Learning with Virtual Verbal Displays: Effects of Interface Fidelity on Cognitive Map Development

Nicholas A. Giudice and Jerome D. Tietz*

University of California, Santa Barbara
Department of Psychology
Santa Barbara, CA 93106-9660
giudice@psych.ucsb.edu

Abstract. We investigate verbal learning and cognitive map development of simulated layouts using a non-visual interface called a virtual verbal display (VVD). Previous studies have questioned the efficacy of VVDs in supporting cognitive mapping (Giudice, Bakdash, Legge, & Roy, in revision). Two factors of interface fidelity are investigated which could account for this deficit, spatial language vs. spatialized audio and physical vs. imagined rotation. During training, participants used the VVD (Experiments 1 and 2) or a visual display (Experiment 3) to explore unfamiliar computer-based layouts and seek-out target locations. At test, participants performed a wayfinding task between targets in the corresponding real environment. Results demonstrated that only spatialized audio in the VVD improved wayfinding behavior, yielding almost identical performance as was found in the visual condition. These findings suggest that learning with both modalities led to comparable cognitive maps and demonstrate the importance of incorporating spatial cues in verbal displays.

Keywords: wayfinding, verbal learning, spatialized audio, interface fidelity.

1 Introduction

Most research investigating verbal spatial learning has focused on comprehension of route directions or the mental representations developed from reading spatial texts [1-4]. Owing to this research emphasis, there is much less known about the efficacy of verbal information to support real-time spatial learning and navigation. What distinguishes a real-time auditory display from other forms of spatial verbal information is the notion of dynamic updating. In a dynamically-updated auditory display, the presentation of information about a person's position and orientation in the environment changes in register with physical movement. For example, rather than receiving a sequential list of all the distances and turns at the beginning of a route, as is done with traditional verbal directions, a real-time display provides the user with context-sensitive information with respect to their current location/heading state as they

* The authors thank Jack Loomis for insightful comments on the manuscript, Maryann Betty for experimental preparation, Brandon Friedman for assistance in running participants, and Masaki Miyanohara for helping with running participants and data analysis. This work was supported by an NRSA grant to the first author, #1F32EY015963-01.

C. Freksa et al. (Eds.): SC 2008, LNAI 5248, pp. 121–137, 2008.

progress along the route. Vehicle-based navigation systems utilizing GPS and speech-based route directions represent a good example of these dynamic displays. Dynamic auditory interfaces also have relevance in navigation systems for the blind, and in this context, they have proven extremely effective in supporting real-time route guidance [see 5 for a review].

Rather than addressing route navigation, the current research uses free exploration of computer-simulated training layouts to investigate environmental learning. The training environments are explored using a non-visual interface called a virtual verbal display (VVD). The VVD is based on dynamically-updated geometric descriptions, verbal messages which provide real-time orientation and position information as well as a description of the local layout geometry [see 6 for details]. A sample output string is: "You are facing West, at a 3-way intersection, there are hallways ahead, left, and behind." If a user executed a 90° left rotation at this t-junction, the VVD would return an updated message to reflect that he/she was now facing South, with hallways extending ahead, left, and right.

We know that geometric-based displays are extremely effective for supporting free exploration (open search) in both real and computer-based layouts [7-9]. However, their efficacy for supporting cognitive map development is unclear. That is, participants who trained using a virtual verbal display to search computer-based environments performed significantly worse on subsequent wayfinding tests in the corresponding real environment [7, 8] than subjects who trained and tested exclusively in real environments [9]. These findings suggest that training with a virtual verbal display results in impoverished environmental learning and cognitive map development compared to use of the same verbal information for searching real building layouts. This deficit cannot be attributed to environmental transfer more generally, as previous studies have demonstrated that learning in virtual environments (VEs) transfers to accurate real-world navigation, even with perceptually sparse visual displays similar to our geometric verbal display [10-12].

The current studies investigate several factors of interface fidelity which may account for problems in spatial knowledge acquisition with the VVD. As described by Waller and colleagues [13], interface fidelity refers to how the input and output of information from the virtual display is used, i.e. how one's physical actions affect movement in the VE and how well feedback from the system supports normal perceptual-motor couplings. These interactions can be distinguished from factors relating to environment fidelity, which refers to how well the information rendered in the VE resembles the real environment, e.g. sensory richness, spatial detail, surface features, and field of view [13]. Our previous work with VVDs dealt with environment fidelity, investigating whether describing more of the layout from a given vantage point, called "verbal view depth," would facilitate learning of global structure and aid subsequent wayfinding behavior. However, the lackluster environmental transfer performance with three levels of verbal view depth, ranging from local to global descriptions, demonstrated that deficits in cognitive map development were not due to availability of environmental information but to the interface itself [7, 8].

The current experiments hold environmental variables constant and manipulate several factors relating to interface fidelity. Experiment 1 compares traditional verbal descriptions, where the message is delivered as a monaural signal to both ears, with spatialized audio descriptions, where the message is heard as coming from a specific direction, e.g. a hallway to the left would be heard as a description emanating from the navigator's left side. Experiment 2 addresses the influence of body-based information, e.g. physical rotation vs. imagined rotation. Experiment 3 follows the same design of the first two verbal studies but uses a visual display as a control. All experiments incorporate training in computer-based layouts and environmental transfer requiring wayfinding in the corresponding real environment. Our focus is on the transfer tests, as they provide the best index of environmental learning and cognitive map development.

2 Experiment 1

In this study, blindfolded participants are given a training period where they use verbal descriptions to freely explore unfamiliar computer-based floors of university buildings and seek out four target locations. At test, they must find routes between target pairs in the corresponding real environment. This design is well-suited for addressing environmental learning, as theories of cognitive map development have long emphasized the importance of free exploration and repeated environmental exposure [14, 15]. The wayfinding test represents a good measure of cognitive map accuracy, as performance cannot be accomplished using a route matching strategy. Since no routes are specified during training, accurate wayfinding behavior requires subjects to form a globally coherent representation of the environment, i.e. the trademark of a cognitive map [16].

Our previous work with virtual verbal displays was based exclusively on spatial language (SL), i.e. consistent, unambiguous terminology for describing spatial relations [17]. The problem with any purely linguistic display is that the information provided is symbolic. A description of a door at 3 o'clock in 10 feet has no intrinsic spatial content and requires cognitive mediation to interpret the message. By contrast, a spatialized audio (SA) display is perceptual, directly conveying spatial information about the environment by coupling user movement with the distance and direction of object locations in 3-D space. For instance, rather than describing the location of the door, the person simply hears its name as coming from that location in the environment.

Several lines of research support the benefit of spatialized auditory displays. Experiments comparing different non-visual displays with a GPS-based navigation system for the blind have shown that performance on traversing novel routes, finding landmarks, and reaching a goal state is superior when guided with spatialized audio versus spatial language [18-20]. Research has also shown that spatialized auditory displays are beneficial as a navigation aid during real-time flight [21] and for providing non-visual information to pilots in the cockpit of flight simulators [22]. It is predicted that spatialized audio displays will have similar benefits on cognitive map development, especially when training occurs in computer-based environments as are used here. Spatial updating and environmental learning is known to be more cognitively effortful in VEs than in real spaces, [23, 24]. However, recent work suggests that SA is less affected by cognitive load than SL during guidance of virtual routes, yielding faster and more accurate performance in the presence of a concurrent distractor task [25]. These findings indicate that the use of SA in the VVD may reduce the working memory demands associated with virtual navigation, thus increasing resources available for cognitive map development.

To address this issue, Experiment 1 compared environmental learning with virtual verbal displays based on Spatial language descriptions about layout geometry [8] with identical descriptions that added spatial information to the signal, e.g. a hallway on the left would be heard in the left ear.

2.1 Method

Participants. Fourteen blindfolded-sighted participants, ages 18-32 (mean = 20.6), balanced equally by gender, ran in the two hour study. Subjects in all experiments were unfamiliar with the test environments, reported normal (or corrected to normal) visual and auditory acuity, gave informed consent, and received course credit for their participation.

Environments and Apparatus. Two simulated floors of the UC Santa Barbara Psychology building, and their physical analogs, were used. The computer-based layouts were rendered to be perceptually sparse, with the verbal messages providing information about the user's facing direction and layout geometry only. The simulated layouts were broken into corridor segments separated by nodes (each segment approximated ten feet in the real space). The floors averaged 100.5 m of hallway extent and 8.5 intersections. Each floor contained four targets which must be found during the search. The name of each target was spoken whenever subjects reached its location in the layout (Ss were told the names of the four targets before starting the trial). Figure 1 shows an illustration of an experimental environment and sample verbal descriptions.

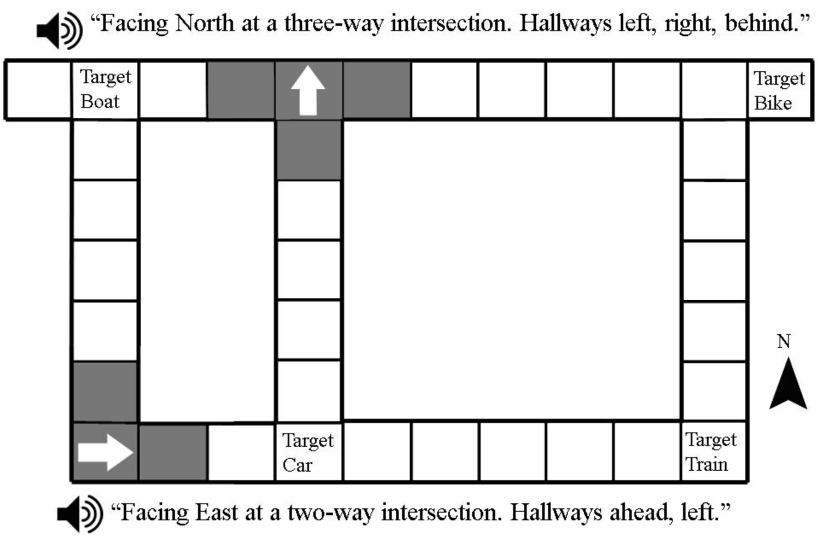

Fig. 1. Experimental layout with target locations denoted. What is heard upon entering an intersection (listed above and below the layout) is depicted in gray. Each arrow represents the orientation of a user at this location.

Participants navigated the virtual environments using the arrow keys on a USB numberpad. Pushing the up arrow (8) translated the user forward and the left (4) and right (6) arrows rotated them in place left or right respectively. Forward movements were made in discrete "steps," with each forward key press virtually translating the navigator ahead one corridor segment (approximately ten feet in the real environment). Left-right rotations were quantized to 90 degrees. Pressing the numpad 5 key repeated the last verbal message spoken and the 0 key served as a "shut-up" function by truncating the active verbal message.

Verbal descriptions, based on a female voice, were generated automatically upon reaching an intersection or target location and rotation at any point returned an updated heading, e.g. "facing north". A footstep sound was played for every forward move when navigating between hallway junctions. Movement transitions took approximately 750 ms. The Vizard 3-D rendering application (www.worldviz.com) was used to coordinate the verbal messages, present a visual map of what was being heard for experimenter monitoring, and to log participant search trajectories for subsequent analyses.

2.2 Design and Procedure

A within subjects design was used with participants running in one spatial language and spatialized audio condition, counterbalanced across the two experimental environments. The experiment comprised three phases. During practice, the movement behavior was demonstrated and participants were familiarized with the speech output from the VVD on a visual map depicting what would be spoken for each type of intersection.

Training Phase. To start the trial, blindfolded participants stood in the center of a one meter radius circle with four three inch RadioShack speakers mounted on tripods (at a height of 152 cm) placed on the circumference at azimuths of 0° (ahead), 90° (right), 180° (behind) and 270° (left). In the SL conditions, the verbal message was simultaneously presented from the left and right speaker only. With the SA conditions, the participant heard the verbal message as coming from any of the four speakers based on the direction of the hallway being described. The spatialized audio messages were generated by sending the signal from the speaker outputs on the computer's sound card (Creative Labs Audigy2 Platinum) to a four-channel multiplexer which routed the audio to the relevant speaker. The input device was affixed via Velcro to an 88 cm stand positioned directly in front of them.

Subjects were started from an origin position in the layout, designated as "start" and instructed to freely explore the environment using the verbal descriptions to apprehend the space and the input device to affect movement. Their task for the training period was to cover the entire layout during their search and to seek out four hidden target locations. Although no explicit instructions were given about search strategy or specific routes, they were encouraged to try to learn the global configuration of the layout and to be able to navigate a route from any target to any other target. The training period continued until the number of forward moves in their search trajectory equaled three times the number of segments comprising the environment. Participants were alerted when 50 % and 75 % of their moves were exhausted.

Testing Phase. Upon completion of the training period, participants performed the transfer tests. Blindfolded, they were led via a circuitous route to the corresponding physical floor and started at one of the target locations. After removing the blindfold, participants were told they were now facing north, standing at target X and requested to walk the shortest route to target Y. They performed this wayfinding task using vision, no verbal descriptions about the environment or target locations were given. Participants indicated that they had reached the destination by speaking the target's name (e.g., "I have reached target dog"). To reduce accumulation of error between trials, they were brought to the actual target location for incorrectly localized targets before proceeding. Participants found routes between four target pairs, the order of which were counterbalanced.

Analysis. Although our focus was on transfer performance, three measures of search behavior were also analyzed from the training phase in all experiments:

1. Floor coverage percent: the number of unique segments traversed during training divided by the total number of segments in the environment.
2. Unique targets percent: ratio of unique targets encountered during training to the total number of target locations (4).
3. Shortest routes traversed: sum of all direct routes taken between target locations during the search period. A shortest route equals the route between target locations with the minimum number of intervening segments.

Two wayfinding test measures were analyzed for all studies during the transfer phase in the real building:

1. Target localization accuracy percent: ratio of target locations correctly found at test to the total number of target localization trials (four).
2. Route efficiency: length of the shortest route between target locations divided by length of the route traveled (only calculated for correct target localization trials).

2.3 Results and Discussion

As predicted, training performance using both VVD display modes revealed accurate open search behavior. Collapsing across SL and SA conditions, participants covered 97.3% of the segments comprising each floor, found 97.3% of the target locations and traveled an average of 9.9 shortest routes between targets. By comparison, the theoretical maximum number of shortest routes traveled during the training period, given 100% floor coverage with the same number of moves, is 14.5 (averaged across floors). Results from the inferential tests provide statistical support for the near identical performance observed between inputs; none of the one-way repeated measures ANOVAs conducted for each training measure revealed reliable differences between SL and SA conditions, all $ps > .1$. Indeed, performance on the training measures was almost identical for all conditions across experiments (see table 1 for comparison of all means and standard errors). These findings indicate that irrespective of training condition, subjects adopted a broadly distributed, near optimal route-finding search strategy.

Table 1. Training Measures of Experiments 1-3 by Condition. Each cell represents the mean (± SEM) on three measures of search performance for participants in experiments 1-3. No significant differences were observed between any of the dependent measures.

Experiment	Condition	Floor Coverage (%)	Unique Targets Hit (%)	Total Shortest Routes Traversed
1 (*N=14*)	Spatialized Audio	98.46(1.35)	98.21(1.79)	10.71(0.98)
	Spatial Language	96.14(2.98)	96.43(3.57)	9.07(1.22)
2 *(N=16)*	Rotation + Spatialized Audio	98.99(0.85)	100.00(0)	12.0625(0.99)
	Rotation + Spatial Language	99.14(0.50)	98.44(1.56)	10.94(1.11)
3 *(N=14)*	Visual Control	97.34(2.46)	96.43(2.43)	11.07(1.29)

Environmental learning/cognitive map development was assessed using a wayfinding test in the physical building. To address the effect of spatialization, one-way repeated measures ANOVAs comparing spatial language and spatialized audio were conducted for the two transfer test measures of target localization accuracy and route efficiency. Participants who trained using spatialized audio in the VVD correctly localized significantly more targets, 76.8% (SE = 6.12) than those who learned with spatial language, 51.8% (SE = 10.63), $F(1,13) = 7.583$, $p = .016$, $\eta^2 = 0.39$. Target localization accuracy in both conditions, average 64.3%, was significantly above chance performance of ~3%, defined as one divided by 33 possible target locations, e.g. a target can be located at any of the 33 segments comprising the average environment, $t(27) = 8.94$, $p < 0.001$.

Route efficiency for correctly localized targets did not reliably differ between conditions, SA, 96.9% (SE = 1.6) and SL, 95.9% (SE = 2.6), ps >.1. Note that route efficiency is calculated for correctly executed routes only. Thus, the near ceiling performance simply means that the routes that were known were followed optimally. It is likely that this measure would be more sensitive to detecting differences between conditions on floors having greater topological complexity.

Experiment 1 investigated the effect of spatialization on environmental learning. Results from the spatial language condition almost perfectly replicated a previous experiment using the same SL condition and near identical design [8]. Both studies showed that training with the VVD led to efficient search behavior but poor wayfinding performance in the real environment, 51.8% target localization accuracy in the current study vs. 51.3% accuracy found previously. These findings support our hypothesis that limitations arising from use of the VVD are not due to problems performing effective searches but to deficits in building up accurate spatial representations in memory.

The SA condition, serving as a perceptual interface providing direct access to spatial relations, was tested here because we believed that it would aid cognitive map development. The confirmatory results were dramatic. Participants who trained in computer-based layouts using spatialized audio demonstrated 50% better wayfinding performance in the corresponding real building than when they trained with spatial language. The 76.8% target localization accuracy in the SA condition was also on par with target localization performance of 80% observed in a previous study after verbal learning in real buildings [9]. This similarity is important as it shows that the same level of spatial knowledge acquisition is possible between learning in real and virtual environments. Our results are consistent with the advantage of spatialized auditory displays vs. spatial language found for route guidance [18-20, 25] and extend the efficacy of spatialized audio displays for supporting cognitive mapping and wayfinding behavior.

3 Experiment 2

Experiment 2 was designed to assess the contribution of physical body movement during virtual verbal learning on cognitive map development. Navigation with our virtual verbal display, as with most desktop virtual environment technologies, lacks the idiothetic information which is available during physical navigation, i.e. body-based movement cues such as proprioceptive, vestibular, and biomechanical feedback. VEs incorporating these cues have greater interface fidelity as the sensorimotor contingencies are more analogous to real-world movement [26]. Various spatial behaviors requiring accessing an accurate cognitive map show improved performance when idiothetic information is included. For instance, physical rotation during VE learning vs. imagined rotation benefits tasks requiring pointing to previously learned targets [27, 28], estimation of unseen target distances [29] and updating self orientation between multiple target locations [30]. Path integration is also better in VEs providing proprioceptive and visual information specifying rotation compared to visual information in isolation [31]. The inclusion of idiothetic information has also led to improved performance on cognitive mapping tasks similar to the current experiment, where VE learning is tested during transfer to real-world navigation [32, 33].

Where the previous work has addressed the role of body-based cues with visual displays, Experiment 2 investigates whether similar benefits for verbal learning manifest when physical body rotation is included in the VVD. As with experiment 1, participants use the VVD to explore computer-based training environments and then perform wayfinding tests in the corresponding real environment. However, rather than using arrow keys to affect imagined rotations and translations during training, participants physically turn in place whenever they wished to execute a change of heading. Translations are still done via the keypad as the benefit of physical translation on VE learning is generally considered nominal. This is consistent with studies in real environments showing that pointing to target locations is faster and more accurate after actual than imagined rotations, whereas errors and latencies tend not to differ between real and imagined translations [34].

We predict that inclusion of idiothetic information in the VVD will yield marked improvements in spatial knowledge acquisition and cognitive map development. In addition to the previous evidence supporting body-based cues, we believe the conversion of linguistic operators into a spatial form in memory is a cognitively effortful process, facilitated by physical movement. Evidence from several studies support this movement hypothesis. Avraamides and colleagues (Experiment 3, 2004) showed that mental updating of allocentric target locations learned via spatial language was impaired until the observer was allowed to physically move before making their judgments, presumably inducing the spatial representation. Updating object locations learned from a text description is also improved when the reader is allowed to physically rotate to the perspective described by the text [35], with egocentric direction judgments made faster and more accurately after physical, rather than imagined rotation [36].

To test our prediction, this experiment adds real rotation to the spatialized audio and spatial language conditions of Experiment 1. If the inclusion of rotational information is critical for supporting environmental learning from verbal descriptions, wayfinding performance during real-world transfer should be better after training with both physical rotation conditions of the current experiment than was observed in the analogous conditions with imagined rotation of Experiment 1. Furthermore, assuming some level of complementarity between rotation and spatialization, the rotation+spatialized audio (R+SA) condition is predicted to show superior performance to the rotation+spatial language (R+SL) condition.

3.1 Method

Sixteen blindfolded-sighted participants, nine female and seven male, ages 18-24 (mean = 19.6) ran in the two hour study.

Experiment 2 employs the same spatial language and spatialized audio conditions as Experiment 1 and adopts the same within Ss design using two counterbalanced conditions, each including a practice, training, and transfer phase. The only difference from Experiment 1 is that during the training phase, participants used real body rotation in the VVD instead of imagined rotation via the arrow keys. Since all intersections were right angle, left and right rotations always required turning 90° in place. An automatically-updated heading description was generated when their facing direction was oriented with the orthogonal corridor. They could then either continue translating by means of the keypad or request an updated description of intersection geometry. Heading changes were tracked using a three degree-of-freedom (DOF) inertial orientation tracker called 3D-Bird (ascension corporation: http://www.ascension-tech.com/products/3dbird.php).

3.2 Results and Discussion

To address the effect of rotation on environmental learning and wayfinding performance, One-way repeated measures ANOVAs were conducted for the transfer tests of target localization accuracy and route efficiency. Results indicated a significant difference for target localization only, with the 78.1% (SE = 5.03) accuracy of the rotation+spatialized audio condition found to be reliably better than the 57.8% (SE = 9.33) accuracy of the rotation+spatial language condition, $F(1,15) = 4.601$, $p<.05$, $\eta^2 = 0.24$. Performance on route efficiency did not differ between conditions, R+SA = 98.5% and R+SL = 100%. As discussed in Experiment 1, the results of this measure are far less interesting than the target localization performance on what they say about cognitive map development. Since we were interested in evaluating whether the physical rotation conditions were better than the same conditions using imagined rotation of Experiment 1, we performed a two-way between subjects ANOVA comparing target accuracy performance between experiments by spatialized and non-spatialized conditions. This between Ss comparison is appropriate as the subject groups in both experiments were similar in age, sex, educational background and spatial ability (as assessed by the Santa Barbara sense of Direction Scale, SBSOD). As can be seen in Figure 2, results showed a main effect of target accuracy by spatialization, $F(1,28) = 11.753$, $p<.05$, $\eta^2 = 0.3$, but the more meaningful experiment by spatialization interaction was not significant, $F(1,28) = .126$, $p = >.1$, $\eta^2 = 0.004$. Likewise, a one-way ANOVA comparing target localization accuracy collapsed across condition between experiments, thereby directly addressing the influence of rotation factoring out spatialization, was not significant, $p > .1$.

The results of Experiment 2 paint a clear, yet surprising picture. The addition of physical rotation in the VVD was predicted to significantly benefit spatial knowledge acquisition and cognitive map development, as "real" movement was thought to be particularly important in converting the symbolic verbal messages into a spatial form in memory. While there was a difference in transfer performance between conditions in this experiment, comparison of the data to analogous conditions of experiment 1 confirm that this difference was driven by the presence of spatialized audio descriptions, not physical rotation. Subjects in the SL condition of experiment 1 found routes between targets in the real building with 51.8% accuracy. The 57.8% accuracy of the R+SL condition of Experiment 2, which is identical to that condition except for the addition of real vs. imagined body turning during training, represents a small, nonsignificant performance improvement, P>.1. Likewise, the absence of reliable differences between the SA condition of Experiment 1 and the same condition with rotation in Experiment 2 (76.8% vs. 78.1% correct target accuracy respectively), demonstrates that the addition of physical rotation did not benefit environmental learning.

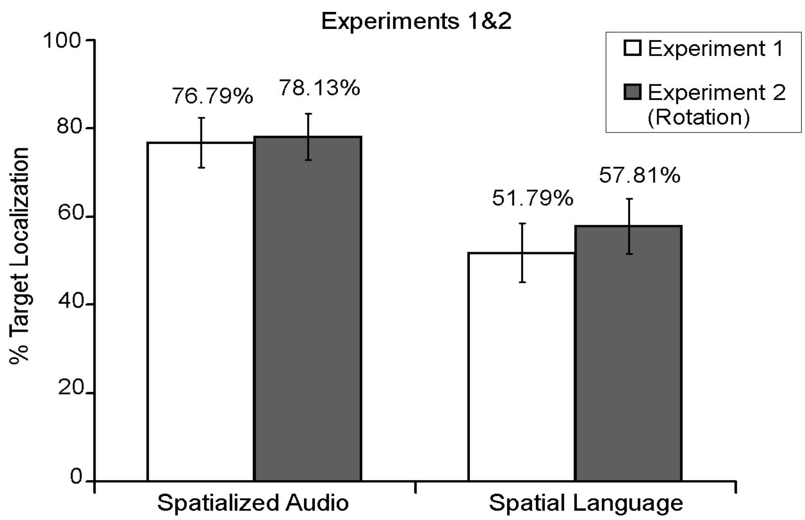

Fig. 2. Comparison of mean target localization accuracy (± SEM) between Experiments 1 and 2. Note: Both experiments compared SL and SA conditions but Experiment 1 used imagined rotation and Experiment 2 (gray bars) used body rotation.

The finding that idiothetic information did not benefit transfer performance was unexpected given previous literature showing that physical body movement during and after verbal learning significantly improves latency and error performance at test [35-37]. Differences in task demands likely contribute to these findings. In the previous studies, subjects learned a series of target locations from text or speech descriptions and then were tested using a pointing-based spatial updating task. The increased weighting of physical movement demonstrated in those studies may be less important with the free exploration paradigm and transfer tests used here, as these tasks do not force updating of Euclidean relations between targets. Thus, the addition of a pointing task between target locations may have shown greater benefit of physical rotation than was evident from our wayfinding task. This needs to be addressed in future experiments as it cannot be resolved from the current data.

4 Experiment 3

Experiment 3 followed the same design of the previous two studies but subjects learned the computer-based training environments from a visual display rather than a verbal display. The main goal of Experiment 3 was to provide a good comparison benchmark with the previous two verbal experiments. Specifically, we wanted to investigate whether learning with verbal and visual displays lead to comparable environmental transfer performance, findings which would provide proof of efficacy of the VVD. Our previous experiments using an almost identical design to the current studies, found that wayfinding performance during environmental transfer was significantly worse after learning from a virtual verbal display than from a visual display [8, Experiment 3, 10]. However, those studies only compared visual learning with a spatial language condition, analogous to that used in experiment 1. By contrast, the significantly improved transfer performance of the spatialized audio conditions are on par with our previous findings with the visual display. Likewise, the SA conditions in the first two experiments provide perceptual information about the direction of hallways which is better matched with what is apprehended from a visual display. Since the visual display and movement behavior in the previous studies differed slightly from the information and movement of the VVD used here, Experiment 3 was run to serve as a more valid comparison.

4.1 Method

Fourteen normally sighted participants, six females and eight males, ages 18-21 (mean = 19.2) ran in the one hour study.

The experimental procedure was identical to the previous studies except that subjects only learned one environment and trained with a visual display instead of the VVD. During training, participants saw the same geometric "views" of the layout on the computer monitor (Gateway VX700, 43.18 cm diagonal) as were previously described with each message from the VVD. The environment was viewed from the center of the monitor and movement was performed via the keypad's arrow keys, as described earlier. Figure 3 shows an example of what would be seen from a 3-way intersection. With each translation, the participant heard the footstep sound and the

Fig. 3. Sample 3-way intersection as seen on a visual display. Information seen from each view is matched to what would be heard in the corresponding message from the VVD.

next corridor segment(s) was displayed with an animated arrow indicating forward movement. With rotations, they saw the viewable segments rotate in place and an updated indication of heading was displayed. In addition, they heard the target names, starting location noise, foot step sound, and percent of training time elapsed via monaural output through the same speakers. This design ensured the visual display was equivalent in information content to what was available in the auditory conditions of experiments 1 and 2.

4.2 Results and Discussion

Performance on the transfer tests after visual learning was quite good, resulting in target localization accuracy of 78.6% (SE = 8.6) and route efficiency of 95.6% (SE = 2.4). Given our interest in comparing learning performance between the visual display and the VVD, independent samples t-tests were used to evaluate how wayfinding performance after visual learning compared to the same tests after verbal learning in Experiments 1 and 2. As the presence or absence of spatialized information was the only factor that reliably affected verbal learning performance, the visual learning data was only compared to the combined performance from the spatial language and spatialized audio conditions of the previous experiments, collapsing across imagined and real rotation. Note that these between-subjects comparisons were based on participants drawn from a similar background and who fell within the same range of spatial abilities as measured by the SBSOD scale. As can be seen in Figure 4, the 78.6% (SE = 8.6) target localization performance observed after visual learning was significantly better than the 54.6% (SE = 5.4) performance of the spatial language conditions, $t(28) = 2.345$, p=.027. By contrast, target localization accuracy in the spatialized audio conditions, 77.5% (SE = 4.1), was almost identical to performance in the visual condition, $t(26) = .116$, p=.908. In agreement with the previous studies, route efficiency was highly insignificant between all conditions, ps > .1.

Experiment 3 was run to benchmark performance with the VVD against visual learning. Replicating earlier work, transfer performance after learning from a visual display was significantly better than learning with spatial language with a VVD [8, Experiment 3, 10]. However, target localization accuracy between the spatialized audio conditions and the visual condition were nearly identical. This finding suggests that learning with a spatialized audio display and an information-matched visual display build up into a spatial representation in memory which can be acted on in a functionally equivalent manner.

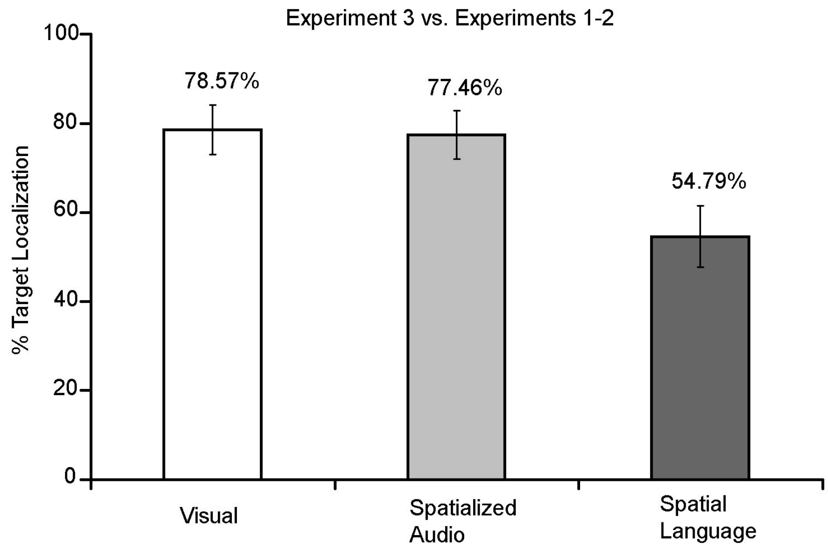

Fig. 4. Comparison of mean target localization accuracy (± SEM) across all experiments. "Spatial language" represents combined data from the two language conditions of Experiments 1 and 2, collapsing across imagined and real rotation. "Spatialized audio" represents the same combined data from the two spatialized conditions of Experiments 1 and 2.

5 General Discussion

The primary motivation of these experiments was to investigate verbal learning and cognitive map development using a new type of non-visual interface, called a virtual verbal display. Previous research has demonstrated that VVDs support efficient search behavior of unfamiliar computer-based environments but lead to inferior cognitive map development compared to verbal learning in real environments or learning in visually rendered VEs. The aim of this research was to understand what could account for these differences. Deficits in spatial knowledge acquisition with the VVD were postulated as stemming from inadequacies of the interface. To address this prediction, two factors influencing interface fidelity, spatialized audio and physical rotation, were compared on a wayfinding task requiring accessing of an accurate cognitive map.

Results showing almost identical performance on the training measures for all conditions across experiments (see Table 1) but widely varying wayfinding accuracy during transfer tests in the real building are informative. Indeed, these findings support the hypothesis that deficits in cognitive map development are related to factors of interface fidelity, rather than use of ineffective search strategies with the VVD. The most important findings from these studies are the results showing that information about layout geometry conveyed as a spatialized verbal description versus from spatial language lead to a dramatic improvement on cognitive map development. These findings are congruent with previous studies showing an advantage of 3-D spatial displays vs. spatial language during route guidance [18-20, 25].

The current results extend the efficacy of spatialized audio for providing perceptual access to specific landmarks in the surrounding environment for use in route navigation to specifying environmental structure during free exploration to support cognitive mapping. Of note to the motivations of the current work, wayfinding performance during transfer after learning in the SA conditions in the VVD was on par with performance after learning with an information-matched visual display, experiment 3, and with verbal learning in real buildings [9]. The similarity of these results suggest that virtual verbal displays incorporating spatialized information can support equivalent spatial knowledge acquisition and cognitive map development. Although comparisons between verbal and visual learning were made between subjects in the current paper, these results are consistent with previous findings demonstrating functionally equivalent spatial representations built up after learning target arrays between the same conditions [38]. Interestingly, the benefit of SA seems to be magnified for open search exploration of large-scale environments vs. directed guidance along routes, as the 50% improvement for spatialized information observed in the current study is much greater than the marginal advantage generally found in the previous real-world route guidance studies. This finding is likely due to the increased cognitive effort known for learning and updating in VEs [23, 24] being offset by the decreased working memory demands of processing spatialized audio vs. spatial language [25].

The effects of including physical rotation vs. imagined rotation in the VVD were investigated in Experiment 2. We expected this factor to have the greatest influence on virtual verbal learning given the importance attributed to idiothetic cues from the inclusion of physical rotation in visually rendered VEs [27, 29, 31, 33], and the importance of physical movement on updating verbally learned target locations [35, 36]. Surprisingly, the inclusion of physical rotation during training with the VVD did not lead to a significant advantage on subsequent wayfinding performance. Indeed, comparison of transfer performance between experiments 1 and 2 shows that conditions employing spatialized descriptions led to the best verbal learning performance and did not reliably differ whether they employed real or imagined rotation. As discussed in Experiment 2, this finding may relate to our experimental design and more research is needed to make any definitive conclusions.

For researchers interested in verbal spatial learning, especially in the context of navigation, dynamically-updated virtual verbal displays represent an excellent research tool. They also have important application to blind individuals for remote environmental learning before traveling to a new place or as part of a multi-modal virtual interface for training sighted people in low-light environments. Until now, their efficacy as a research tool or navigation aid was questionable, as VVD training seemed to lead to deficient cognitive map development. However, the results of this paper clearly demonstrate that the VVD can be used to support these tasks and can be as effective as verbal learning in real buildings or from a visual display when spatialized verbal descriptions are used. These findings have clear implications for the importance of incorporating spatialized audio in dynamically-updated verbal interfaces.

References

1. Taylor, H.A., Tversky, B.: Spatial mental models derived from survey and route descriptions. Journal of Memory and Language 31, 261–292 (1992)
2. Denis, M., et al.: Spatial Discourse and Navigation: An analysis of route directions in the city of Venice. Applied Cognitive Psychology 13, 145–174 (1999)
3. Lovelace, K., Hegarty, M., Montello, D.: Elements of good route directions in familiar and unfamiliar environments. In: Freksa, C., Mark, D.M. (eds.) Spatial information theory: Cognitive and computational foundations of geographic information science, pp. 65–82. Springer, Berlin (1999)
4. Tversky, B.: Spatial perspective in descriptions. In: Bloom, P., et al. (eds.) Language and Space, pp. 463–492. MIT Press, Cambridge (1996)
5. Loomis, J.M., et al.: Assisting wayfinding in visually impaired travelers. In: Allen, G.L. (ed.) Applied spatial cognition: From research to cognitive technology, pp. 179–202. Erlbaum, Mahwah (2007)
6. Giudice, N.A.: Navigating novel environments: A comparison of verbal and visual learning, Unpublished dissertation, University of Minnesota, Twin Cities (2004)
7. Giudice, N.A.: Wayfinding without vision: Learning real and virtual environments using dynamically-updated verbal descriptions. In: Conference and Workshop on Assistive Technologies for Vision and Hearing Impairment, Kufstein, Austria (2006)
8. Giudice, N.A., et al.: Spatial learning and navigation using a virtual verbal display. ACM Transactions on Applied Perception (in revision)
9. Giudice, N.A., Bakdash, J.Z., Legge, G.E.: Wayfinding with words: Spatial learning and navigation using dynamically-updated verbal descriptions. Psychological Research 71(3), 347–358 (2007)
10. Giudice, N.A., Legge, G.E.: Comparing verbal and visual information displays for learning building layouts. Journal of Vision 4(8), 889 (2004)
11. Ruddle, R.A., Payne, S.J., Jones, D.M.: Navigating buildings in "desk-top" virtual environments: Experimental investigations using extended navigational experience. Journal of Experimental Psychology: Applied 3(2), 143–159 (1997)
12. Bliss, J.P., Tidwell, P., Guest, M.: The effectiveness of virtual reality for administering spatial navigation training to firefighters. Presence 6(1), 73–86 (1997)
13. Waller, D., Hunt, E., Knapp, D.: The transfer of spatial knowledge in virtual environment training. Presence 7, 129–143 (1998)
14. Piaget, J., Inhelder, B., Szeminska, A.: The child's conception of geometry. Basic Books, New York (1960)

15. Siegel, A., White, S.: The development of spatial representation of large scale environments. In: Reese, H. (ed.) Advances in Child Development and Behavior. Academic Press, New York (1975)
16. O'Keefe, J., Nadel, L.: The hippocampus as a cognitive map. Oxford University Press, London (1978)
17. Ehrlich, K., Johnson-Laird, P.N.: Spatial descriptions and referential continuity. Journal of Verbal Learning & Verbal Behavior 21, 296–306 (1982)
18. Loomis, J.M., et al.: Personal guidance system for people with visual impairment: A comparison of Spatial Displays for route guidance. Journal of Visual Impairment & Blindness 99, 219–232 (2005)
19. Loomis, J.M., Golledge, R.G., Klatzky, R.L.: Navigation system for the blind: Auditory display modes and guidance. Presence 7, 193–203 (1998)
20. Marston, J.R., et al.: Evaluation of spatial displays for navigation without sight. ACM Transactions on Applied Perception 3(2), 110–124 (2006)
21. Simpson, B.D., et al.: Spatial audio as a navigation aid and attitude indicator. In: Human Factors and Ergonomics Society 49th Annual Meeting, Orlando, Florida (2005)
22. Oving, A.B., Veltmann, J.A., Bronkhorst, A.W.: Effectiveness of 3-D audio for warnings in the cockpit. Int. Journal of Aviation Psychology 14, 257–276 (2004)
23. Richardson, A.E., Montello, D.R., Hegarty, M.: Spatial knowledge acquisition from maps and from navigation in real and virtual environments. Memory & Cognition 27(4), 741–750 (1999)
24. Wilson, P.N., Foreman, N., Tlauka, M.: Transfer of spatial information from a virtual to a real environment. Human Factors 39(4), 526–531 (1997)
25. Klatzky, R.L., et al.: Cognitive load of navigating without vision when guided by virtual sound versus spatial language. Journal of Experimental Psychology: Applied 12(4), 223–232 (2006)
26. Lathrop, W.B., Kaiser, M.K.: Acquiring spatial knowledge while traveling simple and complex paths with immersive and nonimmersive interfaces. Presence 14(3), 249–263 (2005)
27. Lathrop, W.B., Kaiser, M.K.: Perceived orientation in physical and virtual environments: Changes in perceived orientation as a function of idiothetic information available. Presence (Camb) 11(1), 19–32 (2002)
28. Bakker, N.H., Werkhoven, P.J., Passenier, P.O.: The effects of proprioceptive and visual feedback on geographical orientation in virtual environments. Presence 8(1), 36–53 (1999)
29. Ruddle, R.A., Payne, S.J., Jones, D.M.: Navigating large-scale virtual environments: What differences occur between helmet-mounted and desk-top displays. Presence 8(2), 157–168 (1999)
30. Wraga, M., Creem-Regehr, S.H., Proffitt, D.R.: Spatial updating of virtual displays during self- and display rotation. Mem. and Cognit. 32(3), 399–415 (2004)
31. Klatzky, R.L., et al.: Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. Psychological Science 9(4), 293–299 (1998)
32. Grant, S.C., Magee, L.E.: Contributions of proprioception to navigation in virtual environments. Human Factors 40(3), 489–497 (1998)
33. Farrell, M.J., et al.: Transfer of route learning from virtual to real environments. Journal of Experimental Psychology: Applied 9(4), 219–227 (2003)
34. Presson, C.C., Montello, D.R.: Updating after rotational and translational body movements: Coordinate structure of perspective space. Perception 23(12), 1447–1455 (1994)

35. de Vega, M., Rodrigo, M.J.: Updating spatial layouts mediated by pointing and labelling under physical and imaginary rotation. European Journal of Cognitive Psychology 13, 369–393 (2001)
36. Avraamides, M.N.: Spatial updating of environments described in texts. Cognitive Psychology 47(4), 402–431 (2003)
37. Chance, S.S., et al.: Locomotion mode affects the updating of objects encountered during travel: The Contribution of vestibular and proprioceptive inputs to path integration. Presence 7(2), 168–178 (1998)
38. Klatzky, R.L., et al.: Encoding, learning, and spatial updating of multiple object locations specified by 3-D sound, spatial language, and vision. Experimental Brain Research 149(1), 48–61 (2003)