Citation:
Giudice, N.A., Bakdash, J.Z., & Legge, G.E. (2007). Wayfinding with words: Spatial learning and navigation using dynamically-updated verbal descriptions. *Psychological Research*, 71(3), 347-58.

download from: www.vemilab.org
Leave comments on this article using the recommender system at:
www.vemilab.org/node/34
Original content adapted for web distribution.

Psychological Research (2007) 71:347–358
DOI 10.1007/s00426-006-0089-8

ORIGINAL ARTICLE

# Wayfinding with words: spatial learning and navigation using dynamically updated verbal descriptions

**Nicholas A. Giudice · Jonathan Z. Bakdash · Gordon E. Legge**

Received: 12 November 2005 / Accepted: 15 March 2006 / Published online: 16 September 2006
© Springer-Verlag 2006

**Abstract** This work investigates whether large-scale indoor layouts can be learned and navigated non-visually, using verbal descriptions of layout geometry that are updated, e.g. contingent on a participant's location in a building. In previous research, verbal information has been used to facilitate route following, not to support free exploration and wayfinding. Our results with blindfolded-sighted participants demonstrate that accurate learning and wayfinding performance is possible using verbal descriptions and that it is sufficient to describe only local geometric detail. In addition, no differences in learning or navigation performance were observed between the verbal study and a control study using visual input. Verbal learning was also compared to the performance of a random walk model, demonstrating that human search behavior is not based on chance decision-making. However, the model performed more like human participants after adding a constraint that biased it against reversing direction.

N. A. Giudice (✉)
Department of Psychology, University of California,
Office: Bldg 551, Room 2226,
Santa Barbara, CA 93106-9660, USA
e-mail: giudice@psych.ucsb.edu

J. Z. Bakdash
Department of Psychology, University of Virginia,
Charlottesville, VA, USA

G. E. Legge
Department of Psychology, University of Minnesota,
Twin Cities, MN, USA

## Introduction

The current research investigates the use of dynamically updated verbal descriptions, messages whose content changes with respect to the movement of the navigator through the environment, to describe spatial layouts during wayfinding tasks. By addressing environmental learning and wayfinding, this work differs from most of the existing literature, which emphasizes verbal descriptions during route navigation only. While this proof-of-concept study used blind-folded sighted participants, the usefulness of verbal information to serve as an alternative, non-visual mode of environmental access has obvious application to blind wayfinders or navigation in low or no light conditions, such as firefighters maneuvering through smoke-filled buildings.

Much of the research investigating language-based spatial learning does not address actual navigation. Instead, these studies employ some variant of a paradigm where participants read a spatial narrative and are then tested on tasks that attempt to characterize the mental representation built up from reading these texts. In contrast to early theories of text comprehension arguing that the actual words are preserved in memory, the evidence is now clear that the mental representation is based on the spatial relations and conditions described by the texts, see (Zwaan, 1998) for a review. This ability for language to develop into an abstract spatial form in memory, called a ''cognitive map'' (O'Keefe & Nadel, 1978; Tolman, 1948) has been shown using various measurement techniques, such as spatial priming and recall, distance and pointing judgments, mental scanning operations and map reproduction tasks (Denis & Cocude, 1989, 1997; Denis

& Zimmer, 1992; Ferguson & Hegarty, 1994; Franklin & Tversky, 1990; Hirtle & Heidorn, 1993; Johnson-Laird, 1983; Perrig & Kintsch, 1985; Talmy, 1983; Taylor & Tversky, 1992; Wilson, Tlauka, & Wildbur, 1999).

Much less research has directly addressed the use of verbal descriptions during spatial learning and navigation in real environments. The studies that have been conducted generally relate to giving (or interpreting) verbal route descriptions (Allen, 1997; Denis, Pazzaglia, Comoldi, & Bertolo, 1999; Lovelace, Hegarty, & Montello, 1999; Tversky, 1996). Such studies are based on static descriptions, as the information conveyed by the verbal route directions does not change in register with the participants' physical movement through the space. By contrast, a verbal message is termed dynamically updated if the information described is coupled to the navigator's changing position and orientation in the environment (see Tom & Denis, 2003 for an example).

In addition, the use of updated speech displays and virtual acoustic displays, spatialized sound that appears to come from targets in 3d space, coupled with GPS tracking have also been shown to be sufficient for guiding blind and blindfolded sighted participants along routes between target locations (Loomis, Golledge, & Klatzky, 1998; Loomis, Marston, Golledge, & Klatzky, 2005). Similar verbal displays are also commercially available as part of modern in-car navigation systems. These route guidance systems are built on what is dubbed here as a point-based display, as they provide information about the distance and direction of target locations. Point-based displays convey information about discrete landmarks and decision-points rather than attempting to describe geometric information about the environment.

Contrasting with the previous work, the current study employs a geometric-based display composed of verbal descriptions about layout configuration (the network of corridors in large-scale indoor settings). By providing updated geometric descriptions about an environment rather than information about discrete landmarks and employing a free exploration paradigm rather than using route navigation, the research described in this paper extends the investigation of spatial language from performance on sequential route tasks to open search learning, wayfinding behavior and cognitive map development.

Several models of verbal direction giving have been specified for route navigation (Allen, 1997, 2000; Couclelis, 1996) but there are no accepted principles specifying the spatial information that should be conveyed by geometric descriptions to support free exploration. As the current study is based on the latter, it was necessary to develop a set of formal instructions to describe layout geometry through discrete speech-based messages (see Giudice, 2004 for full details). In summary, descriptions of user heading and hallway configuration were given at all intersections and were updated in register with the movement of the participant as they navigated through the space.

Since the geometric descriptions are based on a relatively small number of spatial primitives, we ensure that consistent, unambiguous messages are conveyed to all participants, factors known to be critical for effective use of spatial language (Ehrlich & Johnson-Laird, 1982; Levelt, 1996). In addition, evidence from both the animal literature (Benhamou, 1998; Cheng, 1986; Gallistel, 1990; Poucet, 1993) and from human studies (Hermer & Spelke, 1994) demonstrate that geometric configuration is critical in learning novel layouts.

The issue of spatial scale must also be addressed when using geometric verbal descriptions. It is not known how much environmental information should be described to support large-scale navigation. In this paper, the amount of the environment that is accessible to the navigator from a given vantage point is here termed ''verbal view-depth.'' Three verbal view-depths were used in the current experiment, with each condition describing a different amount of geometric detail ranging from only local information to a global description of layout configuration. The goal was to determine the least complex message that facilitated the highest level of learning. Manipulation of view depths also addresses a theoretical question about whether decreasing spatial integration, by increasing verbal view-depth, facilitates development of an accurate spatial representation.

The current studies use a training period to address learning of unfamiliar layouts through free exploration and several measures of testing performance to evaluate the properties of the resulting knowledge structure. Of particular interest is to compare the training and test data in order to determine whether people are operating on a route-based or map-like spatial representation. If the representation is based on fixed routes as proposed by some models of human spatial development, e.g., the influential landmark, route, survey (LRS) proposal by Siegel & White (1975), we would predict that success in following a route from target A to B in the testing phase would be contingent on travel and rehearsal of the same route during the learning phase. On the other hand, if free exploration leads to the development of a configurational representation, we would expect correct route execution at test regardless of prior experience traveling that route.

## Study 1: learning and navigation using dynamically updated verbal descriptions

This study addressed four primary questions:

1. Does access to dynamically updated verbal descriptions enable people to use free exploration to learn complex indoor environments?
2. How much spatial information should be conveyed by these descriptions to promote accurate learning?
3. What is the structure of the mental representation built up from training with verbal descriptions?
4. Is navigation performance with verbal descriptions comparable to performance on the same tasks carried out using vision?

The experiment is broken into two sub-studies, the first encompassing three verbal learning conditions and the second describing a visual control condition. The verbal learning study required blindfolded, normally sighted participants to freely explore three training environments using three verbal modes. Each verbal mode provided access to a different level of view-depth information about layout geometry. After a fixed training period, in which participants used the verbal descriptions to search through the entire layout and find four target locations, they were tested on their ability to plan and execute routes between target pairs in the same environment.

The purpose of the visual study, using vision rather than verbal descriptions to explore the environment, was twofold. It served as a control for the verbal conditions, representing a measure of baseline performance on the same learning and wayfinding tasks. Also, by comparing the results across studies, it is possible to investigate whether functionally equivalent spatial representations are built up in memory between training with verbal and visual information. There are some minor methodological differences between the verbal and visual studies, noted below, but these do not affect the comparison between conditions.

### Methods

#### *Participants*

Fifteen blindfolded-sighted participants, eight females and seven males between the ages of 18 and 38 (mean age of 22), took part in the verbal study and 13 sighted participants, 7 females and 6 males, between the ages of 18 and 40 (mean age of 22), took part in the visual study. All subjects reported that they had normal or corrected to normal vision. All gave informed consent and were compensated with extra credit in a psychology course or with payment for their participation.

#### *Environments*

Portions of three floors of the Psychology building at the University of Minnesota were used for both studies. The floors were of similar size and complexity but differed in layout topology, making transfer of knowledge between floors unlikely. The layouts used in the verbal conditions averaged 495 feet of corridor length and contained 11.6 intersections (see Fig. 1 for an illustration of the three verbal learning environments). The visual condition used slightly modified environments, as it was part of a larger project designed to incorporate similar, yet different, layouts than had been studied previously. The layout topology was almost identical between verbal and visual conditions and the differences were nominal, averaging ~30 more feet and ~1 additional intersection per floor in the visual condition.

#### *Verbal modes*

Three verbal description modes depicted the geometric structure of the corridor layouts. The black lines represent the information heard in each of the three view-depth conditions (see Fig. 2 for an example).

1. Local verbal mode: Describes layout geometry at the user's current position. ''Facing east, at a two-way intersection, ahead is a hallway, to the right is a hallway.''
2. Maplet verbal mode: Includes the local information and adds a description of the distance and

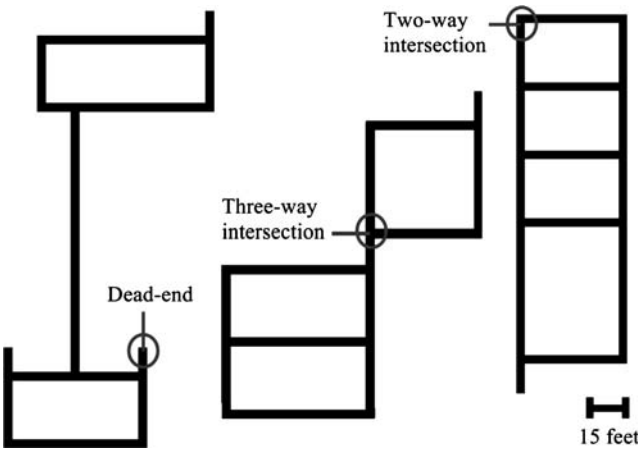


**Fig. 1** Three experimental layouts used in the verbal study with scale and intersection types denoted

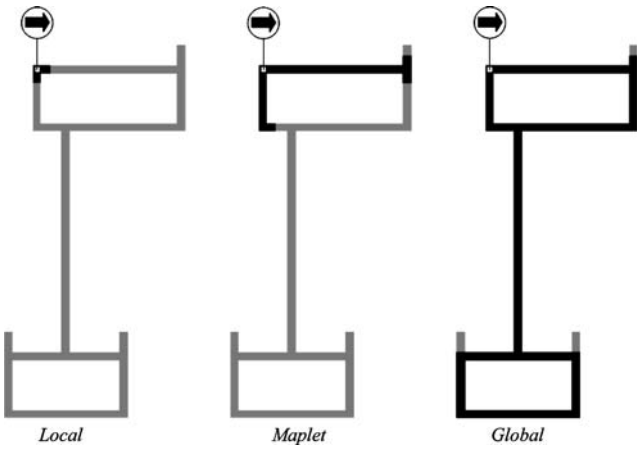


**Fig. 2** The *circle* and *arrow* represent the user's location and orientation in the layout. The *black lines* are described by the verbal language and the *gray lines* symbolize the entire floor

geometry for all adjacent intersections. Note that the information about adjacent intersections is given as if you were walking down the hallway being described. ''Facing east, at a two-way intersection, ahead is a 90 foot hallway ending at a three-way intersection, to the right is a 30 foot hallway ending at a two-way intersection to the left.''

3. Global verbal mode: Includes the Maplet information and adds a general description of the overall geometric structure of the layout. ''This floor can be thought of as two east–west rectangles connected by a north–south hallway. On the north end there is a 75 foot by 30 foot rectangle. On the south end there is a 60 foot by 45 foot rectangle. A 120 foot hallway connects the south west corner of the northern rectangle to the middle of the southern rectangle. You are at the northwest corner of the northern rectangle.'' The global message was immediately followed by a maplet description. (To reduce verbosity, the full global description was only spoken three times, at the beginning, one-quarter, and three-quarters of the way through the training period. The maplet description was given the rest of the time.)

*Movement behavior*

In order to explore and learn the environments, participants in the verbal study walked blindfolded around the floor guided by an experimenter. At each intersection (decision point) they were given a verbal message specifying their orientation and describing the layout geometry at this location; the amount of information provided depended on the verbal mode. Upon receiving the verbal description they could either ask to have the message repeated or could tell the experimenter which direction they wished to walk. Participants in the visual study explored the floor by walking around under normal viewing conditions.

In order to quantify participant's movement behavior, their search trajectory was logged on a laptop computer by an accompanying experimenter using custom software designed in our lab. The software approximated the layout by a map that broke the network of hallways comprising the floor into equal 15 foot corridor segments, each separated by a node. A move was defined as traversal of one segment between two nodes in the map.

*Design and procedure*

The verbal study used a within-subjects design with floor by verbal mode order counterbalanced using a Latin square. All participants were blindfolded during the training and testing experimental phases and none were familiar with the environments. Participants trained and tested in each of the three verbal conditions and the entire experiment took approximately 3 h per subject. The experimental paradigm included three phases: a practice session, a training period and a testing phase. During the practice session, participants were shown examples of each type of intersection and given the corresponding description that would be heard from all three verbal conditions. To ensure that participants fully understood the verbal messages, practice continued until they were able to navigate a sample layout from all three verbal view-depth conditions and describe two examples of all intersection types using the terminology of each.

*Training period: verbal study*

During the training period, participants freely explored three environments, one from each of the three verbal view-depth modes. They were blindfolded, guided to the training environment and started from a random location on the floor. They were instructed to use the verbal information provided by the experimenter to explore the entire layout (cover all parts of the floor) and find four target locations. No explicit route information was given. The targets consisted of high imagery words (e.g. chair, desk, pen and book). Each target's name was spoken upon intersecting its $x$–$y$ location. When found, participants were asked to imagine a picture of the target to help facilitate memory of its location on the floor.

Participants were guided without description along whatever corridor they designated until they reached the next intersection (or until the subject directed the experimenter to stop or deviate from this trajectory). Rotations were permitted at any point of travel and a ''repeat'' query could be requested as many times as necessary for any message. Participants freely searched the training environment for a fixed time period and were alerted when 50 and 75% of their time had elapsed. Training in the Local condition occurred for 20 min; 25 min training was allowed for the Maplet and Global conditions. Pilot testing indicated that these training periods were sufficient to learn the environments. The difference in training time between conditions was designed to equate learning by accounting for differences in the length of the verbal message. This was confirmed by post-hoc analyses showing that the average distance traveled during training did not differ by more than two percent between any of the verbal modes.

*Testing phase: verbal study*

Following the training period, participants were tested on their knowledge of the environment. The method of movement and presentation of verbal information was identical to training except that during testing, rather than freely exploring the floor, participants engaged in a directed search task. For this task, knowledge of the floor was assessed by their ability to plan and execute routes between pairs of targets. Before beginning the testing phase, participants were disoriented by walking them blindfolded along a circuitous route through part of the floor not used during the training period. The disorientation route terminated at one of the target locations, which became the starting point for the experimental trials. Participants were given the name of the target where they were standing and a description of the intersection geometry at this location and then asked to find the shortest path to another target in the environment. Participants had to say ''this is the target,'' once they believed they had navigated to the specified location. If incorrect, they were brought to the correct target location before beginning the next test trial. Participants were required to find routes between four target pairs, the order of which were counterbalanced.

*Training and test for visual study*

This study took approximately 1 h and employed a between-subjects design. Each participant trained and tested on one layout and the subject by floor assignments were randomized across the two experimental environments used. The training and test procedures for the visual study were very similar to those used with verbal learning except that rather than using verbal descriptions participants navigated the environments using vision in both experimental phases. There was also no manipulation of view-depth. As with the verbal experiment, targets were only auditory, with the name spoken when the participant reached the target location. Rather than learning for a fixed number of minutes, as was done in the verbal conditions, training was based on an equivalent number of moves (15 foot segments). The training period consisted of moving three times the total number of segments in the layout. This guaranteed equal training experience across participants, irrespective of walking speed or layout size. Although the learning criterion here was based on movement rather than time, the amount of training was almost identical between conditions, with visual exploration averaging 105 moves on a 35-segment layout, compared to verbal exploration which averaged 103.5 moves on a 34.5-segment layout. The testing phase of the visual study was identical to the verbal study except that participants navigated between requested target pairs using vision rather than verbal descriptions.

## Results and discussion

*Training phase: measures of search behavior*

Four measures were assessed from the training period:

1. Floor coverage percent: expressed as the percentage of unique nodes traversed during search relative to the total number of nodes. This measure provides an indication of how well participants were able to use the verbal descriptions to visit the entire layout (perform an exhaustive search).
2. Unique targets encountered: expressed as a percentage of the number of unique targets visited during training relative to the total number (four) of targets in the layout.
3. Number of shortest paths traversed: the sum of all direct routes taken between target locations during the search period (recall that no explicit instructions were given about routes). Only shortest paths, those with the minimum number of intervening nodes between target pairs, were scored. Multiple traversals of the same route were counted in the total sum and a route traveled in one direction was considered a different route than when traveled in the other direction, e.g. route a, b, c and c, b, a were considered two separate routes.

4. Entropy: used to characterize the distribution of moves during the search. A high entropy score indicates that participants are equally distributing their movement across the entire environment; a low entropy score indicates that they are concentrating their search to specific regions of the layout (Schlicht, 2001). A high floor coverage percentage is necessary to attain high entropy, but by itself it is not sufficient. For example, a participant could traverse the entire floor but concentrate the majority of their moves in a single region, resulting in a high floor coverage percentage with an entropy score well below the maximum value. Entropy of an environment, $H(\mathrm{e})$, is expressed by the following equation: $H(e) = -\sum_{x} p(x)\log^2[p(x)]$. where $e$ = the environment and $x$ is an individual node. The probability the subject visited an individual node is $p(x)$; calculated from the number of times node $x$ was reached divided by the total number of forward moves executed during training.

Although we predicted that search performance would be lowest using local verbal information, no statistically significant differences were observed between verbal modes for any of the training measures. Therefore, the results are collapsed across view depth in the following discussion. Table 1 shows verbal and visual performance for each of the training measures.

The finding that across all view conditions participants covered an average of 96.6% of the floors during their search demonstrates the ability of verbal descriptions to support free exploration tasks. The theoretical minimum floor coverage to reach all of the unique targets was 57.2%. The finding that participants covered almost 97% of the floors demonstrates that the ability to perform an exhaustive search was not simply a consequence of traveling between target locations. Considering that subjects had no a priori knowledge about target locations or connecting routes, it is particularly noteworthy that all participants in the verbal study found 100% of the hidden target locations and traveled an average of 10.89 shortest paths between these locations during the training period. These data suggest that participants were effectively updating their position in the environment and weighting their search toward a route-finding strategy. This hypothesis is further bolstered by the findings of a random walk model, discussed in Study 2, demonstrating that human performance cannot be accounted for by chance decision-making behavior. High entropy scores, about 4.9 were observed for all verbal view depth conditions (by comparison the theoretical maximum value is 5.1). This indicated that participants adopted a distributed search strategy rather than concentrating their movement to particular regions of the floor. Finally, the comparable performance observed across training measures between verbal and visual conditions suggests that search behavior can be as effective from verbal descriptions as from visual input. The training measures are near ceiling by design because we wanted to allow ample training to ensure that the environments were well learned by all participants before moving on to the test phase.

**Table 1** Training measures for the three verbal view-depth conditions and visual control study

| Study | Floor coverage (%) | Unique targets encountered (%) | Number of shortest paths traversed | Entropy |
|---|---|---|---|---|
| Verbal Global | $M$ = 98.0 SE = 0.9 | $M$ = 100.0 SE = 0.0 | $M$ = 10.00 SE = 1.17 | $M$ = 4.91 SE = 0.02 |
| Verbal Maplet | $M$ = 95.4 SE = 1.3 | $M$ = 100.0 SE = 0.0 | $M$ = 10.47 SE = 0.88 | $M$ = 4.89 SE = 0.03 |
| Verbal Local | $M$ = 96.4 SE = 1.4 | $M$ = 100.0 SE = 0.0 | $M$ = 12.20 SE = 1.05 | $M$ = 4.89 SE = 0.03 |
| Verbal Combined | $M$ = 96.6 SE = 0.7 | $M$ = 100.0 SE = 0.0 | $M$ = 10.89 SE = 0.61 | $M$ = 4.89 SE = 0.01 |
| Visual Control | $M$ = 96.8 SE = 1.2 | $M$ = 100.0 SE = 0.0 | $M$ = 7.54 SE = 0.58 | $M$ = 4.86 SE = 0.03 |

Each cell represents the mean and standard error for each of 15 participants in the verbal study and 13 participants in the visual study

The verbal data was compared within subjects by view-depth condition and between subjects with the visual control data for each of the training measures. A one-way repeated measures ANOVA comparing the three levels of verbal view-depth (Global, Maplet, Local) was conducted for floor coverage, $F(2, 28) = 1.667$, $P = 0.207$, entropy, $F(2, 28) = 0.293$, $P = 0.748$, and number of shortest paths traversed, $F(2, 28) = 1.039$, $P = 0.367$. (Comparisons were not made between unique targets encountered, as every participant found all four targets).

Independent sample $t$ tests were performed to compare the three levels of verbal view-depth to the visual control data. Bonferroni correction was used to guard against inflation of the alpha level, requiring $P < 0.017$ to attain significance. Reliable differences were only observed for the shortest paths measure between the visual control group ($M = 7.54$) and the Maplet ($M = 10.47$), $t(26) = 2.663$, $P = 0.013$, and Local ($M = 12.20$), $t(26) = 3.713$, $P = 0.001$, conditions of the verbal study. The significant difference in shortest paths between studies is likely attributed to the relative ease of accessing distal information with

vision, making it unnecessary to travel a path, rather than verbal learning being more amenable than vision for route finding strategies.

*Test phase: measures of spatial ability*

Two measures were assessed in the testing phase (see Table 2):

1. Target localization accuracy: the percentage of target locations that were correctly found during test divided by the total number of target localization trials (four).
2. Route efficiency: expressed as the length of the route executed between target locations divided by the length of the actual route (only applies to correctly localized targets). Route length is defined by the number of intervening nodes along the shortest path between the origin and destination target location.

Performance on this task was high for both studies. Targets were correctly localized in the verbal study ($M = 85.0\%$) and in the visual study ($M = 92.3\%$). Target localization accuracy in both studies was significantly above chance which is ~3%, defined as 1 over the number of possible target locations [verbal: $t(44) = 57.96$, $P < 0.001$ and visual: $t(12) = 44.24$, $P < 0.001$.] A one-way repeated measures ANOVA found no significant difference in target localization accuracy between the three verbal view depths, $F$ (2, 28) = 1.248, $P = 0.303$. Likewise, independent sample $t$ tests revealed no significant differences for target accuracy between the Global, Maplet, and Local verbal view-depth conditions and the visual control group, $P$'s > 0.05.

**Table 2** Test phase percentage measures of target localization and route finding by verbal view-depth and visual control

| Study | Target accuracy at test (%) | Route efficiency at test (%) |
|---|---|---|
| Verbal Global | $M = 85.0$ SE = 6.4 | $M = 98.7$ SE = 0.1 |
| Verbal Maplet | $M = 90.0$ SE = 4.1 | $M = 98.1$ SE = 1.9 |
| Verbal Local | $M = 80.0$ SE = 7.8 | $M = 88.5$ SE = 5.7 |
| Verbal Combined | $M = 85.0$ SE = 6.1 | $M = 95.1$ SE = 2.5 |
| Visual Control | $M = 92.3$ SE = 4.6 | $M = 98.9$ SE = 1.2 |

Each cell for target accuracy represents the mean and standard error of four trials for each of 15 participants (verbal study) and 13 participants (visual control). Route efficiency is based only on the correct trials for each of the Ss.

Although a difference in view-depth was predicted, with the local information expected to yield the worst performance, the results indicate that the increased spatial integration demands associated with access to purely local information do not adversely affect route-finding performance. This finding agrees with the results from the training measures and provides evidence that describing minimal geometric information is sufficient to support verbal learning and wayfinding.

During route finding in the verbal study, the overall mean for optimal path selection was $M = 95.1\%$, indicating that when target localization was successful, efficient routes were learned and executed. For the verbal study, a one-way repeated measures ANOVA was used with a Greenhouse-Geisser correction, which adjusts the degrees of freedom to account for unequal variances. No significant differences were observed between view-depth for route efficiency, $F$ (1.281, 17.938) = 3.675, $P = 0.063$. Likewise, independent sample $t$ tests revealed no significant difference in route efficiency between verbal view-depths (Global, Maplet, and Local) and the visual control group, $P$'s > 0.017. While both mean target accuracy and route efficiency were lowest in the Local condition, these differences were small and failed to reach statistical significance. Nevertheless, it is possible that a subset of participants had a harder time orienting themselves in the environment when using local information but were able to correct for their uncertainty by taking a somewhat longer route to the target. Of greater importance is the lack of reliably different results on any of the test conditions as a function of training modality. Corroborating the findings of the training measures, results from the test phase demonstrate that verbal learning is on par with visual learning. It is likely that with a substantially larger sample for both the verbal and visual groups, a significant difference would have been found in target accuracy. However, given the small observed effect size, $\eta^2 = 0.082$, such a difference would not be particularly meaningful.

*Comparing route traversal at training and test*

The last series of measures was aimed at characterizing the structure of the spatial representation built up from the training period. The goal here was to determine whether cognitive maps were route-based or map-like in nature. To address this question, we compared the

**Table 3** Verbal and visual contingency table of training experience versus target localization accuracy at test

| Training experience/target accuracy performance | Correctly localized | Incorrectly localized | Total count training experience |
|---|---|---|---|
| Route traversed at training | | | |
| Verbal | 40 | 5 | 45 (43.7%) |
| Visual | 10 | 1 | 11 (29.7%) |
| Route not traversed at training | | | |
| Verbal | 48 | 10 | 58 (56.3%) |
| Visual | 24 | 2 | 26 (70.3%) |
| Total count for target accuracy | | | |
| Verbal | 88 (85.4%) | 15 (14.6%) | 103 (100%) |
| Visual | 34 (91.9%) | 3 (8.1%) | 37 (100%) |

Note that the count for routes traversed during training represents only exact matches with the test routes, i.e. the shortest path in the same direction. Only routes with one shortest path were analyzed; multi-path routes were not considered. Thus, the verbal data in the table represent 103 of the total 180 target localization trials given at test and for visual data, 36 out of the 52 target localization trials. Data is summed across participants

proportion of success at test for people who followed the route at training versus the proportion of success at test for people who did not previously experience the route (see Table 3).

Observations of route traversal and target accuracy were not independent within each participant. Thus it was necessary to evaluate if there was a relationship between these two variables using a Cochran–Mantel–Haenszel (CMH) chi-square test, stratified by participant. Using a CMH test, no significant general association was found between route traversal during training and subsequent target localization accuracy for either the verbal study, $\chi^2(1) = 0.106$, $P = 0.744$, or the visual control study, $\chi^2(4) = 2.000$, $P = 0.1573$. These data show that experience with a route at training has no effect on the probability of success at test.

As can be seen from the verbal contingency table, 48 of the total 88 correctly localized targets (54.5%) were found using routes not previously traveled during training. Of the 58 total novel test routes traversed, targets were localized with 82.8% accuracy. A similar pattern of results was found with the visual study, 26 novel test routes were executed with a target accuracy of 92.3%. The significance of these findings are twofold. First, the data demonstrate that people are not simply executing a sequence of distance and turn information to traverse a remembered route between targets but that they are able to plan and infer paths, optimally in most instances, from their cognitive map. Second, following from the above, the similar performance between verbal and visual conditions suggests that both modes of environmental learning lead to the development of functionally similar spatial representations in memory.

## Study 2: comparing human performance to a random walk model

The consistently accurate results from verbal learning found in the first experiment were interpreted as demonstrating the efficacy of updated verbal descriptions in promoting effective wayfinding behavior. However, without an explicit test, it is not known how the search trajectories of our human navigators compare to what would be expected from a search based on random decision-making behavior.

The purpose of the second study was to compare the human training behavior in the verbal conditions of the first study with two versions of a Monte Carlo simulation of a random walk model. By comparing the training trajectories of the models to those of the human participants on the same floors and training measures, we can address whether human search behavior significantly differs from what would be observed by chance, i.e. use of cognitive strategies rather than random decision-making.

### Methods

Each model was given 3,000 random walk trials in the same environments used by human participants in the verbal conditions of Study 1. The models were allowed approximately the same number of moves during ''training'' as human participants. The training periods of 92, 97, and 111 moves for the three layouts used in the simulation were determined by taking the average number of nodes covered per floor across all human participants in the Local condition. Since intersection geometry is the only information available to the models, comparisons were made with human

performance in the Local condition only, as this mode provides a verbal description of essentially the same information.

As with the human participants, each simulation run of the models was started at a random position in the environment and its trajectory logged as it ''explored.'' Since human participants rarely turned around between intersections, e.g. made a 180° turn midway along a hallway, decision-making for both models was limited to intersections. Decisions were made at random, as the model had no prior knowledge about the environment and retained no memory of where it had already been during the simulation. The random walk models differed in that the first, called ''unconstrained,'' made purely random movement decisions about which branching corridor to follow at each intersection. Thus, at a two-way intersection, it would have a 50–50 chance of progressing along each branch. The second random walk model, called ''constrained,'' made a random decision about which corridor branch to follow, excluding the corridor just traversed, i.e., the corridor behind. This means that at a two-way intersection, the constrained model had only one choice. At a three-way intersection, it had two choices, etc. Navigation was restricted to valid paths of travel for both models, so the movement probabilities were contingent on the number of corridor legs making up the intersection. For instance, at a four-way intersection each of the four corridors had an equal probability of being chosen, chance is 25%, by the unconstrained model; chance for the constrained model at the same intersection is 33% as it does not consider the hallway from which it came as a valid decision. The only exception is at a dead-end, where the constrained model is forced to make a 180° rotation. Since human participants tended to weigh their search toward forward movements and rarely reversed their route, this simple constraint was added to determine if it would allow the model to perform more similarly to human behavior.

## Results and discussion

Figure 3 shows training performance for the 15 human verbal learning participants of Study 1 and 3,000 walks for each model (1,000 on each of the three layouts).

The primary rationale for this study was to demonstrate that human search behavior (Study 1) differs from what would be expected from chance performance. As is obvious from Fig. 3, human search performance was vastly better than the unconstrained random walk model for all of the training measures. Although this may seem trivially obvious, due to the unfamiliarity of verbally based wayfinding by most people and the lack of any research precedent on the task, it was not clear how people would perform. If they were completely confused by the verbal descriptions, it is possible that they might wander around the environments in a random fashion. However, these data demonstrate that human behavior cannot be attributed to purely chance decision-making and that they employed specific search strategies to accomplish the tasks. One strategy is not to turn back unless necessary. The purpose of the constrained model was to investigate whether adding this movement restriction would make the model perform more like human behavior. As is apparent from the performance of the constrained model in the figure, the addition of this

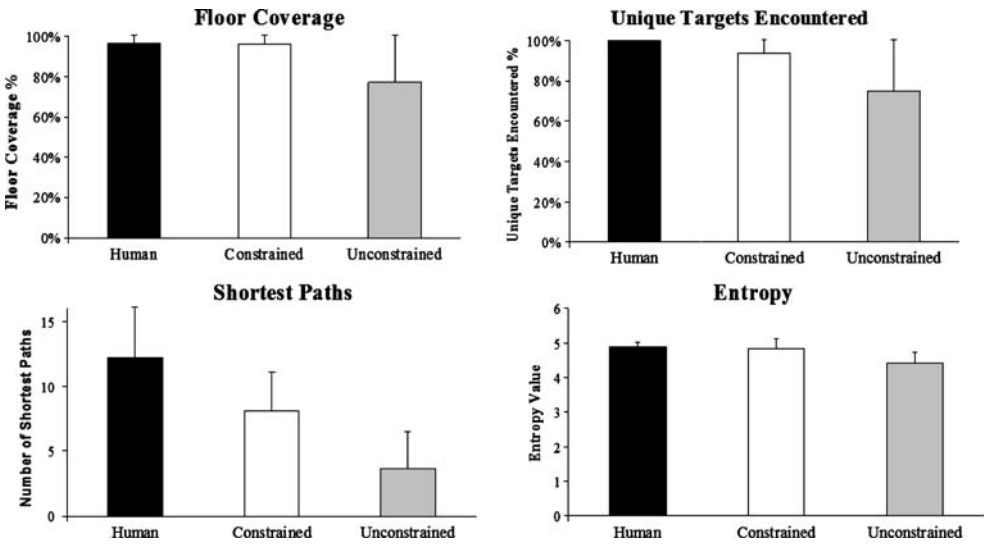


**Fig. 3** Comparison of human verbal performance to random walk models (constrained and unconstrained). *Error bars* represent one standard deviation. Human data has 15 participant trials per cell and model data has 3,000 simulation trials per cell

simple deterministic constraint does indeed close much of the gap between the humans and the unconstrained random walker. This result may imply that a small set of very simple strategies, one already identified and incorporated into the constrained model, can account for general properties of human exploration of novel environments.

Although our interest was between human performance and each of the models, all of the levels were first analyzed with an omnibus *F*-test using Type III sums of squares, which was the appropriate procedure given the unbalanced design. The subsequent post-hoc tests present the focused comparisons between the two groups (humans and models). One-way ANOVAs yielded significant differences for floor coverage, $F(2, 6{,}012) = 1613.299$, $P < 0.001$, number of shortest paths traversed, $F(2, 6{,}012) = 1710.236$, $P < 0.001$, and entropy, $F(2, 6{,}012) = 1841.215$, $P < 0.001$. Statistical comparisons could not be made for ''unique targets encountered,'' since all participants found 100 percent of the targets. Nevertheless, human performance was notably superior to both models. All post-hoc analyses were conducted using Tamhane's T2 which is appropriate for unequal variances and unequal group sizes (Toothaker, 1993). Not surprisingly, human performance was significantly better than the unconstrained model for all measures, *P*'s $< 0.001$. However, human performance was reliably higher than the constrained model for only the percentage of unique targets encountered and number of shortest paths traversed, *P*'s $< 0.001$.

The greatest similarity between humans and the constrained model was observed for entropy and floor coverage. These measures relate to the breadth and depth of the search and may be affected by layout geometry. For instance, the floor coverage exhibited by the constrained model may be somewhat inflated as there were no four-way intersections, and 1/3 of the intersections were dead-end or two-way junctions, both of which leave the constrained model with only one possible movement decision. The similarity between the constrained model and human behavior may be less evident given a more complex floor.

The result that human performance was significantly better than both models on finding the hidden targets and executing shortest paths between these targets provide compelling evidence that human performance cannot be completely attributed to use of a simple forward bias in the local decision statistics. The ability for humans to navigate such a large number of optimal routes between targets implies that they were accessing a spatial representation of the environment from memory and accurately updating their position and orientation as they navigated. Future experiments should address whether adding other simple constraints on the model would lead to greater similarity to the search strategies adopted by human navigators.

## General discussion

The major goals of this research were to establish whether or not access to dynamically updated verbal descriptions of layout geometry support effective searching and wayfinding in complex indoor layouts and if verbal learning performance was comparable to visual learning on the same tasks. Four major findings should be highlighted from this research.

1. Open searching of complex indoor layouts can be accomplished using dynamically updated verbal descriptions. Previous research has shown that the use of point-based verbal displays to provide updated spatial information promoted route navigation (Loomis *et al.*, 1998; Loomis, Golledge, & Klatzky, 2001). The current research, using geometric-based displays, demonstrates that updated verbal information also supports free exploration and environmental learning of novel indoor layouts. The results comparing training performance between the human participants and the random walk models reveal that human search behavior was not based on chance decision making (unconstrained model) but that adding a ''do not turn around'' constraint to the model closed much of the performance gap between humans and the constrained random walker, at least for the floor coverage and entropy measures. In contrast to human participants, the number of shortest paths traveled between target locations was significantly lower for both models (33.9% fewer for the constrained model and 70.1% fewer for the unconstrained model). These results show that the ability to keep track of targets and travel routes between these locations cannot be accounted for by random decision making and suggest that participants were using effective spatial updating strategies to perform these tasks.
2. Use of verbal descriptions during free exploration leads to accurate environmental learning, as evidenced by participant's 85.0% target localization accuracy at test. Furthermore, the finding that the majority of routes between targets at test had not been previously traversed at training

demonstrates that symbolic verbal descriptions develop into an accurate spatial representation. These findings contrast with the traditional view of human spatial knowledge acquisition (Siegel & White, 1975), which postulates a map-like representation of layout configuration is not developed without extensive prior learning and navigation of routes.

3. The highly similar pattern of learning and wayfinding behavior observed between verbal and visual conditions indicate that the spatial representation built up from verbal learning is functionally similar to that developed from visual learning. The comparable performance between modalities demonstrates the efficacy of updated verbal descriptions to support spatial operations normally subserved by visual input.
4. The absence of significant differences between verbal view-depth conditions at test was not expected as the reduced spatial integration demands of the Maplet and Global conditions were predicted to improve knowledge of layout configuration. Although there was a trend towards lower performance in the Local condition for test target accuracy and route efficiency, conclusive evidence cannot be drawn from this small decrement. The lack of a ''view-depth effect'' suggests that minimal geometric information is sufficient to perform searching and wayfinding tasks. The findings from the constrained random walk model corroborate this conclusion, as it also exhibited accurate exhaustive search behavior on the basis of information, which was essentially equivalent to what was available from the local verbal condition. Taken together, the data show that increasing access to layout geometry yields little performance benefit, at least for the tasks employed in these experiments. While the extra information provided by more complex geometric descriptions is not necessary for accurate searching and route finding, access to this information may prove beneficial for performing other kinds of spatial operations, such as tasks requiring explicit knowledge of metric relations or layout topology.

The results of these experiments demonstrate that verbal descriptions can do far more than describe static scenes, specify landmark locations, and provide sequential route directions. Similar to visual apprehension of the environment, updated verbal descriptions are an effective medium for describing environmental relations and supporting nonvisual learning and wayfinding behavior in large-scale unfamiliar layouts.

**Acknowledgments** This research was supported by NIDRR grant H133A011903 and NIH training grant 5T32 EY07133.

## References

Allen, G. L. (1997). From knowledge to words to wayfinding: Issues in the production and comprehension of route directions. In S. C. Hirtle & A. U. Frank (Eds.), *Lecture notes in computer science* (Vol. 1329, pp. 363–372). Berlin Heidelberg New York: Springer.

Allen G. L. (2000). Principles and practices for communicating route knowledge. *Applied Cognitive Psychology, 14*, 333–359.

Benhamou, S. (1998). Place navigation in mammals: A configuration based model. *Animal Cognition, 1*, 55–63.

Cheng, K. (1986). A purely geometric module in the rat's spatial representation. *Cognition, 23*, 149–178.

Couclelis, H. (1996). Verbal directions for way finding: Space, cognition and language. In J. Portugali (Ed.), *The construction of cognitive maps* (pp. 133–153). Dordrecht: Kluwer.

Denis, M., & Cocude, M. (1989). Scanning visual images generated from verbal descriptions. *European Journal of Cognitive Psychology, 1*, 293–307.

Denis, M., & Cocude, M. (1997). On the metric properties of visual images generated from verbal descriptions: Evidence for the robustness of the mental scanning effect. *European Journal of Cognitive Psychology, 9*, 353–379.

Denis, M., Pazzaglia, F., Comoldi, C., & Bertolo, L. (1999). Spatial discourse and navigation: An analysis of route directions in the city of Venice. *Applied Cognitive Psychology, 13*, 145–174.

Denis, M., & Zimmer, H. D. (1992). Analog properties of cognitive maps constructed from verbal descriptions. *Psychological Research, 54*, 286–298.

Ehrlich, K., & Johnson-Laird, P. N. (1982). Spatial descriptions and referential continuity. *Journal of Verbal Learning & Verbal Behavior, 21*, 296–306.

Ferguson, E. L., & Hegarty, M. (1994). Properties of cognitive maps constructed from texts. *Memory & Cognition, 22*(4), 455–473.

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

Gallistel C. R. (1990). *The organization of learning.* Cambridge, MA: The MIT Press.

Giudice, N. (2004). *Navigating novel environments: A comparison of verbal and visual learning.* Unpublished Dissertation, University of Minnesota, Twin Cities, MN.

Hermer, L., & Spelke, S. S. (1994). A geometric process for spatial reorientation in young children. *Nature, 370*, 57–59.

Hirtle, S. C., & Heidorn, P. B. (1993). The structure of cognitive maps: Representations and processes. In T. Garling, &R. G. Golledge (Eds.), *Behavior and environment: Psychological and geographical approaches* (pp. 170–192). Amsterdam: North-Holland.

Johnson-Laird P. N. (1983). *Mental models*. Cambridge, MA: Harvard University Press.

Levelt, W. J. M. (1996). Perspective taking and ellipsis in spatial descriptions. In P. Bloom, M. A., Peterson, M. F., Garrett, & L., Nadel. (Eds.), *Language and space* (pp. 77–107). Cambridge, MA: MIT Press.

Loomis, J., Golledge, R., & Klatzky, R. (1998). Navigation system for the blind: Auditory display modes and guidance. *Presence, 7*, 193–203.

Loomis, J. M., Golledge, R. G., & Klatzky, R. L. (2001). Gps-based navigation systems for the visually impaired. In W. Barfield, &T. Caudell (Eds.), *Fundamentals of wearable computers and augmented reality* (pp. 429–446). Mahwah, NJ: Lawrence Erlbaum Associates.

Loomis, J. M., Marston, J. R., Golledge, R. G., & Klatzky, R. L. (2005). Personal guidance system for people with visual impairment: A comparison of spatial displays for route guidance. *Journal of Visual Impairment & Blindness, 99*, 219–232.

Lovelace, K., Hegarty, M., & Montello, D. (1999). Elements of good route directions in familiar and unfamiliar environments. In C. F. D. M. Mark (Eds.), *Spatial information theory: Cognitive and computational foundations of geographic information science* (pp. 65–82). Berlin Heidelberg New York: Springer.

O'Keefe J., Nadel L. (1978). *The hippocampus as a cognitive map*. London: Oxford University Press.

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

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

Schlicht, E. J., Legge, G.E., Stankiewicz, B.J. & Giudice, N.A. (2001). *Are visual landmarks necessary for effective transfer of navigational knowledge between real and virtual buildings.* Paper presented at the Annual Meeting of the Association of Research in Vision and Ophthalmology, Ft. Lauderdale, FL.

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

Talmy, L. (1983). How language structures space. In H. Pick, & Acredolo, L. (Ed.), *Spatial orientation: Theory, research, and application*. New York, NY: Plenum Press.

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

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

Tom, A., & Denis, M. (2003). Referring to landmark or street information in route directions: What difference does it make? In W. Kuhn, M. F. Worboys, and S. Timpf (Eds.), *Spatial information theory: Foundations of geographic information science* (pp. 384–397). Berlin Heidelberb New York: Springer.

Toothaker L. E. (1993). *Multiple comparison procedures*. Newbury Park, CA: Sage.

Tversky, B. (1996). Spatial perspective in descriptions. In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.). *Language and space* (pp. 463–492). Cambridge, MA: MIT Press.

Wilson, P. N., Tlauka, M., & Wildbur, D. (1999). Orientation specificity occurs in both small- and large-scale imagined routes presented as verbal descriptions. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 25*(3), 664–679.

Zwaan, R.A. (1998). Situation models in language and memory. *Psychological Bulletin, 123*, 162–185.