Max−Planck−Institut
für biologische Kybernetik

Spemannstraße 38 • 72076 Tübingen • Germany

Technical Report No. 82

# Visual Homing is possible without Landmarks

## A Path Integration Study in Virtual Reality

Bernhard E. Riecke[1],
Henricus A.H.C. van Veen[2] &
Heinrich H. Bülthoff[3]

September 30, 2000

While in Tübingen, Hendrik-Jan van Veen was funded by the Max-Planck Society and by the Deutsche Forschungsgemeinschaft (MA 1038/6-1, 1038/7-1).

We thank the authors of (Loomis et al., 1993) and (Péruch et al., 1997) for letting us use their raw data. We especially thank Jack Loomis, Patrick Péruch, and Markus von der Heyde for interesting and helpful discussions.

[1] [3] AG Bülthoff, Max-Planck Institute for Biological Cybernetics, Spemannstrasse 38, 72076 Tübingen, Germany, E–mail: bernhard.riecke@tuebingen.mpg.de
[2] TNO Human Factors, P.O. Box 23, 3769 ZG Soesterberg, The NetherlandsAG

---

This report is available via anonymous ftp at `ftp://ftp.kyb.tuebingen.mpg.de/pub/mpi-memos/pdf/TR-082.pdf` in PDF–format or at `ftp://ftp.kyb.tuebingen.mpg.de/pub/mpi-memos/TR-082.ps.Z` in compressed PostScript–format. The complete series of Technical Reports is documented at: `http://www.kyb.tuebingen.mpg.de/bu/techr/`

# Visual Homing is possible without Landmarks
## A Path Integration Study in Virtual Reality

*Bernhard E. Riecke, Henricus A.H.C. van Veen & Heinrich H. Bülthoff*

**Abstract.** The literature often suggests that proprioceptive and especially vestibular cues are required for navigation and spatial orientation tasks involving rotations of the observer. To test this notion, we conducted a set of experiments in virtual reality where only visual cues were provided.

Subjects had to execute turns, reproduce distances or perform triangle completion tasks: After following two prescribed segments of a triangle, subjects had to return directly to the unmarked starting point. Subjects were seated in the center of a half-cylindrical 180° projection screen and controlled the visually simulated ego-motion with mouse buttons. Most experiments were performed in a simulated 3D field of blobs providing a convincing feeling of self-motion (vection) but no landmarks, thus restricting navigation strategies to path integration based on optic flow. Other experimental conditions included salient landmarks or landmarks that were only temporarily available.

Optic flow information alone proved to be sufficient for untrained subjects to perform turns and reproduce distances with negligible systematic errors, irrespective of movement velocity. Path integration by optic flow was sufficient for homing by triangle completion, but homing distances were biased towards mean responses. Additional landmarks that were only temporarily available did not improve homing performance. However, navigation by stable, reliable landmarks led to almost perfect homing performance. Mental spatial ability test scores correlated positively with homing performance especially for the more complex triangle completion tasks, suggesting that mental spatial abilities might be a determining factor for navigation performance. Compared to similar experiments using virtual environments (Péruch et al., 1997; Bud, 2000) or blind locomotion (Loomis et al., 1993), we did not find the typically observed distance undershoot and strong regression towards mean turn responses.

Using a virtual reality setup with a half-cylindrical 180° projection screen allowed us to demonstrate that visual path integration without any vestibular or kinesthetic cues is sufficient for elementary navigation tasks like rotations, translations, and homing via triangle completion.

## 1 Introduction

Spatial orientation and navigation are, in general, based on a number of different sources of information and perceived through different sensory modalities. Successful spatial orientation and navigation involve a number of different processes, including sensing the environment, building up a mental spatial representation, and using it (e.g. to plan the next steps). During navigation, one needs to update one's mental representation of the current position and orientation in the environment ("spatial updating"). Spatial updating methods can be classified by the type of information used: Position ("position"- or "recognition-based navigation") or velocity and acceleration ("path integration" or "dead reckoning") (Loomis et al., 1993).

**Position- or recognition-based navigation** (also called **piloting**) uses exteroceptive information to determine one's current position and orientation. Such information sources include visible, audible or otherwise perceivable reference points, so-called "landmarks" (i.e., distinct, stationary, and salient objects or cues). Many studies have demonstrated the usage and usability of different types of landmarks for navigation purposes, see Hunt & Waller (1999) for an extensive review. Only piloting allows for correction of errors in perceived position and orientation through reference points and is thus more suited for large-scale navigation.

**Path integration**, on the other hand, is based on integrating the perceived velocity or acceleration vector over time to determine the current position and orientation with respect to some starting point. Velocity signals can stem from exteroceptive signals, like an optic or acoustic flow field or an air current, but also from proprioceptive signals like the muscle effort to maintain a constant speed while riding a bicycle. Proprioceptive and especially vestibular cues play an important role for the perception of linear and rotatory self-acceleration. Path integration is based on the perception of time, velocity, and acceleration, and is therefore susceptible to accumulation errors due to the integration process. It is well suited for small-scale navigation and connecting neighboring landmarks, but uncertainty and error increase exponentially with traveled distance.

As recognition-based strategies are known to provide sufficient information for accurate homing performance in simple navigation tasks (see section 3), we focus here on navigation tasks based solely on path integration, without the aid of external reference points (landmarks).

### 1.1 Triangle completion paradigm and homing vector hypothesis

In most of the experiments described in this paper, we used **Triangle completion**, a paradigm commonly used for navigation tasks without landmarks: Subjects are led along two sides of a given triangle and have to find the shortest way back to the starting position by themselves. Triangle completion is a special form of homing behavior that provides several experimental advantages over studying more complex homing behavior:

- The task is simple, well defined and ecologically inspired (members of most ambulatory species need to find their way back to a specific place at some point in time, e.g., for food, shelter or social purposes).
- Results are usually clear and relatively easy to interpret.
- The task is commonly used for human as well as animal studies. This allows for comparisons among species, (e.g., to test for correlations between brain complexity and navigation strategies - even ants can be trained to perform triangle completion tasks, see Müller & Wehner (1988)).
- More complex navigation behavior can always be decomposed into elementary translations and rotations, which are often studied in conjunction with triangle completion experiments. Triangle completion uses simplest non-trivial combination of translations and rotations.

Probably the easiest strategy for triangle completion is to continuously update a so-called “**homing vector**”, which points from the current position to the origin of travel. Animals are often believed to use this strategy in homing tasks, as it does not require long term memory or a complex spatial representation of the whole environment (Müller & Wehner, 1988). However, it has been demonstrated that some species are able to use more than just one spatial reference frame (Benhamou, 1997). Humans, for example, are known to not only update a homing vector, but also to form more complex representations. Among them are view-graphs (Mallot, Gillner, Steck, & Franz, 1999), occupancy grids or survey representations in the form of a “cognitive map” in the sense of Tolman (1948), allowing for short-cutting to locations other than the starting point (Loomis et al., 1993). For an overview of different path integration models, see Maurer & Séguinot (1995), Collett & Collett (2000), Etienne, Maurer, & Séguinot (1996), Mittelstaedt (2000).

### 1.2 Distinguishing between piloting and path integration

For navigation experiments, one might wish to distinguish between the contributions of piloting and path integration. This can be done by excluding one of the two spatial updating modalities at a time: Path integration can be rather easily excluded by eliminating all velocity and acceleration information, e.g., through a slide-show type presentation. The elimination of recognition-based spatial updating is more difficult and, perhaps, more critical, as landmarks play a dominant role in normal navigation. The difficulty of navigating in heavy fog or snowfall illustrates this dominance.

**Proprioceptive and vestibular cues** reveal normally no information about external landmarks, and are as such well suited for path integration studies. **Auditory cues** from individual sound sources convey some spatial information about the sound source, which can consequently be used as an acoustic landmark. Providing acoustic flow from an abundance of sound sources could in principle be used for path integration studies, but is experimentally difficult to handle (Begault, 1994). **Visual cues** provide information about the location of the objects seen, which can consequently be used for recognition-based navigation. Apart from blindfolding people, the only way to circumvent this navigation-by-landmarks is through displaying optic flow only, (i.e., removing the landmark-character from the visible objects). Methodologically, this can be achieved through presenting an abundance of similar objects that can only be tracked over a short distance, which can be easily implemented using a Virtual Reality setup. The effect is similar to moving through heavy snowfall or flying through clouds that block the vision for all distant landmarks (cf. subsection 2.1.4).

## 1.3 Virtual Reality

### Definition and applications in spatial cognition

Using Virtual Reality (VR) for experiments on orientation and navigation has several advantages over the classic approach: The **real-time interactivity** of VR makes an closed-loop paradigm possible that is important for studying natural behavior. **Data collection and analysis** can be performed easily and on the fly, allowing for immediate feedback if required. The **experimental design is flexible** and could be changed even during the experiment, depending, for example, on the subject's performance. Most importantly, the **experimental conditions are well-defined** and can easily be **reproduced** (Bülthoff & van Veen, 2000; Loomis, Blascovich, & Beall, 1999).

This is often an advantage over navigation experiments performed in real environments, where it is very difficult to control a number of experimental factors. Among them are weather conditions (e.g., sun position, clouds, visibility of landmarks), existence, location and persistence of landmarks (e.g., parked cars, construction work, people walking around, sound sources) and previous knowledge of the environment. To circumvent these issues, experiments on spatial cognition have often used slide shows, film sequences or models/maps of the environment traveled (Goldin & Thorndyke, 1982). All those experiments had in common that they were either highly unrealistic (models and maps) or not interactive (slide shows and film sequences), thus lacking a (possibly important) component of natural navigation (Flach, 1990). The recent evolution of virtual environments technology provides the opportunity to tackle these issues. The number of studies on human spatial cognition and navigation using VR has rapidly increased over the last years and has given rise to a number of interesting results (see Péruch & Gaunet, 1998; Darken, Allard, & Achille, 1998; Christou & Bülthoff, 1998, for extensive reviews).

### Virtual Reality as a tool to disentangle different sensory modalities and render piloting impossible

In addition to the above mentioned properties of VR, we used virtual environments in this study for two specific purposes: To disentangle the different sensory modalities and to render piloting impossible.

The virtual environment was presented only visually, thus excluding all spatial cues from other sensory modalities, especially kinesthetic[1] and vestibular cues from physical motion. To reduce proprioceptive cues from motion control to a minimum (and consequently restrain motor learning), subjects pressed buttons to control their self-motion, instead of using more sophisticated input devices like data gloves or joysticks. However, in previous experiments we have shown that adding proprioceptive cues through the use of a bicycle as a locomotion device only marginally affected homing performance (Riecke, 1998; van Veen, Riecke, & Bülthoff, 1999).

To ensure that subjects rely on path integration only, piloting was rendered impossible through presenting optic flow information only (in a 3D field of blobs, see subsection 2.1.4) or through making landmarks only temporarily visible (through "scene swap", see subsection 2.1.4).

In the remainder of this section, we will review relevant literature that motivated this study and give an outline of our experiments. Closely related triangle completion studies are mentioned in section 1.4 and will be discussed in more detail in section 6.2, where they will be compared to our results. Studies demonstrating the difficulty of updating purely visual or imagined rotations are reviewed in section 1.5. Section 1.6 is dedicated to related literature showing the importance of the visual field of view and spatial reference frames for navigation and spatial orientation in VR.

## 1.4 Triangle completion studies

The simplest experimental paradigm for path integration studies is blind locomotion with ears muffled. Vision and audition are easily excluded by blindfolding subjects and displaying white noise over noise-attenuating headphones. Sauvé (1989), Marlinsky (1999c), Bud (2000, exp. 3), Klatzky, Loomis, Golledge, Cicinelli, Pellegrino, & Fry (1990), and Loomis et al. (1993) showed in triangle completion studies that proprioceptive and vestibular cues from blind walking allow for homing, but lead to strong systematic errors. In all five studies, subjects showed a regression towards standardized responses, for example similar turning angles for different triangle geometries.

Qualitatively similar results were found for purely visual triangle completion without salient landmarks. Both presentation via head-mounted display (HMD) (Duchon, Bud, Warren, & Tarr, 1999; Bud, 2000) and flat projection screen (Péruch et al., 1997; Wartenberg, May, & Péruch, 1998) led to larger systematic errors than in the blind walking studies. The observed general tendency to underturn might be explained by the tendency to undershoot intended turning angles when only visual feedback is available (Péruch et al., 1997; Bakker, Werkhoven, & Passenier, 1999).

---

[1] feedback from muscles, joints, and tendons and motor efferent commands.

The above studies will be discussed in more detail in section 6.2, where they will be compared to the experiments presented in this paper.

### 1.5 Differences between updating translations and rotations

The above mentioned difficulty in updating rotations from visual cues alone is consistent with observed fundamental differences between the updating of rotations and translations: For example, studies by May, Péruch, & Savoyant (1995) and Chance, Gaunet, Beall, & Loomis (1998) revealed that vestibular and kinesthetic cues are more important for the perception of rotations than for translations. Simulated turns presented only visually resulted in a reduced spatial orientation ability compared to physical rotations with the same visual input. Chance et al. (1998) suggest "*the advisability of having subjects explore virtual environments using real rotations and translations in tasks involving spatial orientation.*" However, simply adding physical movements does not necessarily guarantee better spatial orientation performance, as was demonstrated by Bud (2000, cf. section 6.2.2): Response variablity decreased, but subjects were still insensitive to angles turned.

Rieser (1989) and Presson & Montello (1994) found a similar difference between rotations and translations for *imagined* movements: Updating the location of several landmarks during imagined self-rotations (without translations) proved more difficult and error-prone than during translations (without rotations). Klatzky, Loomis, Beall, Chance, & Golledge (1998) proposed that this difficulty in updating rotations is due to the lack of proprioceptive cues accompanying the self-rotation. Comparing visually presented locomotion[2] with and without physical rotations, Klatzky et al. (1998) conclude that "*optic flow without proprioception, at least for the limited field of view of our virtual-reality system, appears not to be effective for the updating of heading*". A question arising here is whether a larger field of view and a higher screen resolution could enable correct updating of heading in the absence of all related vestibular and kinesthetic cues. We will attempt to answer this question in this paper.

### 1.6 Influence of field of view and external reference frame

The studies on triangle completion by Péruch et al. (1997) and Bud (2000) and the turning study by Bakker et al. (1999) all used a physical visual field of view (FOV[3]) that was well below the natural FOV of the human eye (horizontal FOV of 45°, 60°, and 24° respectively, compared to more than 180° for humans). Might their finding that humans can not use visual information for accurate path integration be due to the unnaturally limited FOV and the missing visibility of ones own body and the physical environment, which might serve as a helpful reference frame?

To address these questions, we conducted navigation experiments similar to those by Péruch et al. (1997), but using a half-cylindrical 180° projection screen to visually present a virtual environment. Furthermore, three different environments were used, providing different types of spatial information: reliable and salient landmarks, temporarily available landmarks, and no landmarks at all, (i.e., optic flow only).

It is known that enlarging the FOV results in a more realistic spatial perception and has a positive influence on motion perception (Hendrix & Barfield, 1996). On the other hand, most displays currently used have a rather limited FOV of below 100° (usually below 60°) horizontally, especially for HMD's. Arthur (2000) provides an extensive review on past work as well as several experiments on the influence of FOV in HMD's on task performance. Using a custom-built HMD, he found a significant performance benefit in walking tasks for enlarging the horizontal FOV from 112° to 176°, which is much wider than the FOV of commercially available HMD's.

Riecke (1998, chap. 5.4) found a highly subject- and strategy-specific influence of the FOV and the external reference frame: Subjects performed a triangle completion task based on path integration by optic flow, similar to experiment TOWN&BLOBS, section 2. Repeatedly reducing the horizontal FOV by half, from 180° down to 11.25°, increased the between subject variability and was accompanied by a shift in the navigation strategy: For turns, subjects started to use their hand, head or eyes as a pointer to follow the overall rotational optic flow pattern, and continued this tracking behavior even beyond the edge of the display (i.e.,

---

[2] Locomotion was visually presented through an HMD with a field of view of 44°×33°.

[3] The physical field of view (FOV, somtimes referred to as absolute FOV) is a property of the physical setup; it is defined by the angle (horizonal and vertical) under which the observer sees the display.

The simulated field of view (sFOV) generated by the computer (also referred to as geometric FOV) in contrast is a property of the simulation. It is defined by the geometry of the viewing frustrum, i.e. by the angle (horizonal and vertical) under which the virtual (simulated) eyepoint sees the virtual environment.

For most immersive simulations the physical and simulated FOV are kept identical. sFOV > FOV corresponds to a wide angle effect, sFOV < FOV corresponds to a telescope-like view.

they pointed the where the optic flow pattern would be if the FOV was larger). Thus, the borderline of the visual FOV was used as an external reference frame to better estimate turning angles. This might explain why performance for some subjects did not deteriorate for the reduced FOV. In a final block with the full FOV of 180°, subjects' performance showed smaller variations and errors compared to the initial 180° experiment, especially for the turning response. This indicates a long-term learning effect without any explicit performance feedback.

### 1.7 Summary, motivation, and preview of experiments

Several of the above mentioned studies have shown that homing based solely on visual cues without reliable landmarks leads to strong systematic errors (Péruch et al., 1997; Bud, 2000). Furthermore, optic flow alone has been demonstrated to be insufficient for effective spatial updating during rotations (Bakker et al., 1999; Péruch et al., 1997; Chance et al., 1998), and vestibular cues seem to play a crucial role (Rieser, 1989; Presson & Montello, 1994; May et al., 1995; Chance et al., 1998; Farrell & Robertson, 1998; Bakker et al., 1999; May & Klatzky, 2000).

The literature also suggests that human orientation and navigation abilities are greatly influenced by spatial reference frames and the visual field of view (Darken et al., 1998; Sadalla & Montello, 1989; Alfano & Michel, 1990; Arthur, 2000). Can those factors help overcome the above mentioned limitations of visually based homing by optic flow?

The goal in this study is to test whether an external reference frame and a broad visual field of view provided through a 180° projection screen can help compensate for the lack of landmarks and proprioceptive/vestibular cues in visually based navigation. In short: Is visual homing without landmarks possible?

In the **first experiment** (“**TOWN&BLOBS**”, section 2), we compared homing by optic flow (in a 3D field of blobs) with homing by landmarks that were only temporarily available (town with “scene swap”). There are two primary questions here: First, is optic flow information alone sufficient for accurate homing? Second, do natural-looking landmarks that are only temporarily visible improve homing accuracy?

The **second experiment** (“**LANDMARKS**”, section 3) was a control experiment, in which we intended to establish a baseline for visual homing: Given an abundance of salient landmarks in a natural-looking virtual environment, how good is visually based homing? The results form a baseline for comparison with the other experiments, which investigated visual navigation performance *without* any salient landmarks.

Figure 1: Virtual environment lab with 180° projection screen displaying the town environment. The subject is seated behind the table in the center of the half-cylindric screen. On the table are mouse and keyboard as input devices.

In a **third experiment** (“**TURN&GO**”, section 4), we investigated how well untrained subjects can perform simple rotations and translations, given optic flow information only. Turns and translations constitute the basis for navigation tasks, as all movements can be decomposed into a combination of those elementary operations. All other experiments were triangle completion experiments.

The **fourth experiment** (“**RANDOM TRIANGLES**”, section 5) was designed to investigate the influence of the simplicity of the triangle geometry: How does the homing performance change when each triangle geometry is novel (randomized) instead of isosceles (as in the first experiment)? To our knowledge, so far nobody investigated triangle completion for completely randomized lengths of the first and second segment and the enclosed angle.

Finally, we conducted two standard **mental spatial abilities tests** to investigate whether mental spatial ability might be a determining factor for this type of navigation performance (subsection 6.1).

## 2 Experiment 1: “TOWN&BLOBS”

### 2.1 Methods

#### 2.1.1 Participants

For all experiments described in this paper, subjects had normal or corrected-to-normal vision. Participation was always voluntarily and payed at standard rates. Four of all the subjects had to be excluded from the analysis as they had extreme difficulties with the experiment: Their behavior showed no correlation with the requirements of the particular trials. Additionally, they took much longer to complete the training phase. Only one of the subjects experienced symptoms of simulation sickness (general discomfort, nausea, dizziness, and vertigo) and preferred not to finish the experiment.

Ten female and ten male subjects participated in the TOWN&BLOBS experiment, 17 of them were students. Ages ranged from 17 to 30 years (mean: 24.2 years, SD: 3.4 years).

### 2.1.2 Visualization

Experiments were performed on a high end graphics computer (SGI ONYX2 3-pipe Infinite Reality) using C++ , Vega, and Performer applications. The experiment took place in a completely darkened room. Subjects were seated in the center of a half-cylindrical projection screen (7m diameter and 3.15m height, see fig. 1), with their eyes at a height of 1.25m. Three neighboring color images of the virtual environment were rendered at an update rate of 36 Hz and projected non-stereoscopically side by side, with a small overlap of 7.5° smoothed by Soft-Edge-Blending. The resulting image had a resolution of about 3500 × 1000 pixel and subtended a physical field of view of 180° horizontally by 50° vertically. Physical and simulated field of view (used for the image rendering) were always identical. A detailed description of the setup can be found in van Veen, Distler, Braun, & Bülthoff (1998).

### 2.1.3 Interaction

Subjects could freely move through the virtual environment, using three mouse buttons as an input device. Pressing the middle button initiated forward translations that lasted as long as the button was being pressed. Releasing the button ended the motion. Similarly, the left or right button produced left or right rotations, respectively. Pressing or releasing a button resulted in a short acceleration or deceleration phase, respectively, with a constant maximum velocity in between. The maximum velocity was $v_0 = 5m/s$ for translations and $\dot{\alpha}_0 = 40°/s$ for rotations. Motion parameters were chosen to prevent subjects from getting simulator sickness. Combined rotations and translations were possible, but, interestingly enough, were hardly used by the subjects.

### 2.1.4 Scenery

The experiments were performed in two different environments, which were generated with MultiGen® and Medit 3D modeling software: A simple 3D field of blobs and a more complex town environment (see fig. 2).

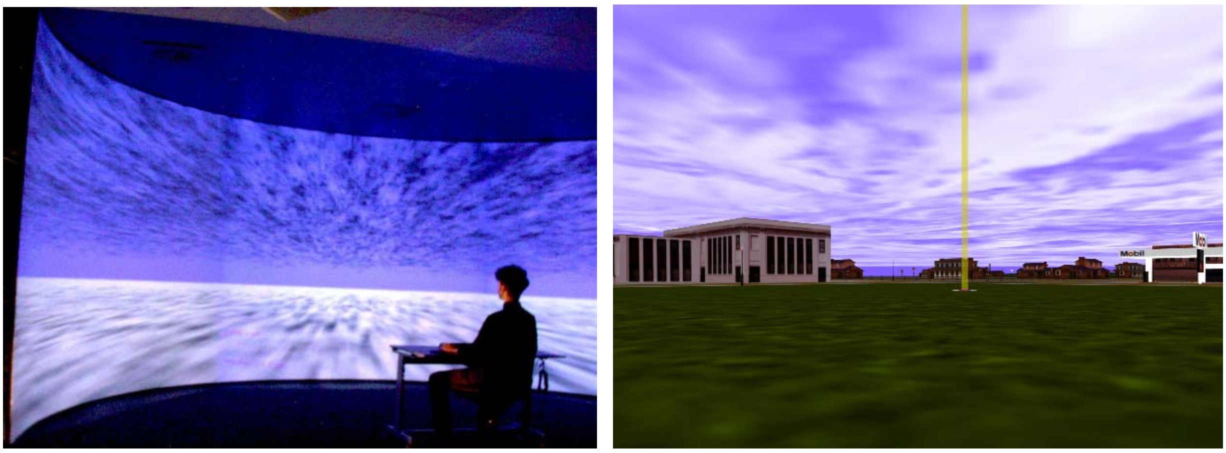
**3D field of blobs** The blob environment consists of a ground plane and four semi-transparent upper planes, all textured with randomized blob patterns (see fig. 2(a)). The blob environment was designed to create a compelling feeling of self-motion (vection) using high optic flow. The individual, similar looking blobs became blurred for viewing distances larger than about 10m, thus providing no salient landmarks that could be used for position-based navigation strategies. Consequently, subjects had to rely on path integration, which was our second design goal.

**Town environment** The town environment consists of a photorealistic models of a small town, with a green open square for the navigation task (see fig. 2(b)), surrounded by an abundance of distinct landmarks (streets, trees, houses etc.).

**“Scene swap”** To exclude object recognition and scene matching as a possible homing strategy, all landmarks (houses, streets etc.) in the scene were repositioned or replaced by others during a brief dark interval just before the onset of the return path (“scene swap condition”). The changed landmarks were arranged to form a different-looking, green square of about twice the original size, with the subject located at its center. In the field of blobs environment, all blobs were randomly repositioned before the return path.

After the “scene swap”, there were no objects left indicating where the starting point was. Hence, subjects had to resort to path integration for homing, e.g. by using a homing vector or some other form of landmark-independent homing strategy.

This scene swap paradigm is an extension of methods used in animal navigation by Müller & Wehner (1988), where desert ants were trained to perform triangle completion tasks: Just before the return path, the ants were picked up, put into a dark box, and carried to a different location, from which they had to complete the homing task. In both experiments, the usage of landmarks and scene matching was excluded through what we call “scene swap”. For humans, the usage of virtual environments was the easiest viable implementation of this scene swap paradigm. Using scene swap in the town environment, subjects could use piloting during the excursion (to build up a mental spatial representation), but not for the homing task. If this temporal availability of landmarks is helpful in navigation, we would expect a better homing performance in the town environment. Else we would not expect any performance differences between the two scenes.



(a) cloud of blobs

(b) town environment

Figure 2: Views of the 3D field of blobs (a) and town environment (b). The yellow cylinder represents the first goal, i.e., the first corner of the triangle to be traveled.

### 2.1.5 Procedure

After reading the experimental instructions, subjects participated in a two-phase training session that lasted about 40 minutes. Goal of the training phase was to familiarize the subjects with navigation in a virtual environment and the task requirements. Only during the training phase was the experimenter present to answer possible questions and to ensure that subjects correctly understood the instructions. The training phases were similar to the actual experiment, but used different triangle geometries and additional feedback about the current position and orientation of the observer. Both training phases consisted of ten homing trials each.

**Training phase** In the first training phase, compass directions (N, S, E, W) were overlaid on the display to provide a global orientation aid, where "north" was defined by the initial heading for each trial. Additionally, a top down (orthographic) view of the scene was presented on an extra monitor placed next to the subject (see fig. 3). The current position and orientation of the subject was displayed (symbolized by a white arrow) as well as the triangle corners currently visible (symbolized by the vertical "light beams").

In the second training phase, the orientation aids were switched off during the navigation phase. Only after completing each trial, the orthographic view was briefly presented (for 2s) to provide feedback.

The training phase was designed to help unexperienced subjects overcome initial disorientation, to ensure a comparable level of proficiency in VE navigation and to avoid the influence of initial learning effects. Preexperiments had shown that some subjects had initially orientation problems in virtual environments without those additional orientation aids. These findings are in accordance with studies by Darken & Sibert (1996) and Ruddle, Payne, & Jones (1997), who showed that disorientation in VE can be overcome by additional orientation aids.

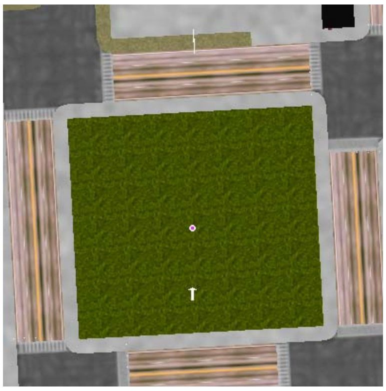


Figure 3: Top down orthographic view (here of the town environment) displayed on an auxiliary screen during training phase 1.

**Test phase** Each subject performed two experimental blocks in separate sessions on different days, one in the 3D field of blobs, the other one in the town environment, in balanced order. The first block began directly after the training session as described above, the second block was preceded by a shortened training session of $2 \times 5$ trials.

For each trial, subjects had the following tasks:

1. **Excursion**: At the beginning of each trial, subjects were positioned and oriented randomly in the virtual environment, facing the first goal, i.e., the first corner of the triangle, symbolized by a semi-transparent yellow "light beam" (see fig. 2). Subjects moved to the yellow light beam which disappeared upon contact. Then the second goal appeared, i.e., the second corner of the triangle, symbolized by a blue light beam. As the second goal could be outside of the current visual field, the proper turning direction was indicated at the bottom of the projection screen. Subjects turned towards the second goal and moved to it. Like the first goal, it disappeared upon contact.
2. **Homing task**: After reaching the second goal, the whole scene was faded out into darkness for 2s. During that brief dark interval, all objects in the environment were exchanged and repositioned. Only in the "LANDMARKS" experiment did the scene remain unchanged. The actual task was now to turn and move directly to the non-marked starting point, as accurately as possible. Pressing a designated button recorded the homing endpoint and initiated the next trial.

### 2.1.6 Experimental design and measurands

**Experimental design** A repeated-measures within-subject design was used (see tab. 1). For each block, each subject completed 60 triangles in random order, corresponding to a factorial combination of 6 repetitions for 5 different angles of the first turn and 2 turning directions varied within a block, and 2 scenes varied across blocks. The order of the within-block conditions (angles and turning direction) was randomized, the order of the between-block conditions (scenes) was counterbalanced across subjects. There was no time limit for completing the tasks and no feedback about performance accuracy during the test phase. Typically, the test phase lasted about one hour.

The triangles with $\alpha$ = 60°, 90°, and 120° correspond to a subset of the triangles used for triangle completion experiments by Péruch et al. (1997) in virtual environments and Loomis et al. (1993) on blind locomotion (see section 6.2 for a comparison).

**Nomenclature** The nomenclature used for the triangle is depicted in figure 4. At the beginning of each trial, subjects are positioned at the starting position $x_0$, the origin of the coordinate system facing north (corresponding to the positive y-axis). For the excursion, they move along $s_1$ to $x_1$, the first goal, then turn by $\overline{\alpha} := (180° - \alpha)$, where $\alpha$ denotes the internal angle, and move along $s_2$ towards $x_2$, the second goal. For the homing task, they turn by $\overline{\beta_m}$, the measured turning angle, where $\overline{\beta_c}$ would have been the correct turning angle. Then they move along $s_{3_m}$ to $x_3$, the measured endpoint of the trajectory where the subject assumes the starting point to be, whereas $x_0$ would have been the correct starting point.

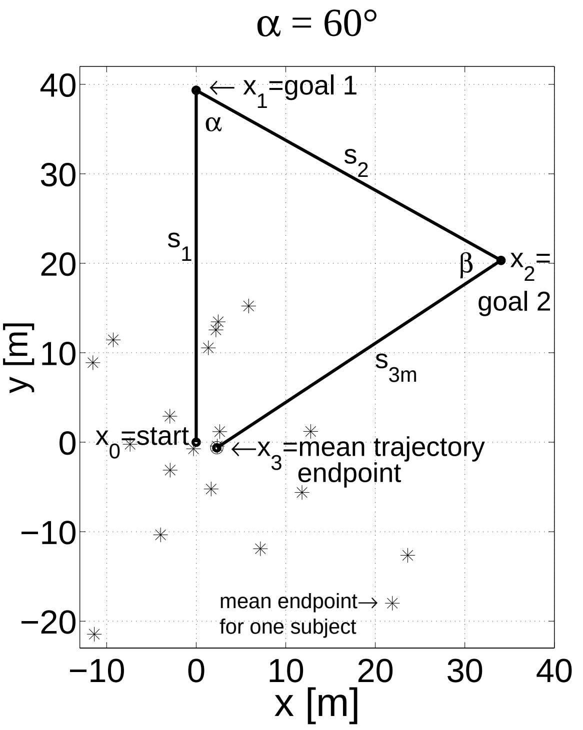


Figure 4: Nomenclature of a triangle to be traveled, as described in the text. The asterisks denote the homing trajectory endpoints for each subject, pooled over turning direction (left/right) and scenery (town/blobs).

**Dependant variables** To quantify the homing error, we used the signed error for turning angle and distances traveled, i.e., the deviation from the correct value for rotations ($\overline{\beta_m} - \overline{\beta_c}$) and for distances traveled ($s_{3_m} - s_{3_c}$) (see fig. 4). We chose those measurands instead of the distance between the homing endpoint and the starting position ($|\vec{x_3} - \vec{x_0}|$), as they better reflect the behavioral response of the subject: They decided and acted first on the turning angle $\overline{\beta_m}$, then on the distance to be traveled $s_{3_m}$. Furthermore, this analysis better disambiguated between different homing strategies and sources of systematic errors.

**Elimination of outliers** Some subjects reported not having payed attention for some trials or having accidentally terminated a trial too early. To reliably eliminate those outliers for all subjects, we developed the following criterion: There were always six repetitions per experimental condition. If one of the six endpoints

| independent variable | levels | values |
|---|---|---|
| $\alpha$=turning angle at 1st corner | 5 | $\alpha \in \{30°, 60°, 90°, 120°, 150°\}$ |
| turning direction | 2 | left or right |
| scene | 2 | 3D field of blobs or town environment |

Table 1: Experimental design for the TOWN&BLOBS experiment.

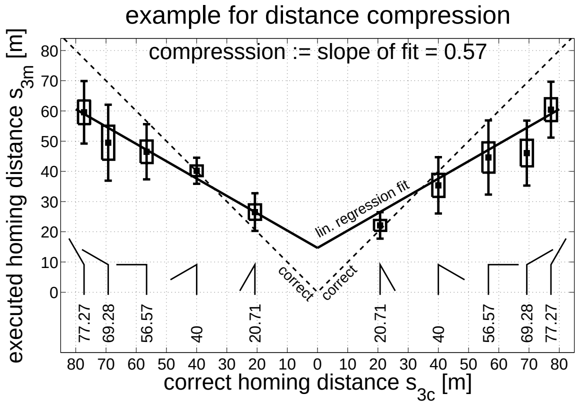


Figure 6: Behavioral response of one representative subject for the town environment. Actual values for distance traveled to complete the triangle ($s_{3_m}$, see fig. 4) is plotted over its corresponding correct values ($s_{3_c}$), left for left turns, right for right turns. The symmetry of the plot illustrates the similarity of the response for left and right turns. The mean values over the six repetitions are plotted for each of the ten triangle geometries (symbolized by the little icons below). The boxes refer to the standard error of the mean, the "whiskers" depict one standard deviation. A linear regression line was fitted through the data and captures nicely the main aspects of the data. The slope ("compression rate") of the linear fit is 0.57, well below the value for perfect response (slope 1), indicated by the dashed black line going straight through the origin.

of the trajectory came to lie outside of a 4.5 $\sigma$ standard ellipse around the five other endpoints, it was eliminated from the further analysis. An average of four of the 60 data point per experimental block were eliminated.

### 2.2 Results and discussion

Homing errors were analyzed using two separate repeated measures 3-way ANOVA's (5 angles $\alpha$ × 2 turning directions × 2 scenes) for the signed error of the two dependant variables turning angle and distances traveled. The ANOVA's revealed a highly significant main effect of the triangle geometry (angle $\alpha$) on distance error (F[4,76] = 32.5, $p < 0.0005$). None of the other factors or any of the interactions came close to significance ($p > 0.25$ in all other cases).

This implies that neither the turning direction nor the scenery used had a significant influence on homing performance. For the further analysis, the data were consequently pooled over both left and right turns and over the two scenes unless indicated differently.

To get a first impression of the homing results, the pooled data are visualized in figure 5. We can see that, averaged over all subjects, the turning error is small, whereas the main effect of triangle geometry on distance error is obvious[4]: The shortest homing distance is typically overshot (left plot), whereas larger homing distances are undershot (right plots).

To quantify that behavior, the data are plotted differently in figure 6. It shows one representative experimental block by one subject for the town environment. The homing distance actually traveled is plotted over its corresponding correct value. As for all subjects, a linear regression line fits nicely to the data and captures its main aspects: The slope ("compression rate") in this example is 0.57, whereas perfect performance would result in a slope of 1, indicating no compression.

For both measurands (turning error and distance error) and all subjects, the linear regression line fitted the data well and captured the main aspects of homing performance. The same was true for the later experiments. Consequently, we were able to summarize homing performance by only four parameters: The compression rate and the signed error[5] for both turning angle and for distance traveled. Those four parameters are plotted in figure 12 and will be used for later comparison among the different experiments.

Averaged over all subjects, the distance compression was 0.60 ± 0.07 (standard error of the mean, SE), indicating a general tendency to overshoot short

---

[4] The **95% confidence ellipse** is a 2D analogue of the confidence interval (mean ± two standard errors of the mean). It covers the population center with a probability of 95% and decreases with $1/\sqrt{N}$ with sample size $N$. The **standard ellipse** is a 2D analogue of the standard interval (mean ± one standard deviation). It is used to describe the variability of the data and covers roughly 40% of the data (see Batschelet, 1981, p. 141).

[5] The standard parametrization of the regression line would be $y(x) = a \cdot x + b$, where $a$ is the slope (compression rate), and $b$ the y-axis offset. Instead of using the y-axis intercept $b$ which has no simple interpretation for the experiments, we decided to use the more meaningful signed error, which is mathematically aquivalent.

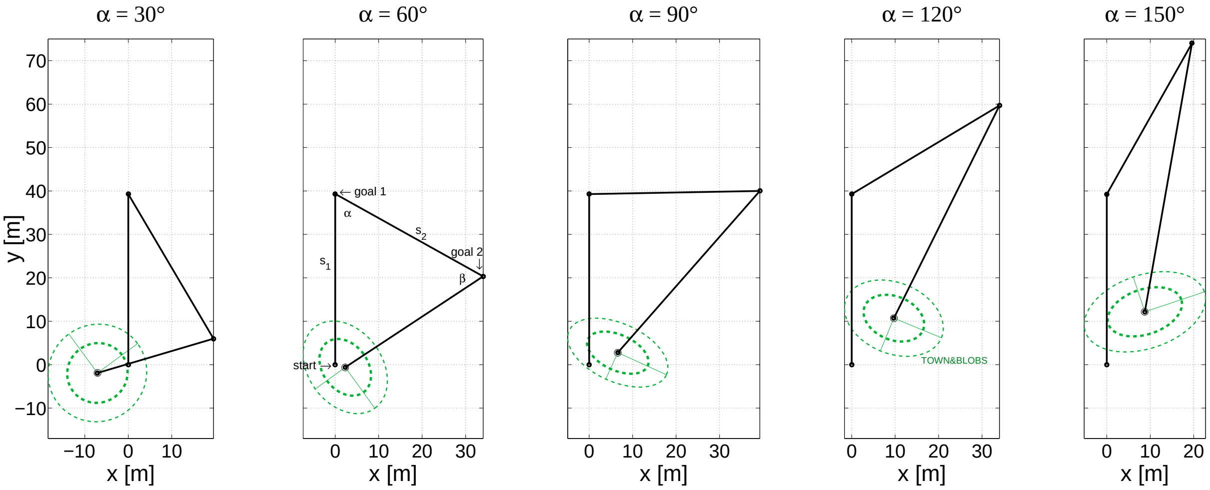


Figure 5: Homing performance in the TOWN&BLOBS experiment. The data is pooled over the independent variables turning direction (left/right) and scenery (town/blobs), as they had no significant influence on homing performance. Plotted are the mean (centroid), the 95% confidence ellipse[4] (inner ellipse with thick dashed line) and the standard ellipse (outer ellipse with thin dashed line) for the homing endpoints.

distances and undershoot long distances (see fig. 12). This tendency proved highly significant (two-tailed t-test, t(19)=5.6, $p < 0.0005$).

The compression rate for turning angles was 0.91 ± 0.08, which is not significantly below the correct value of 1 (t(19)=1.0, p=0.32). This indicates that, averaged over subjects, there was no systematic over- or undershooting of turning angles.

The signed errors for turns and distances are -2.82° ± 2.95° and -0.91m ± 1.61m, respectively, indicating a slight tendency to undershoot both turns and distances. However, neither of the two tendencies came close to statistical significance (t(19)=0.96, p=0.35 and t(19)=0.56, p=0.58).

The lack of performance differences between the blobs and town environment suggests that subjects were not able to take advantage of natural-looking landmarks that are only temporarily available. On the other hand, path integration based solely on optic flow proved to be sufficient for correct turn responses for almost all subjects. However, most subjects had a tendency to overshoot short distances and undershoot long distances, a phenomenon commonly found in the literature. The variability between subjects was rather pronounced, though (see fig. 4), which might be due to different navigation strategies used.

## 3 Experiment 2 : “LANDMARKS”

The second experiment was designed to establish a baseline for visual homing, for comparison with the other experiments, which investigated visual navigation performance *without* any stable, salient landmarks: What is the accuracy of visually based homing, if an abundance of salient landmarks in a natural-looking virtual environment are available to be used as navigation aids?

Our hypothesis was that homing performance should be more accurate and more consistent if subjects were able to take advantage of reliable landmarks. This implies that systematic errors and variability should be smaller than in the first experiment.

### 3.1 Methods

#### 3.1.1 Participants

Five male and two female subjects participated in the LANDMARKS experiment. All of them had earlier completed the TOWN&BLOBS experiment. Ages ranged from 23 to 30 years (mean: 26.5 years, SD: 2.6 years).

#### 3.1.2 Procedure

The training phase was omitted, as all subjects were familiar with the task. The procedure for the test phase was essentially the same as in the second block of Experiment TOWN&BLOBS for the town environment, with one mayor difference: Scene swap was omitted during the brief dark interval and subjects completed the homing task within the same, unchanged town environment. The goal here was to find a baseline of how well subjects perform in a visual homing task if they are allowed to use stable landmarks and scene matching like in natural situations.

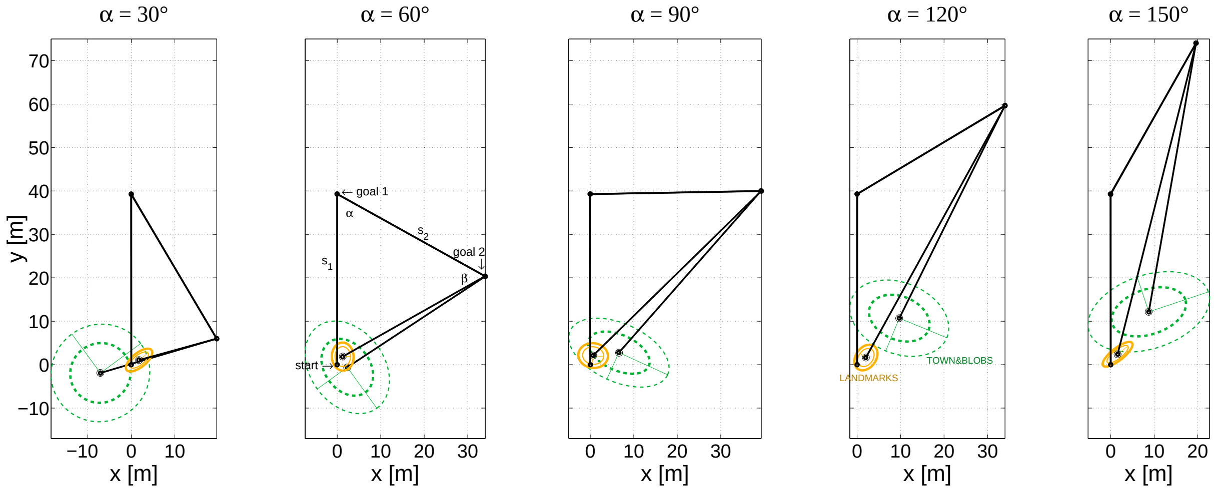


Figure 7: Homing performance in experiment LANDMARKS (smaller ellipses with solid line) as compared to experiment TOWN&BLOBS (larger ellipses with dashed line). Legend as in figure 5. Due to the smaller number of subjects in experiment LANDMARKS, the corresponding standard ellipse is smaller than the 95% confidence ellipse. The ellipses for the LANDMARKS experiment are smaller and include the origin, indicating less variability and more accurate homing performance than in experiment TOWN&BLOBS *with* scene swap. Non-overlapping 95% confidence ellipses indicate significant performance differences between the experiments (Batschelet, 1981).

## 3.2 Results

Homing errors were analyzed using two separate repeated measures 2-way ANOVA (5 angles $\alpha \times 2$ turning directions) for the signed error of the two dependant variables turning angle and distances traveled. None of the factors or any of the interactions came close to significance ($p > 0.24$ in all cases). For the further analysis, the data were consequently pooled over left and right turns. This pooled data are graphically represented in figure 7, together with the results from experiment TOWN&BLOBS for comparison. We can see that omitting the scene swap leads to smaller homing errors and a smaller variability, compared to experiment TOWN&BLOBS.

To quantify that behavior, we again used the compression rate and the signed error for both measurands (see fig. 12): The mean distance error was -1.9m ± 0.6m (SE) which is significantly below the correct value (two-tailed t-test, t(6)=3.1, p = 0.02). This means that subjects generally undershot the correct homing distance by 1.9m. Turning error, as well as the compression rate for turns and distances, did not differ significantly from its correct value (see fig. 12).

To examine the influence of omitting the scene swap, we compared the results with those from experiment TOWN&BLOBS. If subjects were able to take advantage of the reliable landmarks, homing performance should be more accurate and more consistent than in experiment TOWN&BLOBS.

### 3.2.1 Accuracy

As can be seen from figure 12, the only significant difference between the sample means for the two experiments was in terms of distance compression: Omitting the scene swap in the LANDMARKS experiment got rid of the pronounced distance compression observed in experiment TOWN&BLOBS. This implies that the availability of stable landmarks removed the otherwise strong regression towards mean homing distances.

### 3.2.2 Consistency

To quantify the consistency of homing performance, an F-test was used to compare sample variance in the two experiments. For all four variables plotted in figure 12, the difference in variance between the two experiments was highly significant (F[19,6]=59.9, p<0.0001 for turning error, F[19,6]=19.9, p=0.0013 for distance error, F[19,6]=25.4, p=0.00065 for angular compression and F[19,6]=188.8, p<0.0001 for distance compression).

## 3.3 Discussion

The availability of reliable landmarks in experiment LANDMARKS largely improved overall homing performance as expected. Only the mean distance error was slightly, but insignificantly, larger than in experiment TOWN&BLOBS.

The small variability of the homing response both within and between subjects indicates that subjects

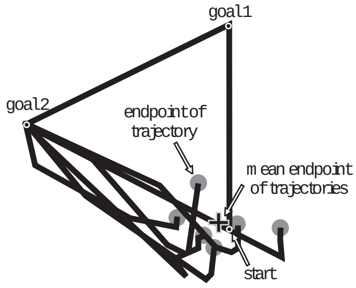


Figure 8: Examples of trajectories for subject *sime* indicating snapshot matching: For the homing task, the subject drove south of the assumed starting point, then turned north and approached it “from behind”, until the current view matched the original view from the starting spot. The non-straight trajectories further suggest that piloting is the dominant navigation mechanism, whereas path integration played only a minor role.

were much more confident in using stable landmarks (Exp. LANDMARKS) than in using optic flow or transient landmarks (Exp. TOWN&BLOBS). This is in accordance with the subjects’ ratings of difficulty for the two experiments. When asked about the strategies they used for homing in experiment LANDMARKS, all subjects reported using configurations of landmarks (scene matching). Some subjects even used snapshot matching as a homing strategy: They approached the assumed starting point from “behind” and moved north until the current view matched the initial view from the starting point (see fig. 8 for an example).

We conclude that piloting and especially scene matching led to almost perfect homing performance, played the dominant role in navigation, and was used whenever possible. However, homing performance was not quite perfect, which might be due to the lack of salient objects close enough to be able to identify the starting position uniquely. We assume that homing accuracy could have improved further, had we provided more salient, nearby landmarks like a location-specific ground texture, and added visibility of the virtual floor directly beneath the subjects via a floor display.

## 4 Experiment 3: “TURN&GO“

### 4.0.1 Purpose

Rotations and translations constitute the basis for all navigation behavior, as all movements can be decomposed into a combination of those elementary operations. The third experiment (“TURN&GO”) was designed to investigate how well untrained subjects are able to perform simple turns and translations, given optic flow information only.

### 4.0.2 Hypotheses

This experiment was designed to disambiguate different sources of error in the other experiments on triangle completion. Fujita, Klatzky, Loomis, & Golledge (1993) developed a “encoding-error model of pathway completion without vision”, to analyze potential origins of the systematic homing errors found in the blind triangle studies by Loomis et al. (1993). This model is based on the axioms that the subjects’ mental representation satisfies the Euclidean axioms and that there is no systematic error in either computing or executing the homeward trajectory. Hence, all systematic errors are attributed to errors in mentally encoding the distances walked and angles turned.

The present experiment examines whether the axioms of the encoding error model are satisfied: If subjects were able to execute intended turns with relatively small systematic errors and variance, we could argue that “turn execution errors” play only a minor role in the other experiments, too. Consequently we could argue that the systematic turn errors observed in the triangle completion experiments should be ascribed to systematic errors in encoding or mental “computation” of the homeward trajectory.

If subjects were able to reproduce traveled distances with relatively small systematic errors and variance, we could at least argue that encoding and execution errors, if present, cancel each other out. That would suggest that the systematic distance errors observed in triangle completion experiments have to be attributed to the mental representation or geometric reasoning. This contradicts the assumptions of the encoding error model, which would consequently not be applicable.

Furthermore, if both rotation and translation performance were excellent, the observed systematic errors and the rather large variance for turns and distances in the triangle completion experiments would be caused by errors in mental representation and geometric reasoning and possibly also errors in encoding turns, but *not* simply by an execution error.

## 4.1 Methods

### 4.1.1 Participants

A new group of six female and three male naive subjects participated in experiment TURN&GO and later also in experiment RANDOM TRIANGLES. Ages ranged from 20 to 36 years (mean: 26.6 years, SD: 4.4 years), eight of them were students. A tenth subject had to be excluded from the analysis, as she had misunderstood the instructions.

### 4.1.2 Procedure

Experiments were performed in the 3D field of blobs (see section 2.1.4). Interface and VR-setup remained unchanged. Before the actual experiment, a handout with a graphical representation of the turning angles was shown to the subjects. To ensure that they understood the turning instruction properly, subjects were standing and asked to turn physically by angles indicated by the experimenter.

The experiment consisted of 96 trials with three phases each:

1. **Distance encoding phase**
   Subjects were positioned randomly within the 3D field of blobs, facing a yellow "light beam" at a given distance $s_1$. By pressing the middle mouse button, they moved to the light beam where they stopped automatically upon contact. Turning was disabled during phase 1 and 3.
2. **Turn execution phas**e
   Subjects were requested to turn by an angle $\alpha_c$ and in the direction specified by the instructions displayed as usual at the lower part of the screen. Translation was disabled during this phase.
3. **Distance reproduction phase**
   Subjects were asked to reproduce the distance $s_1$ from the first phase by traveling that distance in the current direction.

Before the actual experiment, subjects performed six test runs to get accustomed with the interface and the task requirements. Subjects were never given any feedback about their performance or accuracy. Just as for the other experiments, there was no time limit for fulfilling the task. Typically, an experiment lasted about one hour.

### 4.1.3 Experimental design

The experimental design is summarized in table 2. Each subject completed 96 trials, corresponding to a factorial combination of 8 distances × 6 turning angles × 2 turning directions. The range of distances corresponds to the range of homing distances $s_3$ in the previous triangle completion experiments, the range of turning angles is considerably larger.

To test the influence of movement velocity, translational and rotational velocities were randomized independently for each trial and each segment, within an interval centered around the velocity used in the previous experiments (cf. tab. 2). If subjects were able to properly use path integration by optic flow to derive angles turned and distances traveled, we would expect no correlation between movement velocity and turns executed or distances traveled. A significant correlation on the other hand would suggest the usage of a timing strategy (like counting seconds to estimate distances) or general problems with path integration by optic flow.

## 4.2 Results

### 4.2.1 Elimination of outliers

For a few trials, subjects accidentally pressed the confirm button before completing the trial or turned in the wrong direction. To eliminate those outliers consistently from the data, we used the following criterion: A trial was removed if the subject didn't turn at all or if the signed turning error was outside its 4 $\sigma$ interval for that value of $\alpha_c$ (i.e., if $\alpha_m = 0$ or $|\alpha_m - \alpha_c| > 4 \cdot stdDev(\alpha_m - \alpha_c)$). A total of 15 trials or 1.7% of the trials was removed from further analysis.

### 4.2.2 Signed errors and compression rates

The typical distance reproduction and turn execution performance is displayed in figure 9 for one representative subject. The general results are summarized in figure 12. The overall performance was excellent. Only the angular compression differed significantly from its correct value (two-tailed t-test, t(8)=3.1, p= 0.015). However, this angular compression of 0.98 ± 0.01 was only marginal.

### 4.2.3 Correlation analysis

To investigate the influence of the independent variables individually, we performed pairwise correlation tests between the signed and absolute errors for distances $(s_{3_m} - s_{3_c})$ and turns $(\alpha_m - \alpha_c)$ and the independent variables (cf. tab. 3). The Fisher r-to-Z transformed values of the coefficients of correlation were tested against zero using a two-tailed t-test. The results are summarized in table 3: The translational velocity ratio $(v_{s_2}/v_{s_1} = gain_{s_2}/gain_{s_1})$, the turning direction and the turning velocity $\dot{\alpha}$ revealed no correlation with signed or absolute errors for distances or turns (p>0.15 for all cases).

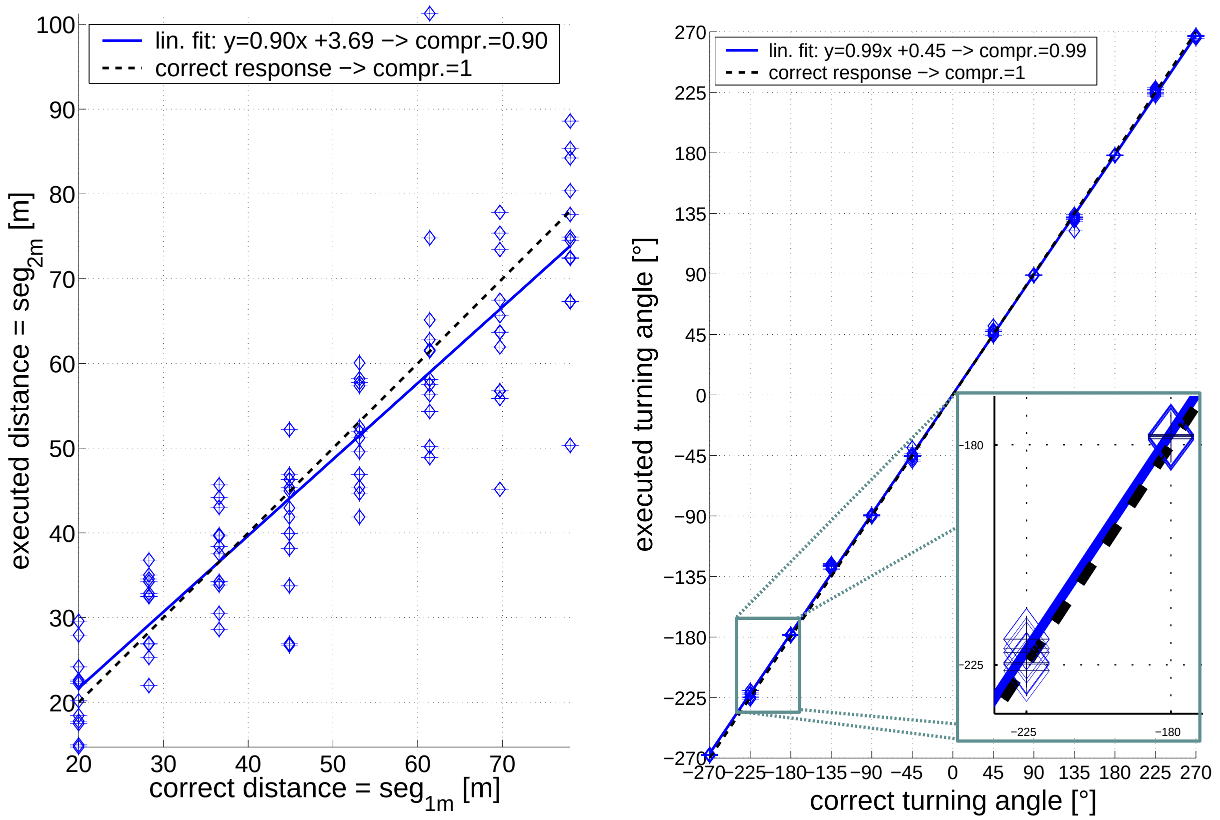
The distance to reproduce $(s_1)$ was negatively correlated with the signed distance error and highly positively correlated with the absolute distance error (cf. tab. 3). A linear regression for the signed distance error[6] reveals an overshoot for small distances and a

---

[6] The regression equation for the signed distance error yields $s_2 - s_1(s_1) = -0.092 \cdot s_1 + 6.3$.

|  | independent variable | levels | values |
|---|---|---|---|
| translations | distance $s_1$ | 8 (equally spaced) | $s_1 \in \{20, \ldots, 78\}$ |
|  | velocity<br>$v_{s_1} = gain_{s_1} \cdot v_0$ | continuously randomized | $gain_{s_1} \in [0.75, 1.5]$ |
|  | velocity<br>$v_{s_2} = gain_{s_2} \cdot v_0$ | continuously randomized | $gain_{s_2} \in [0.75, 1.5]$ |
| rotations | turning angle $\alpha_c$ | 6 (equally spaced in 45° steps) | $\alpha_c \in \{45°, \ldots, 270°\}$ |
|  | turning direction | 2 | left and right |
|  | rotational vel.<br>$\dot{\alpha} = gain_{\alpha} \cdot \dot{\alpha}_0$ | continuously randomized | $gain_{\alpha} \in [0.5, 2]$ |

Table 2: Experimental design and results from the correlation analysis for the TURN&GO experiment. $v_0 = 5m/s$ and $\dot{\alpha}_0 = 40°/s$ are the movement velocities used in the previous experiments. Further explanations in the text.



(a) distance reproduction performance

(b) turning performance

Figure 9: Typical distance reproduction (a) and turn execution performance (b) from one subject. The left and right graphs show the executed distance respectively turning angle, plotted versus their corresponding correct values. The distance and angular compression are 0.9 and 0.99, respectively, as is indicated in the top inset of each figure. The enlargement in (b) illustrates the extremely small within-subject variability and error for turns, indicating the ease with which the task was performed.

| independent variable | correlated with dependent variable | | | |
|---|---|---|---|---|
| distance $s_1$ | $s_2 - s_1$ : | $r = -0.16$ | $r^2 = 2.5\%$ | t(8)=2.4, p=0.04 |
| | $\lvert s_2 - s_1 \rvert$ : | $r = 0.31$ | $r^2 = 9.7\%$ | t(8)=5.5, p=0.0005 |
| | $s_2$ : | $r = 0.82$ | $r^2 = 66.7\%$ | t(8)=8.9, p<0.0001 |
| velocity ratio $v_{s_2}/v_{2_1} = gain_{s_2}/gain_{s_1}$ | n.s. | | | |
| turning angle $\alpha_c$ | $\alpha_m - \alpha_c$ : | $r = -0.30$ | $r^2 = 8.8\%$ | t(8)=6.7, p=0.0002 |
| | $\lvert \alpha_m - \alpha_c \rvert$ : | $r = 0.17$ | $r^2 = 2.9\%$ | t(8)=3.0, p=0.017 |
| | $\alpha_m$ : | $r = 0.999$ | $r^2 = 99.8\%$ | t(8)=9.6, p<0.0001 |
| turning direction | n.s. | | | |
| rotational vel. $\dot{\alpha} = gain_{\alpha} \cdot \dot{\alpha}_0$ | n.s. | | | |

Table 3: Results from the correlation analysis for the TURN&GO experiment. Explanations in text.

slight undershoot for large distances, indicating a compression of the distance response range, which can also be seen from figure 12. A linear regression for the absolute distance error[7] reveals that response variability increased for larger distances $s_1$.

A similar trend was apparent for the turning response: The angle to turn $(\alpha_c)$ was highly negatively correlated with the signed turning error and highly positively correlated with the absolute turning error (cf. tab. 3). A linear regression for the signed turning error[8] reveals a slight overshoot for small turns and an undershoot for large turns, indicating a compression of the turn response range (cf. fig. 12). A linear regression for the absolute distance error[9] reveals that response variability slightly increased for larger turning angles $\alpha_c$.

To test how well the correct distance or turning angle predict the observed distance and turning angle, respectively, we performed a similar correlation analysis on them. As expected, the correlation was highly significant for both distances and turns (cf. tab. 3). A $r^2$ value of 0.67 for distances implies that 67% of the variance in the distance traveled $(s_2)$ can be explained by the distance to reproduce $(s_1)$. For the turning angles, almost the whole variance $\left(r^2 = 99.8\%\right)$ in angles turned $(\alpha_m)$ can be explained by the angle to turn $(\alpha_c)$, indicating an excellent turning response and a negligible execution error.

### 4.2.4 Variability

Figures 9 and 12 illustrate the relatively small within- and between-subject variability, especially for the turning response. To quantify this, we compared the variance of the four variables with the results from experiment TOWN&BLOBS, using a pairwise F-tests. The decrease in variance for the distance error from experiment TOWN&BLOBS to experiment TURN&GO did not turn out significant (F[19,8]=2.1, p=0.29). The decrease in distance compression variance was marginally significant (F[19,8]=3.8, p=0.058). The decrease in variance for both turning error and angular compression proved highly significant (F[19,8]=19.3, p=0.00022 and F[19,8]=432.4, p<0.0001, respectively).

## 4.3 Discussion

Subjects were able to accurately integrate velocity and acceleration information derived from optic flow to estimate angles turned and distances traveled, irrespective of movement velocity.

The small variance and systematic errors for executing turns imply an almost negligible execution error for rotations: Untrained subjects are able to perform simple turns with almost perfect performance, given optic flow information only. This suggests that the observed systematic errors and variance in experiment TOWN&BLOBS cannot be caused by an execution error for turning angles.

The observed distance compression of 0.9 proved not significant (cf. fig. 12), and is significantly smaller than the strong distance compression of 0.6 observed in experiment TOWN&BLOBS (two-tailed two sample t-test, t(27)=2.9, p=0.008). That implies that distance encoding and execution errors, if present, widely cancel each other out. This suggests that the strong distance compression observed in Experiment TOWN&BLOBS cannot be fully explained by the systematic distance errors observed in this experiment, but have to be partly attributed to the mental representation or geometric reasoning. This contradicts the

---

[7] The regression equation for the absolute distance error yields $|s_2 - s_1|\,(s_1) = 0.151 \cdot s_1 + 3.2$.

[8] The regression equation for the signed turning error yields $\alpha_m - \alpha_c(\alpha_c) = -0.042 \cdot \alpha_c + 4.6$.

[9] The regression equation for the absolute turning error yields $|\alpha_m - \alpha_c|\,(\alpha_c) = 0.024 \cdot \alpha_c + 3.4$.

axioms of the encoding error model, which is consequently inapplicable. Some of the variability in the homing response of experiment TOWN&BLOBS, however, could be caused by the distance variability observed in this experiment.

# 5 Experiment 4: "RANDOM TRIANGLES"

### 5.0.1 Purpose

Experiment TOWN&BLOBS demonstrated that homing by optic flow or transient landmarks is possible and allows for decent homing performance, apart from a rather pronounced distance compression. A question that remained unanswered was how the simplicity of the triangle geometry (only isosceles triangles with angles $\alpha$ in 30° steps) might have influenced homing performance. To tackle this question, we used the triangle completion paradigm with the 3D field of blobs again, but with novel triangles of completely randomized geometry for each trial.

### 5.0.2 Hypothesis

If subjects had been able to take advantage of a simple, repetitive, isosceles triangle geometry in experiment TOWN&BLOBS, we would now expect a clear deterioration in homing performance: Subjects should be less certain about the correct homing response and therefore be more conservative in their response, leading to a more pronounced compression rate and an increase in variability.

## 5.1 Methods

Participants were the same ten subjects as in experiment TURN&GO. The experimental procedure was the same as in experiment TOWN&BLOBS using the 3D field of blobs, but using different triangle geometries for each trial.

### 5.1.1 Experimental design

The experimental design is summarized in table 4. Each subject completed 60 trials. For each trial, a value for the length of the first segment, the second segment and the enclosed turning angle was drawn independently and randomly from a set of 60 equally spaced values each. Additionally, the turning direction was chosen randomly. There was no repetition of conditions, which ensured that subjects could not memorize individual triangle geometries and utilize them directly in a later trial, as might have been possible in experiment TOWN&BLOBS.

## 5.2 Results

### 5.2.1 Signed errors

To allow for a direct comparison with the previous experiments, we used the same four variables (turning error, distance error, angular and distance compression) and plotted them as usual in figure 12. Mean turning error and distance error were remarkably small and did not differ significantly from zero (two-tailed t-test against zero, t(9)=0.79, p=0.44 and t(9)=0.35, p=0.73, respectively) or from the results from experiment TOWN&BLOBS (see fig. 12). However, the variance of the distance error was significantly increased, compared to experiment TOWN&BLOBS (F-test for comparison of variances, F[9,19]=5.0, p=0.0032), whereas the variance of the angular error remained unchanged (F[19,9]=1.7, p=0.42).

### 5.2.2 Compression rates

Both angular and distance response showed an obvious compression of 0.76 and 0.85, resprectively, which was significantly below the correct value of one (2-tailed t-test, t(9)=5.0, p=0.00073 and t(9)=3.9, p=0.0037, respectively). The angular compression was slightly, but insignificantly more pronounced than in experiment TOWN&BLOBS (two-tailed two sample t-test, t(28)=1.3, p=0.22). In contrast, the distance compression was significantly smaller than in experiment TOWN&BLOBS (two-tailed two sample t-test, t(28)=2.6, p=0.016). Interestingly enough, the variance of both angular and distance compression was significantly reduced, compared to experiment TOWN&BLOBS (F-test, F[19,9]=6.0, p=0.0089 and F[19,9]=6.5, p=0.0069, respectively).

### 5.2.3 Correlation analysis

To investigate the influence of the independent variables and to attempt to explain the observed variance in the data, we performed pairwise correlation analyses. The details and results of the correlation analysis are summarized in table 5. The analysis revealed a strong correlation between the independent variable $s_1$, $s_2$ and $s_{3_c}$ and the observed distance error. For increasing values of $s_1$ and $s_2$ and $s_{3_c}$, the distance response shifted from an overshoot to an undershoot, indicating a tendency of the subjects to produce medium-sized triangles, which indicates a regression towards mean homing distances for differently sized triangles. The influence of $s_1$ and $s_2$ on turning error is best understood by looking at the influence of their ratio ($s_2/s_1$) or difference ($s_2 - s_1$): For triangles with a shorter second segment ($s_2 < s_1$), turning angles are increasingly overshot. Conversely, turning angles are increas-

| independent variable | levels | values |
|---|---|---|
| $s_1$= length of segment 1 | 60 (equally spaced) | $s_1 \in \{20m, \ldots, 73m\}$ |
| $s_2$= length of segment 2 | 60 (equally spaced) | $s_2 \in \{20m, \ldots, 73m\}$ |
| $\alpha$=turning angle at 1st corner | 60 (equally spaced) | $\alpha \in \{20°, \ldots, 160°\}$ |
| turning direction | 2 | left or right |

Table 4: Experimental design for the RANDOM TRIANGLES experiment.

ingly undershot for triangles with a longer second segment ($s_2 > s_1$). This highly significant correlation explains about $r^2 = 11.4\%$ of the variance in homing errors. However, distance and turning errors were not independent from each other: Distance error increased with increasing turning error. Interestingly enough, the turning angle $\alpha$ between the first and second segment did not show any systematic influence on the pattern of homing errors. The strong correlation observed between distance error and correct homing distance $s_{3_c}$ and between turning error and correct homing angle $\beta_c$ expresses the distance and turn compression described above.

### 5.3 Discussion

The most striking results from this experiment are the relatively small between-subject variability of compression rates and the less pronounced distance compression, compared to experiment TOWN&BLOBS. This is all the more astonishing, as the variability in distance error was significantly increased.

The correlation analysis revealed a regression towards “standard” responses: For “extreme” triangles (i.e., extreme values of $s_1$, $s_2$, $s_{3_c}$, $s_2 - s_1$ and $\alpha_c$), subjects responded as if those values weren’t as extreme. This could be interpreted as a tendency to opt for the “safe bet” for difficult triangle geometries.

However, there was no clear performance deterioration as compared to experiment TOWN&BLOBS. This suggests that neither motor learning, nor direct learning transfer between trials, nor the simplicity of isosceles triangles was a determining factor for homing accuracy in experiment TOWN&BLOBS.

## 6 General discussion

The negligible systematic execution error for turns found in experiment TURN&GO suggests that for all four Experiments, the observed turning angle directly reflects the turning angle intended by the subject. The same is true for distances traveled, but with a reduced precision. Hence, we can use the observed navigation behavior to infer about the intended navigation behavior and the underlying mental representation.

We can use this information to interpret the observed angular and distance compression in experiment RANDOM TRIANGLES, which was significantly below its correct value: Subjects generally overshoot small turning angles or distances and undershoot large ones. Consequently, their mental representation of the triangle should also show this compression. The mental representation of the triangle traveled must have been distorted such that small turning angles $\overline{\beta}$ and homing distances $s_3$ were overestimated and large ones underestimated. Whether this is due to a systematic error in the encoding of the path traveled or a systematic error in mental spatial reasoning or representation cannot be answered from experiments TURN&GO and RANDOM TRIANGLES alone. To attempt to answer this question, we’ll combine the results from all experiments and the interviews of the subjects following the experiments.

There is some rather anecdotal evidence suggesting that encoding error for turns might be quite small. In general, subjects were able to estimate turns well even when not actively controlling the motion, e.g., when the experimenter turned and they just observed. Many subjects were even able to pinpoint the exact angles $\overline{\alpha}$ turned in experiment TOWN&BLOBS or during the training phases, indicating a negligible encoding error for turns.

There is no direct evidence on systematic encoding errors for distances traveled, as distances cannot be queried without referring to an absolute or relative scale. However, experiment TURN&GO demonstrated that subjects can reproduce distances fairly well, suggesting that the distance traveled gives a rough estimate of the distance mentally represented and intended to travel.

Using that information, we’ll try now to understand the origin of the strong distance compression (0.60 ± 0.07) observed in experiment TOWN&BLOBS. Most subjects realized soon that $s_1$ and $s_2$ were equal and held constant. This suggests that $s_1$ and $s_2$ should have been encoded to the same, constant value, irrespective of $\alpha$. On the other hand, experiment TURN&GO showed that intended turns can be executed with negligible systematic errors. As the turning response in experiment TOWN&BLOBS was essentially correct, and most subjects knew they were traveling isosceles triangles, we can conclude that they had an essentially

| correlation between | | $r$ | $r^2$ | t(8) | p |
|---|---|---|---|---|---|
| dist. error | $s_1$ | -0.310 | 9.6% | 5.9 | 0.00027 |
| dist. error | $s_2$ | -0.176 | 3.1% | 4.4 | 0.0017 |
| dist. error | $\alpha$ | -0.007 | 0.0% | 0.28 | 0.78 |
| dist. error | $s_2/s_1$ | 0.095 | 0.9% | 2.0 | 0.073 |
| dist. error | $s_2 - s_1$ | 0.086 | 0.7% | 2.0 | 0.080 |
| dist. error | $s_{3_c}$ | -0.256 | 6.6% | 4.0 | 0.0031 |
| dist. error | $\beta_c$ | 0.015 | 0.0% | 0.25 | 0.80 |
| turn error | $s_1$ | 0.263 | 6.9% | 5.8 | 0.00027 |
| turn error | $s_2$ | -0.224 | 5.0% | 3.9 | 0.0037 |
| turn error | $\alpha$ | -0.030 | 0.1% | 1.3 | 0.21 |
| turn error | $s_2/s_1$ | -0.290 | 8.4% | 5.5 | 0.00039 |
| turn error | $s_2 - s_1$ | -0.338 | 11.4% | 5.4 | 0.00045 |
| turn error | $s_{3_c}$ | -0.044 | 0.2% | 0.74 | 0.48 |
| turn error | $\beta_c$ | 0.357 | 12.8% | 4.3 | 0.0020 |
| turn error | dist.err. | 0.126 | 1.6% | 1.9 | 0.087 |

Table 5: Correlation table for correlation between the error for distances and turns (first column) and the parameters in the second column. The Pearson correlation coefficient, r, and $r^2$, the coefficient of determination were computed by performing a correlation for each subject's data individually, transforming the resulting r-values (via a Fisher r-to-Z transformation) into Z-values, taking their mean, and transforming the mean back via the inverse transformation (Fisher Z-to-r transformation) into mean r-values. To test whether the correlation coefficients differ significantly from zero ("not correlated"), a two-tailed t-test was calculated for the r-to-Z transformed r-values of the individual subject's data. The resulting significance level is displayed in the last column.

correct mental representation of the triangle geometry, apart from a possible overall scaling factor.

The question arising now is where the rather pronounced distance compression in experiment Town&Blobs stems from, given that the mental representation was an isosceles triangle with approximately the correct angle $\alpha$.

An explanation we favor is that subjects experienced problems in determining the correct homing distance from the mental representation, even though they had all the information needed. Most subjects were apparently unable to mentally compute or somehow infer the correct homing distance from a known triangle geometry. This is also the main difference between the distance reproduction task in experiment Turn&Go and the triangle completion tasks in experiment Town&Blobs and Random Triangles: For the latter experiments, subjects had to use mental geometric or spatial reasoning, with experiment Random Triangles requiring more complex spatial reasoning.

## 6.1 Spatial imagination abilities test scores correlate with homing performance

To investigate whether mental spatial abilities might be a determining factor for homing accuracy, we performed two standard, paper and pencil spatial imagination abilities tests with the subjects from experiment Town&Blobs and Random Triangles and correlated the results with the homing performance. Test 1 was a "Schlauchfiguren-Test" (Stumpf & Fay, 1983), where subjects saw one picture of a tube folded into a

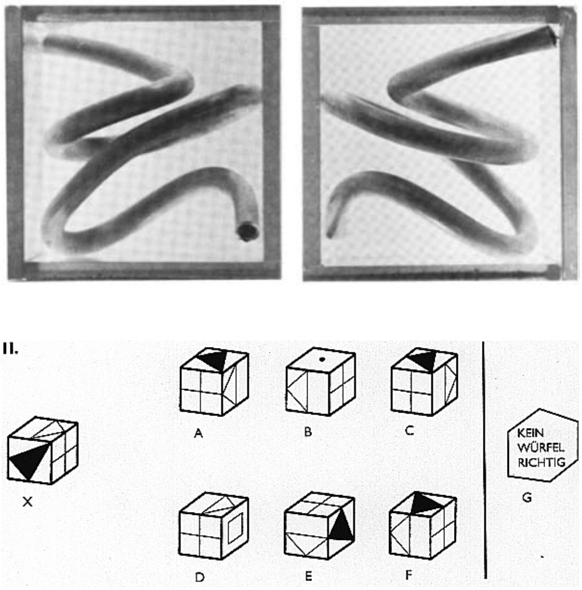


Figure 10: Sample stimulus from spatial imagination abilities test 1 (top) and test 2 (bottom).

| measurand | spatial imagination test | $r$ | $r^2$ | $t$ | $p$ |
|---|---|---|---|---|---|
| TOWN&BLOBS | | | | | |
| abs. turn. error | test 2 | -0.42 | 17% | t(12)=1.6 | 0.070 |
| abs. dist. error | test 2 | -0.36 | 13% | t(12)=1.3 | 0.10 |
| \|signed turn. error\| | test 2 | -0.55 | 30% | t(12)=2.3 | 0.021 |
| RANDOM TRIANGLES | | | | | |
| abs. dist. error | test 1 | -0.67 | 45% | t(8)=2.6 | 0.016 |
| abs. turn. error | test 2 | -0.48 | 23% | t(8)=1.5 | 0.081 |
| abs. dist. error | test 2 | -0.79 | 62% | t(8)=3.6 | 0.0035 |
| \|signed turn. error\| | test 1 | -0.48 | 23% | t(8)=1.6 | 0.080 |
| \|signed dist. error\| | test 1 | -0.66 | 43% | t(8)=2.5 | 0.019 |
| \|signed turn. error\| | test 2 | -0.54 | 29% | t(8)=1.8 | 0.0532 |
| \|signed dist. error\| | test 2 | -0.70 | 49% | t(8)=2.8 | 0.012 |

Table 6: Results of the correlation analysis between homing performance in experiments TOWN&BLOBS and RANDOM TRIANGLES and the number of correct trials in two mental spatial abilities test. Only correlations that were significant on at least a p=15% level are displayed.

transparent cube, and had to decide from which viewpoint the second picture of the same object was taken (cf. fig. 10, top pictures). The second test was a "Würfel Erkennen Test", part six of the "Intelligenz Struktur Analyse Test" (ISA, 1998), in which subjects had to judge on the identity of cubes seen from different directions (cf. fig. 10, bottom picture).

A correlation analysis was conducted between the test result (% correct responses) and the absolute error and absolute value of the signed error for turns, distances, angular compression, and distance compression. We used 14 of the 20 subjects from experiment TOWN&BLOBS and all 10 subjects from experiment RANDOM TRIANGLES. Our hypothesis was that, if mental spatial abilities played an important role for homing performance, at least one of the error measures should be negatively correlated with the test performance and no one positively. Additionally, we expect a higher correlation for experiment RANDOM TRIANGLES, which required more complex spatial reasoning. To test these hypotheses, one-sided t-tests for $r < 0$ were conducted. The results that had a $p < 0.15$ are displayed in table 6.

Five error measures were significantly correlated ($p < 0.05$), five more were approaching significance ($p < 0.1$). All of those correlations were negative, indicating that a good test result coincided with a small error measure and hence a good homing performance. For experiment RANDOM TRIANGLES, which required more complex mental spatial reasoning, both test results correlated nicely especially with the distance error measures, and were able to explain up to $r^2 = 62\%$ of the rather large variance (cf. tab. 6).

We conclude that mental spatial ability, as assessed by both tests, correlates positively with homing performance, especially for the more complex task in experiment RANDOM TRIANGLES. This suggests that mental spatial ability might be a determining factor for homing performance in triangle completion experiments based on path integration. This finding fits nicely into our explanation of the homing errors proposed earlier. However, the number of subjects participating in this study was rather limited, and further experiments are needed to corroborate this hypothesis.

### 6.2 Comparison with previous work

Here we will discuss the relationship between the present results and previous work on path integration and triangle completion. Section 6.2.5 contains a detailed reanalysis of the triangle completion data by Loomis et al. (1993) and Péruch et al. (1997) and a comparison with our results.

#### 6.2.1 Influence of proprioceptive cues for updating self-rotations

Bakker et al. (1999) investigated the relative contribution of visual, vestibular, and kinesthetic cues on path integration in virtual environments. Subjects were immersed in a forest of randomly positioned trees and asked to turn specific angles under different combinations of visual, vestibular, and kinesthetic cues.

Subjects were either seated on a computer-driven turntable to provide isolated vestibular cues, or were standing and using their legs to turn around the vertical axis, thus getting both vestibular and kinesthetic cues. Visual cues were presented through a stereoscopic Head-mounted Display (virtual I/O HMD with a FOV of 24° × 18° and 180,000 pixel resolution).

Subjects generally undershot the instructed angle by a constant factor, with the strongest underturn of 41%

for the pure visual condition, suggesting a considerable overestimation of the visually perceived turning velocity by a factor of 1.7. The purely visual condition "*resulted in the largest (absolute) errors, with largest standard deviations, and the lowest subjective confidence*" (Bakker et al., 1999, p. 50). Additional vestibular cues (turning on the turntable) only slightly improved performance. Best performance and smallest undershoot was found for blind turning while standing, most likely due to the kinesthetic feedback from the legs. The authors conclude that "*proprioceptive feedback, particularly kinesthetic, can be used quite effectively for orientation based on path integration. The perception of orientation from optic flow or vestibular feedback alone is inaccurate and may lead to disorientation*" (Bakker et al., 1999, p. 51).

A similar effect of undershooting an instructed turning angle with proprioceptive feedback is found by some experimenters (Bles, Dejong, & Dewit, 1984; Sholl, 1989; Loomis et al., 1993), others find an overshoot (Klatzky et al., 1990) or a compression towards reference angles of 90° (Sadalla & Montello, 1989).

Similar to the undershooting found by Bakker et al. (1999) using a HMD for the visual display, Péruch et al. (1997) found an undershoot of purely visually displayed rotations by 16% using a flat projection screen: When required to turn by 180°, subjects responded by a turn of only 150.4° ± 0.9°.

In contrast to the above mentioned studies, experiment TURN&GO demonstrated that untrained subjects can turn by visual path integration with a high accuracy and without systematic over- or underturning (cf. section 4). Moreover, both within- and between subject variability in experiment TURN&GO was much smaller than for nonvisual turning (Marlinsky, 1999b; Klatzky et al., 1990; Bakker et al., 1999) and turns presented only visually through a HMD (Bakker et al., 1999). This striking performance difference might be caused by the display being a half-cylindrical projection screen with a wide FOV, as will be discussed in section 6.3.

#### 6.2.2 Triangle completion experiments with head mounted display

Bud (2000) and Duchon et al. (1999) conducted similar triangle completion experiments in a virtual environment consisting of a large round room with uniformly textured walls and floor. In one condition, egomotion was controlled using a joystick and visually presented via a non-headtracked head mounted display with a horizontal FOV of 60°. Subjects' homing performance was sensitive to changes in segment length of the triangle, suggesting that they were able to integrate optic flow from translations to yield the distance traveled. In contrast, subjects' mean homing response reflected no sensitivity to variations in turning angle $\overline{\alpha}$: For isosceles triangles with angles $\alpha \in \{60°, 90, 120°\}$, subjects produced the same response regardless of actual triangle geometry, acting as if traveling an equilateral triangle. Subjects seemed to be unable to use the rotational optic flow to extract the turning angle. This effect was not found in the present experiments or the experiments by Péruch et al. (1997), all of which used projection screens. This suggestes that the type of display (HMD versus projection screen) might influence the sensitivity to turning angles.

In another condition in Bud's (2000) subjects wore a headtracked HMD and physically walked triangles, with the triangle corners being indicated visually as before. The homing results showed a reduced variability reflecting a higher subjective confidence. However, subjects still gave the same stereotyped response irrespective of the turning angle $\overline{\alpha}$.

Compared to the tendency to *under*turn by 7.1° (SD: 35.9°) for purely visual navigation in the first condition, physical walking led to a general *over*turning by 19.9° (SD: 27.1°). Removing all visual information except the poles denoting the triangle corners hardly altered subjects' responses, indicating that the proprioceptive cues from walking dominated over optic flow information. This overturning and lack of stimulus response for physical rotations was not found by blind walking experiments by e.g. Loomis et al. (1993), Marlinsky (1999c) and can hence not be simply attributed to proprioceptive cues from walking. Consequently, the effect seems to be caused by the visual display presenting the triangle to be traveled (cf. section 6.3).

Technological advances in display quality might help to overcome some of the problems observed for most displays: Using a custom-built HMD, Arthur (2000) demonstrated that navigation time can benefit from enlarging the horizontal FOV beyond 112°. A lighter HMD with lower latency and better visual quality, however, resulted in a similar performace benefits, even though it had a limited FOV of only 48°. Further studies are needed to pinpoint the exact influences and interactions of the visual display parameters on navigation and spatial orientation abilities.

#### 6.2.3 Non-visual navigation experiments based on path integration

The simplest experimental paradigm for path integration studies is blind locomotion with ears muffled.

Vision and audition are easily excluded by blindfolding subjects and displaying white noise over noise-attenuating headphones. Using this paradigm, Sauvé (1989), Klatzky et al. (1990) and Loomis et al. (1993) showed in triangle completion studies that proprioceptive and vestibular cues from blind walking allow for homing, but lead to strong systematic errors. Blind and blindfolded subjects were led along two sides of given triangles and had to walk back directly to the origin ("triangle completion task"). In all three studies, subjects showed a regression towards standardized responses: Subjects overturned for small turning angles (<90°) and underturned for large turning angles (>90°). The same compression towards stereotyped responses was found for distances traveled: Short distances were overshot, large distances undershot. This bias is a commonly found trend in psychophysical experiments (Poulton, 1979; Stevens & Greenbaum, 1966). Loomis et al. (1993), in accordance to Klatzky et al. (1990) conclude that "*not only were there significant signed errors for the average of all subjects but also no single subject came close to exhibiting negligible errors over the 27 triangles. It appears that even for the short paths over which subjects were passively guided here* [2, 4, and 6m segment length, remark by the author], *the proprioceptive and vestibular cues were inadequate for accurate path integration.*"

In similar blind walking triangle completion experiments, Marlinsky (1999c) found a compression for distances, but not for turns. However, a small sample size and large within- and between-subject variability might have covered possible effects. Transporting subjects passively along the outbound path (instead of guided walking) increased distance compression when overall triangle size was varied, but not for changes in triangle shape (angle $\alpha$). Turning angles, however, remained always unaffected by the movement type.

Attempts to predict an individual subject's homing responses by using their distance reproduction and turn execution data from previous experiments (Marlinsky, 1999a, 1999b) revealed a qualitative aggreement for the one subject tested. However, a quantitative measure used on more subjects is needed before one could reliably ascribe the individual subjects' systematic errors in homing performance to their individual systematic errors in distance reproduction and turn execution. For the visual triangle completion experiments described in this paper, the observed systematic errors could not be explained by the (rather small) systematic errors in distance reproduction and turn execution observed in experiment TURN&GO.

Movement velocity also seems to affect path integration performance. Blind triangle completion experiments by Mittelstaedt & Glasauer (1991) revealed an influence of the walking speed: Subjects overshot distances for walking speeds faster-than-normal and undershot distances for slow walking speed.

The tendency to turn too far and intersect the outbound path, which is commonly found in animal triangle completion experiments for desert ants (e.g., Müller & Wehner, 1988) spiders (Görner, 1958), bees (Bisetzky, 1957), hamsters (Seguinot, Maurer, & Etienne, 1993) and dogs (unpublished data from Séguinot, mentioned by Etienne et al. (1996)), was not found for human triangle completion (e.g., Klatzky et al., 1990; Loomis et al., 1993).

### 6.2.4 Triangle completion experiments with projection screens

Loomis et al. (1993) and Klatzky et al. (1990) have shown that kinesthetic and vestibular cues from blind walking are inadequate for accurate path integration as assessed by triangle completion experiments. Péruch et al. (1997) conducted comparable triangle completion experiments in virtual environments to investigate human path integration ability based on visual information (optic flow). The experiments by Péruch et al. (1997) (reported also in Wartenberg et al. (1998)) will be described here in more detail, as our experiments are an extension of their's.

Subjects used a joystick to move within an area surrounded by 16 identical cylinders equally spaced on a circle of 60m diameter. The simulated ego-motion was displayed on a planar projection screen subtending a physical field of view (FOV) of 45° horizontal × 38° vertical. Subjects had to complete 27 triangles corresponding to a factorial combination of 3 values for the simulated field of view (horizontal sFOV = 40°, 60° and 80°) × 9 triangle geometries (3 angles × 3 lengths of the second segment).

Interestingly enough, the sFOV had no significant effect on homing performance. Subjects showed a general undershoot for both turning angles and distances traveled (see fig. 11). Results also revealed a strong regression towards stereotyped values for turning angles and distances traveled, especially for isosceles triangles (see fig. 12). All those effects were stronger than in the blind walking studies by Loomis et al. (1993) and Klatzky et al. (1990), suggesting that path integration by optic flow is inferior to path integration by kinesthetic and vestibular cues. We will argue that this is not true in general and heavily depends on the experimental setup (cf. section 6.2.5 and 6.3).

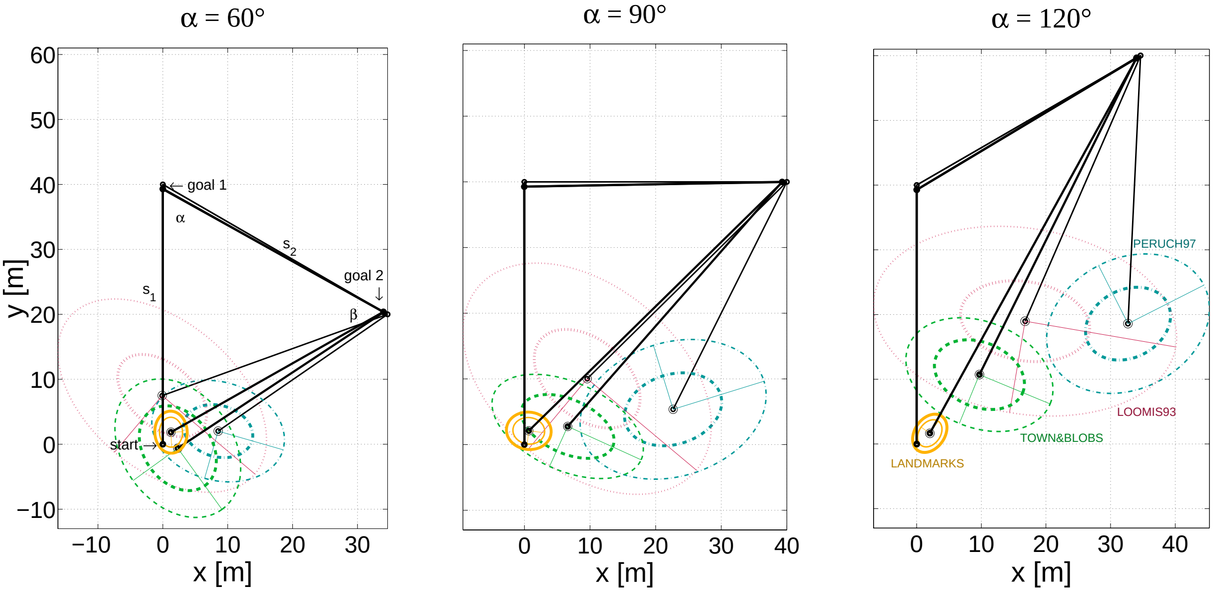


Figure 11: Homing performance under different conditions, plotted as in figure 5 and 7. Dashed lines represent results from experiment Town&Blobs, solid lines for experiment Landmarks without scene swap, dotted lines for visual triangle completion within a circle of equal cylinders (reanalysis of data from Péruch et al. (1997), data from experiment 1 and 2 pooled) and dash-dotted lines for blind walking (reanalysis of data from Loomis et al. (1993), Experiment 1, triangles with $s_1 = s_2 = 4m$). Data from Péruch et al. and Loomis et al. is scaled to fit the triangles used in our experiments. Non-overlapping 95% confidence ellipses (inner ellipses with thick lines) indicate significant differences between the experiments (Batschelet, 1981).

### 6.2.5 Comparison to triangle completion studies by Loomis et al. (1993) and Péruch et al. (1997)

Here we will compare the results from experiment Town&Blobs and Landmarks with the results from visual and blind walking triangle completion experiments by Péruch et al. (1997) (labeled "Peruch97" as in figure 11 and 12, cf. section 6.2.4) and Loomis et al. (1993) ("Loomis93", cf. section 6.2.3), respectively.

For a first qualitative comparison, we plotted the mean homing endpoints for the triangle geometries common to all experiments in figure 11. Homing endpoints for experiment Landmarks (with reliable landmarks) were quite close to the starting point, irrespective of the triangle geometry, suggesting that landmarks were used for homing and allowed for highly accurate homing. The other experiments were all based on path integration, and we observe a similar influence of triangle geometry on homing performance: Larger turning angles and homing distances are increasingly undershot. This tendency is strongest in Exp. Peruch97, less pronounced in Exp. Loomis93 and smallest in Exp. Town&Blobs. The inter-subject turning variability is largest for Exp. Loomis93, indicated by the standard ellipses covering about 40% of the mean subjects' homing endpoints.

To examine the difference between the path integration experiments more quantitatively, the results for all experiments are displayed in figure 12, for the measures turning error, distance error, angular and distance compression. Additionally, we performed two-tailed t-tests and F-tests to compare means and variances of interest, respectively.

The mean turning errors for the experiment by Loomis et al. were close to zero, but showed a rather large variance which was significantly larger than for experiment Town&Blobs (F[36,19]=3.6, p=0.004 for isosceles triangles and F[36,19]=3.7, p=0.003 for all triangles). All other measures from the experiments by Loomis et al. and Péruch et al. were substantially below their correct value, indicating general undershooting and biases towards stereotyped responses (two-tailed t-test, t(36)>4.2 and t(25)>6.2, respectively, p<0.0005 for all cases).

Compared to experiment Town&Blobs, homing performance for blind walking (Loomis93 all) reveal a similar, small mean turning error and a slightly stronger distance compression. Distance error and angular compression were much more pronounced (t(55)=5.4, p<0.0001 and t(55)=3.4, p=0.001, respectively). Intersubject variance for isosceles triangles (Loomis isosc.) was larger, especially for compression rates. Several subjects showed even negative compression rates.

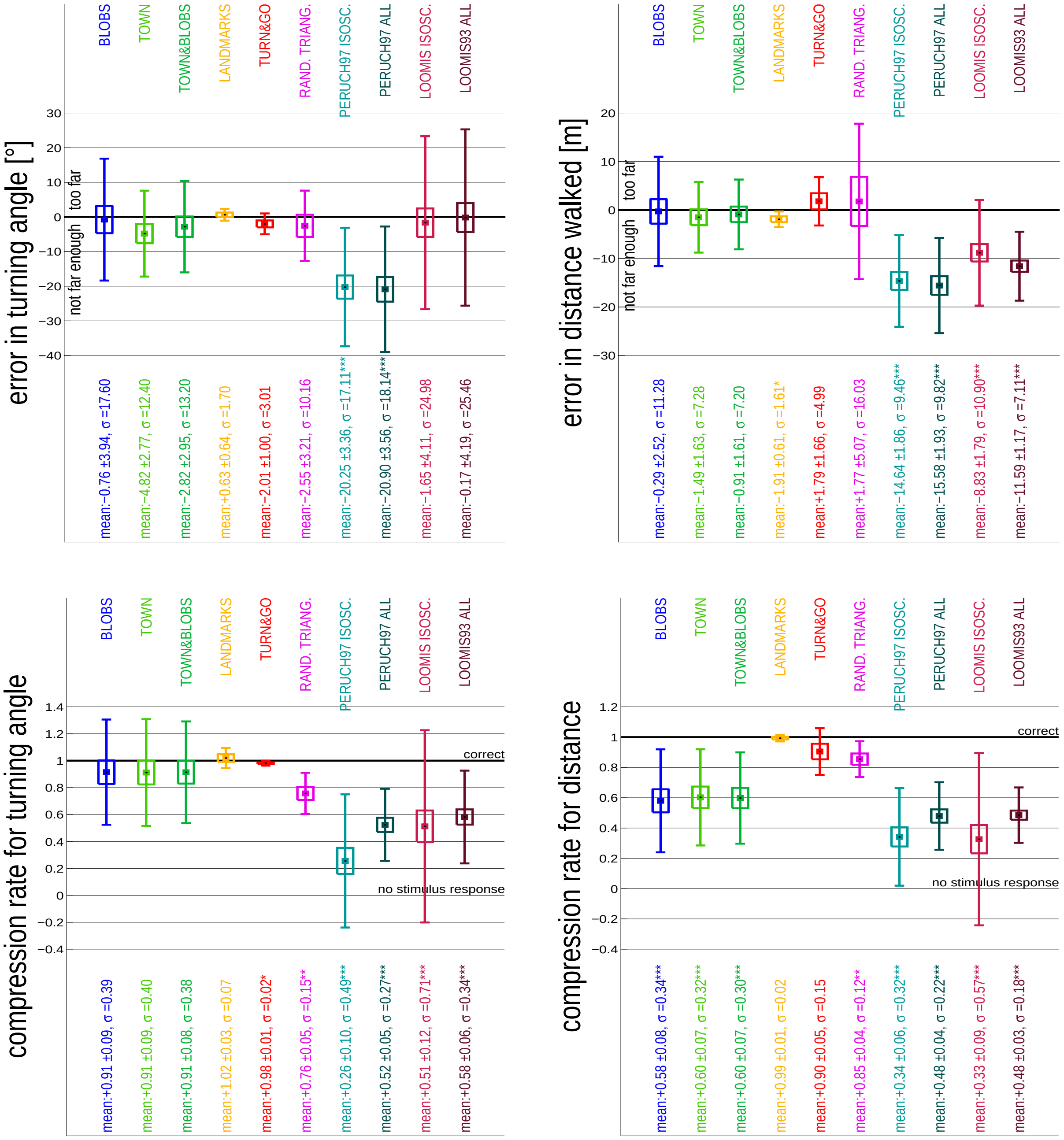


Figure 12: Comparison of navigation performance for the different experimental conditions. At the top of each plot, the experimental conditions are displayed (from let to right): Exp. 1, using the 3D field of blobs (“Blobs”), the town environment (“Town”) and data from both experiments pooled together (“Town&Blobs”); Exp. 2 (“Landmarks”) with reliable landmarks; Exp. 3 (“Turn&Go”); Exp. 4 (“Random Triangles”); reanalysis of data from Péruch et al. (1997) on visual triangle completion within a circle of equal cylinders for isosceles triangles only (“Peruch97 isosc.”) and for all triangles (“Peruch97 all”); reanalysis of data from Loomis et al. (1993) on blind walking triangle completion, again for isosceles triangles with $s_2 = 4m$ only (“Loomis93 isosc.”) and for all triangles (“Loomis93 all”). Data from Péruch and Loomis is scaled to match the triangle size used in our experiments.
Below are the plots of the four measures, the center indicating the arithmetic mean. Boxes represent intervals of one standard error of the mean, whiskers represent one standard deviation. The compression rate was defined as the slope of the linear regression fit.
At the bottom of each plot, the numeric values of the mean, standard error, and standard deviation are displayed. The stars ’*’ indicate whether the mean differs significantly (on a 5%, 0.5% or 0.05% significance level) from the corresponding correct value, depicted as a thick horizontal line.

The most obvious difference in homing results between experiments LOOMIS93 and PERUCH97 is the general undershooting of turning angles observed in experiment PERUCH97, but not LOOMIS93. This might be related to the turn execution error observed by Péruch et al.: When asked to turn around by 180°, subjects responded by turning only 150.4°, corresponding to a underturn by 16%. A similar general underturning of 15% or 20.3° was observed for isosceles triangles. Could this execution error of underturning by 16% explain the underturn of 15% observed for triangle completion, rather than a encoding error?

While distance compression was only slightly greater in experiment PERUCH97 than in experiment TOWN&BLOBS, turning error, distance error and angular compression were all much more pronounced ($t(44)=3.8$, $p=0.0005$, $t(44)=5.6$, $p<0.0001$ and $t(44)=4.1$, $p=0.0002$, respectively). The question arises as to where the obvious performance difference between experiment TOWN&BLOBS and PERUCH97 stem from, which were both based on visual path integration. The execution error of underturning observed by Péruch et al. (1997) can only explain only the differences in signed turning errors. The remaining performance differences might be caused by the different experimental procedures (training phase, number of triangles). They might also be due to differences in the VR-setup: Péruch et al. (1997) used a joystick and a planar projection screen with non-matched simulated and physical FOV, whereas mouse-button based navigation and a half-cylindrical projection screen with matched simulated and physical FOV was used for experiment TOWN&BLOBS. Further experiments might provide a more definitive answer to this question.

#### 6.2.6 “Encoding error model” is not applicable

To analyze potential origins of the systematic homing errors, Loomis et al. (1993) and Péruch et al. (1997) applied an “encoding error model”. This model was initially proposed by Fujita et al. (1993) to explain their blind walking data, and attributes all systematic errors to errors in mentally encoding the distances walked and angles turned. Loomis et al. (1993) and Péruch et al. (1997) concluded that a compression in the encoding of turns and distances is the only source of the observed systematic errors. Péruch et al. (1997) argued for a nonlinear compression according to a power function with exponents below 1, where as Fujita et al. (1993) and Loomis et al. (1993) used a simple linear compression.

It’s not clear that the encoding error model is applicable to Péruch’s data. There is some evidence that the assumption of no execution error might not be met: Péruch et al. reported a significant systematic undershooting by 16% (30°) for requested simple 180° turns. This indicates a turn execution error, which in turn violates the axioms of the encoding error model. Regardless, as we argued before (cf. section 4.3 and 6), the encoding error model alone was not able to explain the observed systematic errors in our experiments. Furthermore, we found evidence that the mental determination of the homeward trajectory not free of systematic errors. Hence, the axioms of the encoding error model were not satisfied, making it inapplicable. Attempts to nevertheless apply this encoding error model to our data produced nonsensical results.

### 6.3 General conclusion

The experiments reported here were aimed at investigating human navigation ability based solely on visual path integration. We found that untrained subjects were able to reproduce distances and perform turns with negligible systematic errors, irrespective of movement velocity. The systematic errors and variance both within-subject and between-subject was strikingly small, especially for rotations. This finding is in sharp contrast with results from turning experiments by Bakker et al. (1999), who demonstrated that visual information displayed via a head-mounted display leads to systematic undershooting of turning angles by 41% and a large variability. Using a flat projection screen, Péruch et al. (1997) found an undershoot of purely visually displayed rotations by 16% (cf. section 6.2.4). This trend suggests that the half-cylindrical projection screen used in the present study might be the determining factor for the excellent turning performance observed there. However, the large FOV of 180° does not seem to be the sole determining factor for turning accuracy, even though it facilitates navigation (Arthur, 2000): Systematically reducing the FOV in previous triangle completion experiments only slightly decreased homing performance (Riecke, 1998, Exp. 4). This suggests that the half-cylindrical reference frame provided by the projection screen and the visibility of one’s own body might also play a critical role for navigation performance (cf. section 1.6). Most subjects experienced little difficulties determining egocentric angles between objects presented on the screen: The half-cylindrical reference frame might facilitate the estimations of egocentric angles by suggesting a polar coordinate system. This hypothesis is corroborated by the fact that we did not find the strong bias towards stereotyped turn responses typically observed for triangle completion experiments (Péruch et al., 1997; Bud, 2000; Klatzky et al., 1990; Loomis et al., 1993). Further experiments are planned to investigate this issue by blocking out the vision for parts of the projection screen, thus removing the semi-cylindrical reference frame.

Path integration using solely optic flow proved to be sufficient for basic navigation tasks (homing by triangle completion, rotations and translations). However, homing distances were compressed towards stereotyped responses. In trying to understand the origin of the systematic homing errors observed, we suspect that mental spatial abilities might be a determining factor, as execution and encoding errors at least for turns were rather small and could not explain the observed data. Results from two standard mental spatial abilities tests confirmed our hypothesis: Especially for the more difficult triangle completion tasks in experiment RANDOM TRIANGLES, mental spatial ability correlated positively with homing performance. This suggests that subjects with good mental spatial abilities had less problems determining the correct homing response from the information available. However, further experiments are needed to test for a *causal* relation between mental spatial ability and navigation performance.

Contrary to our expectation, most subjects were *not* able to take advantage of natural-looking landmarks if they were only temporarily visible. The reasons for this remain unclear. Longer exposure to virtual reality and the experimental procedures might allow subjects to develop more efficient strategies, as was demonstrated in Riecke (1998, Exp. 4), cf. page 4. Conversely, triangle completion experiments with stable, reliable landmarks demonstrated that piloting by salient landmarks and visual scene-matching plays a dominant role in visual navigation, is used whenever possible and leads to almost perfect homing performance.

It is often claimed that kinesthetic and vestibular cues are necessary for spatial orientation tasks involving rotations of the observer (Klatzky et al., 1998; May et al., 1995; Chance et al., 1998; Bakker et al., 1999, cf. section 1.5). It might well be that purely visually displayed movements do not allow for the rapid, obligatory spatial updating found by Rieser (1989), Wang & Simons (1999), May & Klatzky (2000), Farrell & Robertson (1998) for physical movements. However, the lack of all nonvisual cues including vestibular and kinesthetic cues in the present experiments did not prevent subjects from executing turns, reproducing distances and performing triangle completion tasks with rather small systematic errors. Extended exposure to virtual environments, initial feedback training, unlimited response time, the non-immersive VR setup, and the spatial reference frame and large FOV provided by the half-cylindrical projection screen might all contribute to the relatively good overall navigation performance. From the present data, we can only speculate about the determining factors. Further experiments will be needed to pinpoint the critical factors for good spatial orientation in virtual and real environments.

Using a virtual reality setup proved to be a powerful method to investigate human navigation abilities and infer about the underlying mental spatial processes. The “scene swap” paradigm and a 3D field of blobs allowed us to reduce possible navigation mechanisms to purely visual path integration without any landmarks. Applying this paradigm, we were able to demonstrate that purely visual path integration is indeed sufficient for basic navigation task like rotations, translations, and homing.

## References

Alfano, P. L. & Michel, G. F. (1990). Restricting the field of view: Perceptual and performance effects. *Perceptual and Motor Skills*, **70**, 35–45.

Arthur, K. W. (2000). *Effects of Field of View on Performance with Head-Mounted Displays*. Ph.D. thesis, Department of Computer Science, University of North Carolina, Chapel Hill. [Online] Available: http://www.cs.unc.edu/ arthur/diss/.

Bakker, N. H., Werkhoven, P. J., & Passenier, P. O. (1999). The effects of proprioceptive and visual feedback on geographical orientation in virtual environments. *Presence-Teleoper. Virtual Env.*, **8**(1), 36 – 53.

Batschelet, E. (1981). *Circular statistics in biology*. London: Acad. Pr.

Begault, D. R. (1994). *3-D sound for virtual reality and multimedia*. Boston: Academic Press Professional.

Benhamou, S. (1997). On systems of reference involved in spatial memory. *Behavioural Processes*, **40**(2), 149–163.

Bisetzky, A. R. (1957). Die Tänze der Bienen nach einem Fußweg zum Futterplatz. *Z. vergl. Physiol.*, **40**, 264–288.

Bles, W., Dejong, J. M. B. V., & Dewit, G. (1984). Somatosensory compensation for loss of labyrinthine function. *Acta Oto-Laryngol.*, **97**(3-4), 213 – 221.

Bud, M. J. (2000). Path Integration from Optic Flow and Body Senses in an Homing Task. Master’s thesis, Department of Cognitive & Linguistic Sciences, Brown University.

Bülthoff, H. H. & van Veen, H. A. H. C. (2000). Vision and Action in Virtual Environments: Modern Psychophysics in Spatial Cognition Research. In L. R. Harris & M. Jenkin (Eds.), *Vision and Attention*chap. 12. New York: Springer-Verlag. in press.

Chance, S. S., Gaunet, F., Beall, A. C., & Loomis, J. M. (1998). Locomotion mode affects the updating of objects encountered during travel: The contribution of vestibular and proprioceptive inputs to path integration. *Presence: Teleoperators & Virtual Environments*, **7**(2), 168–178.

Christou, C. G. & Bülthoff, H. H. (1998). Using Realistic Virtual Environments in the Study of Spatial Encoding. In C. F. et al. (Ed.), *Spatial Cognition II: Integrating Abstract Theories, Empirical Studies, Formal Methods, and Practical Applications*, Vol. 1849 of *Lecture notes in computer science: Lecture notes in artificial intelligence* (pp. 317 – 332). Berlin Heidelberg: Springer.

Collett, M. & Collett, T. S. (2000). How do insects use path integration for their navigation? *Biol. Cybern.*, **83**(3), 245 – 259.

Darken, R. P., Allard, T., & Achille, L. B. (1998). Spatial orientation and wayfinding in large-scale virtual spaces: An introduction. *Presence: Teleoperators & Virtual Environments*, **7**(2), 101–107.

Darken, R. P. & Sibert, J. L. (1996). Navigating Large Virtual Spaces. *International Journal of Human-Computer Interaction*, **8**(1), 49–71.

Duchon, A., Bud, M., Warren, W. H., & Tarr, M. J. (1999). The role of Optic Flow in Human Path Integration. Talk presented at the 40th Annual Meeting of the Psychonomics Society, CA, Nov 18–21 1999.

Etienne, A. S., Maurer, R., & Séguinot, V. (1996). Path Integration in Mammals and its Interaction with Visual Landmarks. *The Journal of Experimental Biology*, **199**(1), 201–209.

Farrell, M. J. & Robertson, I. H. (1998). Mental rotation and the automatic updating of body-centered spatial relationships. *J. Exp. Psychol.-Learn. Mem. Cogn.*, **24**(1), 227 – 233.

Flach, J. M. (1990). Control with an eye for perception: Precursors to an active psychophysics. *Ecological Psychology*, **2**(2), 83.

Fujita, N., Klatzky, R. L., Loomis, J. M., & Golledge, R. G. (1993). The Encoding-Error Model of Pathway Completion without Vision. *Geographical Analysis*, **25**(4), 295–314.

Goldin, S. E. & Thorndyke, P. W. (1982). Simulating Navigation for Spatial Knowledge Acquisition. *Human Factors*, **24**(4), 457–471.

Görner, P. (1958). Die optische und kinästhetische Orientierung der Trichterspinne *Agelena labyrinthica*. *Z. vergl. Physiol.*, **41**(8), 111–153.

Hendrix, C. & Barfield, W. (1996). Presence within virtual environments as a function of visual display parameters. *Presence: Teleoperators & Virtual Environments*, **5**(3), 274–289.

Hunt, E. & Waller, D. (1999). Orientation and Wayfinding: A Review. [Online] Available: http://depts.washington.edu/huntlab/vr/pubs/huntreview.pdf.

ISA (1998). *Intelligenz-Struktur-Analyse*. Frankfurt: Sweets & Zeitlinger B. V., Sweets Test Services. ITB Institut für Test- und Begabungsforschung GmbH, Bonn.

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

Klatzky, R. L., Loomis, J. M., Golledge, R. G., Cicinelli, J. G., Pellegrino, J. W., & Fry, P. A. (1990). Acquisition of route and survey knowledge in the absence of vision. *Journal of Motor Behavior*, **22**(1), 19–43.

Loomis, J. M., Blascovich, J. J., & Beall, A. C. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behav. Res. Methods Instr. Comput.*, **31**(4), 557 – 564.

Loomis, J. M., Klatzky, R. L., Golledge, R. G., Cicinelli, J. G., Pellegrino, J. W., & Fry, P. A. (1993). Nonvisual navigation by blind and sighted: assessment of path integration ability. *J. Exp. Psychol. Gen.*, **122**(1), 73–91.

Mallot, H. A., Gillner, S., Steck, S. D., & Franz, M. O. (1999). Recognition-triggered response and the view-graph approach to spatial cognition. In C. Freska & D. M. Mark (Eds.), *Spatial Information Theory*, Vol. 1661 of *Lecture Notes in Computer Science* (pp. 367 – 380). Berlin: Springer.

Marlinsky, V. V. (1999a). Vestibular and vestibulo-proprioceptive perception of motion in the horizontal plane in blindfolded man - I. Estimations of linear displacement. *Neuroscience*, **90**(2), 389 – 394.

Marlinsky, V. V. (1999b). Vestibular and vestibulo-proprioceptive perception of motion in the horizontal plane in blindfolded man - II. Estimations of rotations about the earth-vertical axis. *Neuroscience*, **90**(2), 395 – 401.

Marlinsky, V. V. (1999c). Vestibular and vestibulo-proprioceptive perception of motion in the horizontal plane in blindfolded man - III. Route inference. *Neuroscience*, **90**(2), 403 – 411.

Maurer, R. & Séguinot, V. (1995). What is modelling for?: A critical review of the models of path integration. *Journal of Theoretical Biology*, **175**(4), 457–475.

May, M. & Klatzky, R. L. (2000). Path integration while ignoring irrelevant movement. *J. Exp. Psychol.-Learn. Mem. Cogn.*, **26**(1), 169 – 186.

May, M., Péruch, P., & Savoyant, A. (1995). Navigating in a virtual environment with map-acquired knowledge: Encoding and alignment effects. *Ecological Psychology*, **7**(1), 21–36.

Mittelstaedt, H. (2000). Triple-loop model of path control by head direction and place cells. *Biol. Cybern.*, **83**(3), 261 – 270.

Mittelstaedt, M.-L. & Glasauer, S. (1991). Idiothetic Navigation in Gerbils and Humans. *Zool. Jb. Physiol.*, **95**, 427–435.

Müller, M. & Wehner, R. (1988). Path integration in desert ants *Cataglyphis fortis*. *Proceedings of the National Academy of Sciences*, **85**, 5287–5290.

Péruch, P. & Gaunet, F. (1998). Virtual environments as a promising tool for investigating human spatial cognition. *Current Psychology of Cognition*, **17**(4–5), 881–899.

Péruch, P., May, M., & Wartenberg, F. (1997). Homing in virtual environments: Effects of field of view and path layout. *Perception*, **26**(3), 301–311.

Poulton, E. C. (1979). Models for biases in judging sensory magnitude. *Psychological Bulletin*, **86**, 777–803.

Presson, C. C. & Montello, D. R. (1994). Updating after rotational and translational body movements: Coordinate structure of perspective space. *Perception*, **23**(12), 1447–1455.

Riecke, B. E. (1998). Untersuchung des menschlichen Navigationsverhaltens anhand von Heimfindeexperimenten in virtuellen Umgebungen. Diploma thesis, Fakultät Physik, Eberhard-Karls-Universität Tübingen, Available: http://www.kyb.tuebingen.mpg.de/bu/people/bernie/.

Rieser, J. J. (1989). Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **15**(6), 1157–1165.

Ruddle, R. A., Payne, S. J., & Jones, D. M. (1997). Navigating Buildings in 'Desk-Top' Virtual Environments: Experimental Investigations Using Extended Navigational Experience. *Journal of Experimental Psychology: Applied*, **3**(2), 143–159.

Sadalla, E. K. & Montello, D. R. (1989). Remembering changes in direction. *Environ. Behav.*, **21**(3), 346 – 363.

Sauvé, J. P. (1989). *L'orientation spatiale: formalisation d'un modèle de mémorisation égocentrée et expérimentation chez l'homme.* Ph.D. thesis, Université d'Aix-Marseille II.

Seguinot, V., Maurer, R., & Etienne, A. S. (1993). Dead reckoning in a small mammal: the evaluation of distance. *J. Comp. Physiol. A.*, **173**, 103–113.

Sholl, M. J. (1989). The relation between horizontality and rod-and-frame and vestibular navigational performance. *J. Exp. Psychol.-Learn. Mem. Cogn.*, **15**(1), 110 – 125.

Stevens, S. S. & Greenbaum, H. B. (1966). Regression effect in psychophysical judgement. *Perception & Psychophysics*, **1**, 439–446.

Stumpf, H. & Fay, E. (1983). *Schlauchfiguren – Ein Test zur Beurteilung des räumlichen Vorstellungsvermögens*. Göttingen: Verlag für Psychologie Hogrefe.

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

van Veen, H. A. H. C., Riecke, B. E., & Bülthoff, H. H. (1999). Visual Homing to a Virtual Home. *Invest. Ophthalmol. Vis. Sci.*, **40**(4), #4200B3. (ARVO Abstract).

van Veen, H. A. H. C., Distler, H. K., Braun, S. J., & Bülthoff, H. H. (1998). Navigating through a virtual city: Using virtual reality technology to study human action and perception. *Future Generations Computer Systems*, **14**, 231–242.

Wang, R. X. F. & Simons, D. J. (1999). Active and passive scene recognition across views. *Cognition*, **70**(2), 191 – 210.

Wartenberg, F., May, M., & Péruch, P. (1998). Spatial Orientation in Virtual Environments: Background Considerations and Experiments. In C. Freska, C. Habel, & K. F. Wender (Eds.), *Spatial Cognition: an interdisciplinary appraoch to representing and processing spatial knowledge*, Vol. 1404 of *Lecture notes in computer science: Lecture notes in artificial intelligence* (pp. 469–489). Berlin Heidelberg: Springer.