Can People Not Tell Left from Right in VR? Point-to-origin Studies Revealed Qualitative Errors in Visual Path Integration

Bernhard E. Riecke* Jan M. Wiener†

Max Planck Institute for Biological Cybernetics, Tübingen, Germany

ABSTRACT

Even in state-of-the-art virtual reality (VR) setups, participants often feel lost when navigating through virtual environments. In psychological experiments, such disorientation is often compensated for by extensive training. The current study investigated participants' sense of direction by means of a rapid point-to-origin task without any training or performance feedback. This allowed us to study participants' intuitive spatial orientation in VR while minimizing the influence of higher cognitive abilities and compensatory strategies. After visually displayed passive excursions along one- or two-segment trajectories, participants were asked to point back to the origin of locomotion "as accurately and quickly as possible". Despite using a high-quality video projection with a 84° × 63° field of view, participants' overall performance was rather poor. Moreover, six of the 16 participants exhibited striking qualitative errors, i.e., consistent left-right confusions that have not been observed in comparable real world experiments. Taken together, this study suggests that even an immersive high-quality video projection system is not necessarily sufficient for enabling natural spatial orientation in VR. We propose that a rapid point-to-origin paradigm can be a useful tool for evaluating and improving the effectiveness of VR setups in terms of enabling natural and unencumbered spatial orientation and performance.

Keywords: ego-motion simulation, human factors, navigation, point-to-origin, psychophysics, spatial orientation, spatial updating, triangle completion, Virtual Reality.

Index Terms: H.1.2 [Models and Principles]: User/Machine Systems—Human factors, Human information processing; H.5.1 [Information Interfaces and Presentation, (e.g. HCI]: Multimedia Information Systems—Artificial, augmented, and virtual realities; J.4 [Social and Behavioral Sciences]: Psychology

1 INTRODUCTION

Most modern Virtual Reality (VR) simulators suffer from a grave malady: Severe disorientation [5, 8, 17, 27]. This strong tendency to easily get lost when navigating in VR can be overcome if people (a) are allowed to physically perform the simulated actions (e.g., though physical walking or at least turning [4, 13, 11, 10, 37], (b) are provided with useful visual landmarks or a well-known visual scene [4, 10, 24, 25], and/or (c) are given sufficient time to employ higher cognitive processes like mental spatial reasoning and/or receive extensive feedback training on the task [7, 12, 24, 35].

This stands in striking contrast to the real world, where spatial orientation and spatial updating typically operate automatically and effortlessly, requiring few if any cognitive resources [6, 19, 26]. Thus, most VR simulation paradigms do not empower people to use their "normal", evolutionary-developed spatial orientation abilities. Instead, VR users often seem to resort to cognitively more demanding and computationally more expensive strategies. This might be related to the lack of robust and effortless spatial updating observed in many VR situations.

In order to determine what critical aspects of the real world are not being captured in modern VR systems, we developed an experimental paradigm that mitigates the influence of higher cognitive abilities and strategies. There are two main elements to the experimental paradigm. First, a simple and ecologically plausible task is used – **rapid pointing to the origin of locomotion** after visually displayed passive excursions consisting of a linear translation, a subsequent rotation, and, in some cases, a second linear translation. In a way, one could picture this task as providing the indication of a "homing vector" that points from the current position and orientation back to the starting position [13, 11]. When performed in the real world using physical walking, pointing back to the origin of travel after one- or two-segment excursions is usually perceived as quite easy and not requiring much cognitive effort or computationally demanding strategies, even when performed with limited or no visual cues [10, 28, 30]. Using a *rapid* pointing paradigm has the strong advantage that it neither provides the time nor the feedback necessary to develop or use higher cognitive abilities (e.g., spatial reasoning) or strategies [25]. It is important to note that participants in the present study never received any performance feedback. Second, by presenting only optic flow information using a uniformly textured ground plane, **visual landmarks and other navigation aids are eliminated** from the virtual environment, further restricting the possible influence of high-level strategies.

Rapid pointing after simple excursion paths is quite trivial to perform in the real world, even when all visual and auditory spatial cues and landmarks are excluded (e.g., using blindfolds and headphones displaying broad-band noise). Due to an "automatic spatial updating" of our egocentric mental spatial representation of our immediate surroundings while walking, we maintain a natural and intuitive knowledge of where we are with respect to the environment during shorter periods of travel [6, 19, 26]. When visual and auditory cues are excluded, vestibular, proprioceptive, and kinaesthetic cues are still sufficient for enabling automatic spatial updating. We may not be perfectly accurate and precise due to accumulating path integration errors during the locomotion, but the task is relatively easy to perform in the sense that it does not require noticeable cognitive effort – we just seem to automatically "know" where we are with respect to immediate objects of interest. This is typically reflected in the subjective ease of performing the task, a minimal cognitive load, a lack of qualitative errors like left/right reversals, and rather short overall response times (typically below 2s) with little or no dependence on the angle turned or distance traveled [6, 26].

When comparable tasks are performed in a virtual environment where only path-integration based visual cues (optic flow) are provided and participants are not allowed to physically move, overall response errors increase and participants typically think more before responding [4, 10, 7, 18]. For simple spatial orientation tasks like triangle completion or estimation of turning angles, both systematic and variable errors seem to depend considerably on the display device used, with head-mounted displays and flat projection screens yielding the largest systematic and random errors, and large, curved projection screens yielding the lowest errors

---

*e-mail: bernhard.riecke@tuebingen.mpg.de

†e-mail: jan.wiener@college-de-france.fr

IEEE Virtual Reality Conference 2007
March 10 - 14, Charlotte, North Carolina, USA
1-4244-0906-3/07/$20.00 ©2007 IEEE

[9, 18, 23, 29]. It seems, though, that some kind of feedback training is often critical for enabling acceptable performance in VR, even for spatial orientation tasks as simple as pointing to the origin of locomotion after short excursions. In the following, we will discuss three relevant VR-based point-to-origin studies in more detail.

Lawton and Morrin displayed simple computer-simulated rectangular mazes on a desktop monitor and asked participants to point back to the origin of travel after excursions of 3, 5, or 7 segments using a compass-like pointer [12]. Despite the simple geometry of the maze and path layout (constant segment lengths with 90° in-between turns), pointing performance showed considerable errors even for the simplest condition: Mean absolute pointing errors averaged around 40° for men and 60° for women and increased for increasing number of path segments. Participants who maintained some kind of "feeling" of the relative direction of the origin (similar to a homing vector) performed significantly better than those who did not. Conversely, remembering the sequence of left and right turns had detrimental effects on pointing accuracy. When participants were repeatedly asked to indicate the homing vector during the excursion, pointing errors decreased by about 10° for both men and women. Providing pointing feedback only at the end of the excursion did not, however, improve pointing performance significantly. The data suggests that *continuously* maintaining a representation of the direction toward the origin of travel (similar to a homing vector) was critical for good pointing performance at the end of the trajectory.

In a recent point-to-origin task performed in desktop VR [7], participants followed a visually displayed uniformly textured tunnel consisting of straight and curved segments and were asked at the end of the excursion to indicate the direction to the origin of travel (homing vector) by adjusting a simulated 3D arrow using mouse buttons. Participants were given repeated feedback about the correct pointing direction, which might have contributed to the relatively low absolute pointing errors (10°– 25°). Differences between initial and final heading never exceeded 60°, which largely reduces the range of sensible pointing directions and might also have contributed to the good overall performance. To obviate this limitation, the current experiment was designed to maximize the range of correct pointing directions to span the whole range from small angles (as low as ±10°) to large angles (±180°).

In order to investigate the influence of path complexity on visual path integration performance, Wiener and Mallot used a joystick-based point-to-origin paradigm in a simple virtual environment consisting of a uniformly textured ground plane presented on a flat back-projection screen (90° × 60° FOV) [35]. Given sufficient feedback during an initial training phase, participants were able to perform the purely visual point-to-origin tasks with reasonable accuracy (20°– 35° absolute pointing error), even when the excursion path included up to 4 turns (albeit always in the same direction). Performance was moreover independent of the number of turns. Response times were, however, always above two seconds, suggesting that the task was not perceived as simple. This was corroborated by subjective reports of participants and the amount of errors during the training phase. That is, instead of using quick and robust, automatic spatial updating as in the real world, participants apparently had to resort to different, computationally more demanding strategies.

The three above-mentioned VR-based studies all used extensive feedback training and unlimited response times. This allowed for fairly accurate pointing performance. In the present study, however, we aimed at investigating how well participants perform when they are *never* provided with any performance feedback and are asked to respond as "accurately and quickly as possible" – factors that are critical for the overall acceptance and usability of VR.

2 METHODS

Sixteen naive participants (gender-balanced, with a mean age of 23.75 years) completed the experiment. Participation was voluntary and payed at standard rates. All participants had normal or corrected-to-normal vision.

2.1 Stimuli and apparatus

Participants were seated at a distance of 89cm from a flat projection screen (1.68m width × 1.26m height, corresponding to a field of view of about 84° × 63°), as illustrated in Figure 1. The virtual environment was quite simple and consisted of a textured flat ground plane that did not contain any absolute orientation or distance cues. The ground plane texture was selected to contain both a broad range of spatial frequencies and a high contrast in order to provide strong optic flow about the distance traveled and angles turned. Note, however, that the virtual scene did *not* contain any useful landmark information that participants could have used for determining their position or orientation relative to the origin of locomotion. Visual stimuli were projected non-stereoscopically using a JVC D-ILA DLA-SX21S video projector with a resolution of 1400 × 1050 pixels.

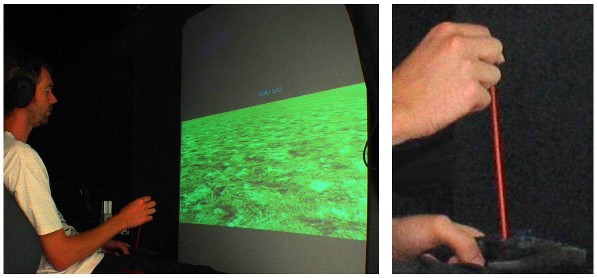

Figure 1: VR system showing a participant with the pointing device (modified game-pad) seated in front of the projection screen displaying the textured ground plane devoid of any landmarks.

In a **high immersion condition**, participants wore active noise canceling headphones (Sennheiser HMEC 300) playing several mixed layers of flowing water. In addition, the curtains on both sides of the projection screen were closed, such that participants could neither see nor hear the surrounding laboratory (see Fig. 1). In a second, **low immersion condition**, participants wore no headphones, and the curtains on both sides of the projection screen were opened, such that the surrounding lab was visible. Care was taken to adjust the light level in the lab to be similar to the curtains in the high immersion condition. We hypothesized that the high-immersion condition might help to reduce the conflict between the simulated virtual scene (depicting a simulated self-motion) and the real world (i.e., the VR setup and surrounding lab, which was stationary) [22], and thus indirectly facilitate spatial orientation relative to the virtual scene rather than the real world (see also [20]).

2.2 Procedure and experimental design

Each trial consisted of a passive motion phase, a pointing phase, and a fixed inter-trial interval. The motion phase consisted of a translation along a first segment s_1(8 m/s maximum translational velocity, with a brief acceleration and deceleration phase to avoid motion sickness), followed by a rotation (30°/s), and a subsequent translation along a second segment s_2 (same velocity as for s_1). For the one-segment experiments, the second translation was omitted. Upon arriving at the end of the trajectory, participants were asked to point "as accurately and quickly as possible" to the origin of locomotion as if they had physically moved (pointing phase). The inter-trial interval consisted of a 3s period where the screen was blanked, and a 2s interval where participants were instructed to prepare for the next trial. The turning direction was alternated between trials to reduce the occurrence of potential motion aftereffects and motion sickness, but was not analyzed separately.

Pointing was performed using a modified game pad as an input device (see Figure 1). The pointing device consisted of a 18cm long thin (2mm) plastic rod that replaced the knob on a standard game pad. Participants were instructed to hold the top end of the pointer with their index finger and thumb of their preferred hand, and the bottom end of the pointer with the index finger and thumb of the opposite hand as indicated above. Pre-tests had shown that this allows for more accurate pointing than simply using a joystick (which is often used in pointing studies), most likely because ones uses both hands and a precision grip on a simple, straight rod that is rotationally symmetric. Participants indicated that the pointing device was easy and intuitive to use. Note that participants were never provided with any performance feedback throughout the experiment to mitigate the usage of cognitive strategies or re-calibration. Furthermore, participants were asked to point "as accurately and quickly as possible" to reduce the likelihood of them building up any abstract geometric representations – like, for example, a top-down view of the path geometry, as was observed in experiments were participants were given unlimited response time [24]. This was important for the purpose of the experiment, as we were interested in testing if participants were able to orient themselves naturally, i.e., quickly and intuitively (most likely through automatic spatial updating), without any need for feedback training and/or computationally expensive processing. Previous studies had shown that participants can indeed perform triangle completion and point-to-origin tasks in VR relatively well if given unlimited response time and sufficient feedback training [7, 12, 35]. The experiment consisted of four parts: A practice phase, followed by a two-segment familiarization experiment, a two-segment main experiment, and a final one-segment experiment[1].

Real-world practice phase Before the main experiment, participants received written and oral instructions and were given a demonstration by the experimenter. Furthermore, participants were asked to walk physically with eyes closed along at least five 2-segment paths in the actual lab and use the pointing device (which was for that purpose detached from the computer) to point back to the origin of locomotion. Pointing back to the origin of locomotion after a 2-segment real-world excursion was perceived as rather trivial, but served well to familiarize participants with the experimental task and pointing device without providing them with any specific feedback that could be used in the actual experiment. In fact, none of the participant showed any problems or qualitative errors (like left/right confusions) in the practice phase from the very beginning, and quantitative errors were minimal. Once participants indicated that they did not need any more practice trials and clearly understood the instructions, experimental procedures, and task requirements for the VR test, they proceeded with the familiarization experiment. For all VR conditions, participants were instructed to treat the visual motion simulation as if it originated from an actual self-motion, and to respond as if they had actually moved (just like in the real-world practice phase).

Two-segment familiarization experiment In order to reduce the impact of learning effects on the main experiment, all participants first performed a familiarization experiment. The familiarization experiment consisted of 22 trials, consisting of a factorial combination of 2 lengths of s_1(16m, 24m) × 5 turning angles γ (45°, 75°, 105°, 135°, 165°) × 2 turning directions (left, right; alternating), plus 4 additional baseline trials without any rotation ($\gamma = 0°$, 2 lengths of s_1(16m, 24m) × 2 repetitions). The turning angles were selected to be different from those used in the main experiment in order to avoid potential direct leaning transfer or memorization of turning angles. For each participant, the immersion condition for the familiarization experiment matched that of the first session of the main experiment.

Two-segment main experiment After completing the practice phase and familiarization experiment, participants performed a two-segment main experiment which was split into two sessions ("high immersion" and "low immersion" condition) in balanced order. Each of the two sessions of the main experiment was composed of 52 trials, consisting of a factorial combination of 2 lengths of s_1 (16m, 24m; randomized) ×6 turning angles γ (30°, 60°, 90°, 120°, 150°, 170°; randomized)×2 turning directions (left, right; alternating) ×2 repetitions per condition, plus 4 baseline trials (randomly interspersed) without any turns between the two segments (2 lengths of s_1 (16m, 24m; randomized) ×2 repetitions for $\gamma = 0°$)).

One-segment experiment A large amount of variation across different studies seems to be caused by problems in perceiving and encoding visually simulated turns [10, 4, 24, 23]. This naturally raises the question as to whether some of the errors observed in the main experiment above are caused by problems in veridically perceiving and encoding the visually presented turning angles. To control for this possibility, all participants performed a subsequent one-segment experiment. The task was simply to point back to the origin of locomotion after being presented with a visually simulated passive forward translation ($s_1 = 16m$) followed by a passive rotation with angle γ, but no additional second translation. As in the main experiment, each participant performed two sessions ("high immersion" and "low immersion") in balanced order (same order before). The one-segment experiment consisted of 28 trials per session: a factorial combination of 6 turning angles γ (30°, 60°, 90°, 120°, 150°, 170°; randomized)×2 turning directions (left, right; alternating) ×2 repetitions per condition, plus 4 baseline trials (randomly interspersed) without any turn after the translation (4 repetitions for $\gamma = 0°$).

One-segment encoding control experiment In order to rule out the possibility that the new pointing device induced a systematic measurement error and to further investigate the potential influence of misperceiving the visually displayed rotations, we ran a new set of seven naive but psychophysically experienced observers (lab members, all male) in a modified version of the one-segment experiment (low immersion condition). Unlike in the previous experiment, participants were now given explicit advance information about the upcoming turn. That is, participants were told verbally about the exact turning angle and turning direction (e.g., "120° left") prior to the onset of each trial (and thus, in principle, had all the information they needed to determine the location of the origin). This procedure should essentially eliminate all errors from the *encoding phase* (building up an internal representation of the angle turned and trajectory traveled), such that all remaining errors should stem from problems with determining the proper response (*mental spatial reasoning phase*) and/or problems in actually performing the intended pointing response (*execution phase*). See [24] for a discussion of these three different phases in the context of a triangle completion task in VR.

Dependent variables Apart from a graphical analysis, pointing performance was quantified in terms of response time and mean angular deviation: The **response time** was calculated as the time until the pointer was deflected by 95%. The **mean angular deviation** can be conceived as the circular statistics counterpart of the standard deviation and is a measure of the variability or consistency of the pointing data per participant and condition [3].

3 RESULTS AND DISCUSSION

The pointing data were quantified using repeated measures ANOVAs for both dependent measures and the factors immersion

---

[1] A subset of the data has previously been presented as a conference abstract (poster) [reference omitted to keep manuscript anonymous]

condition, length of the first segment s_1, and turning angle γ. The no-turn condition ($\gamma = 0°$) was excluded from the ANOVA as it was intended as a baseline condition. Surprisingly, the immersion condition did not show any significant main effects at all for any of the dependent measures. Thus, it seems as if immersion either did not play an important role for the point-to-origin task used, or the manipulation was too subtle to be effective. For the further analysis, the data was pooled over the two immersion conditions and the turning directions (which were not the focus of the current study). The pooled data are summarized in Figure 2, 3, 4, 5, and 7.

3.1 Response times

Mean response times in the familiarization experiment were already quite short (1.31s, with 1.46s standard deviation (SD)) and decreased further in the main experiment (0.97s on average, 0.50s SD, see Fig. 7). Response times in both two-segment experiments did not show any significant relation to the turning angle. Unexpectedly, response times in the one-segment experiment were 1.37s on average (SD: 0.80s) and thus noticeably *higher* than in the main experiment. As more processing time was apparently needed directly after a rotation, one might argue that participants might have perceived the rotations as more difficult to update and needed more processing time for rotations than translations. This bears some resemblance to spatial updating studies where rotations are typically found to be harder to imagine than translations [26, 14, 19]. Interestingly, the one-segment experiment revealed a marginally significant tendency toward decreasing response times for larger turning angles (F(5, 75) = 2.25, p = 0.058), as if turns became easier to update with increasing turning angles.

3.2 Pointing errors

As can be seen in Figures 2, 3, 4, and 5, the pointing data were rather noisy and showed considerable variability both within-subjects and between-subjects. The pointing data for the smaller turning angles show an interesting feature that is to the best of our knowledge not known from blindfolded walking studies (see, however, discussion in subsection 4.1): The pointing responses for 30° rotations (see Fig. 2, left subplot) show for example that ten of the 16 participants pointed leftward, which is at least roughly the direction toward the origin, whereas the pointing directions for the other six participants (right subplot) seem to be mirrored with respect to the current observer orientation in the virtual scene. Careful analysis of all the experimental conditions revealed that the participant population clustered indeed into two distinct groups that exhibited *qualitatively different* overall pointing behavior: For turns to the left, the proper pointing direction is always to the left and vice versa. Ten of the 16 participants pointed indeed consistently in the correct overall direction, that is, leftwards for left turns and rightwards for right turns (at least for turning angles $\gamma \leq 90°$). The other six participants pointed, however, consistently into the *wrong* direction (see Fig. 2-4). That is, when the excursion path contained a counterclockwise (left) turn, they pointed consistently to the right instead of to the left and vice versa, even though left turns should always result in leftwards pointings for turning angles <180°[2]. This group of participants will in the following be termed “left-right inverters”. The left-right errors are most clearly visible for smaller turning angles. For larger turns, pointing directions are more noisy and left-right side errors might be confounded with the large misestimations of the actual turning angle indicated in Figure 5. If the presented turning angles are overestimated, a 150° left turn might, for example, be perceived as a 200° left turn, and the resulting pointing direction would then be rightwards (consistent with a 200° turn, as for participant ID 13) and not leftwards (consistent with a 150° turn). Due to the observed qualitative pointing errors (left-right inversions), we refrained from analyzing the signed or absolute pointing errors and analyzed only the response time and mean angular deviation, which are both unaffected by the left-right inversion.

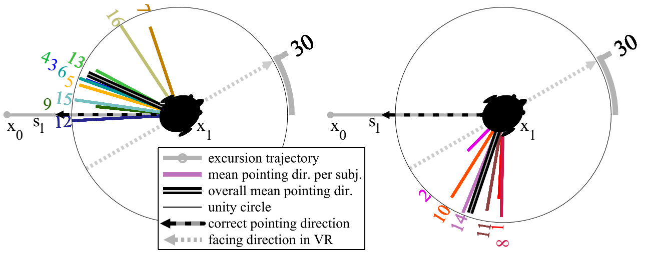

Figure 2: Sample data for the 30° condition of the one-segment experiment, split up into the six participants who showed systematic left-right errors (right subplot) and the remaining ten participants who did not show such systematic left-right errors (left subplot). Plotted is a top-down schematic view of the excursion path (in solid gray) from the start point x_0 to the endpoint x_1 and the subsequent turn by 30°. The mean pointing direction of each participant is indicated by the colored bars and subject IDs. The length of the mean pointing vector indicates the consistency of the individual pointing directions: Shorter mean pointing vectors indicate higher angular deviations between the individual pointing (e.g., participant 2), whereas mean pointing vectors close to the surrounding black unity circle indicate high consistency and low angular deviations between the individual pointings (e.g., participant 14 and 16) [3].

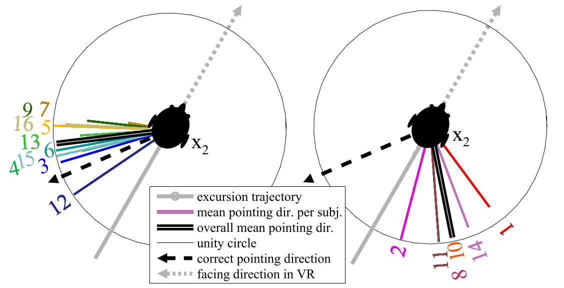

Figure 3: Sample data for the 60° condition of the two-segment main experiment illustrating the systematic left-right errors for six of the 16 participants (right subplot). The data is plotted as in Figure 2 and represents a close-up of the endpoint of the 60° trajectories (cf. Fig. 4).

Mean angular deviation The mean angular deviation showed a clear influence of the turning angle γ ($F(5,75) = 6.88, p < .0005^{***}$ for the two-segment main experiment and $F(5,75) = 3.90, p = .003^{**}$ for the one-segment experiment). As can be seen in Figure 7, the mean angular deviation was lowest in the no-turn condition and 30° condition and gradually increased with increasing turning angles. This implies that participants were less consistent (more variable) in their pointing responses for increasing turning angles. This might be related to the increased task difficulty for larger turning angles and/or increasing uncertainties in the estimation of the turning angles.

Correlation between left-right inversion and post-experimental data Even though the reasons underlying the observed left-right inversions are not fully understood yet (see subsection 4.1), it is interesting to note that left-right inversion was associated with lower spatial abilities as measured using

---

[2] Across the three experiments, there were 18 conditions that contained turns $\gamma \leq 120°$. Participants were categorized as left-right inverters if their mean pointing directions were left-right inverted for at least 15 of those 18 conditions. Three participants showed left-right inversions in all 18 conditions (participant ID 1, 8, & 11), two participants in 17 of the 18 conditions (participant ID 10 & 14), and one participant in 16 of the 18 conditions (participant ID 2).

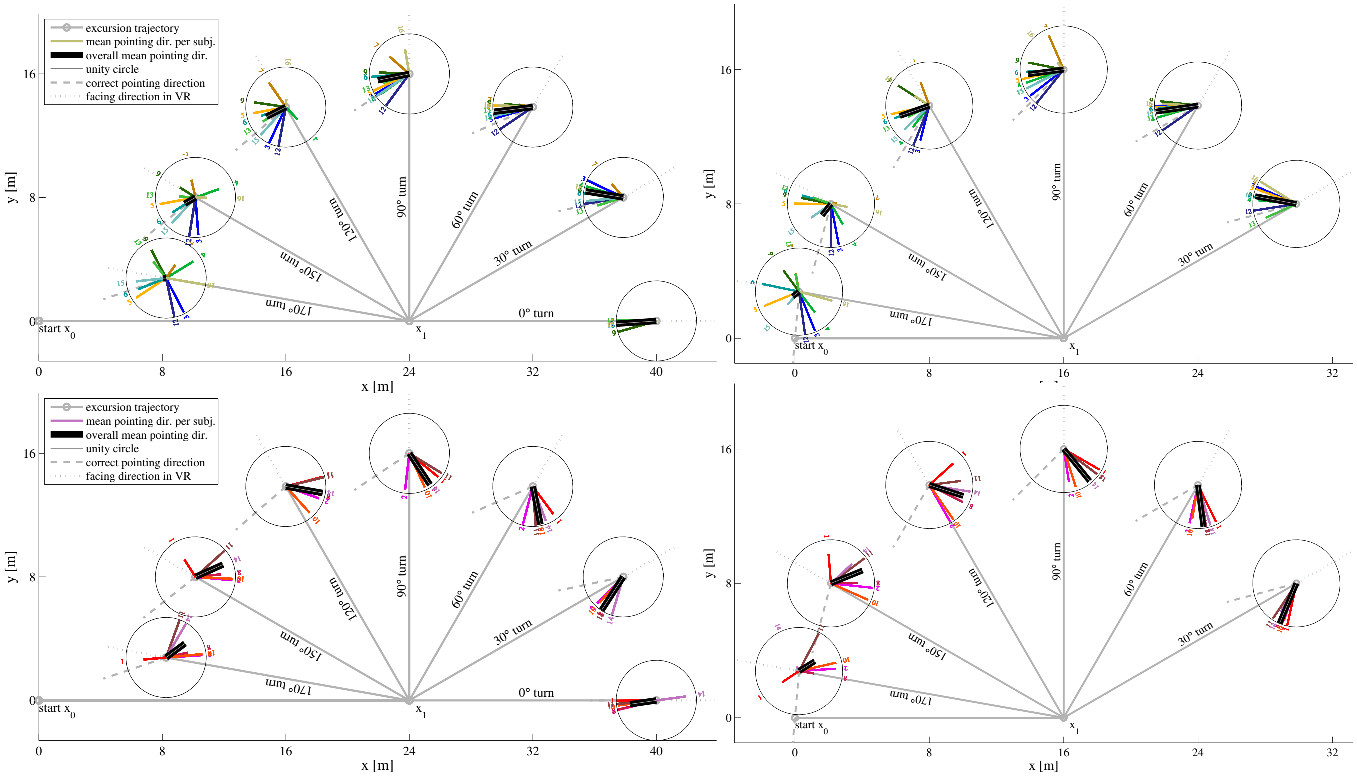

Figure 4: Mean pointing directions for the different turning angles γ and lengths of $s_1$of the 2-segment main experiment, plotted as in Figure 3. The bottom subplots represent data from the six left-right inverters (depicted in reddish colors), the top subplots shows data from the ten non-inverters.

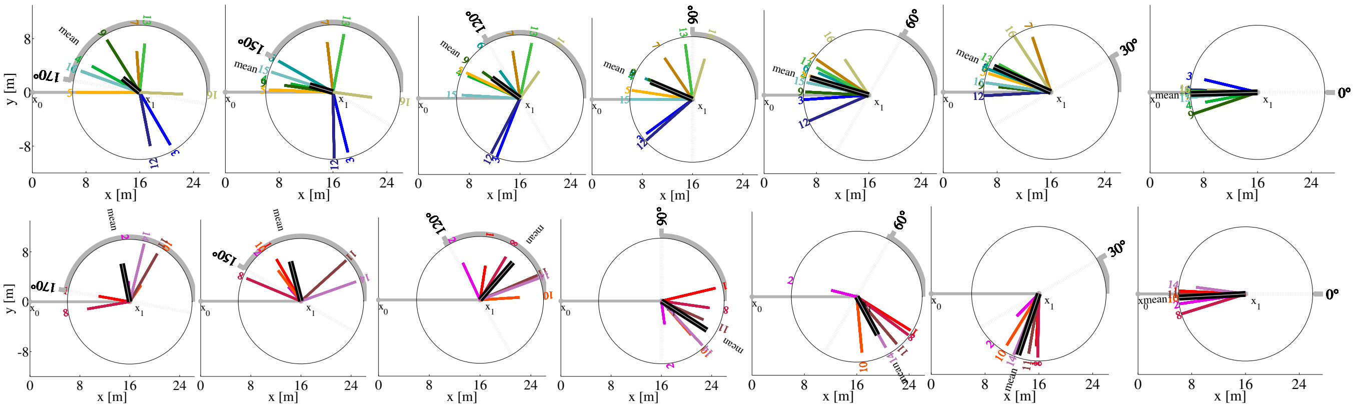

Figure 5: Mean pointing directions for the different turning angles γ of the one-segment experiment, plotted as in Figure 2 and separated into inverters (bottom) an non-inverters (top). Note the increasing absolute pointing errors and within- and between-subject pointing variability for increasing turning angles.

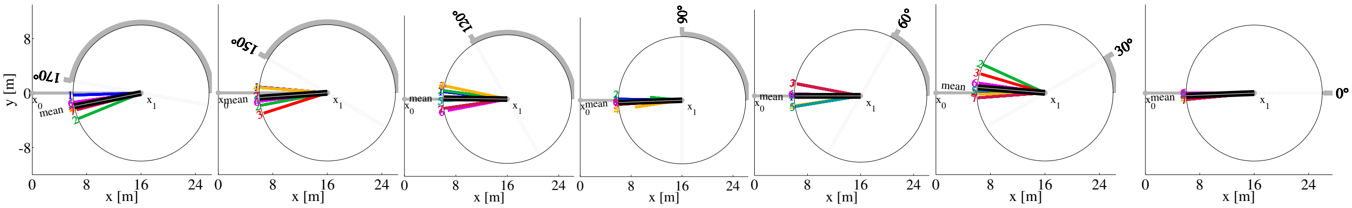

Figure 6: Mean pointing directions for the one-segment encoding control experiment, plotted as in Figure 5. Note the high pointing accuracy (low systematic pointing errors) and precision (low within- and between-subject variability, indicated by the mean pointing vector length close to 1).

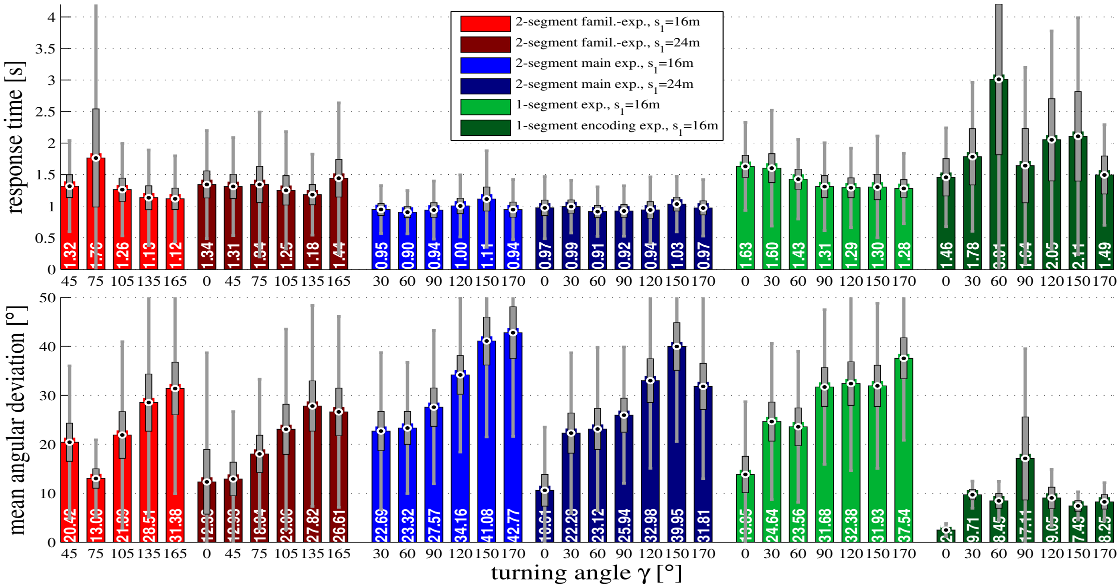

Figure 7: Summary of the arithmetic means of the response time (top) and mean angular deviation (bottom). Boxes and whiskers indicate one standard error of the mean and one standard deviation, respectively.

a standard paper-and-pencil test [31] (8.7 points for the inverters vs. 15.2 for the non-inverters; $t(14) = 4.72, p < .0005^{***}$. Furthermore, it turned out that five of the six left-right inverters were female, whereas the non-inverters were predominately male (7/10, $\chi^2(1, N = 16) = 4.27, p = .039^{*}$). General computer usage did not show a significant difference between the two groups (1.52 vs. 2.36 hours/day on average for the inverters and non-inverters, respectively; $t(14) = 1.14, p = .22$). 3D computer game experience showed, however, a significant effect: None of the inverters had any 3D computer game experience, whereas six of the ten non-inverters did ($\chi^2(1, N = 16) = 5.76, p = .016^{*}$).

3.3 One-segment encoding control experiment

Data from the one-segment encoding control experiment are summarized in Figures 6 and 7. Compared to the previous one-segment experiment, the encoding experiment where participants had advance knowledge about the upcoming turning angle showed a considerable decrease in pointing errors and pointing variability (both within- and between-subjects). In fact, systematic pointing errors were minimal. This suggests that the large pointing errors found in the previous experiments might be largely caused by an encoding error or a mental computation error: that is, participants might have misperceived the turning angle, and/or might have been unable to infer or mentally compute the proper response from the given information. Note, however, that differences in the participant population might also have contributed. Overall response times were somewhat higher than in the previous experiments, which suggests that participants needed more cognitive resources to perform the task. This is consistent with participants stating that they perceived the task as “unexpectedly extremely difficult”.

The high overall pointing accuracy in the encoding control experiment suggests that the pointing device and procedure induced little systematic execution errors. In fact, despite using a rather simple pointing device, response times, pointing accuracy, and mean angular deviations in the encoding control experiment were roughly comparable to previous spatial updating experiments that used a similar rapid pointing paradigm but a technically more advanced, two-handed, position-tracked pointing wand [25, 22]. Furthermore, all participants of the current study reported that pointing device was easy and intuitive to use. Taken together, this suggests that the pointing methodology was appropriate and did not contribute to the qualitative errors overall and the observed difficulty of the point-to-origin task.

4 GENERAL DISCUSSION AND CONCLUSIONS

4.1 Origin of left-right inversions

The consistent left-right inversions observed for six of the 16 participants in the present study bears resemblance to differences in pointing strategies observed by Gramann et al. using point-to-origin experiments in desktop VR [7]: In their study, only 20 of the 43 participants (accumulated over three experiments) updated their heading according to the visual turns (so-called “turners”). The other 23 participants responded as if still facing the original direction (north) (“non-turners”). These left-right inversions observed in VR studies resemble real-world data from imagined walking experiments using two-segment excursion [2, 10]. Instead of using a pointing method, participants in these real world studies were asked to turn to face the origin as if they had actually walked the excursion trajectory. Participants failed to update heading changes during the imagined rotation between the first and second segment, and responded as if standing at the to-be-imagined location, but still facing the initial orientation (north). Similar failures to update rotations were observed by Klatzky and colleagues when only optic flow information presented on a HMD indicated the trajectory or when participants watched another person walk the excursion path [10]. Only when participants actually walked the path or at least physically executed the turn between the two segments did they properly incorporate the rotation, which corroborates the often posited importance of physical motion cues for automatic spatial updating [6, 16, 19, 26, 37].

Gramann et al. [7] argued that the turners in their study used an egocentric strategy, whereas the non-turners used an allocentric strategy. Following this line of reasoning, one might be tempted to conclude that the six participants who consistently produced left-right inversion errors in the current experiment were “non-turners” and hence did not update their facing direction according to the visual stimulus at all. The other twelve participants would then be categorized as turners. To give an example, if participants are presented with a two-segment isosceles path containing a 60° left turn, the turners would point 150° to the *left*, whereas the non-turners would be expected to point 150° to the *right* (see Fig. 4). As we will argue below, this interpretation is consistent with most, but not all, of the data from the two-segment experiments of the present study: For “non-turners”, the origin of locomotion is always *behind* them (unless $s_2 > s_1$, which was never the case for the current studies). Hence, they should always point backward (i.e, egocentric pointing angles $\alpha > 90°$), and not forward ($\alpha < 90°$). One of the left-right inverting participants in the current study did, however, point consistently forward for the 150° and 170° turns (see Fig. 4, participant ID 1). This suggests that the left-right inversion observed in the current study cannot be fully explained by a simple failure to update rotations that are not physically performed.

We are currently running a control experiment similar to the encoding control experiment, but using psychophysically inexperienced, naive participants and advance information about the turning angle but not direction. When debriefed, all those participants could correctly name the simulated turning directions for a given trajectory, even though some of them showed left-right inversions in their pointing responses. This suggests that left-right inversion is not just caused by a misperception of one’s turning direction indicated by the visual stimulus. Moreover, several of those left-right inverters pointed consistently forward ($\alpha < 90°$) for turns $\geq 150°$, which corroborates our earlier argument that simple failure to update a visually presented turn might not be sufficient to explain the current data.

In sum, we have so far admittedly no conclusive explanation for the origin of the observed consistent left-right inversions in VR. Failure to update rotations might explain some, but not all, of our data. We can only speculate that updating/computing a homing vector might be difficult or confusing in VR because of the lack of physical motion cues, the absence of landmarks, and/or the fact that the pointing target remains mostly outside of the field of view of the visual display. We are currently planning experiments to test these hypothesis. Furthermore, the observed qualitative errors might simply disappear if a verbal response is used instead of a bodily response like pointing, as was the case in the imagined walking study by Avraamides et al. [2]: According to a recently proposed “sensorimotor interference hypothesis”, the perceived discrepancy between the physical orientation (which remained unchanged) and the to-be-imagined or visually simulated orientation creates an interference at the response level if (and only if) a bodily response like pointing or turning one’s body is used [14, 15, 22, 33, 36].

Hence, VR users might well be able to update some kind of “cognitive heading” deliberately, which allows for verbal or other responses that are somewhat detached from one’s bodily reference frame. Whenever a more embodied response like pointing, grasping, or turning is required, however, VR systems that provide neither physical motion cues nor useful landmarks might be insufficient. This might severely limit the effectiveness and user acceptance of VR at large, in particular for tasks where robust and effortless (i.e., natural) spatial orientation is essential.

4.2 Automatic spatial updating versus effortful cognitive processing

The small overall response times and the lack of an increase in response time for increasing turning angles and number of path segments suggests that participant performed some mental spatial computations like updating a homing vector already *during* the simulated motion, and did not wait until the end of the trajectory was reached. The overall large mean angular deviations and the striking qualitative errors (consistent left-right inversions observed for six of the 16 participants) provide, however, strong evidence that spatial updating was by no means automatic and effortless, despite apparently occurring online during the simulated motion. This lack of automatic spatial updating was corroborated by participants’ subjective ratings of task difficulty: Only one participant rated the task as easy, whereas eight participants rated the task as medium difficult and seven as quite difficult. Some participants even mentioned explicitly that they were often unsure which direction (left or right) they had to point to – something that is not observed for real-world tasks, even with eyes closed. The fact that the left-right inverters also showed significantly lower mental spatial abilities further suggests that participants were not able to use automatic spatial updating, but instead had to resort to more cognitive strategies like abstract mental spatial reasoning. Even in the one segment encoding control experiment, where experienced psychophysical observers were provided with explicit verbal information about the turning angle and direction, the task was rated as “extremely difficult”.

4.3 Conclusions and guidelines for VR system designers

In conclusion, this study suggests that even an immersive, high-resolution video projection setup is not necessarily sufficient for enabling quick and intuitive spatial orientation and automatic spatial updating when only optic flow cues without any landmarks are available. The fact that observers in the current study did not actively control the visually displayed motions can presumably be excluded as a contributing factor, as previous studies demonstrated that actively executing a motion is not required for automatic spatial updating, at least for physical motions [34, 37]. It is interesting to note that immersion did not seem to play any significant role in the task used, even though one might imagine that a higher degree of immersion might be able to reduce the interference between the physical and visually simulated orientation and thus facilitate the updating of heading [22].

One issue that should be carefully considered when attempting to improve spatial orientation in VR is the type and properties of the visual display device. Even though the current study was not designed to investigate this issue, the literature suggests that not only the visual FOV, but also the display type and geometry are critical factors for improving the effectiveness of VR systems [1, 23, 29, 32]. Tan and colleagues demonstrated, for example, that using a physically larger display ($193 \times 145cm$ vs. $36 \times 26cm$) can enhance performance in a variety of spatial tasks, even though the physical FOV subtended by the two displays was identical ($31° \times 24°$) [32]. Compared to the larger display of their study, the current display was slightly smaller ($168 \times 126cm$), but provided a much larger FOV ($84° \times 63°$) and a higher resolution (1400×1050 pixel compared to 1024×768). A review of all the display factors affecting spatial orientation in VR would go beyond the scope of the current paper; as a rough guideline, though, human spatial orientation in VR typically benefits from a large FOV, and physically large projection screens (in particular if they are curved around the observer) tend to outperform flat or desktop displays and head-mounted displays (HMDs) [1, 5, 17, 24, 25, 23, 27, 29, 32].

A different approach to improve spatial orientation in VR is to train the users on the task. Recent studies showed that extensive feedback training can be employed to increase accuracy for a given task [7, 12, 24, 35]. The relatively long response times, and the high cognitive demand and rated task difficulty that are often observed in those studies suggest, however, that participants still do not have a robust and effortless spatial orientation comparable to real-world performance, despite the training. This is in agreement with spatial updating studies that found a lack of automatic spatial updating whenever simulated self-rotations where not physically performed [4, 10, 37] and/or only optic flow stimuli without useful landmarks were presented [10, 22, 37]. That is, as a general guideline, VR users should be allowed to physically locomote through the environment or at least physically rotate whenever feasible [4, 10, 37]. Razzaque and colleagues suggested that a limited walking area could still convey much larger virtual spaces if the virtual scene is automatically and imperceptibly rotated such that users are guided away from the physical walls of the lab [21].

Even if no physical motion cues are provided at all, visual cues alone *can*, under certain conditions, still be sufficient for enabling good spatial orientation and automatic spatial updating if a naturalistic, user-known visual stimulus that includes useful landmarks is used [24, 25, 22]. In general, however, it remains an open question and a considerable technological and scientific challenge to determine whether visual information devoid of any landmarks might in principle be capable for enabling robust and effortless, natural spatial orientation in simulated environments. If so, then this would be of substantial impact both for our understanding of human spatial orientation and for the design of human-centered, lean yet effective ego-motion simulators.

Using a rapid point-to-origin paradigm in VR proved to be a powerful method for investigation human spatial orientation under carefully controlled and repeatable stimulus conditions, and we propose that measuring the occurrence of qualitative errors might be a simple yet effective way of evaluating the perceptual quality of a given VR system and, in particular, its capability in enabling natural and unencumbered spatial behavior and performance.

ACKNOWLEDGEMENTS

This research was funded by the Max Planck Society and the EU grant POEMS-IST-2001-39223 (see www.poems-project.info)

REFERENCES

[1] K. W. Arthur. *Effects of Field of View on Performance with Head-Mounted Displays*. PhD thesis, Department of Computer Science, University of North Carolina, Chapel Hill, 2000.

[2] M. N. Avraamides, R. L. Klatzky, J. M. Loomis, and R. G. Golledge. Use of cognitive versus perceptual heading during imagined locomotion depends on the response mode. *Psychological Science*, 15(6):403–408, June 2004.

[3] E. Batschelet. *Circular statistics in biology*. Acad. Pr., London, 1981.

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

[5] R. P. Darken and B. Peterson. Spatial orientation, wayfinding and representation. In K. M. Stanney, editor, *Handbook of Virtual Environments*, chapter 24, pages 493–518. Lawrence Erlbaum, 2002.

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

[7] K. Gramann, H. J. Muller, E. M. Eick, and B. Schonebeck. Evidence of separable spatial representations in a virtual navigation task. *Journal of Experimental Psychology-Human Perception and Performance*, 31(6):1199–1223, Dec. 2005.

[8] S. C. Grant and L. E. Magee. Contributions of proprioception to navigation in virtual environments. *Human Factors*, 40(3):489–497, 1998.

[9] M. J. Kearns, W. H. Warren, A. P. Duchon, and M. J. Tarr. Path integration from optic flow and body senses in a homing task. *Perception*, 31(3):349–374, 2002.

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

[11] R. L. Klatzky, J. M. Loomis, and R. G. Golledge. Encoding spatial reresentations through nonvisually guided locomotion: Test of human path integration. In D. Medin, editor, *The psychology of learning and motivation*, volume 37, pages 41–84. Acad. Press, 1997.

[12] C. A. Lawton and K. A. Morrin. Gender differences in pointing accuracy in computer-simulated 3D mazes. *Sex Roles*, 40(1-2):73–92, 1999.

[13] J. M. Loomis, R. L. Klatzky, R. G. Golledge, and J. W. Philbeck. Human navigation by path integration. In R. G. Golledge, editor, *Wayfinding Behavior: Cognitive mapping and other spatial processes*, pages 125–151. Johns Hopkins, Baltimore, 1999.

[14] M. May. Cognitive and embodied modes of spatial imagery. *Psychologische Beiträge*, 38(3/4):418–434, 1996.

[15] M. May. Imaginal perspective switches in remembered environments: Transformation versus interference accounts. *Cognitive Psychology*, 48(2):163–206, Mar. 2004.

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

[17] P. Péruch and F. Gaunet. Virtual environments as a promising tool for investigating human spatial cognition. *Cah. Psychol. Cogn.-Curr. Psychol. Cogn.*, 17(4-5):881–899, aug-oct 1998.

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

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

[20] J. D. Prothero. *The Role of Rest Frames in Vection, Presence and Motion Sickness*. PhD thesis, University of Washington, 1998. Available: www.hitl.washington.edu/publications/r-98-11/.

[21] S. Razzaque, Z. Kohn, and M. C. Whitton. Redirected walking. In *Proceedings of Eurographics 2001*, pages 289–294. Eurographics Association, 2001.

[22] B. E. Riecke, D. W. Cunningham, and H. H. Bülthoff. Spatial updating in virtual reality: The sufficiency of visual information. *Psychological Research*, 2007. (in press).

[23] B. E. Riecke, J. Schulte-Pelkum, and H. H. Bülthoff. Perceiving simulated ego-motions in virtual reality - comparing large screen displays with HMDs. In *Proceedings of the SPIE*, volume 5666, pages 344–355, San Jose, CA, USA, 2005.

[24] B. E. Riecke, H. A. H. C. van Veen, and H. H. Bülthoff. Visual Homing Is Possible Without Landmarks: A Path Integration Study in Virtual Reality. *Presence - Teleoperators and Virtual Environments*, 11(5):443–473, 2002.

[25] B. E. Riecke, M. von der Heyde, and H. H. Bülthoff. Visual cues can be sufficient for triggering automatic, reflex-like spatial updating. *ACM Transactions on Applied Perception (TAP)*, 2(3):183–215, 2005.

[26] J. J. Rieser. Access to knowledge of spatial structure at novel points of observation. *J. Exp. Psychol.-Learn. Mem. Cogn.*, 15(6):1157–1165, Nov. 1989.

[27] R. A. Ruddle and D. M. Jones. Movement in cluttered virtual environments. *Presence - Teleoperators and Virtual Environments*, 10(5):511–524, Oct. 2001.

[28] E. K. Sadalla and D. R. Montello. Remembering changes in direction. *Environ. Behav.*, 21(3):346–363, May 1989.

[29] J. Schulte-Pelkum, B. E. Riecke, M. von der Heyde, and H. H. Bülthoff. Influence of display device and screen curvature on perceiving and controlling simulated ego-rotations from optic flow. Technical Report 122, MPI for Biological Cybernetics, 2004.

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

[31] H. Stumpf and E. Fay. *Schlauchfiguren - Ein Test zur Beurteilung des räumlichen Vorstellungsvermögens*. Hogrefe, Göttingen, 1983.

[32] D. S. Tan, D. Gergle, P. Scupelli, and R. Pausch. Physically large displays improve performance on spatial tasks. *ACM Trans. Comput.-Hum. Interact.*, 13(1):71–99, March 2006.

[33] R. F. Wang. Beyond imagination: Perspective change problems revisited. *Psicologica*, 26(1):25–38, 2005.

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

[35] J. M. Wiener and H. A. Mallot. Path complexity does not impair visual path integration. *Spatial Cognition and Computation*, 6(4):333–346, 2006.

[36] M. Wraga. Thinking outside the body: An advantage for spatial updating during imagined versus physical self-rotation. *Journal of Experimental Psychology-Learning Memory and Cognition*, 29(5):993–1005, Sept. 2003.

[37] M. Wraga, S. H. Creem-Regehr, and D. R. Proffitt. Spatial updating of virtual displays during self- and display rotation. *Memory & Cognition*, 32(3):399–415, Apr. 2004.