6) that *S(North Korea, China) > S(China, North Korea)*.

Holyoak and Mah (1982) interpreted the question, *How close is X to Y?* (and its variants) as being analogous to *How similar is X to Y* except that only geographic features are considered when answering *how close*. Figure 10 presents their argument graphically. Imagine that our knowledge of the geographic region around a landmark X extends to knowledge about the area immediately around the landmark, including related points Y, Y', and Y''. On the other hand, knowledge about a non-landmark point may extend only to information about the point itself, the adjacent landmark, and the path between X and Y. Therefore equation 6 applies, and psychologically, non-landmark Y is closer to landmark X than landmark X is to non-landmark Y.

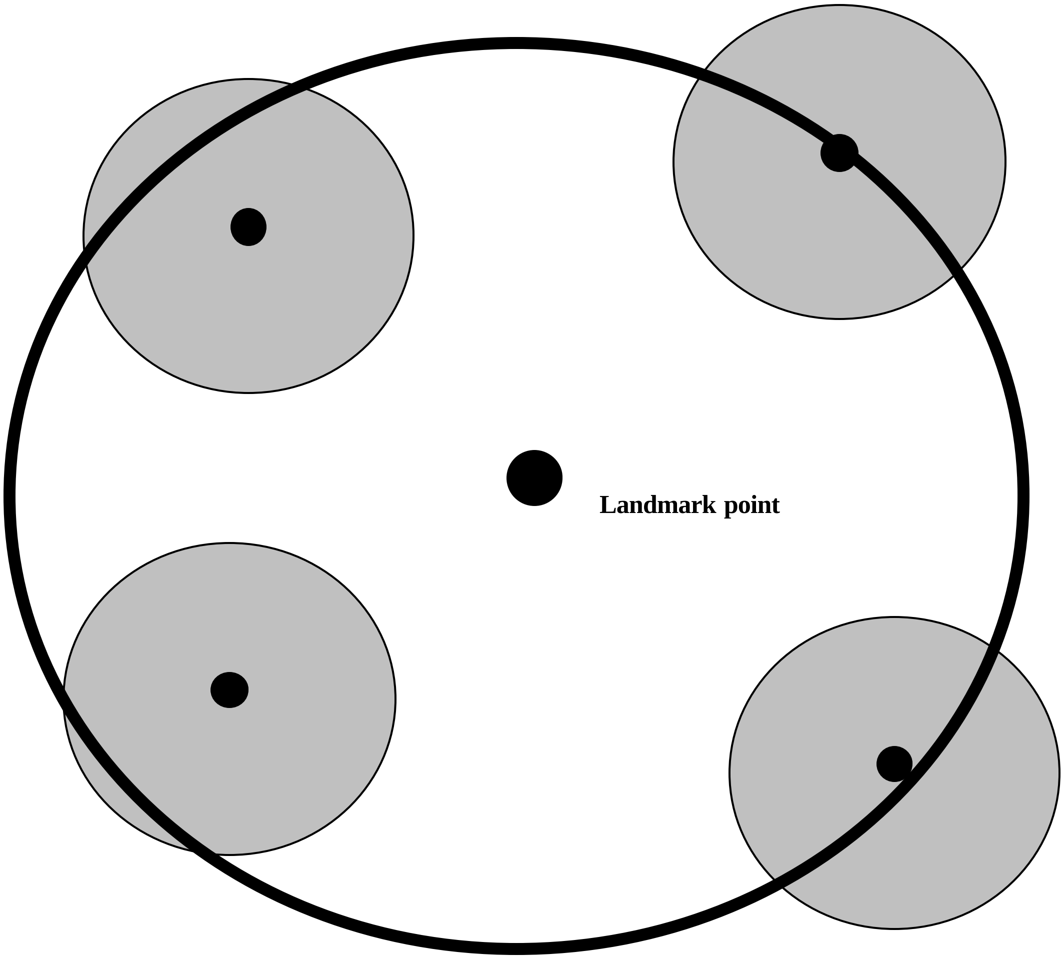


Figure 10. A graphic presentation of the A. Tversky (1977) model, as applied to distance judgements. The open circle indicates features associated with the landmark point. The gray areas indicate the information associated with non-landmark points within the neighborhood of the landmark. The fraction of points associated with a non-landmark point that are also associated with the landmark is greater than the fraction of points associated with the landmark that are also associated with any one non-landmark.

A similar explanation can be offered to explain why two points in the same neighborhood or region are generally considered to be closer than two equidistant points, each in its own neighborhood or region.

Points within the same region should share any features that are associated with the region itself. This would lead us to anticipate (although I have not varied my conjecture) that people would think that Seattle, Washington is closer to Spokane, Washington (in the same state) than Seattle is to Eugene, Oregon. In fact, the two distances are about equal.

Individual differences in orienting skills.

The discussion so far has focused on what 'people in general' do, for the results given were, in general, averaged across participants in an experiment. The study by Bovet (1994) illustrates this. He took university students on a several kilometer bus trip away from their home campus, in daylight and with the windows open. They then got out of the bus and pointed back to the campus. Figure 11 shows the results. While the majority of the estimates were within 30° of the correct direction, some people were conceptually 'turned around,' pointing away from the target location.

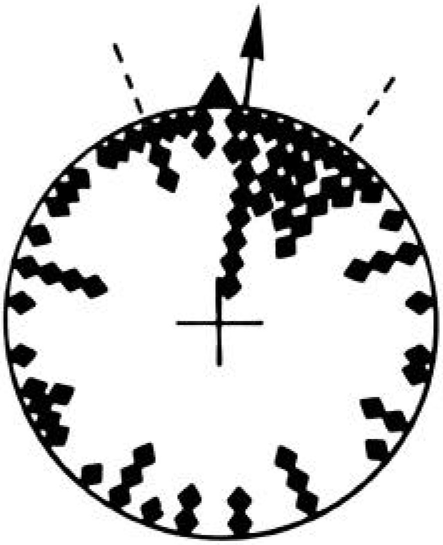


Figure 11. Results from Bovet's (1994) study in which students were taken on a several kilometer drive and then asked to point back to their home campus. The dots indicate individual data points for directions pointed. The small triangle indicates the correct bearing , the large arrow indicates the mean bearing chosen by the students. Although the mean bearing is quite close to the true bearing, a substantial number of the choices of direction were more than 90° away from the true bearing.

Results such as Bovet's are not surprising, given the dimensions of intellectual variation observed in the psychometric literature. Spatial-visual reasoning is one of the basic dimensions of intelligence, as defined by psychometric tests (Carroll, 1993). It is also true, at least on a statistical basis, that there is substantial evidence for a general intelligence factor. Therefore it makes sense to ask how spatial-visual reasoning abilities, in general, and orientation in particular, relate to general intelligence.

Individuals with low to abnormally low IQs (IQ < 70) display poor spatial-visual reasoning, but this is probably a concomitant of their poor general reasoning abilities. It is now well established that general intelligence (g) is more pervasive at the low end of the IQ scale than at the high end (Deary et al., 1996; Detterman & Daniel, 1989). Therefore we would expect someone with abnormally low verbal skills to have poor spatial skills. The converse is not true, people with high verbal skills do not necessarily have high spatial skills.

Turner's syndrome presents a major exception to these general statements. Turner's syndrome is a genetic abnormality in which a woman has a single X chromosome, instead of the normal complement of XX or XY sex chromosomes.[4] Women with Turner's syndrome show normal or near-normal verbal IQ scores but perform very poorly on tests of spatial-visual reasoning. They are particularly weak on tasks that require mental transformations and on tasks that require analysis of configurations of objects. Object recognition does not seem to be impaired (Money, 1993; Murphy et al., 1994 and references therein). The skills that are weak in Turner's cases seem to be those to keep track of locations of objects in a surround. Turner's syndrome patients also show deficits in attention. This may be related to their deficits in spatial reasoning, for it has been suggested that spatial reasoning is an attention-demanding task. There are numerous reports of failure of attention in Turner Syndrome women. I have been unable to locate any studies that show, directly, that Turner's syndrome patients are poor wayfinders, although the literature certainly suggests that this is the case.

Age differences in spatial orientation

Children under ten easily become lost. Thereafter spatial orientation skills evidently develop very rapidly. By age 12 children learn as much as adults do from a guided walk through a new environment (Cornell et al., 1994) It is not clear whether this is due to the ontological progression of orientation skills as such, or whether it is due to maturation in the ability to control attention. Cornell et al. remark (pg. 637) that younger children became tired of orienting tasks after about an hour. Siegel (1981) reported that children's performance improved markedly if their attention was directed to landmarks that adults had selected rather than landmarks other children had selected.

As people enter late adulthood and old age, their performance on tests of spatial-visual ability deteriorates (Salthouse, 1991). Popular lore suggests that the elderly are poor wayfinders, although objective reports are lacking. As was the case with Turner's syndrome, most of the research is based on performance on paper and pencil psychometric tests said to relate to spatial orientation, rather than tests in the field. Since the correlations between test and field performance are far below one, it is possible that the deficiency in test performance is not carried over into the field. It is clear that there are very large individual differences in the loss of spatial abilities in the aged. Spatial disorientation is frequently observed in "middle-old" (i.e. post 70) adults who show other signs of mental deterioration. This can range from the major deficiencies associated with Alzheimer's dementia to substantial but remediable problems in becoming familiar with a new environment. Quite a different picture is obtained from observations of healthy "young-old" groups, especially those who regularly take part in activities that demand spatial skill. Elderly hunters recover from losing their bearings as well or better than younger hunters (Hill, 1992). Many senior commercial aviators are in their fifties, and the Air Line Pilots association has waged a campaign to raise the enforced retirement age from 60 to 65. These observations suggest that age-related declines in spatial orientation may at least partly due to specific injuries and disease processes that are statistically associated with aging, rather than being a concomitant of the aging process itself.

The influence of gender on spatial orientation

There has been a great deal of discussion of male-female differences in wayfinding. As in the case of aging, many of the assertions about male superiority in orientation actually refer to superiority in performance on paper and pencil tests said to measure spatial orientation. However several studies have demonstrated male superiority in the acquisition of configurational information outside of a test setting. Settings studied include retention of information after examining a scene (Arthur, Hancock, & Chrysler, 1997), exploring an area (Matthews, 1987), and traversing a pre-specified route (Anooshian & Young, 1981; Lawton, 1994, Lawton, Charleston, & Zieles, 1996). The male advantage in acquiring configurational information may at least partly be due to a difference in the strategy used during wayfinding. Men report noticing bearings to landmarks, while women report strategies that depend on describing control points and noticing cues to the route, such as street signs (Lawton, 1994, 1996b). In the terms used earlier, women tend to use strategies appropriate to tracking and piloting, while men use strategies appropriate for navigation. The observed strategy difference is an explanation of male-female differences in orientation at one level; if men and women are attending to different cues during a traverse they will pick up different information about the neighborhood.

Individual differences in strategies exert powerful effects on scene perception. Kearins (1981) reported an interesting demonstration of this point for memories of visual scenes. Grade school children from two different cultural groups, Australian aboriginal children and Australian children of European descent viewed a spatial layout of small objects, and then attempted to reproduce it from memory. In general, the Aboriginal children were better at the task than the European-Australians[5]. Kearins attributed this to the Aboriginal children's adapting a strategy of staying very still and concentrating on forming (and remembering) a clear visual image of the scene. By contrast, the European-descent children moved about and named the objects, thus adopting a verbal description strategy that did not work very well for the particular scenes that Kearins used.

Moving from the Australian outback to modern cities, McDonald and Pellegrino (1992) offer an anecdote about a teen-ager who never learned the route to school although he was driven by his parents along this route almost daily. This anecdote suggests, at the least, that it is possible to experience a route without even acquiring piloting knowledge of it. This may be extreme and, after all, the anecdote is only an anecdote. However we have already seen well-documented reports that people can spend years in a building without acquiring configurational knowledge of it (Moeser, 1988; Thorndyke and Hayes-Roth,1982). On the other hand, if people are instructed to concentrate on locating landmarks (distinctive pictures) within a building they acquire some configurational knowledge with very little exposure (Anooshian, 1996). It may well be that some people (mostly men) have developed the habit of noticing the sorts of cues that lead to good orientation, while other people have not.

As the reader may have noted, most of the literature on individual differences in spatial orientation has concentrated on associations between wayfinding ability and other characteristics of the individual, such as age, sex, or cultural background. In a doctoral study conducted in our own laboratory, Infield (1991) examined specific spatial-visual capacities and strategies that were associated with superior wayfinding. Infield contrasted the behavior of very good wayfinders to the behavior of 'average' wayfinders, students enrolled in introductory psychology courses. Infield's "very good" wayfinders were competitors in the sport of orienteering racing, in which the runners race from one control point to another, following a map route rather than one marked on the ground. [6] Infield identified two characteristics of good, compared to average, way finders. Good wayfinders were markedly superior to average wayfinders in their ability to quickly co-ordinate spatial information contained in separate views of an environment, as evaluated by the Guilford-Zimmerman (1948) orientation test. Figure 12 presents test scores for different groups of examinees, ranging from undergraduate Psychology students, who are presumably average wayfinders, to the outstanding wayfinders who compete in international races. We do not know the extent to which the orienteers' perspective taking performance is innate or due to their having practiced perspective tasks, although both innate ability and practice are certainly involved.

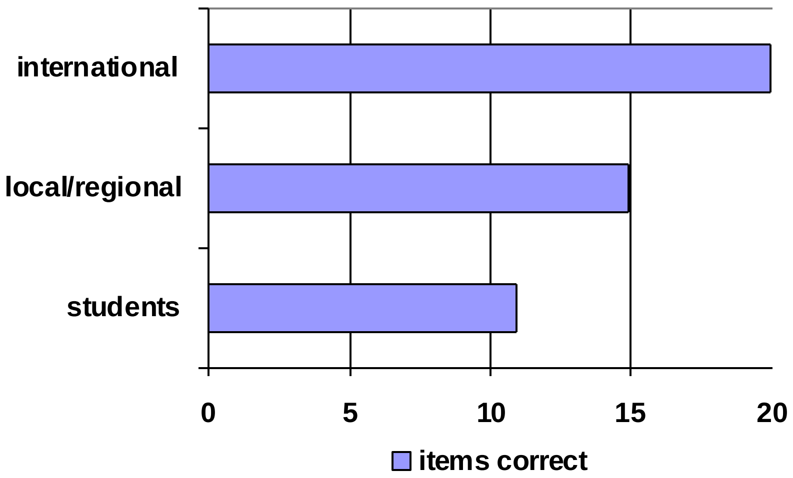


Figure 12. Scores on the Guilford-Zimmerman test of spatial orientation obtained by undergraduate psychology students, and sports orienteers at regional and international competitive levels (Infield, 1991; experiment 4). The scores have been corrected for guessing.

In addition, Infield found one orienteering trait that is almost certainly learned. Orienteers and psychology students were shown a drawing of a crowded street scene, and then tested for their memory of it. The orienteers reported more location invariant items (e.g. street lamps as opposed to dogs) than the Psychology students did. This was not due to the orienteers' having better visual memory than the students, for the two groups were no different in their ability to recall a still-life picture of objects on a shelf.

Artifacts

Donald Norman's book (1993) *Things that make us smart*, describes a variety of artifacts for enhancing thinking, ranging from computers to notepads. Some are used to enhance orientation. We will consider three of them. Two -- maps and language -- are ancient. The third, the virtual environment, is a product of the computer age.

Maps

Maps come in two broad types. *Plan views* use the conventional birdseye view of a region as seen from above, whereas *perspective views* show how things might look to an imaginary observer following a route. Both perspective and plan maps have a long history. The fourth century astronomer, Ptolemy, drew a credible map of the countries surrounding the Mediterranean (Olson, 1994).

Maps use visual properties of a drawing to represent abstract geographical and cultural information. For example, alignment is represented such that 'Up' on the paper represents a particular bearing on the ground. Most modern maps follow one of two alignment conventions. Up is either North or, in the case of 'you are here' maps that can only be viewed from a certain position up may be aligned with the reader's line of sight. Bearing information is maintained on a map if for any three arbitrary points, A, B, and C, the angle <ABC on the map is identical to angle <ABC on the ground. This preserves shape information. Metric (distance) information is maintained if the length of the line AB between arbitrarily chosen points A and B on the map is proportional to the distance between locations A and B on the ground. Strictly speaking, it is not possible to maintain both bearing and distance on a flat map, since the Earth is (almost) a sphere. However the distortions are so small that they can be disregarded for distances less than several hundred kilometers.

Lloyd (1993) observed that the interpretability of a map depends upon how well visual features of the map transmit information about the region that the map depicts. Interpretability and detail of information can conflict. Bearing and distance information are much more clearly represented in plan than perspective views. On the other hand, it may be easier to associate perspective views with visual scenes experienced on the ground. Probably for this reason, preschoolers prefer perspective views to plan views (Liben & Yekel, 1996).

There are three related ways of using maps. Most obviously, maps can be used to assist in exploring. Maps can also be used as study devices, to acquire information about a region without actual exploration. By convention, psychologists refer to this as obtaining geographic information from a secondary source. Finally, a map can be used as a reference in order to give verbal directions to someone else. We look at each in turn.

Maps as guides to exploration. The first step in using a map as a guide is to establish a correspondence between position and bearing on the map and location and direction on the ground. The convention in our culture is that up on a map means north on the ground. However this can be bad advice to give a wayfinder. Beginning orienteers are told, 'hold the map so that up on the map corresponds to forward on the ground (forward = up).' Establishing correspondence is much easier when this is done (Levine, 1982; Levine, Marchon, & Hanley, 1984). The reason is illustrated in Figure 13. If the map is correctly aligned, bearings from the top of the map (e.g. 45° to the right of up) correspond with a right or left angle on the ground (45° to the right, from the direction you are facing while reading the map). To the extent that the map is misaligned, the reader must transform a direction on the map into direction on the ground. Laboratory studies of simplified you-are-here maps (Levine et al., 1982; Shepard & Hurwitz, 1984) have shown that this is a time consuming and error-prone operation. Infield (1991) found that orienteers were faster and more accurate than controls in solving contra-alignment problems.

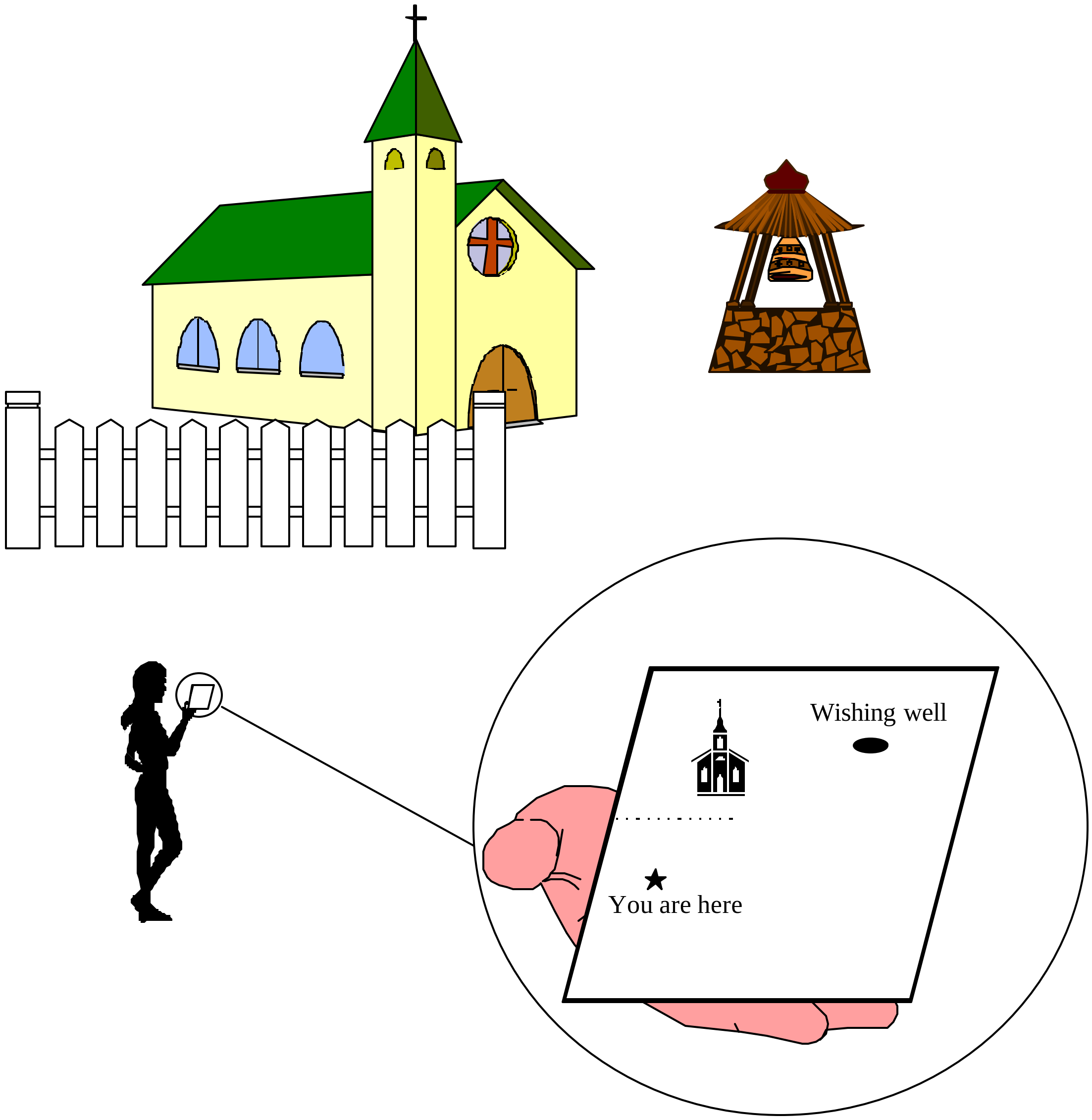


Figure 13. A properly aligned you-are-here map. Up on the map corresponds to forward on the ground, and the absolute frame of reference of the map is aligned with the egocentric relative frame of reference of the map reader.

Map reading also involves interpretation of abstract codes representing different ground features. These include the use of contour interval lines to represent height and the use of false color images to emphasize features vegetation levels in aerial photographs. Learning to use these codes is no mean feat. During the 1990's, the U.S. Marines developed a course in topographical map reading for non-commissioned officers. The course was effective, but it took fifteen hours (Tkacz, 1998).

Maps as substitutes for exploration. Maps are frequently used as secondary sources of geographic information. Sometimes this is done as a matter of convenience; it may be useful to study a neighborhood map even if you plan to go there. Other times maps are the only possible source of spatial information, either because of the inaccessibility of the neighborhood (e.g. planning a military raid or rehearsing a fire fighting procedure) or because the size of the region precludes personal exploration (e.g. the Western Hemisphere). Maps are not only useful for this purpose, they are in some ways superior to exploration because they highlight configurational information. We have already seen that people who have considerable experience with a neighborhood may still not acquire configurational knowledge of it (Moeser, 1988; Thorndyke & Hayes-Roth, 1982). On the other hand, only minimal map study can produce good configural information.

Lloyd (1989) conducted a study that showed how strong the map reader's advantage can be. University of South Carolina students made distance and bearing judgments between landmarks in Columbus, S.C., the city in which the university is located. All participants in this phase of the experiment had lived in Columbus for more than two years. A second group of South Carolina students studied a map of a city that they were told was Fargo, North Dakota., In fact the map they studied was a map of Columbus, reflected around its horizontal axis, and with landmarks renamed (e.g. the center of the campus was labeled a bus stop). After a few minutes study the 'Fargo' group made distance and direction judgments based upon their memory of the map. The results of the Fargo and Columbus groups can be compared, for every direction and bearing judgment in one task can be mapped to a complementary judgment in the other task. The 'Fargo' judgments were markedly more accurate than the Columbus judgments. A few minutes of map study produced more accurate configural knowledge than years of first hand experience!

While Lloyd's study shows the value of a good map, it does not show that people would be better wayfinders if they studied maps instead of exploring neighborhoods. The explorers would become acquainted with landmarks and tracks, and thus would have piloting information that would not be available to people who only knew the neighborhood from maps. Just this effect has been observed. People who have learned about a building from exploration are faster at following routes through it than people who have studied a building plan (Witmer, Bailey, Knerr, & Parsons 1996).

The alignment effect that applies to map reading also applies to memory for maps. When people have learned a neighborhood layout from maps they make directional judgments more easily if the pointing direction on the ground is aligned with the pointing direction on the map, presented in its normal orientation (Evans & Pezdek, 1980). However there are strong individual differences in this respect. Some people seem to be able to ignore alignment effects (Rossano, Warren, & Kenan, 1995). The relation between this ability and other measures of individual differences in visual-spatial abilities is not known. What we can say is that when people interact with maps (including imaging maps from memory) they are interacting with a medium that highlights configural information. McDonald and Pellegrino's (1993) principle of interaction applies both to the space itself and to its secondary representation in a drawing. What people remember about a space depends upon how they interact with it.

This paean to cartography should not bias us to some defects in acquiring information from maps. One, the alignment effect, has already been mentioned. B. Tversky (1981) has documented another. Map readers remember geographic regions as having their major axes roughly parallel to the sides of the map, and remember different regions as having their axes aligned with each other. American college students think of the coast of California as being more North-South than it actually is, and think that North and South America lie along the same longitude, when actually most of South America is considerably east of North America. The reported distortions indicate both that adjacent regions have been aligned and that directions have been distorted to line up with the sides of the map. This is not just a matter of “Americans not knowing geography,” for similar effects have been observed in Israel, using maps of that rather small country (Glicksohn, 1994).

These errors in memory are consistent with what seems to be fairly common errors in geographic knowledge. Just as we remember neighborhoods as being more regular than they are, we regularize maps in our memory. For instance, U.S. students claim that Reno, Nevada is to the east of San Diego, California (Stevens & Coupe,1978). Actually it is northwest of San Diego. This error could be explained by assuming that in their memories, students align the state of Nevada with the state of California, and straighten out the turn taken by both the California coastline and the southern part of the California-Nevada border. However there is an alternative explanation, based on Huttenlocher, Hedges, and Duncan's (1991) argument that when people make judgments about locations they blend what they know about the properties of specific locations with what they know about the properties of regions containing those locations. Therefore, since Nevada does generally lie to the east of California, it is easy to confuse the longitudes of San Diego and Reno.

One other distortion should be mentioned. Thorndyke (1982) showed that when people judge distances from maps they are sensitive to the number of intervening points; distances seem longer if there are more points. This may be the reason why American students are surprised to find that the distance from New York to Washington D.C., (with an almost continuous line of intervening cities, including the metropolises of Philadelphia, and Baltimore) is considerably less than the distance from San Francisco to Los Angeles.

<u>Maps as the basis for directions</u>. The third way we use maps is to give directions to other people. Direction giving shows a good deal of commonality between map use and actual wayfinding. Below age 10, children are generally confused when asked to give directions from a map; however, by age 12 they approximate the accuracy of adult performance (Blades & Medlicott, 1992). When adults give route instructions from maps they vary markedly in the extent to which they emphasize navigational instructions (directions and bearing) and the extent to which they emphasize piloting instructions, specifying locations, paths, and turning. The difference appears to be partly due to the extent to which the map-readers consciously code and rehearse

configural patterns, such as a triangle formed by three intersecting roads. As would be expected, readers who encode this way are generally more accurate than those who do not (Thorndyke & Stasz, 1980).

There are male-female differences in the way that maps and map routes are described. Women are more likely than men to give piloting rather than navigational instructions. This is what would be expected given male-female differences in wayfinding itself (Dabbs et al., 1998; Galea & Kimura, 1993; Ward, Newcombe, & Overton, 1986). The effects are so pervasive that they have been given the ultimate recognition in our society: discussion in a comic strip.[7]

Maps are useful ways of representing environmental information. Indeed, for the purposes of acquiring configural information they may actually be preferable to exploration. However maps are not perfect. The tasks of establishing correspondences between own location and bearing and map location and bearing are not trivial operations. When we operate from our memories of maps we find that they, like all visual objects, are subject to distortions in both perception and memory. This can influence wayfinding decisions. Nevertheless, no sensible person would to do away with maps. If we did, people would have to rely on verbal descriptions of regions. These have their own costs and benefits.

Representing spatial information with language

Language has been described as a device by which a speaker builds a mental space in a listener's mind (Fauconnier, 1997). Here is an excerpt from a modern detective story, in which the narrator describes a traffic accident:

> *We fishtailed, then skidded to a stop crossways in the road. Brakes screeching, wheels shuddering, the hay truck halted further down the hill, and as it did so the second trailer rolled slowly onto its side like a great beast lying down for a nap. The truck and both trailers ended up blocking the highway, the cab tipped on two wheels and twisted at an angle so that should the driver open his door, he would fall six vertical feet to the pavement. Dozens of exploded hay bales mulched the scene. The wheels on the upended trailer continued to spin. A hundred yards up the hill two cars braked carefully on the slick pavement.*
>
> Emerson (1995, pg. 13)

The passage evokes a description of an accident. But just what information does the image contain? Suppose a reader was asked, *At the end of the incident, was the narrator's car between the truck and the two cars?* The answer is implied by but not contained in the passage. All the passage says, directly, is that

> The narrator's car stopped in the road.
> The truck stopped further down the hill.
> The two cars stopped further up the hill.

In separate approaches that have reached much the same conclusion, Walter Kintsch (1998; Van Dijk & Kintsch, 1983) and Philip Johnson-Laird (1983; Johnson-Laird & Byrne, 1991) have equated comprehension of a text with the creation of a mental model of the situation being described. This is another situation in which the distinction between early computation and late computation applies. Readers create models as they read a text. Later, when asked a question about the text, the reader recalls the model from memory. Therefore the difficulty the reader has in answering the question will be a function of the ease of retrieving the required information from the model. In the case of the accident passage, a reader could answer a question about 'betweeness' directly if, at the time of reading, the reader had created a mental model that contained information about the relative locations of the cars. The question would have to be answered indirectly, by computations initiated by the question itself, if the reader's mental model did not contain configural relations but did contain information from which configural relations could be inferred.

Readers obviously have considerable latitude in their choice of the mental model that they construct. Their choices determine the extent to which questions are answered by early or late computation processes. Descriptions are propositional statements about the space in question, as in the illustration just given. The person receiving the description (the *comprehender*) could store this information as a set of propositions (including some inferred propositions). This minimizes the amount of early computation, but may present a problem when late computations are required to answer a question. The problem would be acute if a key proposition had been forgotten. Alternatively a person could construct images from the description and remember them. This increases the early computation burden, but, at some later time, the images could be used to infer configural information. Finally, a listener could compute a configurational model of a situation as it was described. This maximizes the amount of early computation but, if the correct information is included in the configurational model, it may minimize the amount of late computation. In the automobile accident illustration just given an observer who calculated a configurational model of own position, truck position, and car position could answer 'betweeness' questions by reading the information directly from the model.

This principle can be illustrated by considering one of the first experiments offered to support the mental models approach to comprehension (Bransford, Barclay, & Franks, 1972). College students heard sentences such as:

Three turtles rested on a floating log and a fish swam beneath them or

Three turtles rested beside a floating log and a fish swam beneath them.

Subsequently the students were presented a test sentence:

Three turtles rested beside a floating log and a fish swam beneath it.

And asked if they had heard it.

Suppose that listeners had constructed a mental model from the first sentence. If the original sentence was “on the log” the model would have show that a fish had swum under the log, for how else would it get under the turtles? On the other hand, if the original sentence had been “beside the log” the model would be neutral about whether or not the fish swam under the log. Computationally, in the first case the test sentence could be verified by inspecting the model, in the second case some sort of inference about probability would be required. The data conforms to this, for people mistakenly classified the test sentences as having been heard in the “on the log” case, but rejected in the “beside the log” case.

The situations studied by Bransford et al. were at about the level of complexity found in a book to be read to kindergarten or first grade children. Barbara Tversky and her colleagues extended the paradigm to the more complex task of determining what sort of models are developed when people hear descriptions of surrounds and neighborhoods. One of their studies (Franklin & Tversky,1990) has already been mentioned as an analog of Hintzman et al.’s (1981) work on memory for surrounds. We now look at it in more detail.

Students first read descriptions of complicated surrounds. An example is this brief description of a ‘night at the opera;’

> …you are standing next to the railing of a wide, elegant balcony, overlooking the first floor. Directly behind you, at your eye level, is an ornate lamp attached to the balcony wall. The base of the lamp, which is attached to the wall, is gilded in gold. Straight ahead of you, mounted on a nearby wall beyond the balcony, you see a large bronze plaque…

After reading the description the students were asked to imagine themselves facing one object, and then to identify the objects to their front, back, right, left, or above or below them. The situation was analogous to Hintzman et al.’s (1981) study, with the addition of an above-below dimension. Franklin and Tversky obtained the same M shaped reaction time profile that Hintzman had obtained when observers actually were in a surround. Objects to the front (in their imagined orientation) were located faster than those behind. The slowest responses were obtained to queries to simulate virtual motion. Objects that fell along the above-below dimension, which had not been included in the Hintzman et al. study, were even more rapidly located. This pattern has been replicated in other studies (Bryant, Tversky, & Franklin, 1992; Franklin, Tversky, & Coon, 1992), and appears to be quite reliable. Tversky (1991) has argued that the effect is due to the front-back dimension falling into two strongly marked linguistic categories: the usual case of front and the exception of back. She further claims that the up-down dimension is privileged, because it is constant in most situations, correlated with gravity, and aligned with the major axis of the body.

In one of the most recent studies in this series Bryant, Tversky, and Lanca (in press) compared location responses based on memories to location responses made in a perceptual condition, in which the subject was actually in the surround. They found that the M shaped pattern of response times only appeared when the observer seemed to be making a response from memory. If a participant actually looked about (‘scanned’), response times increased as a function of the angle turned. Therefore, responses to 'back' were slower than responses to 'left' or 'right' on scanning trials. It follows from Bryant et al.'s work that the spatial properties of mental models are not identical to the properties of the space being modeled. This is an important conclusion. Therefore, while there is no reason to be unduly suspicious of the Bryant et al. result, replication studies would be highly desirable.

One of the advantages of reading over being there is that the reader can be asked to view the same area from multiple perspectives, while the observer is constrained to his or her own position. The following description of a location is paraphrased from Franklin et al. (1992)

> *Ted is to your left. To your front, a pitchfork lies on top of some crates. To Ted's left a loosely wound coil of rope droops from a wood peg in the barn wall.*

The narrative switches back and forth from 'you' to 'Ted.' Franklin et al. found that people who read such descriptions develop a single spatial model including information received from each perspective. This shows acceptance of a rather heavy early computation burden in order to avoid the burden on memory that would be associated with storing information about the same space, seen from the perspectives of different observers.

Text can be used to describe neighborhoods as well as surrounds. The descriptions can be written in a way that emphasizes either relative or absolute frames of reference. Returning again to Washington, D.C., the Mall area could be described as

> *Standing on the Capitol steps you can see the see the National Gallery of Art. Walking by it you can see the Air and Space museum to your left front…*

or

> *The Mall is bounded on the east by the Capitol. Just west of the Capitol is the National Gallery of Art. Beyond it the Air and Space museum is the easternmost of a line of buildings on the south edge of the Mall, that includes…*

In theory, a listener could determine the propositions stated in the text, relate them to each other, and memorize the resulting structure. Kintsch (1998; Van Dijk & Kintsch, 1983) refers to this step in comprehension as the development of a text based model. If comprehension stopped at this point, questions about the neighborhood would be answered either by finding the appropriate proposition in the text base model (early computation) or by computing inferences of the propositions (late computation). The alternative, which is what both Kintsch and Johnson-Laird claim people actually do, is to develop a situation model describing the neighborhood in spatial terms. The evidence for this is that people answer questions requiring inferred knowledge about as quickly as they answer questions requiring knowledge stated directly in the text (Perrig & Kintsch, 1985; Taylor & Tversky, 1992a). An example in the Mall description would be the inferred fact that the Air and Space Museum is west of the Capitol, compared to the directly stated fact that the National Gallery is west of the Capitol. This is taken as evidence that the inferred knowledge was incorporated into a situation model during the early computation stage.

Does it make any difference what sort of model a description invites? The evidence is somewhat equivocal. Perrig and Kintsch (1985) reported that inferences were made more quickly if the form of the question (route or survey) matched the form of the original description. Generalizing their results “Looking from the Capitol steps, is the Air and Space Museum on your right or left?” would be answered more quickly if the participant had read a route description. A similar study by Taylor and Tversky (1992a) did not find the same statistically reliable interaction between description type and inference type, but examination of the pattern of their data shows that it is not inconsistent with Perrig and Kintsch’s results. A safe, although not strong, conclusion is that route descriptions probably lead to route representations, and survey descriptions to survey representations, but that more definitive data is needed. Individual and situational differences may also play a role. We return to this point below.

Language can be used to describe maps as well as to describe routes. Denis (1996) found that when people read descriptions of maps, their reactions to questions are similar to those of people who answer questions from memory after studying actual maps. Kosslyn, Ball, and Reiser (1978) have found that when people are asked to attend in succession to two locations on an imaged map (e.g. New York, then St. Louis) the time taken to shift attention increases with the distance of the shift. A person imaging a map can shift attention from New York to St. Louis more rapidly than shifting attention from New York to Los Angeles. Denis and his colleagues showed that the same phenomenon occurs in situations in which a person reads or hears a description of a map, but never actually sees it.

Just as texts can be used to establish mental models of a space, texts can be used to guide access to information in a mental model. We have already seen a rudimentary example. McNamara’s studies of priming showed that asking a question about one building in a familiar neighborhood facilitates access to information about a nearby building. In this work the ‘text’ involved was rudimentary, questions were presented in isolation, as single sentences. Morrow, Greenspan, and Bower (1987) have developed a widely used paradigm for studying how people access spatial information while comprehending more complex texts. Participants first study the floor plan of a hypothetical building. After the participant has fixed the map information into memory he or she reads a story about a character going through same building. Periodically, the story is interrupted and two object names are shown. To illustrate, suppose that the story was about a character named ‘Joe.’ A participant might read

> *Joe turned away from the window and entered the Green Room. There he*
> *ARMOR – PAINTING*

The task is to say, as quickly as possible, whether or not the objects are in the same room. Morrow et al. found that positive decisions were made more rapidly if the objects were in the room that the character was in at the time of the question. This result has been replicated. (See Zwann & Radavansky (1998) for a detailed review).

Morrow et al.'s results and related research can be used to argue that readers utilize prior knowledge to develop a model of the situation implied by a text. No one would dispute this. On the other hand, these results cannot be used to argue that the reader constructs an image of a person going through a space. There are two reasons to doubt that imaging invariably occurs. One is logical. It is possible to simulate these results by passing activation through items in a connectionist network, in which objects within a room are tied logically to a single superordinate node. Spatial distance is not represented within this model, although in-room relationships are (Haenggi, Kintsch, & Gernsbacher, 1995). The second reason for doubting the imaging model is empirical. When a readers' attention is directed to one of several items within the same surround (e.g. a room) the reader is not especially quick to respond to questions about nearby items. If the readers were inspecting a mental model you would expect them to be quick to respond to a question about an item next to the item that they are currently looking at in their image. However, they are not. (Langston, Kramer, & Glenberg, 1998).

How do people use language to transmit spatial information? One strategy is to tell the interrogator what they would see if they were there. This is what New Yorkers did when asked to describe their apartments (Linde & Labov, 1975). They took the experimenters on a 'gaze tour,' in which they described what a person would see when they entered and looked about a room. This finding is not trivial. Although people could describe an apartment by using an absolute frame of reference, the floor plan, they clearly preferred the relative frame of reference.

The same principle can be applied to descriptions of neighborhoods, but somewhat different results are found. How the neighborhood is described depends upon how it has been experienced. When people have learned about neighborhoods from maps their descriptions show the familiar hierarchical organization, first one small region is described and then another. The analogs to gaze tours and plans of a room are route and survey descriptions. When neighborhoods have been learned from maps they are described with a mixture of route and survey terms. People will say something like "The town hall is north of the statue, and if you turn right at the statue you will see the museum." On the other hand, when people are asked to describe a neighborhood that they know from personal exploration they tend to give route descriptions (Taylor & Tversky, 1992b; 1996).

These results, taken together, suggest the following general principles. In general, when people are asked to describe regions they respond by trying to give the questioner the experience of exploring the region. If this is the (implicit) goal it is only sensible to describe the region from the viewpoint of a traveler moving through it. On the other hand, if the learning situation has emphasized an absolute frame of reference, as examining a map does, people can provide survey representations. It is also worth remembering that if people have learned about a space from personal experience they may not have developed a surveyor's model of it! Such a person would have to respond by describing the space from an egocentric frame of reference.

Once again McDonald and Pellegrino's (1993) axiom holds true. If you want to know what people have learned about an environment, examine the way in which they interact with it. In fact, we can expand the principle. How people will describe an environment depends upon what they think the interrogator needs to know. Any conclusion about how 'people' deal with interactions between spatial orientation and language has to be qualified by statements about who the people are and what they are required to do.

This conclusion suggests that people have a good deal of flexibility in dealing with mental representations of space. There is considerable evidence documenting individual differences in willingness to use various forms of description, possibly related to different abilities to deal with visual-spatial information. In addition flexibility can be found within a single individual.

Both proclivities and flexibility can be demonstrated using a very simple visual-spatial situation. MacLeod, Hunt, and Mathews (1978) showed college students descriptions of rudimentary visual scenes, such as *The star is above the plus* or *The plus is not below the star.* Subsequently the students were shown a picture of a plus above a star or a star above a plus, and asked whether or not the sentence correctly described the picture. People with high visual-spatial ability, as determined by psychometric tests, would image the picture they expected to see, and then respond if the actual picture matched the image. People with high verbal abilities preferred to memorize the sentence, and then, on seeing the picture, determine whether or not the sentence described the picture. This trait was not a fixed characteristic of people. A subsequent study showed that college students could switch strategies upon request (Mathews, Hunt, & MacLeod,1980). Proclivities toward particular forms of models for text can be found in tasks involving large scale space. Denis (1996) found that the ability to abstract spatial representations from descriptions intended to induce a survey representation was related to spatial-visual ability. Haenggi et al. (1995) report a similar finding when people are asked to "follow the character" during Morrow et al.'s perspective taking experiment.

The spatial-visual tests used in these studies all involved the imaging of movement (e.g. a spatial rotation test). These are the tests that are most closely related to facility in manipulating visual images (Poltrock & Brown, 1980). This does not mean that 'high spatial' people are producing conscious images, although in some cases they may do so. The finding does complement findings showing that the higher order visual processing system is heavily involved in both imaging and spatial orientation. Since tests involving mental imaging show considerable male-female differences, in favor of males (Voyer, Voyer, & Bryden, 1995) it is not surprising to find that there are male-female differences in descriptions of neighborhoods. Perrig and Kintsch (1985) reported that women preferred to receive route than survey descriptions, while men were indifferent. Women are also more likely to give route descriptions than men are. Statistical analyses have linked the male-female difference in style of reporting to male-female differences on tests of spatial imagery skills (Dabbs et al., 1998).[8]

These results all speak to the issue of capability; people develop the sorts of spatial mental model that they find easy to use. Intentions also count. McKoon and Ratcliff (1992), among others, have pointed out that when people read a passage they do not draw all the inferences that they could, nor do they activate all prior knowledge that could possibly be relevant. In the terms that have been used here, early computation is restricted to those inferences that the reader thinks are likely to be of use later. In most of the experiments discussed here, the procedure was such that participants had every reason to expect questions about location, and therefore would have been rational to stress spatial information during the early computation stage. However there are many situations in which the spatial information contained in a text is not central to comprehension. A series of studies (cited by Zwann & Radavansky, 1998) have shown that spatial models are developed when they are relevant to a story, but may not be developed when they are not. It would be nice to be able to go beyond these general principles. Unfortunately no one has systematically explored how changes in the mental model of space developed from a text are related to the purpose with which the text is read.

Finally, it is possible that biases for particular types of spatial models are built into the language that a person uses. Spatial terms do not always translate across from one language to another. For instance, in some circumstances the English terms *in* and *on* both map to the Spanish *en*. Conversely, Spanish uses several terms to represent distance from the speaker (*aqui, alli', alla'*) whereas English is restricted to *here* and *there*. More subtly, languages may differ in the extent to which they encourage the use of relative, internal, or absolute frames of reference. In Quecha, a language spoken in the Andes, intrinsic frames of references are based upon body parts. Quecha speakers refer to the ‘hand’ or ‘face’ of a mountain. Hunt and Agnoli (1990) have pointed out that using this system for defining requires agreement about which way the mountain is facing. Therefore it would be difficult to give directions to a first-time traveler to the mountain. See Levinson (1996a) for many other examples. According to modern views of the Whorfian hypothesis (Hunt & Agnoli, 1990) these differences in language should be reflected in behavior. But are they?

Levinson (1996b) has reported an elegant series of studies showing how language controls spatial reasoning in two different linguistic communities: Dutch and Tzeltal, a central American Amerindian language. Tzeltal does not use a relative frame of reference; it has no terms for ‘left’ and ‘right.’ It does have an absolute reference system, based on a vertical dimension-up or down the mountain, that can also be applied to flat ground. Dutch, like English, uses both relative and absolute reference frames. In Levinson’s study Dutch and Tzeltal speakers sat facing a table that had an arrow placed on it. In English (and in Dutch) we would say that the arrow pointed to the right or left of the observer. The observer then turned 180^o, to face a second table with two arrows on it, one pointing to the observer’s right and the other to the observer’s left. The task was pick out the arrow on the second table that “pointed in the same direction” as the arrow on the first. Dutch speakers interpreted this to mean “the direction with respect to me,” and picked the arrow on the second table that pointed to the same side of their body (but in a different absolute direction) as did the arrow on the first table. The Tzeltal speakers interpreted “same direction” as ‘same absolute direction,’ and picked the arrow that the Dutch speakers rejected.

So how do language and space interact? The key word seems to be *flexibly*. Regular relations between language and spatial behavior can be found for specific situations, but the same regularities do not always apply across situations or people. People certainly can use language to extract and convey spatial information. The details of the exchange depend upon the cognitive traits of the speaker, why the exchange is taking place, and the facilities of the language in which the exchange is to be conducted.

Virtual environments (Virtual reality)

The word “virtual” has become so much of a buzzword in our society that the term is in danger of losing any meaning. The popular press talks of virtual cities, virtual certainties, virtual communities, and, as a staggering oxymoron, virtual reality. Here we shall be more precise. The term *virtual environment* will refer to computer generated displays that resemble the views a person would experience if he or she moved through an analogous surround or neighborhood. *Virtual reality* (VR) will be used to refer to the various technologies used to create virtual environments. Note that this limited use of the phrase avoids the oxymoron, for the software and hardware technologies for virtual reality do indeed exist.

Rheingold's (1991) popular book describes a number of the early steps toward the creation of virtual environments up to 1990. Most of these efforts presented sequences of pictures or segments of movies that essentially took people on tours of an area. The effect was analogous to that in the handful of movies in which everything is filmed from the perspective of a narrator, with one essential difference. The viewer could alter the sequence of views. For example, Rheingold describes an M.I.T. production in which a data bank was created containing video segments of pictures taken as a car drove around Aspen, Colorado. Each segment consisted of a driver’s eye view of a drive along one block of the city. At the end of each segment the participant chose whether to drive through an intersection or turn to the right or left. The choice of route determined what block-long segment would be shown next.

Examples like this permitted people to interact with the computer generated display at specific choice points. The same techniques were used in early computer games. Vision could be augmented by sound, but each computer character only spoke a fixed set of sentences. The advent of faster computers and programming techniques changed virtual environment programs in a fundamental way. Instead of having pre-computed sequences of actions programs could contain models of the space being represented. These models are analogous to an architect’s floor plan or an extremely detailed surveyor’s map. The participant begins at a fixed position, and then “navigates” to new positions using an input device to indicate the desired motion and direction of gaze. Keystrokes and electronic pointing devices, such as a mouse or a data glove, allow the participant to signal the computer to simulate virtual motion. Sensors on the legs or body have also been used for navigational input. Direction of gaze can be controlled either by the pointing device or by tracking head movements. The computer uses its data base to generate the view that a person would have were he/she to be standing at the indicated spot, looking in the indicated direction.

While the resulting displays are far more compelling than the displays generated by the interactive movies of the 1980’s, they are also far short of actual experience. Therefore, when we talk about virtual environments, we have to be precise about just what we mean. In doing so, it is useful to distinguish between *representation*s of situations and *simulations* of those same situations. A representation of a situation contains enough information so that certain aspects of the original situation can be re-created, providing that one knows the translation algorithm. However there is only minimal, if any, perceivable similarity between the way the information is held in a representation and the way that it exists in the represented situation. Contour maps are representations; contour lines describe the topography of a hill but no one would ever mistake a contour line for a hill. In simulations, an attempt is made to maintain some perceptual similarity between the form of the representation and the thing being represented. A painting or a photograph of a hill is a simulation in this sense.

The distinction between representations and simulations carries with it an important psychological distinction. Interpreting a representation is a cognitive operation. Interpreting a simulation depends upon cognitive and perceptual operations, with the balance shifting toward perception as the fidelity of the simulation increases. Neither the cognition-perception distinction or the low fidelity-high fidelity distinction are exact, but the relation between them is a sensible one.

With these distinctions in mind, let us look at virtual environments, as they existed going into the third millennium.

Virtual environments depend on three related but distinct technologies. The most common is *desktop VR*. In desktop VR the observer sees scenes on a conventional computer screen. Strictly speaking, what the participant sees is analogous to the scene before the operator of a remote-controlled vehicle with a TV camera mounted on it (*teleoperating*). However the participant is usually invited to imagine that he or she is in the environment rather than being a teleoperator. The observer controls motion through the environment by keystrokes or by the use of an electronic mouse, joystick, or similar pointing device. Sound can be presented using speakers or stereo headsets. (In practice this is little done.) Thus desktop VR, as used in most scientific experiments, is analogous to and often somewhat less detailed than the displays used in sophisticated electronic arcade games.

Desktop VR is so simple that many popularizers might not want to call it VR at all. However, as we shall show, considerable useful research on virtual environments can be done with desktop VR systems. This is important, because desktop VR systems cost perhaps a tenth as much as the more elaborate systems that we now describe.

Immersive VR systems are systems in which nothing intrudes upon the visual field except a scene in the simulated environment. It is generally the case that the scene is not as broad as the visual field, but the periphery is simply occluded. Geometrically, immersive systems, as of the late 1990’s, offered the view one would get while wearing glasses with blinders on the side, rather than the view of a person who was actually in the environment. As of 1999 immersive VR was generally achieved by having the participant wear a head-encompassing helmet that occluded all sights of the room except the scene developed by the computer. There were a few sophisticated systems that projected separate displays to each eye. This made it possible to produce binocular disparity, and hence stereoscopic displays, even though the actual visual object (the screen) was only a few inches from the eye. Most of the helmets were fairly clumsy, and certainly restricted motion. However goggle systems have been developed, and there is every reason to believe that systems as light a conventional pair of eyeglasses can be developed using turn-of-the-millennium technology.

Keyboard and mouse-like navigation interfaces generally do not work well with immersive VR, because the participant can only see the display. Therefore a number of special interfaces have been developed. Head (helmet) movements are used to indicate changes in direction of gaze. Fields of view are usually less than 100^{o}, which is considerably less than the field of view of normal vision, but is much wider than the receptive area of the fovea. As a result, it is possible to shift one's gaze by eye movements, while maintaining a fixed head position, in both normal vision and in immersive VR. However the angular width of the field that can be covered in this way is greater in the real than in virtual environments. This may be important because of the reduced field of vision reduces optic flow, which is a cue to movement. The participant may use hand motions to indicate direction, either by wearing a specially wired data glove or holding a transmitting device (wand) whose position can be sensed by the computer. In a few applications whole body motion is sensed, by having the participant wear a specially instrumented slipper, stand on a balancing platform, or walk on a treadmill. In these cases leg motion can be used to indicate movement in the virtual environment, but the motion is most definitely not normal walking, trotting, or running.

Caves constitute the third, and perhaps least known, VR technique. An electronic cave consists of one or more wall-sized computer-controlled display systems. The observer in a cave is actually in a room looking at life-sized projections of houses, people, etc. on the large displays. The user's position and orientation are transmitted by sensors set on the head, body, and by the use of instrumented slippers and devices analogous to wands. For example, a VR cave developed for infantry squad training by the Department of Defense senses the location of the participant and an instrumented 'rifle' that he carries.

Insofar as I have been able to determine, no psychological experiments using VR caves have been reported in the usual scientific literature. This is understandable, because VR caves require expensive hardware and software. It is interesting to note that simulators used in the aircraft industry are very similar to VR caves. For instance, in a commercial aircraft simulator pilots see full "out of cockpit" views of what they would see as they land an aircraft at a specified airfield. In this case, of course, own motion stands for itself, and is restricted to motion in the cockpit. The visual effects created by an aircraft's motion in reference to its rest frame are determined by the settings a pilot makes on an actual control panel. The only difference is that the control settings are transmitted to the simulation program controlling the display, rather than being transmitted to the engines and control surfaces of an aircraft.

As of 1999, most of the displays used in VR systems are about the artistic quality of television cartoons. They show clearly recognizable, highly stylized drawings of houses, trees, cars, mountains, and other common objects. This is partly due to computational limits. When a person 'moves' through a virtual environment the change in perspective view has to be calculated for each successive location on the person's traverse. This process is very computationally expensive, and at present imposes a severe limitation on the appearance of virtual environments. Given a few years, though, this technological limit will disappear. Another limit will then be apparent. In order to compute all possible perspective views a computer has to have the three dimensional specifications of every object in the environment. Consider a tree. A reasonable, recognizable symbolic tree can be displayed by calculating the observer's perspective view of a green sphere on top of a brown cylinder. If someone wanted to change the picture from a stylized tree to a real one it would be necessary to have a three-dimensional model of every trunk, branch, twig, and leaf. Obviously this complicates software development, for someone has to write down the model. Developments in software as well as hardware will be needed before VR becomes an easy-to-use laboratory tool.

What does all this mean psychologically? We will answer this question in two phases, first by considering how virtual environments might alter the information flow underlying spatial orientation in humans, and then by examining some specific experiments.

In order to represent an environment people have to be able to sense directions and distances from their own position to remote (distal) objects. In addition, they have to be able to update their own position with respect to a presumably stable environment, or *rest frame*. Let us consider each in turn. Direction estimates in desktop VR are typically divorced from direction estimates in the world. The observer maintains his/her gaze on a computer screen, 'turning' by using a mouse or keyboard. The situation is somewhat better in immersive VR, especially when head movements are used to indicate gaze. In almost all cases, though, the ability to change direction by the hand produces an unnatural method of determining direction. In terms of the representation-simulation distinction, desktop VR is clearly a representation, while immersive VR provides simulations at various levels of fidelity.

The issue of distance is more complex. Disregarding sound (which can be simulated exactly in virtual environments), humans use a variety of cues to estimate distance. Cutting (1997) has made the useful point that each of these cues can be described by the scaling relation between the information provided by the cue and actual distance and the range of distance over which the cue provides a reasonable level of accuracy. For example, occlusion (i.e. "in front of") is ordinally related to distance. If object A is in front of object B than object A has to be closer to the observer than object B, but there is no way of knowing how much closer. Occlusion also has the useful property of working at all distances. On the other hand, we cannot rely exclusively on occlusion since objects can vary in distance without the near one occluding the further one. Binocular disparities between the views presented to each eye are always available and, in theory, provide enough geometric information for triangulation. Therefore binocular disparity is an indicator of absolute distance. The utility of this cue depends upon the observer's ability to detect the disparities, which decrease in absolute size as the target recedes. Based on a review of the literature, Cutting concluded that for humans, binocular disparity provides an accurate cue to absolute distance for targets within about 30 meters.

Table 1extends Cutting's argument by showing the nine cues Cutting considered, the type of scale they provide, their range of accuracy, and whether or not these cues can be represented by desktop and immersive VR. The reader may note that certain commonly discussed cues to depth -- texture, linear perspective, and brightness -- have been omitted. Cutting justifies this on the grounds that these are not unitary cues, but rather systems built from the nine basic cues. Therefore the extent to which the relevant system of cues can be presented in virtual environments depends upon the extent to which the individual cues are simulated in the supporting VR apparatus.

Examination of the table suggests several conjectures. In general, VE studies using immersive VR technology should lead to better distance estimates than studies based on desktop VR, for immersive VR can support more cues for depth than desktop VR can. The table also suggests that distance estimation may be particularly poor at close range, since convergence and accommodation cues drop out at this point. This point has to be modified for immersive VE experiences, since binocular disparity cues, which are very accurate in this range, can be represented more easily in an immersive than a desktop system. If separate views are projected to each eye in an immersive system, binocular cues can be provided by programming. Desktop VR cues can be produced by arranging a display analogous to the 19th century stereoscopes, but, as was the case in the old stereoscopes, natural appearing depth cues are hard to arrange. In most cases this would be a disadvantage. However, there might be some cases in which magnified binocular disparity cues improved distance estimation for objects more than thirty meters away.

The problem of achieving accurate distance estimation is of practical as well as scientific interest. As of 1999, the U.S. Navy was exploring a virtual environment training for ship handling. Distance estimation at ranges varying from 30 to 1,000 meters is an important part of this task. Proposals have been made to use VE training for surgery, where distance estimates on the order of millimeters would be crucial.

Cue	Scale Provided	Limit of accuracy	Status in VR systems
Occlusion	Ordinal	None	Veridically represented
Height in visual field	Absolute in theory, in practice ordinal	30-35 meters	Veridical in immersive VR. Resolution restricted in desktop VR
Relative size	Unanchored ratio, assuming planar surface	None	Veridical in immersive VR. Resolution reduced in desktop VR
Relative density of objects	Ordinal in appropriate situations	None	Texture discrimination is reduced in VR displays
Aerial perspective (Indistinctness of objects at greater distance)	Ordinal	Over 500 and under 3000 meters	Only a few gradients of blurring can be represented in current displays
Binocular disparity	Absolute	30 meters	Represented in immersive but not desktop VR. Could be enhanced in VR.
Accommodation (change of shape of eyes)	Ordinal providing that individual is less than 40 years old.	1-1.5 meters	Not represented in VR (objects are all at screen distance or directly focused on retina.
Convergence	In theory absolute. In practice ordinal	1 – 1.5 meters	Not represented in VR, as in case of accommodation
Motion perspective (changes in position of objects at fixed locations as the observer moves)	Absolute	30-35 meters	Represented very poorly in desktop VR. Can be simulated with considerable accuracy in immersive VR

Table 1. A summary of the visual cues to distance estimation, showing their scale relation to actual distance, the range over which they provide accurate estimation, and the applicability of a cue to VR systems.

An observer's own motion can produce visual cues to distance. Self motion is sensed by three different systems: the visual system (motion perspective); the tactile-kinesthetic system, which provides feedback from the musculature and from contact with the ground and other solid objects; and the vestibular system, which senses accelerative forces applied to the head and, by inference in normal motion, the rest of the body. Desktop VR represents rather than simulates all aspects of self motion. Typing or moving an electronic mouse is not the same as walking. Motion perspective occurs within the computer display window while the window itself and the room around it remain in sight and stationary.

Immersive VR systems can be very good at accounting for motion perspective, especially if the computer can update the display rapidly enough to avoid jerkiness. This was much more of a problem in VE applications developed prior to 1995 than it is in more modern systems. Systems that use wands or data gloves to command the computer to change a scene are obviously representing movement rather than simulating it. A few experimental VE applications have been developed that sense whole body or leg-foot motion, but these have not been tested very much beyond the laboratories that constructed them. The problem of providing other than very limited vestibular stimulation is essentially unsolved for the general walking problem, although devices to decouple vestibular simulation from motion perspective have been built.[9] Since these result in a very unusual experience for a (virtual) wayfinder they will not be further discussed.

These remarks suggest that VE's are at best poor substitutes for real environments. In fact, this is not the case. Several experiments have shown a strong resemblance between psychological representations established by real and virtual environments. Such resemblance can be established in several ways. The most common measure among VR enthusiasts, but probably least acceptable to cognitive psychologists, is a subjective report of 'sense of presence.' Although this concept is frequently mentioned in the VR and VE literature precise definitions seem to be lacking. Often, the sense of presence is measured either by participants' informal comments about how much they felt that 'they were there' or, more rarely, by asking the participant to fill out a check list indicating their feelings more specifically. There have been surprisingly few systematic efforts to relate scores on such checklists to different parameters of the VE experience itself (but see Witmer & Singer, 1998).[10]

A more reliable method of investigating VE's is to assume that a VE experience is like exploration of a real environment. The same measurement procedures can then be used to determine the participant's representation of each experience. For example, there are studies (cited below) in which people were asked to do pointing and distance estimation tasks following exploration of virtual environments. Two extensions of this method have been used. One is to examine the transfer of information from a virtual to a real environment. In addition to answering interesting scientific questions, this method addresses a practical problem. To what extent can people be trained to perform in a virtual environment as a way of rehearsing performance in a real environment? For instance, can firefighters train in a 'virtual hotel' in order to be prepared to fight a fire in a real one? Finally, aids that are known to improve performance in real environments can be introduced into virtual environments, to see if the same effects are observed.

People can learn to wayfind in virtual environments. In an early study using immersive VR equipment that was primitive by the standards of five years later, Regian and Shebilske (1992) showed that people learn to take efficient routes through a virtual building. By the mid 1990's several computer games were based on wayfinding through a maze. Even technically limited virtual environments can produce a mental representation of a spatial layouts.

An experiment by Jacobs, Thomas, Laurance, and Nagel (1998) addressed the more difficult issue of whether or not the spatial representations acquired in virtual environments resemble those acquired in actual environments. College students were placed in a VE circular arena, with four marked walls in sight outside the arena. Following Jacob et al. we will refer to them as N, E, S, W although this is simply a convention. In our terms, the participant was in a surround with four directional markers, each at a fixed (virtual) distance. The task was to locate a target that appeared when the participant moved to its location. Jacobs et al. point out that this situation is analogous to the water maze, an animal learning paradigm in which a rat has to learn to swim to a submerged platform. Rats learn to do this, and the people in Jacobs et al.'s experiment did too. Following learning, up to three of the four cues were removed. Both rats and people can locate their respective targets under these circumstances. If the cues are scrambled, as in the configuration N, W, S, E, both animals and people have difficulty finding the previously located target. This last finding is hardly surprising, for in this situation previously learned landmarks are unreliable. To appreciate this, imagine that you had learned to navigate in a desert by keeping track of your bearing to four distant mountains. It would be possible to navigate if one or two of the mountains was shrouded in clouds, but navigation would be very difficult if someone moved the mountains.

These results are consistent with the idea that both animal and human navigators learn to associate a place with the bearing from it to distinctive landmarks. As noted in the earlier discussion of the physiology of orientation, single-unit recording studies have shown that this association is reflected by neural activity in the rat hippocampus. The behavioral similarity between humans, in the virtual maze, and rats, in the water maze, suggests a commonality of mental representations. This is not surprising. What is more interesting is that the appropriate cognitive map was established using desktop VR equipment. Therefore the directional cues normally available (e.g. gaze direction) were never present. A cognitive map was established by the way that the observers interpreted a desktop display, rather than by a simulation of an environment.

Two studies show that the cognitive map established in a virtual environment appears to resemble the cognitive map developed by experience with analogous real environments. In one of these studies (Ruddle et al., 1997) people used desktop VR to explore a virtual building modeled after the building that Thorndyke and Hayes-Roth (1982) had used in their studies of mental representations of a real building. Following exploration, the participants indicated bearings and distances from one point to another. As in the Thorndyke and Hayes-Roth study, the explorers' responses were compared to the responses of people who had examined plans of the building. Map readers were more accurate than either Ruddle et al.'s virtual explorers or Thorndyke and Hayes-Roth's real ones.

Ruddle et al.'s study shows that virtual wayfinders are inferior to map readers in the same way that actual wayfinders are. Virtual wayfinders also show a superiority over map readers that mimics a superiority found in actual wayfinders. Recall the orientation dependency effect defined earlier. Map readers have difficulty with pointing tasks when the direction they must point in is not identical to the direction on the map, held in its normal orientation. People who have explored an environment do not show orientation dependency. Neither do people who have explored virtual environments (Tlauka & Wilson, 1996; Rosanno & Moak, 1998).

These results suggest that the representation established by a desktop virtual environment is like the representation established by actual experience with a neighborhood or surround. If this is true, training in a virtual environment should produce positive transfer when people are asked to explore the analogous real environment. Such transfer has been shown, both for explorations of buildings (Wilson, Foreman, & Tlauka, 1997; Witmer et al., 1996) and a specially designed, human-sized maze (Waller et al., 1998).

Does the type of VE make any difference? Wilson et al. (1997) used desktop VR , Witmer et al. (1996) used immersive VR, and both obtained positive results. The fact that Wilson et al.'s participants performed in a manner that at least approximated Thorndyke and Stasz's (1982), suggests that they had reached near asymptotic performance using desktop VR. The Waller et al. study was one of the few studies that measured learning over repeated trials in both desktop and immersive VR conditions. There was very little difference between the desktop and immersed VR conditions. After only a few trials, both VR groups, a group that studied maps of the maze, and a group that had repeated practice in the maze itself all showed near equivalent performance. These results certainly raise questions about the cost-effectiveness of expensive VR training for the purpose of learning about environments. Before we leap to this conclusion, though, it would be prudent to repeat the design of the Waller et al. study using a real environment rather than the artificial maze environment.

There is one final parallel between spatial learning in virtual and real environments. We have seen that, on average, men outperform women in spatial orientation tasks. The evidence for male-female differences in virtual environments is mixed. Most studies either say "no effects are observed" or make no mention of the topic at all. These negative results should be interpreted with caution, since in each case the number of participants has been so small that the contrast lacked statistical power. One of the larger studies (Witmer et al., 1996; 30 men and 34 women, half in VE and half in real training) did find differences in configurational knowledge, in favor of men. The work in our own laboratory, using rather more subjects than in the other studies, has consistently indicated that women have more trouble with virtual environment training than men do. In the Waller et al. (1998a) study, women took longer to traverse a real maze learned via VE training than men did. However, there was no male-female difference in the groups that learned by exploring the real maze (Waller et al., Figure 9). Subsequently we conducted a pointing study using a maze similar to that used in the earlier work (Waller, Hunt & Knapp 1998b). Participants explored a real or a virtual maze, and then were asked to point toward unseen objects when they were in the real maze. Figure 14 shows the results. As is typical in pointing studies, on some trials people were ‘turned around,’ in the sense that they pointed more than 90^o away from the correct bearing. When people had received VE training, all the ‘turned around’ pointings were made by women. (The converse is not true. All pointings made by women were not turned around.) Our results suggest that there are very strong male-female differences in the ability to benefit from VE training. Recent work in our lab has suggested that most of the effect of gender in VE spatial learning is statistically associated with differences in spatial ability (as assessed by paper-and-pencil tests) and proficiency with the navigational interface (Waller, 1999). Presumably, appropriate pretraining can reduce the gender difference in interface proficiency; however, there is surprisingly little evidence that gender differences in psychometrically-assessed spatial ability can be reduced by training.

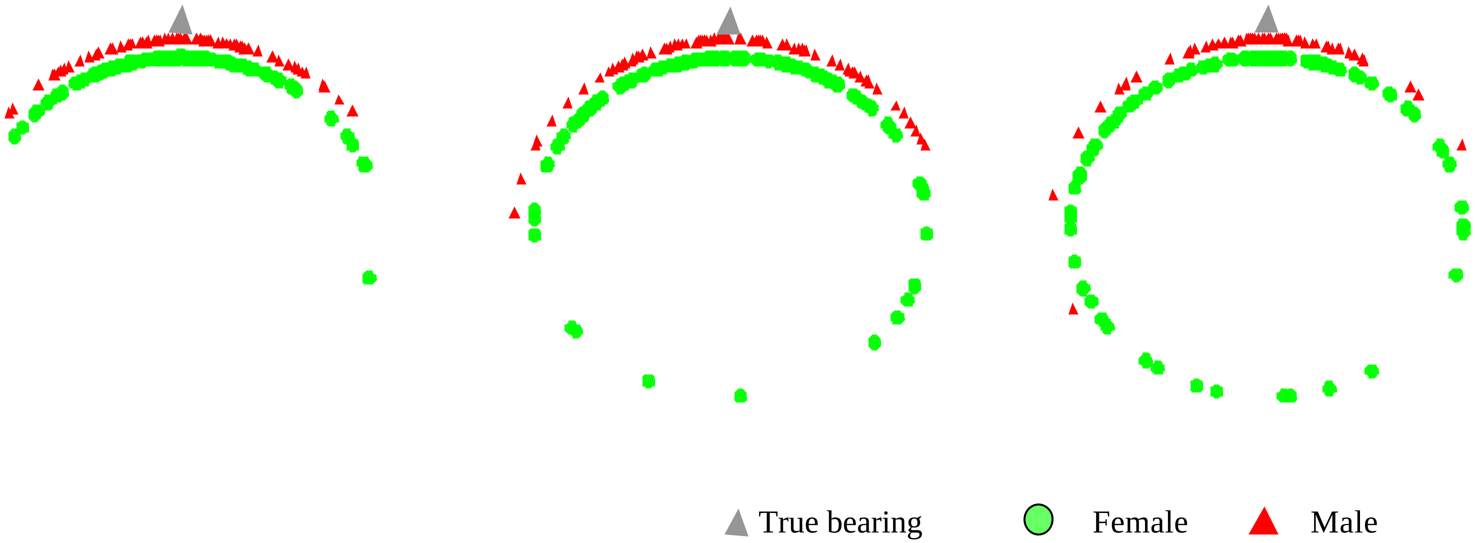


Figure 14. Male-female differences in pointing after three different types of experience with a maze environment. The Real condition refers to pointing in a real situation after practice in that situation. The Virtual condition refers to pointing in a virtual environment after practice with that environment. The Transfer condition refers to pointing in a real environment after practice in an analogous virtual environment. Triangles represent pointing by male participants; circles show pointing by female participants. Virtually all points that are more than 90° from the correct orientation were made by females. Data from Waller, Hunt, & Knapp 1998b.

A summary comment on artifacts.

In discussing virtual environments, a distinction was made between representations and simulations of an environment. The results obtained with maps, written or spoken descriptions, and desktop virtual environments all point to one conclusion. The acquisition of configural information depends upon the conceptual interpretation of information about the environment, not perceptual interpretation of that information. This result has implications for theories of spatial orientation. It also has an implication for the design of training in spatial orientation. For the purposes of environmental learning we may not need a Star Trek® holodeck -- a desktop computer will do just fine. This conclusion runs counter to the considerable effort to create realistic (and expensive) immersive VR, on the grounds that the sense of involvement accompanying an immersive VR experience should produce better learning (Held & Durlach,1992; Rheingold, 1991). Therefore some further comment is in order.

Acquiring configurational knowledge is a controlled task. This means that it is guided by cognitive rather than perceptual interpretation. Accordingly, if people can interpret a representation of an environment they can acquire the information contained in it. This is the why map learning works as well or better than exploration, in the sense of improving a person's configurational knowledge. It is also probably the reason that map learning is less successful than actual or immersive VE exploration in helping people avoid wrong turns when they are on the ground (Witmer et al., 1996). Put in a slightly different way, configurational knowledge is always declarative. As McDonald and Pellegrino (1993) pointed out, route knowledge can be either procedural or declarative. Representations of information work well for establishing declarative information. Actual experience or high-fidelity simulations are required to develop procedural knowledge.

We suggest, then, that immersive VR training may be worth its costs in two classes of situations. One to develop skills in going through a route very quickly, essentially in a situation in which a person must rely on procedural knowledge. The other is when training is intended to influence automated reactions that may interfere with performance in a real situation. For instance, immersive VR systems have been used to desensitize individuals suffering from phobias ranging from spiders to fear of heights. In such situations the emotional reaction has to be elicited so that it can be extinguished, and so that the extinction will be transferred to the real situation. There is evidence that a sense of presence is required to do this (Regenbrecht, Schubert, & Friedman, 1998).

Learning from a representation implies an ability to interpret that representation. Immersive VR may be a superior training device for individuals who cannot read maps or who have difficulty understanding a verbal description of an environment. The same thing is true of desktop VR, which is somewhere in-between a representation and a simulation. Today we have a variety of ways of presenting environmental information. The 1999 prices are about $250,000 for a really good immersive VR system (including software and models), $3,000 for a desktop system, about $50 for a good atlas, and $5.00 for a tour guide. The cost gradient is sharp enough so that research on the benefits is called for.

Where are we at in the study of spatial orientation?

The field has come a long way since Lynch's (1960) study of the city and Siegel and White's (1975) prescient theoretical analyses. These early papers correctly highlighted some major facts about wayfinding. They are

1. Psychologically, space is organized into a hierarchy of regions.
2. These regions are seen as more regular than they are. People seem to have a psychological demand for straight-line route segments and right-angled turns.
3. The distinction between survey and route oriented mental representations is a real one.

Subsequent research has amplified upon these findings in certain key areas. The metric properties of mental space are clearly much different from those of actual space. The fact that the symmetric property fails would probably amaze Euclid. What is equally surprising is how long people can remain in a space without developing configural knowledge of it. Is this perhaps related to the brain mechanisms that we use in spatial orientation?

Neither Lynch nor Siegel and White dealt with the physiological basis of spatial orientation. Since their papers were written, major developments in the neuroscience of vision have occurred. Today we know that spatial processing involves two separate brain mechanisms. The elaborate thalamo-cortical system subserving higher order vision interprets the input as we experience surrounds and neighborhoods, while a hippocampal system stores maps of the regions we have explored. There is a temptation to say that these systems are designed to record route knowledge, making configural knowledge a cognitive add-on, dependent on the development of mental models supported by the frontal and prefrontal areas. This works well as an explanation for primate (including human) learning. It is not as attractive as an explanation for learning in animals with smaller forebrains, but the extent to which these animals acquire configurational knowledge of an area is not clear.

What is somewhat surprising is that the original input need not come from the normal sensory systems -- vision and the vestibular and kinesthetic systems. Humans are surprisingly capable of maintaining orientation with reduced sensory input. Indeed, we can even learn about space from artificial representations, such as maps and texts. This indicates that the brain contains a system that interprets spatial orientation information, regardless of its origin.

Moving back to the behavioral level, more needs to be done to explore the dual roles of attention and intention in learning. Declarative spatial knowledge seems to require intentional learning. On the other hand, it seems that some procedural knowledge about routes can be established without a conscious intent to learn. This issue needs more study. At the gross level, it is clear that how one interacts with an environment determines what one learns about it. It would be nice to sharpen this general principle. This point is likely to be especially important when we consider the effects of artifactual experiences of space: maps, texts, and more recently, virtual environments. These media each invite certain types of interactions, and emphasize certain types of information about space. We need a better catalog of their effects.

So how does one give direction to further research on orientation? All that can be said is “Forward!” A relative frame of reference is all that can be expected, since we can never be sure of the absolute direction that further studies will take.

References

Aguirre, G. K., & D'Esposito, M. (1997). Environmental knowledge is subserved by separable dorsal/ventral areas. *Journal of Neuroscience, 17 (7),* 2513-2518 .

Allen, G. L., & Kirasic, K. C. (1985). Effects of the cognitive organization of route knowledge on judgements of macrospatial distance. *Memory and Cognition, 13 (3),* 218-227.

Allen, G. L., Siegel, A. W., & Rosinski, R. R. (1978). The role of perceptual content in structuring spatial knowledge. *Journal of Experimental Psychology: Human Learning and Memory, 1978, 4 (6),* 617-630.

Anderson, R. A. (1988). The neurobiological basis of spatial cognition. Role of the parietal lobe. In J. Stiles-Davis, M. Kritchevsky, & U. Bellugi. *Spatial Cognition: Brain bases and development* (pp. 57-80). Hillsdale, NJ: Erlbaum.

Annoshian, L. J. (1996). Diversity within spatial cognition: Strategies underlying spatial knowledge. *Environment and Behavior, 28 (4),* 471-493.

Anooshian, L. J., & Young, D. (1981). Developmental changes in cognitive maps of a familiar neighborhood. *Child Development, 52 (1),* 341-348.

Arthur, E. J., Hancock, P. A., & Chrysler, S. T. (1997). The perception of spatial layout in real and virtual worlds. *Ergonomics 40 (1),* 69-77.

Baddeley, A. D. (1987). *Working memory*. New York: Clarendon.

Baddeley, A. D., & Lieberman, K. (1980). Spatial working memory in humans. In R. S. Nickerson (Ed.) *Attention and Performance VIII* (pp. 521-539). Hillsdale, NJ: Erlbaum.

Blades, M., & Medlicott, L. (1992). Developmental differences in the ability to give route directions from a map. *Journal of Environmental Psychology, 12,* 175-185.

Bovet, J. (1994). Homing in humans: A different look. *Behavioural Processes, 32,* 197-208.

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence Memory: A constructivist versus interpretive approach. *Cognitive Psychology, 3 (2),* 193-209.

Bryant, D. J., Tversky, B., & Franklin, N. (1992). Internal and spatial frameworks for representing described scenes. *Journal of Memory & Language, 31 (1),* 74-98.

Bryant, D. J., Tversky, B., & Lanca, M. (in press, 1998). Retrieving spatial relations from observation and memory. In E. Van der Zee & U. Nikanne (Eds.) Configural structure and its interfaces with other modules of representation. ** PUBLISHER UNKNOWN**

Butler, D. L., Acquino, A. L., Hissong, A. A., & Scott, P. A. (1993). Wayfinding by newcomers in a complex building. *Human Factors, 35 (1),* 159-173.

Carroll, J. B. (1993). *Human Cognitive Abilities.* Cambridge: Cambridge U. Press.

Chapius, N., & Varlet, C. (1987). Short cuts by dogs in natural surroundings. *Quarterly Journal of Experimental Psychology, 39B,* 49-64.

Cornell, E. H., Heth, C. D., & Alberts, D. M. (1994). Place recognition and way finding by children and adults. *Memory and Cognition, 22 (6),* 633-643.

Cutting, J. E. (1997). How the eye measures reality and virtual reality. *Behavior Research Methods, Instruments, & Computers, 29* (1), 27-36.

Dabbs, J. M., Jr., Lee, C. E., Strong, R. A., & Milun, R. (1998). Spatial ability, navigation strategy, and geographic knowledge among men and women. *Evolution and Behavior, 19 (2),* 89-98.

Darken, R. P., & Sibert, J. (1996). Navigating large virtual spaces. *International Journal of Human-Computer Interaction, 8 (1),* 49-71.

Deary, I. J., Egan, V., Gibson, G. J., Austin, E. J., Brand, C. R., & Kellaghan, T. (1996). Intelligence and the differentiation hypothesis. *Intelligence, 23* (2), 105-132.

Denis, M. (1996). Imagery and the description of spatial configurations. In M. DeVega, M. J. Intons-Peterson, P. N. Johnson-Laird, M. Denis, & M. Marschark (Eds.) *Models of Visual-Spatial Cognition* (pp. 128-197). Oxford: Oxford U. Press.

Detterman, D. K., & Daniel, M. H. (1989). Correlations of mental tests with each other and with cognitive variables are highest in low IQ groups. *Intelligence, 13*, 349-360.

Diwadkar, V. A., & McNamara, T. P. (1997). Viewpoint dependence in scene recognition. *Psychological Science, 8 (4)* , 302-307.

Emerson, E. (1995). *The vanishing smile* (pp. 13). New York: Ballantine.

Evans, G.W., & Pezdek, K. (1980). Cognitive mapping: Knowledge of real-world distance and location information. *Journal of Experimental Psychology: Human Learning and Memory, 6 (1),* 13-24.

Fauconnier, G. (1997). *Mapping in thought and language*. Cambridge: Cambridge U. Press.

Foley, J. E. & Cohen, A. J. (1984). Working mental representations of the environment. *Environment and Behavior, 16,* 713-729.

Franklin, N., Henkel, L. A., & Zangas, T. (1995). Parsing surrounding space into regions. *Memory & Cognition, 23 (4),* 397-407.

Franklin, N., & Tversky, B. (1990). Searching imagined environments. *Journal of Experimental Psychology: General, 119 (1),* 63-76.

Franklin, N., Tversky, B., & Coon, V. (1992). Switching points of view in spatial mental models. *Memory and Cognition, 20 (5),* 507-518.

Gale, N., Golledge, R. G., Pellegrino, J. W., & Doherty, S. (1990). The acquisition and integration of route knowledge in an unfamiliar neighborhood. *Journal of Environmental Psychology, 10,* 3-25.

Galea, L. A. M., & Kimura, D. (1993). Sex differences in route learning. *Personality and Individual Differences, 14 (1)* ,53-65.

Glicksohn, J. (1994). Rotation, orientation, and cognitive mapping. *American Journal of Psychology, 107 (1),* 39-51 .

Guilford, J. P., & Zimmerman, E. S. (1948). The Guilford-Zimmerman aptitude survey. *Journal of Applied Psychology, 32 (1).*

Held, R., & Durlach, N. (1992). Telepresence. *Presence, 1,* 109-112.

Haenggi, D., Kintsch, W., & Gernsbacher, M. A. (1995). Spatial situation models and text comprehension. *Discourse Processes*, *19,* 173-199.

Hill, K. A. (1992). Spatial competence in elderly hunters. *Environment and Behavior, 24,* (6), 779-794.

Hintzman, D. L., O'Dell, C. S., & Arndt, D. R. (1981). Orientation in cognitive maps. *Cognitive Psychology, 13,* 149-206.

Hirtle, S. C., & Jonides, J. (1985). Evidence of hierarchies in cognitive maps. *Memory & Cognition, 13 (3),* 208-217.

Holyoak, K. J., & Mah, W. A. (1982). Cognitive reference points in judgments of symbolic magnitude. *Cognitive Psychology, 14,* 328-352.

Hunt, E., & Agnoli, F. (1991). The Whorfian hypothesis: A cognitive psychology perspective. *Psychological Review, 98*(3), 377-389.

Huttenlocher, J., Hedges, L.V., & Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review, 98 (3),* 352-376.

Huttenlocher, J., Newcombe, N., & Sandberg, E. H. (1994). The coding of spatial locations in young children. *Cognitive Psychology, 27 (2),* 115-147.

Infield, S. E. (1991). An investigation into the relationship between navigation skill and spatial abilities. (Doctoral dissertation, University of Washington, Seattle). *Dissertation Abstracts International*, 52(5-B), 2800.

Jacobs, W. J., Thomas, K. C. F., Laurance, H. E., & Nagel, L. (1998). Place learning in virtual space: II Topographical relations as one dimension of stimulus control. *Learning and Motivation, 29,* 288-308 .

Johnson-Laird, P. N. (1983). *Mental Models.* Cambridge, MA. Harvard U. Press.

Johnson-Laird, P. N., & Byrne, M. J. B. (1991). *Deduction.* Hillsdale, NJ: Erlbaum.

Kearins, J. M. (1981). Visual spatial memory in Australian aboriginal children of desert regions. *Cognitive Psychology, 13 (3),* 434-460.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* Cambridge: Cambridge U. Press.

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

Klatzky, R. L., Loomis, J. M., Beall, A. C., Chance, S. S., & Golledge, R. G. (1998). Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychological Science, 9 (4),* 293-298.

Kosslyn, S. M., Ball, T. M., & Reiser, B. J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 47-60.

Kritchevsky, M. (1988). The elementary spatial functions of the brain. In J. Stiles-Davis, M. Kritchevsky, & U. Bellugi. *Spatial Cognition: Brain basis and development* (pp. 111-140). Hillsdale, NJ. Erlbaum.

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

Langston, W., Kramer, D. C., & Glenberg, A. M. (1998). The representation of space in mental models derived from text. *Memory & Cognition, 26 (2),* 247-262.

Lawton, C. A. (1994). Gender differences in way-finding strategies: Relationship to spatial ability and spatial anxiety. *Sex Roles, 30 (11-12),* 765-779.

Lawton, C. A. (1996). Strategies for indoor wayfinding: The role of orientation. *Journal of Environmental Psychology, 16 (2),* 137-145.

Lawton, C. A.. Charleston, S. I., & Zieles, A. S. (1996). Individual and gender-related differences in indoor wayfinding. *Environment and Behavior, 28 (2),* 204-219.

Levine, D., Warach, J., & Farah, M. (1985). Two visual systems in mental imagery: Dissociation of “what” and “where” in imagery disorders due to bilateral posterior cerebral lesions. *Neurology , 35,* 1010-1018.

Levine, M. (1982). You are here maps: Psychological Considerations. *Environment and Behavior, 14,* 221-237.

Levine, M., Marchon, I., & Hanley, G. (1984). The placement and misplacement of you-are-here maps. *Environment and Behavior, 16,* 139-157.

Levine, M., Janovic, I., & Palij, M. (1982). Principles of spatial problem solving. *Journal of Experimental Psychology: General, 111,* 157-175.

Levinson, S. C. (1996a). Language and Space. *Annual Review of Anthropology, 25,* 353-382.

Levinson, S. C. (1996b). Frames of reference and Molyneux’s question: Cross-linguistic evidence. In P. Bloom, L. Peterson, L. Nadel, M. Garrett (eds.) *Language and Space* (pp. 109-169). Cambridge MA. MIT Press.

Liben, L. S., & Yekel, C. A. (1996). Preschoolers’ understanding of plan and oblique maps: The role of geometric and representational correspondence. *Child Development, 67,* 2780-2796.

Linde, C., & Labov, W. (1975). Spatial networks as a site for the study of language and thought. *Language, 51,* 924-939.

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

Lloyd, R. (1993). Cognitive processes and cartographic maps. In T. Garling & R.G. Golledge (Eds.) *Behavior and Environment: Psychological and Geographic Perspectives* (pp. 141-169) N. Holland: Elsevier.

Loomis, J. M., Klatzky, R. L., Golledge, R. G., Cicinelli, J. G., Pellegrino, J. W., & Fry, P. A. (1993). Nonvisual navigation by blind and sighted: Assessment of path integration ability. *Journal of Experimental Psychology: General*, 122 (1), 73-91.

Lynch, K. (1960). *Image of the city*. Cambridge, MA. MIT Press.

MacLeod, C. M., Hunt, E., & Mathews, N. N. (1978). Individual differences in the verification of sentence-picture relationships. *Journal of Verbal Learning and Verbal Behavior. 17,* 493-507.

McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review, 99,* 440-466.

McDonald, T. P., & Pellegrino, J. W. (1993). Psychological perspectives on spatial cognition. In T. Garling & R.G. Golledge (Eds.) *Behavior and Environment: Psychological and Geographic Perspectives* (pp. 47-82). N. Holland: Elsevier.

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

McNamara, T. P. (1992). Spatial representation. *Geoforum, 23,* 139-150.

McNaughton, B. L., Barnes, C. A., Gerrard, J. L., Gothard, K., Jung, M. W., Knierim, J. J., Kudirimoti, H., Quin, Y, Skaggs, W. E., Suster, M., & Weaver, K. L. (1996). Deciphering the hippocampal polyglot: The hippocampus as a path integration system. *Journal of Experimental Biology, 199,* 173-185.

Maguire, E. A., Burgess, N., Donnett, J. G., Frackowiak, R. S. J., Frith, C. D., & O'Keefe, J. (1998). Knowing where and getting there: A human navigation network. *Science,* 180 (8 Mar), 921-924.

Maguire, E. A., Frackowiak, R. S. J., & Frith, C. D. (1997). Recalling routes around London: Activation of the right hippocampus in taxi drivers. *Journal of Neuroscience, 17 (18),* 7103-7110.

Mathews, N. N., Hunt, E., & MacLeod, C. M. (1980). Strategy choice and strategy training in sentence-picture verification. *Journal of Verbal Learning and Verbal Behavior, 19,* 531-538.

Matthews, M. H. (1987). Sex differences in spatial competence: The ability of young children to map "primed" unfamiliar environments. *Educational Psychology, 7 (2),* 77-90.

Moar, I., & Bower, G. H. (1983). Inconsistency in spatial knowledge. *Memory and Cognition, 11 (2),* 107-113.

Moeser, S. D. (1988). Cognitive mapping in a complex building. *Environment and Behavior, 20, (1),* 21-49.

Money, J. (1993). Specific neurocognitional impairments associated with Turner (45,X) and Klinefelter (47, XXY) Sydromes: A review. *Social Biology, 40,* 145-151.

Montello, D. R. (1991). Spatial orientation and the angularity of urban routes: A field study. *Environment & Behavior, 23*(1), 47-69.

Morrow, D. G., Greenspan, S. L., & Bower, G. H. (1987). Accessibility and situation models during narrative comprehension. *Journal of Memory and Language, 26,* 165-187.

Murphy, D. G. M., Allen, G., Haxby, J. V., Largay, K. A., Daly, E., White, B. J., Powell, C. M., & Schapiro, M. B. (1994). The effects of sex steroids, and the X chromosome, on female brain function: A study of the neuropsychology of adult Turner syndrome. *Neuropsychologia, 32 (11),*1309-1323 .

Norman, D. A. (1993). *Things that make us smart: Defending human attributes in the age of the machine.* Reading, MA. Addison-Wesley.

Olson, D. R. (1994). *The world on paper: The conceptual and cognitive implications of writing and reading.* Cambridge: Cambridge U. Press.

Perrig, W., & Kintsch, W. (1985). Propositional and situational representations of text. *Journal of Memory and Language, 24, 503-518.*

Poltrock, S. E., & Brown, P. (1980). Individual differences in visual imagery and spatial ability. *Intelligence, 8,* 93-138.

Poucet, B. (1993). Spatial cognitive maps in animals: New hypotheses on their structures and neural mechanisms. *Psychological Review, 100 (2),* 163-182.

Presson, C. C., DeLange, N., & Hazelrig, M. D. (1989). Orientation specificity in spatial memory: What makes a path different from a map? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 15 (5),* 887-897.

Presson, C. C., & Hazelrig, M. D. (1984). Building spatial representations through primary and secondary learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition. 10,* 716-722.

Prothero, J. D. (1998). The Role of Rest Frames in Vection, Presence and Motion Sickness. Doctoral dissertation, University of Washington.

Regenbrecht, H. T., Schubert, T. W., & Friedman, F. (1998). Measuring the sense of presence and its relation to fear of heights in virtual environments. *International Journal of Human-Computer Interaction, 10 (3),* 235-249.

Regian, J. W., & Shebilske, W. L. (1992). Virtual Reality: An instructional medium for visual-spatial tasks. *Journal of Communication, 42 (4),* 136-149.

Rheingold, H. (1991). *Virtual Reality.* New York: Summit Books.

Rhodes, G., Carey, S., Byatt, G., & Proffitt, F. (1998). Coding spatial variations in faces and simple shapes: a test of two models. *Vision Research, 38,* 2307-2321.

Rieser, J. J. (1989). Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 5* (6), 1157-1165

Rieser, J. J., Hill, E. W., Taylor, C. R., & Bradfield, A. (1992). Visual experience, visual field size, and the development of non visual sensitivity to the spatial structure of outdoor neighborhoods explored by walking. *Journal of Experimental Psychology: General*, 121 (2), 210-221.

Roskos-Ewoldson, B., McNamara, T. P., Shelton, A. L., & Carr, W. (1998). Mental representations of large and small spatial layouts are orientation dependent. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 24 (1),* 215-226.

Rossano, M. J., & Moak, J. (1998). Spatial representations acquired from computer models: Cognitive load, orientation specificity and the acquisition of survey knowledge. *British Journal of Psychology, 89,* 481-497.

Rossano, M. J., Warren, D. H., & Kenan, A. (1995). Orientation specificity: How general is it? *American J. of Psychology, 108 (3),* 359-380.

Ruddle, R. A., Payne, S. J., & Jones, D. M. (1997). Navigating buildings in "desk-top" virtual environments: Experimental investigations using extended navigational experience. *Journal of Experimental Psychology: Applied, 3 (2),* 143-159.

Sadalla, E. K., Burroughs, W. J., & Staplin, L. J. (1980). Reference points in spatial cognition. *Journal of Experimental Psychology: Human Learning & Memory, 6*(5), 516-528.

Salthouse, T. A. (1991). *Theoretical perspectives on cognitive aging*. Hillsdale, NJ. Erlbaum.

Samsonovich, A., & McNaughton, B. L. (1997). Path integration and cognitive mapping in a continuous attractor neural network model. *The Journal of Neuroscience, 17 (5),* 5900-5926.

Satalich, G. (1995). Navigation and wayfinding in virtual reality: finding the proper tools and cues to enhance navigational awareness. Unpublished M.S.E. thesis, University of Washington (School of Engineering).

Schiffman, S. S., Reynolds, M. L., & Young, F. W. (1981). *Introduction to multidimensional scaling: Theory, methods, and applications.* Orlando, FL. Academic Press.

Shelton, A., & McNamara, T. P. (1997). Multiple views of spatial memory *Psychonomic Bulletin & Review, 4 (1),* 102-106.

Shepard, R. N., & Hurwitz, S. (1984). Upward direction, mental rotation, and the discrimination of left and right turns in maps. *Cognition, 18,* 161-193.

Sholl, M. J. (1987). Cognitive maps as orienting schema. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 13* (4), 615-625.

Sholl, M. J. (1996). From visual information to cognitive maps in J. Portugali, Ed. *The construction of cognitive maps* (pp. 157-186). The Netherlands: Kluwer.

Sholl, M. J., & Nolin, T. L. (1997) Orientation specificity in representations of place. *Journal of Experimental Psychology: Learning, memory and cognition, 23 (6),* 1494-1507.

Siegel, A. W. (1980). The externalization of cognitive maps in children and adults: In search of ways to ask better questions. In Lynn S. Liben, Arthur H. Patterson, Nora Newcombe. (Eds.) *Spatial representation and behavior across the life span*. New York: Academic, pp. 167-193.

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. 10,* 9-55. New York: Academic Press.

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

Taylor, H. A., & Tversky, B. (1992a). Spatial mental models derived from survey and route descriptions. *Journal of Memory and Language, 31, 261-292.*

Taylor, H. A., & Tversky, B. (1992b). Descriptions and depictions of environments. *Memory and Cognition, 20,* 483-496.

Taylor, H. A., & Tversky, B. (1996). Perspective in spatial descriptions. *Journal of Memory and Language*, *35*, 371-391.

Thorndyke, P. W. (1982). Distance estimation from cognitive maps. *Cognitive Psychology*, *13*(4), 526-550.

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

Thorndkye, P. W., & Stasz, C. (1980). Individual differences in procedures for knowledge acquisition from maps. *Cognitive Psychology, 14 (4),* 560-589.

Tlauka, M., & Wilson, P. N. (1996). Orientation free representations from navigation through a computer-simulated environment. *Environment and Behavior*, *28 (5),* 647-664.

Tkacz, S. (in press,1998). Learning map interpretation: Skill acquisition and underlying abilities. *Journal of Environmental Psychology*.

Tolman, E. C. (1948). Cognitive Maps in Rats and Men. *Psychological Review, 55*, 189-208.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84,* 337-352.

Tversky, B. (1981). Distortions in memory for maps. *Cognitive Psychology*, *13 (3),* 560-589.

Tversky, B. (1991). Spatial mental models. In G. H. Bower (Ed.) *Attention and Performance, Volume 27* (pp. 109-145). San Diego: Academic Press.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of Discourse Comprehension*. New York: Academic Press.

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin, 117(2),* 250-270.

Waller, D. (1999). An assessment of individual differences in spatial knowlege of real and virtual environments. Doctoral dissertation, Univerisity of Washington.

Waller, D., Hunt, E., & Knapp, D. (1998a). The transfer of spatial knowledge in virtual environment training. *Presence, 7 (2),* 126-139.

Waller, D., Hunt, E., & Knapp, D. (1998b). Measuring spatial knowledge in a virtual environment: Distance and angles. Paper presented at the 39th annual meeting of the Psychonomics Society, Dallas, Tx, 21 Nov. 1998.

Ward, S. L., Newcombe, N., & Overton, W. F. (1986). Turn left at the church, or three miles North: A study of direction giving and sex differences. *Environment and Behavior, 18 (2),* 192-213.

Wiest, W. M. & Bell, B. (1985). Stevens's exponent for psychophysical scaling of perceived, remembered, and inferred distance. *Psychological Bulletin, 98*(3), 457-470.

Wilson, P. N., Foreman, N., & Tlauka, M. (1997). Transfer of spatial information from a virtual to a real environment. *Human Factors, 39 (4),* 526-531.

Witmer, B. G., Bailey, J. H., & Knerr, B. W., & Parsons, K.C. (1996). Virtual spaces and real world places: Transfer of route knowledge. *International Journal of Human-Computer Studies, 45,* 413-428.

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

Zwann, R. A., & Radavansky, G. A. (1998). *Psychological Bulletin, 123 (2),* 162-185 .

Endnotes

---

[1] These figures are excerpted from Reiser et al., 1992, Table 1.

[2] Although there is little doubt about the generality of the alignment effect some individuals appear to be immune to it (Rossano, Warren, & Keenan, 1995). This is consistent with the general finding that there are wide individual differences in spatial-visual reasoning.

[3] Individual features could be weighted by their importance. This generalization of Tversky's model will be ignored in the interests of simplifying the presentation

[4] Technically, Turner's syndrome cases have a 45,X karotype.

[5] . By the late 20th century few, if any, Australian aboriginal groups lived totally apart from Australia's dominant, western-oriented culture. Kearins' participants included rural (but not completely traditional) children living in the Desert of Australia and city dwelling Aboriginal children. Spatial memory performance was comparable for both groups, and exceeded the spatial memory performance of Australian children of European descent.

[6] In an orienteering race participants are given a map with from six to ten control points marked on it. The locations of the control points typically define a 6 to 10 km. route, and are chosen so that a person standing at one control point cannot see any other points. The competitors, who start at fixed intervals and thus cannot see each other, must visit each control point in order, but can choose their own route between control points. At the higher levels control points are purposely chose so that they are difficult to locate. The resulting races are face-valid and challenging tests of wayfinding ability.

[7] *Sally Forth*, by Greg Howard. The Sunday, Sept. 3, 1995 strip contained a discussion between a husband and wife, in which the husband said "Go North" and the wife replied "Tell me a direction. North is a compass point, right is a direction." I thank Sharon Tkacz for calling this to my attention.

[8] Taylor and Tversky (1992b) did not find male-female differences in the use of route or survey descriptions. However Taylor and Tversky do not give details of this negative finding, and their studies used relatively small samples in any one experiment. The Dabbs et al. result was based on data from more than 200 participants. The negative findings by Taylor & Tversky may have been a simple case of low statistical power.

[9] For example, Prothero (1998) had people view an immersive VR while seated in a barber's chair. By rotating the chair in the direction opposite from the direction of motion implied by the VE display Prothero decoupled two sources of information that, in the normal world, are perfectly correlated.

[10] There is something of a conflict here between science and entrepneurship. If your goal is to sell a VR system sense of presence is in itself a valid measure of consumer acceptance. In addition, there are certainly ample precedents for the use of self rating systems to evaluate psychological well being. An elaborate technology has been developed for the design and analysis of rating devices. It appears to have been little used in VE and VR research.