Visual Cues can be Sufficient for Triggering Automatic, Reflexlike Spatial Updating

BERNHARD E. RIECKE, MARKUS VON DER HEYDE and HEINRICH H. BÜLTHOFF
Max Planck Institute for Biological Cybernetics, Tübingen, Germany

"Spatial updating" refers to the process that automatically updates our egocentric mental representation of our immediate surround during self-motions, which is essential for quick and robust spatial orientation. To investigate the relative contribution of visual and vestibular cues to spatial updating, two experiments were performed in a high-end Virtual Reality system. Participants were seated on a motion platform and saw either the surrounding room or a photorealistic virtual model presented via head-mounted display or projection screen. After upright rotations, participants had to point "as accurately and quickly as possible" to previously learned targets that were outside of the current field of view (FOV). Spatial updating performance, quantified as response time, configuration error, and pointing error, was comparable in the real and virtual reality conditions when the FOV was matched. Two further results challenge the prevailing basic assumptions about spatial updating: First, automatic, reflexlike spatial updating occurred without any physical motion, i.e., visual information from a known scene alone *can*, indeed, be sufficient, especially for large FOVs. Second, continuous-motion information is not, in fact, mandatory for spatial updating—merely presenting static images of new orientations proved sufficient, which motivated our distinction between *continuous* and *instant-based* spatial updating.

Categories and Subject Descriptors: 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*; H.5.2 [**Information Interfaces and Presentation (e.g. HCI**]: User Interfaces—*Input devices and strategies, Interaction styles*; J.4 [**Social and Behavioral Sciences**]: Psychology

General Terms: Experimentation, Human Factors;

Additional Key Words and Phrases: Spatial updating, virtual reality, spatial orientation, ego-motion simulation, human factors, psychophysics, multimodal, visuovestibular cue integration

1. INTRODUCTION

Being mobile species, humans as well as most animals constantly change their position and orientation in space. Due to this ego-motion, the spatial relationships between the observer and the surrounding environment change constantly in a rather complex manner. As long as all relevant objects are directly and constantly visible that might not pose any problems. As soon as some objects are occluded by others, however, it would make sense to have a process that ensures that we still know where everything is even though we cannot currently perceive it directly. The situation becomes even worse if vision is completely excluded. Imagine for example that you are at home at night when the main fuse blows. You

Author's addresses: Bernhard E. Riecke, Max Planck Institute for Biological Cybernetics, Spemannstrasse 38, 72076 Tübingen, Germany; e-mail: bernhard.riecke@tuebingen.mpg.de; Markus von der Heyde, SCC, Bauhaus-Universität Weimar, Steubenstrasse 6a, R 302, 99425 Weimar, Germany; e-mail: markus.von.der.heyde@scc.uni-weimar.de; Heinrich H. Bülthoff, Max Planck Institute for Biological Cybernetics, Spemannstrasse 38, 72076 Tübingen, Germany; e-mail: heinrich.buelthoff@tuebingen.mpg.de;

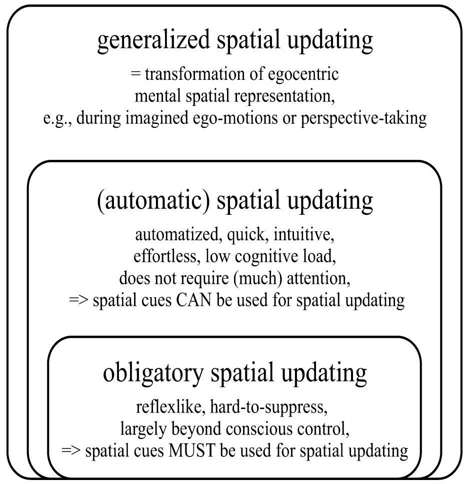


Fig. 1. At the most general level, **generalized spatial updating** refers to all spatial transformations of our egocentric mental spatial representation. This includes mental perspective-taking or consciously updating our egocentric representation during imagined ego-motions. **Automatic spatial updating**, which is often referred to as simply spatial updating, is a more specific subset and refers to the largely automatized transformations of our mental egocentric representation. This automaticity ensures that the attentional and cognitive demands are minimal, if not zero [e.g., Rieser 1989], such that spatial updating does not interfere with other cognitive or noncognitive tasks. **Obligatory spatial updating** is a subset of the more general (automatic) spatial updating. It refers to the reflexlike, hard-to-suppress, and, thus, cognitively almost impenetrable phenomenon of perceived spatial cues triggering spatial updating, whether we want to or not.

will have to find your way around in complete darkness until you find candles or the fuse box. Even though you know your home and its general spatial layout quite well, this knowledge won't help you much if you were not able to create a mental spatial representation of where everything is from your *current* position, and easily update it during all motions, even in darkness.

Having to consciously keep track of all potentially relevant objects in the surround is virtually impossible and would increase the cognitive load to a level that would not allow for any further complex task. As spatial orientation is a rather frequent and essential task, it would be quite inefficient to allocate general cognitive resources for this task. That is, we need a process that automatically keeps track of where relevant surrounding objects are while we locomote, without much cognitive effort or mental load.

This largely automatized and seemingly effortless process is typically referred to as "automatic spatial updating" or just "spatial updating" [Amorim and Stucchi 1997; Farrell and Robertson 1998; Farrell and Thomson 1998; Hollins and Kelley 1988; Klatzky et al. 1998; Loomis et al. 1996; Presson and Montello 1994; Rieser et al. 1982]. It is this process that allows us to locomote in darkness without much cognitive load or constantly bumping into obstacles, by providing quick and intuitive knowledge of where surrounding objects are, even during complex motions. Automatic spatial updating allows us to quickly and accurately grasp objects and point or look into their direction, even when they are not directly visible. Compared to other forms of mental transformation processes like mental perspective-taking, automatic spatial updating can be seen as a specific subset, referring to largely automatized transformation processes acting on our egocentric mental representation. This is illustrated in Figure 1.

As it might be rather hazardous if we somehow forgot to update our representation under situations of high stress or high cognitive load, it seems sensible to propose a spatial updating process that is always operative, irrespective of our focus of attention or any conscious decision. Moreover, as it does not seem to make any sense to *not* update our egocentric representation during ego-motions, this process should ideally be reflexlike and beyond conscious control.

Under full-cue conditions, spatial updating does indeed seem so tightly coupled to the motion cues that we cannot help but update the "world inside our head" when moving through its real counterpart [Farrell and Robertson 1998; May and Klatzky 2000]. Due to this spatial updating automatism, any perceptually signaled movement seems to be mandatorily incorporated into our representation of our current position and orientation and cannot simply be excluded by volition. Only with great effort can we cognitively compensate for this compulsory updating and try to ignore motions toward a different position and orientation in space and, e.g., imagine that we still are at the original location [Farrell and Robertson 1998; May and Klatzky 2000]. That is, if we perform an ego-turn, proprioceptive, vestibular, visual, and auditory cues somehow automatically initiate a corresponding counter-rotation of the world inside our head, whether we want them to or not. Hence, spatial updating can, under those conditions, be considered mandatory or obligatory in the sense of being hard-to-suppress and, thus, to a large degree, cognitively impenetrable. To reflect this reflexlike phenomenon, we introduced the term "**obligatory spatial updating**" as a specific subset of the more general "automatic spatial updating" (cf. Figure 1). It seems as if the "world inside our head" has some kind of inertia that makes it stay in alignment with the outside world and prevents it from moving with the head. It thus acts just like a gyrocompass, not only for rotations but also for translations.

Spatial updating performance can, however, be slightly impaired by high cognitive loads, such as counting backward in steps of seven or three or verbalizing nonsense syllables [Yardley and Higgins 1998; May and Klatzky 2000]. However, even under those conditions, spatial updating still seems to be obligatory in the sense that the mental spatial representation is continuously being updated according to the ego-motions, even though it might not stay perfectly aligned due to accumulating errors induced by the high cognitive load. To make this distinction, we prefer to use the term "obligatory spatial updating" instead of the more general "automatic spatial updating" whenever referring explicitly to that reflexiveness, even though this distinction has, to our knowledge, never been made in the literature.

Rotations are typically as easy to update as translations, but considerably harder to ignore or imagine [Easton and Sholl 1995; Farrell and Robertson 1998; Klatzky et al. 1998; May and Klatzky 2000; May 1996, 2004; Presson and Montello 1994; Rieser 1989]. This was one reason for us to focus solely on rotations in this paper. When asked to ignore or imagine a self-rotation, participants' errors and response latencies typically increase with turning angle, suggesting that participants perform some kind of mental ego-rotation, with a limited rotational velocity [Farrell and Robertson 1998; May 2004; Rieser 1989].

Vestibular and kinesthetic motion cues proved to be sufficient to trigger automatic and obligatory spatial updating during rotations as well as translations in blindfolded participants [Easton and Sholl 1995; Farrell and Robertson 1998; May and Klatzky 2000]: Participants had no problem updating motions and respond according to the new location. However, instructions to ignore a motion and respond as if still being in the previous location resulted in considerable errors, indicating the difficulty of consciously influencing or suppressing spatial updating. Those errors were much greater than those induced by verbal distractions, indicating the inability to ignore physical movements during path integration. This suggests that kinesthetic and vestibular cues from blind locomotion are sufficient to trigger obligatory spatial updating.

Conversely, spatial updating is typically impaired when proprioceptive and vestibular cues in particular are missing [Klatzky et al. 1998; May and Klatzky 2000; Presson and Montello 1994; Rieser 1989; Simons and Wang 1998; Wang and Simons 1999; Wang and Spelke 2000; Wraga et al. 2004]. Qualitative errors seem to occur most often when kinesthetic and/or vestibular cues about ego-turns are missing. Klatzky et al. [1998] and May and Klatzky [2000], for example, found that participants completely forgot to update ego-rotations that were not physically performed, i.e., when the corresponding vestibular and proprioceptive cues were missing. Such results lead to the prevailing opinion that vestibular and proprioceptive cues are absolutely required for triggering the spatial updating automatism.

In this study, we questioned this notion by performing two spatial updating experiments where different combinations of visual and vestibular cues were compared. In order to be able to independently control vestibular and visual cues, a Virtual Reality (VR) setup was used, including a motion simulator (6 degree-of-freedom motion platform) for passively moving participants, and a head mounted display (HMD; both experiments) as well as a projection system (control experiment) for displaying visual stimuli. To obtain a baseline performance of "optimal" spatial updating, the main experiment compared VR performance with real-world performance in the corresponding real environment.

If visual cues alone would prove to be inferior to combined visuovestibular cues, all VR setups that rely heavily on visual cues for simulating ego-motions, while omitting vestibular cues (i.e., most of the existing and affordable VR setups), might face the same problem: Namely, that they do not allow for "normal" and effortless navigation, as they do not sufficiently enable spatial updating. There is, indeed, a number of studies showing that spatial orientation abilities largely deteriorate when nonvisual sensory modalities are excluded, reduced, or only insufficiently simulated [Chance et al. 1998; Bakker et al. 1999; May and Klatzky 2000; Péruch and Gaunet 1998; Sholl 1989; Simons and Wang 1998; Wang and Simons 1999; Wraga et al. 2004]. We suspected, however, that this apparent insufficiency of visual cues might, in fact, be largely due to insufficiencies of the visual simulation, namely, a lack of a naturalistic scene presented with sufficient detail, resolution, and field of view. This motivated us to use detailed photorealistic replica of real scenes in the current study.

2. MAIN EXPERIMENT

2.1 Methods

Twelve naive participants completed the main experiment, with ages ranging from 19 to 33 years (mean: 26.3). For both experiments presented in this paper, participants had normal or corrected-to-normal vision and no signs of vestibular dysfunction. Participation was voluntary and paid at standard rates.

2.1.1 *Stimuli and Apparatus*

2.1.1.1 *Scenery and Visualization.* The pointing stimuli consisted of twelve target objects (numbers 1–12, arranged in a clock-face manner) attached to the walls of the Motion-Lab at eye height (see Figure 2). Participants saw either the real room or a photorealistic virtual replica of it presented nonstereoscopically through a position-tracked head-mounted display [HMD Kaiser ProView XL50; see Figure 2(b)]. The HMD had a resolution of 1024×768 pixels and subtended a physical field of view (FOV) of $40^{\circ} \times 30^{\circ}$.

2.1.1.2 *Vestibular Stimuli and Apparatus.* For vestibular stimulation, participants were seated on a 6 degree-of-freedom Stewart motion platform (Motionbase Maxcue, see Figure 2(c) and von der Heyde [2000]). For the experiment, however, only rotations around the earth-vertical axis (yaw) were used, as these are the behaviorally most relevant rotations for spatial orientation on the earth's surface. Furthermore, translations seem to be rather easy to spatially update (even for imagined motions) and are, hence, less interesting for our purpose [Easton and Sholl 1995; May and Klatzky 2000; May 1996; Presson and Montello 1994; Rieser 1989].

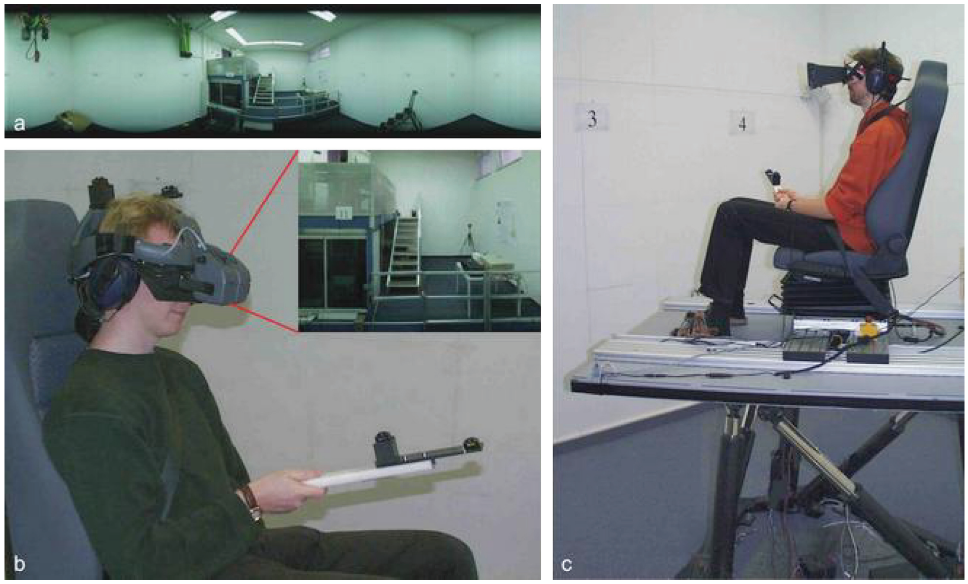


Fig. 2. **(a)** A photorealistic virtual replica of the real Motion-Lab was created from a 360° roundshot (4096 × 1024 pixel) of the real room. **(b)** Participant pointing with the position-tracked pointer and wearing a position-tracked head-mounted display (HMD) and active noise cancellation headphones. **(c)** Experimental setup displaying a participant seated on the motion platform and wearing headphones and purpose-designed blinders (vision delimiting cardboard goggles) reducing the FOV to that of the HMD (40° × 30°). Note the targets on the wall.

2.1.1.3 *Vibrations.* Additional broad-frequency vibrations were applied to the participants' seat and floor plate during all physical motions in order to yield a more compelling feeling of ego-motion and to mask motion-specific microvibration induced by the step motors moving the platform's legs.

2.1.1.4 *Auditory Stimuli.* Instructions during the experiment were given by a computer-generated voice and were presented via active noise-canceling headphones (Sennheiser HMEC 300). Additional broad-band noise (mixed river sounds) was continuously presented at a low level to effectively eliminate all spatial auditory cues of the surround without adding unnecessary discomfort. Furthermore, platform-masking sound was displayed during all simulated motions to mask all auditory motion-specific cues induced by physical platform motions.

2.1.2 *Interaction (Pointing).* After each rotation around the earth vertical axis, the participants' task was to point "as accurately and quickly as possible" to four targets announced consecutively via headphones. Participants were instructed to keep their head still and facing forwards by leaning it against the head rest. The pointing targets were randomly selected to be outside of the FOV of the HMD or the cardboard blinders and within a comfortable pointing range ($|\alpha_{pointer} - \alpha_{straight-ahead}| \in [20°, 99°]$).

Pointing was performed using a purpose-built, 6 degree-of-freedom position-tracked pointing wand (see Figure 2). After each pointing, participants raised the pointer to an upright position, indicating to the computer that the experiment can proceed. This upright default position ensured that there was no directional bias and participants had similar pointing response times for all directions, a problem that is often not accounted for in studies using compasslike pointers (e.g., Wraga et al. [2004]).

This rapid pointing metaphor—much like shooting—has the advantage of allowing the participant only very limited time to perform complex spatial reasoning and utilize abstract mental or geometric strategies, as is often observed in navigation and spatial orientation experiments (e.g., Riecke et al.

Table I. Summary of the Six Different Stimuli (Cue Combinations)

cue combinations (block)	Field of View (FOV)	Useful Visual Cues	Useful Vestibular Cues	Cue Conflict
Block A: "Real-world full FOV"	unrestricted	yes	yes	no
Block B: "Real-world w/ blinders"	$40^\circ \times 30^\circ$	yes	yes	no
Block C: "HMD vis. + vest. cues"	$40^\circ \times 30^\circ$	yes	yes	no
Block D: "HMD just vis. cues"	$40^\circ \times 30^\circ$	yes	no (no motion)	yes
Block E: "HMD constVis. + vest. cues"	$40^\circ \times 30^\circ$	no (static image)	yes	yes
Block F: "Blindfolded just vest. cues"	—	no (blindfolded)	yes	no

[2002a]). Thus, rapid pointing allows us to investigate the expectation of where participants think they are by measuring where they expect objects in their close surround to be with respect to their current position. Via triangulation, we can then backtrack where participants thought they were.

2.1.3 *General Procedure.* After a training phase, each participant completed a test phase consisting of six blocks of different cue combinations (see Table I), split into two sessions. In a repeated-measures, within-subject design, four typical spatial updating conditions were used in each block of this experiment: The 30 trials of each block were divided into 12 UPDATE trials and 6 trials each for of the CONTROL, IGNORE, and IGNORE BACKMOTION conditions in pseudo-randomized order.

1. **UPDATE**: From the current orientation, participants are simply rotated to a different orientation. From there, they have to point consecutively to four targets announced via headphones. If the available cues are sufficient for enabling automatic spatial updating, UPDATE performance should not depend on the angle turned.
2. **CONTROL**: Participants are rotated to a new orientation and immediately back to the original one before being asked to point. This is a baseline condition yielding optimal performance: If the available spatial updating cues are sufficient, UPDATE performance should be about as good as CONTROL performance ("**automatic spatial updating**").
3. **IGNORE**: Participants are rotated to a different orientation, but asked beforehand to *ignore* that motion and "respond as if you had not moved." If the available spatial cues are more powerful in triggering spatial updating and, hence, turn the world inside our head (even against our conscious will), those turns should be harder to IGNORE than to UPDATE. Spatial updating would then be "obligatory" or "reflexlike" in the sense of largely beyond conscious control and consciously hard-to-suppress ("**obligatory spatial updating**").
4. **IGNORE BACKMOTION**: After each IGNORE trial, participants are rotated back to the previous orientation. The main purpose of this condition is to avoid potential disorientation that might have been induced by the previous IGNORE trial.

Due to limitations of the platform turning range, the maximum heading deviation from straight ahead (12 o'clock) was $\pm 57^\circ$. Three different turning angles were used ($\pm 9.5^\circ$, $\pm 19^\circ$, $\pm 28.5^\circ$ there-and-back in the CONTROL conditions, and $\pm 19^\circ$, $\pm 38^\circ$, $\pm 57^\circ$ in the other ones). Movement time was always set to seven seconds, resulting in peak angular velocities of 5.4, 10.9, and 16.3°/s. Each trial consisted of the following three parts:

1. **Auditory announcement** indicating whether the upcoming spatial updating condition was an IGNORE trial, an IGNORE BACKMOTION trial, or a "normal" trial (UPDATE or CONTROL trial).
2. **Motion phase**, which always lasted seven seconds and started as soon as the pointer was in the default (upright) position. The velocity profile was Gaussian, with a peak velocity of twice the mean velocity.

3. **Pointing phase**, consisting of four repetitions of (a) the auditory target announcement of the different targets (e.g., "Object 9"), (b) subsequent pointing, and (c) raising the pointer to the upright (default) position.

In the test phase, each participant was presented with six stimulus conditions (blocks A–F, ca. 15 min. each) in pseudo-balanced order, with different degrees of visual and vestibular information available (see Table I for a comparison). Blocks A and B used the real environment under full-cue conditions as a baseline for optimal performance. Blocks C–F are the four sensible combinations of useful visual cues (yes/no), useful vestibular cues (yes/no), and resulting visuovestibular cue conflict (yes/no). Block D was the only one where participants were not turned physically, and asked to just use visual information. In blocks A–C, the amplitudes of the visual and vestibular (physical) turns were equal.

The pointing data were analyzed in terms of five dependent variables, revealing different aspects of spatial updating (see below). As pointing data are inherently directional (circular) data, we used circular statistics for computing the dependent variables (see, e.g., Batschelet [1981] for an introduction).

1. **Response time**: How easy and intuitive (fast) is the access to our spatial knowledge? Participants showed consistent differences in their mean response time even for the baseline (CONTROL) condition, ranging, for example, from 0.36 to 1.38 s in the CONTROL condition for block A (real-world full FOV). Those subject-specific overall response time differences were corrected for by computing the *relative* response time, which we define as the response time for that participant and block, divided by the ratio between that participant's mean response time in the CONTROL condition of that block and the mean response time across all participants in the CONTROL condition of that block. Using this procedure, the mean response time per condition, averaged over all participants, remains the same, but the between-subject response time differences are effectively removed.
2. **Configuration error = pointing variability**: How consistent is our spatial knowledge of the target configuration? That is, are the angles between landmarks reported consistently? The pointing variability is calculated as the mean angular deviation of the signed error, taken over the 4 pointings.
3. **Absolute pointing error**: How accurately do we know where we are with respect to our surround or specific objects of interest?
4. **Absolute ego-orientation error per trial**: Did participants misperceive their ego-orientation? Parts of the absolute pointing error might be confounded with a general misperception of the perceived ego-orientation and might be explained by the latter. For example, if participants somehow misperceive their orientation by 10°, this might already explain up to 10° of their absolute pointing error. The perceived ego-orientation per trial was estimated by taking the circular mean of the four signed pointing errors per trial [Batschelet 1981, chap. 1.3].
5. **Ego-orientation error in turning direction**: Did participants misperceive their ego-orientation typically in the direction of motion? If they would, that might be explained by some kind of "representational momentum," which describes the systematic tendency for observers to remember an event as extending beyond its actual ending point.

2.2 Results and Discussion

To get a first impression of the results, the data for block A ("Real-World full FOV") are plotted in Figure 3 for the five dependent variables. Figure 3 clearly shows the typical response pattern for automatic as well as obligatory spatial updating: UPDATE performance is comparable to CONTROL performance (implying automatic spatial updating), whereas IGNORE performance is considerably worse (implying obligatory spatial updating). IGNORE BACKMOTION performance was as good as UPDATE performance in all five dependent variables, indicating that participants were properly reanchored to the surround and no longer disoriented by the IGNORE trial beforehand.

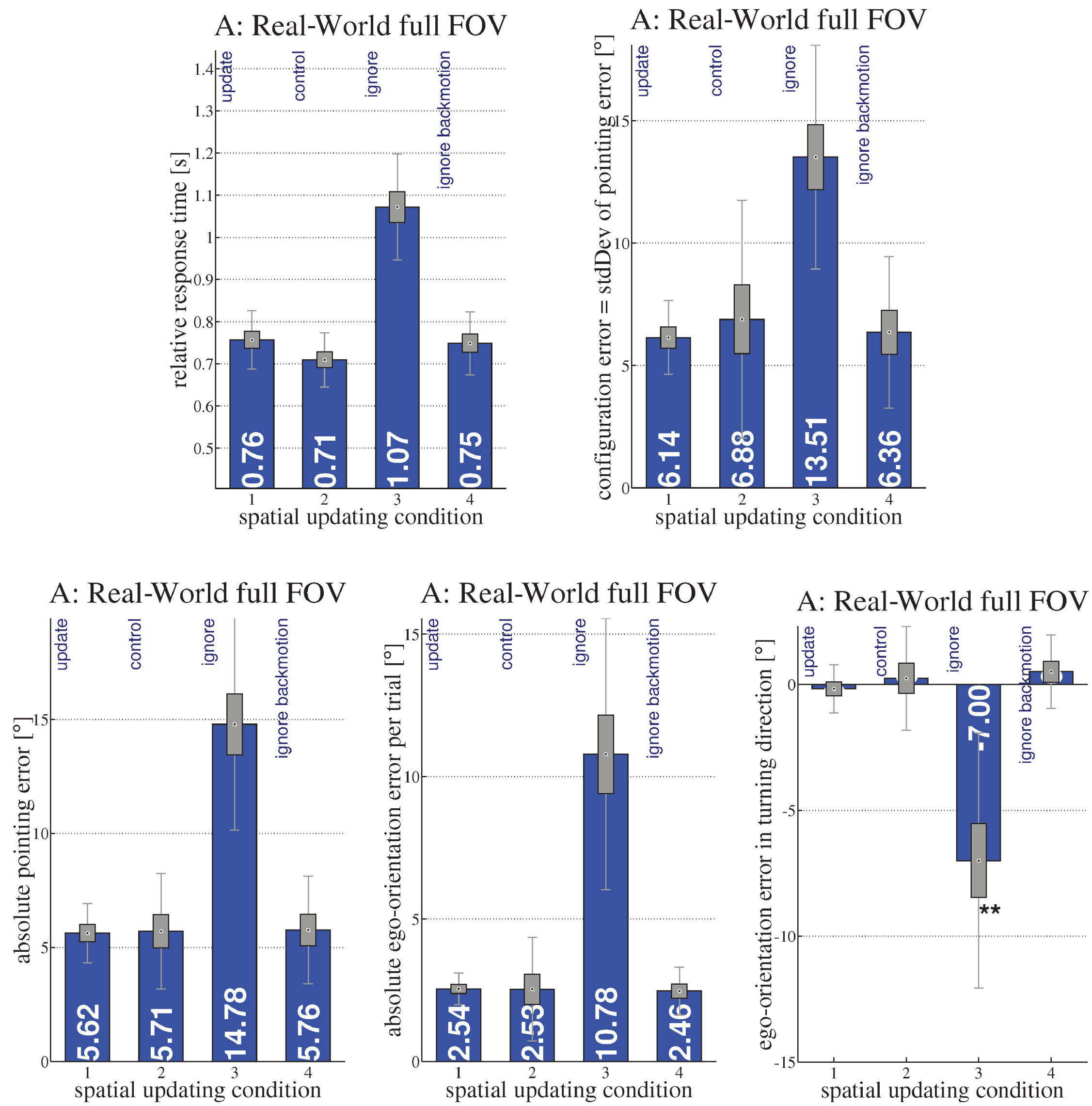


Fig. 3. Pointing performance showing the typical response pattern for spatial updating: UPDATE performance is comparable to CONTROL performance, whereas IGNORE performance is considerably worse. Performance in block A (real-world full FOV) is plotted for the five dependent variables, each for the four different spatial updating conditions. The bars represent the arithmetic mean, which is also numerically indicated by the white numbers at the bottom of each bar. Boxes and whiskers denote one standard error of the mean and one standard deviation, respectively. The asterisks "*" in the right plot indicate whether the mean differs significantly from zero (on a 5, 0.5, or 0.05% significance level, using a two-tailed t-test).

2.2.1 *Baseline* (CONTROL) *Performance.* The there-and-back motion of the CONTROL condition is simple enough that spatial updating of the motion is more or less trivial. Hence, potential performance differences between the different cue combinations (blocks) should indicate differences in the usability of the currently available *static* spatial information without too much influence from the *dynamic* motion cues.

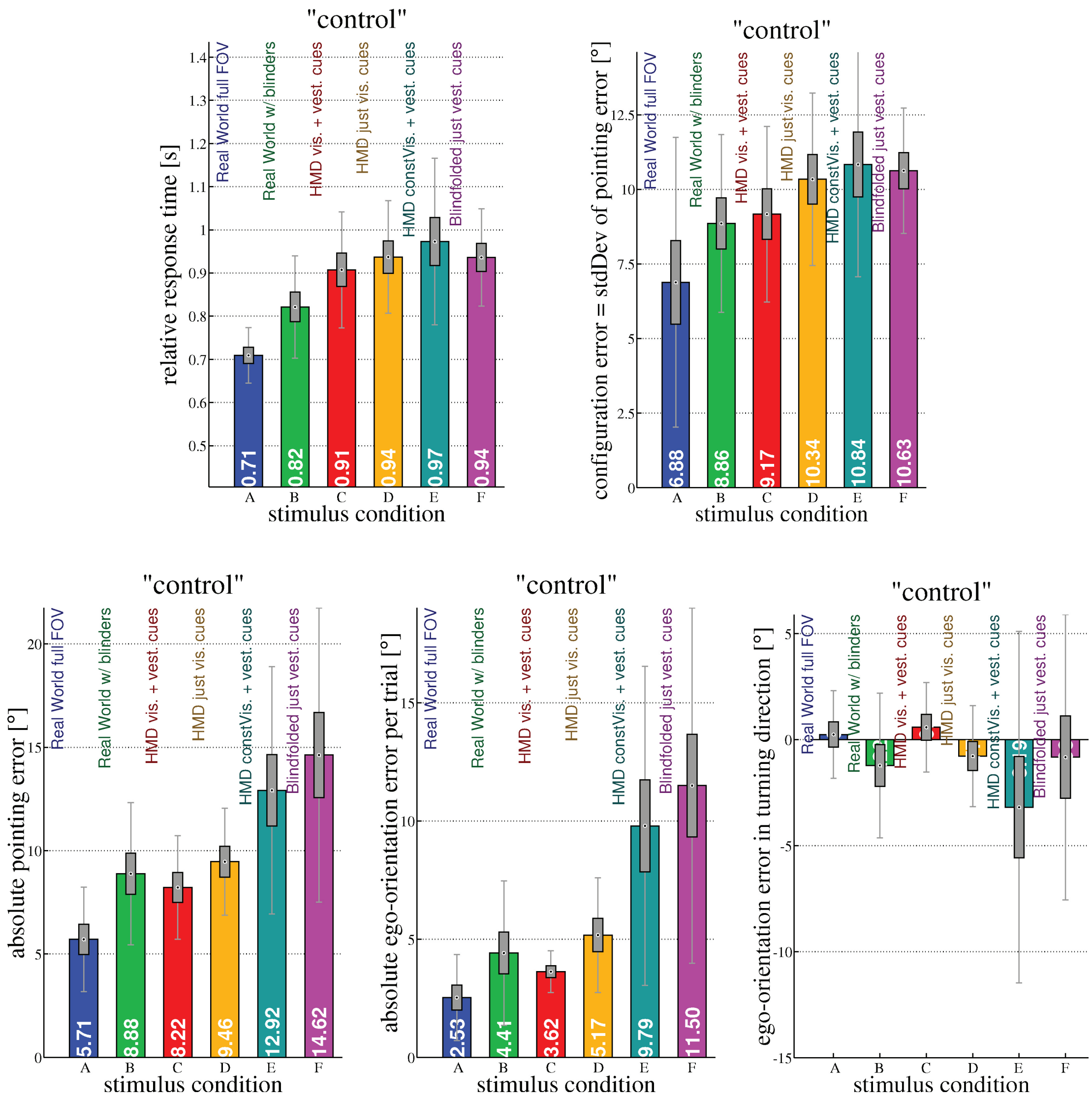


Fig. 4. Baseline (Control) spatial updating performance plotted for the five dependent variables, each for the six different cue combinations. Note the FOV effect even for the simple baseline task (block A versus B).

The Control data are summarized in Figure 4; the corresponding t-tests are compiled in Table II. Different questions guided the choice of cue combinations and will be discussed in the following subsections by comparing Control performance between adjacent blocks.

2.2.1.1 *Influence of FOV (Block A Versus B).* Comparing real-world performance with unrestricted vision (block A) versus constrained FOV (block B, see Figure 4) reveals a clear performance decrease for limiting the FOV to $40^\circ \times 30^\circ$ even in the rather simple baseline (Control) task.

2.2.1.2 *Real World Versus Virtual Reality Performance (Block B Versus C).* Participants in block B saw the real-world through a restricted FOV, whereas they saw in block C the same view presented

Table II. Tabular Overview of the Paired Two-Tailed t-Tests for the Different Comparisons

		Response Time		Configuration Error		Absolute Pointing Error		Absolute Ego-Orientat. Err.		Ego-Orient. Err. in Turn Dir.	
		$t(11)$	p	$t(11)$	p	$t(11)$	p	$t(11)$	p	$t(11)$	p
CONTROL											
influence of FOV	A vs. B	−2.65	**.022**	−1.35	.2	−3.46	**.0054**	−2.13	*.057*	1.27	.23
real-world vs. VR	B vs. C	−1.56	.15	−.62	.55	.802	.44	.971	.35	−1.55	.15
influence of vest. turn cues	C vs. D	−.481	.64	−2.1	*.06*	−2.18	*.052*	−2.13	*.056*	1.3	.22
const. vest. vs. const. vis.	D vs. E	−.427	.68	−.323	.75	−1.79	.1	−2.12	*.057*	.953	.36
const. vis. vs. no vis.	E vs. F	.49	.63	.159	.88	−.679	.51	−.704	.5	−.638	.54
just vis. vs. just vest.	D vs. F	.0344	.97	−.241	.81	−2.53	**.028**	−2.84	**.016**	.0272	.98
UPDATE-CONTROL											
influence of FOV	A vs. B	−.549	.59	.0658	.95	.927	.37	−.569	.58	−1.16	.27
real-world vs. VR	B vs. C	.263	.8	−1.63	.13	−2.54	**.027**	−.925	.37	.999	.34
influence of vest. turn cues	C vs. D	−.0606	.95	.926	.37	.745	.47	.915	.38	−1.82	*.096*
const. vest. vs. const. vis.	D vs. E	−1.6	.14	−3.03	**.011**	−3.41	**.0058**	−1.98	*.073*	.285	.78
const. vis. vs. no vis.	E vs. F	.957	.36	.787	.45	.433	.67	.0564	.96	.928	.37
just vis. vs. just vest.	D vs. F	−1.65	.13	−1.8	*.099*	−1.35	.2	−1.31	.22	1.59	.14
IGNORE-UPDATE											
influence of FOV	A vs. B	−.256	.8	.946	.36	.532	.61	.479	.64	−.387	.71
real-world vs. VR	B vs. C	1.93	*.079*	.00579	1	.839	.42	1.2	.26	−.756	.47
influence of vest. turn cues	C vs. D	−.599	*.56*	.35	.73	.00649	.99	−1	.34	−.243	.81
const. vest. vs. const. vis.	D vs. E	4.95	**.00043**	2.9	**.014**	3.1	**.01**	2.33	**.04**	−.521	.61
const. vis. vs. no vis.	E vs. F	−2.6	**.025**	−.694	.5	−.808	.44	−.618	.55	−2.41	**.035**
just vis. vs. just vest.	D vs. F	4.46	**.00097**	3.21	**.0083**	2.83	**.016**	1.93	*.08*	−2.21	**.049**

t-values are displayed with 3-digit precision, p-values for $\alpha = 0.05\%$ with 2-digit precision. Trailing zeros are omitted. Significant differences ($p < .05$) are typeset in bold, marginally significant differences ($0.1 > p > 0.05$) are typeset in italics.

through a HMD with the same FOV as the blinders. Figure 4 (top left) reveals a small but insignificant response time increase of approximately 90 ms for using the HMD in block C (cf. Table II). This suggests that the HMD condition might be perceived as slightly harder than the real-world condition. Some of the response-time difference, however, might also be caused by small visualization delays in the HMD condition. All other measures showed essentially the same performance and did not differ significantly, indicating that information displayed via HMD allows for the same spatial accuracy and ego-orientation perception.

2.2.1.3 *Influence of Vestibular Turn Cues (Block C Versus D).* Omitting all vestibular turn information and just displaying visual turn cues in block D reduced performance slightly (for three of the five dependent measures) compared to block C which provided vestibular turn cues (see Table II). These differences were, however, only marginally significant ($0.1 > p > 0.05$). Nevertheless, even the tendency toward reduced pointing-performance when only visual motion is available is intriguing, as one could argue that vestibular turn cues do not add any useful information for the simple CONTROL trials where participants had to point from exactly the same orientation as in the previous trial. From the current data, we can only speculate about the underlying reasons: Even though the vestibular turn cues would not have *added* any useful information, it is conceivable that instead the *lack* of matching vestibular cues might have caused the slight performance decrease, as it represents a visuovestibular cue conflict.

2.2.1.4 *Influence of (Missing) Useful Visual Cues (Block D–F).* Compared to block D where *only visual* motion cues were available, providing *only vestibular* turn cues while having to ignore the quasistatic visual cues in block E increased response time, configuration error, and ego-orientation error in turning direction only slightly and insignificantly (see Table II, D versus E). The absolute pointing error and absolute ego-orientation error, however, were considerably increased, indicating that participants tended to lose track of their correct ego-orientation without useful visual cues. This effect was slightly but insignificantly more pronounced for block F where participants were blindfolded (see Table II, D versus F). The lack of useful reference points in conditions E and F can explain the increased absolute ego-orientation error, as participants were constrained to using path integration, and, hence, lost track of their correct ego-orientation after several consecutive turns. For larger overall turning angles and orientation ranges, these ego-orientation errors would most likely be considerably larger.

2.2.1.5 *Summary and Conclusions.* Taken together, the results of the CONTROL condition demonstrated the importance of a large FOV and of useful reference points for quick and accurate knowledge of where the surrounding target objects were. Removing visual landmark information in block E and F reduced the available cues to path integration by vestibular cues, which led as expected to considerable misjudgments of the proper self-orientation. Maybe most critical for the further analysis and experiments, VR performance was—apart from a slightly increased response time—as good as real-world performance, provided that the FOV was matched. This validates our approach of using VR technology for studying spatial tasks and demonstrates the generalizability to comparable real-world situations.

2.2.2 *Automatic Spatial Updating.* In this subsection, automatic spatial updating will be investigated by analyzing the difference between UPDATE and CONTROL performance for the different cue combinations (see Figure 5). Subtracting CONTROL performance from UPDATE performance is an attempt to separate *dynamic* effects (i.e., UPDATE effects due to spatial updating) from baseline (CONTROL) differences most likely due to differences in the *statically* available information. In this manner, we compare spatial updating to different orientations to the supposed-to-be trivial updating there-and-back to the same orientation.

The literature on blindfolded spatial updating suggests a slight response-time increase of approximately 100 ms for UPDATE trials [Farrell and Robertson 1998; May 2000], and a considerable increase in pointing error (e.g., from 15° to 24° in the study by Farrell and Robertson [1998]). Such a pointing error increase might be explained by path integration errors, which should be compensated for by the useful landmarks in the visual conditions (A–D) of this experiment. Hence, we do not expect any major pointing error increase in those conditions.

2.2.2.1 *Conditions with Useful Visual Information (Block A–D).* For all blocks with useful visual landmarks (A–D), response times in the UPDATE trials were consistently increased by approximately 50 ms, compared to CONTROL performance [see Figure 5 (left)]. This difference was significant for the two real-world conditions (blocks A and B, $p < 0.05$), but only marginally significant for the two HMD conditions (blocks C and D, $p < 0.1$). The effect size is less than the 100 ms expected from the literature, indicating that updating to new orientations is almost as easy as updating there-and-back to the same orientation. This, in turn, suggests that spatial updating using landmark-rich visual cues is rather easy and automatic. The response-time increase of 50 ms is lower than the value typically found in the literature for *nonvisual* spatial updating, suggesting that the uncertainty of nonvisual path integration might have contributed to the increased response time there. The differences in terms of configuration error, absolute pointing error, as well as absolute ego-orientation error were all less than 1°, indicating that visually assisted updating to new orientations was virtually as accurate as baseline performance.

2.2.2.2 *Conditions Without Useful Visual Information (Block E and F).* For the blindfolded condition (block F), the response pattern changed somewhat: Response times for UPDATE trials were significantly increased by more than 100 ms. This is about the amount expected from the literature [Farrell and Robertson 1998; May 2000], corroborating the hypothesis that blindfolded spatial updating to new locations is not as quick and easy as for there-and-back motions.

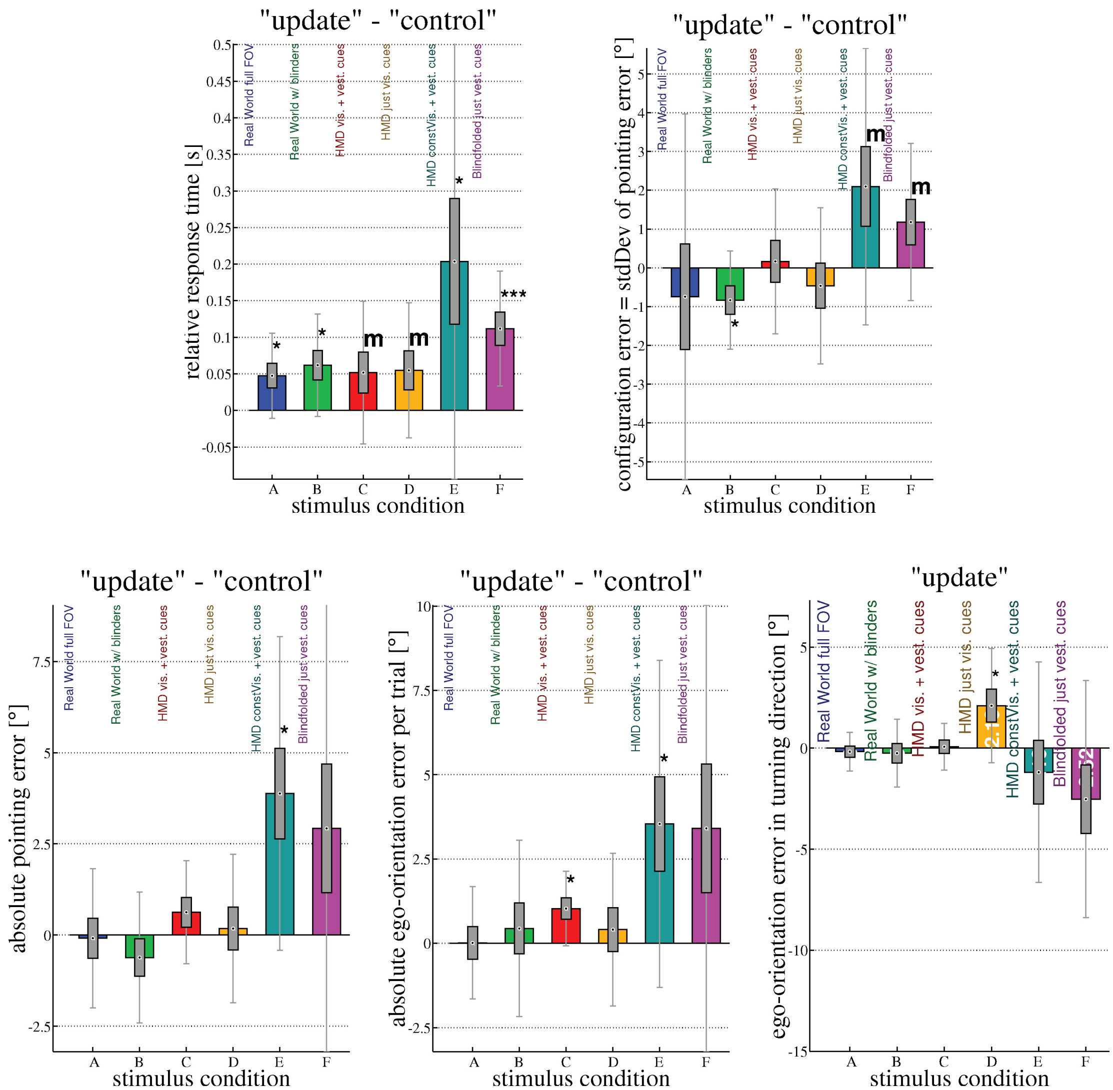


Fig. 5. Automatic spatial updating performance, quantified as the difference between Update and Control performance (except bottom right plot). If updating to new orientations is harder than for baseline there-and-back motions, Update performance should be worse than Control performance, resulting in a positive offset from zero in the above difference plots. This was the case for both conditions that relied on vestibular cues (blocks E and F). A zero or small offset, conversely, indicates that the available dynamic motion cues and static visual cues are sufficient to enable automatic spatial updating. This was observed for all conditions where participants could rely on visual landmark cues (blocks A–D).

Configuration error, absolute pointing error, and absolute ego-orientation error were only marginally increased, indicating that the consistency of the mental spatial representation did not suffer from the nonvisual ego-motion. The absolute error measures for the Update condition were only about a fourth higher than for the Control task, indicating that the main cause of the absolute pointing and ego-orientation errors is the accumulating path-integration error from the consecutive turns and not so much the updating to new orientations. That is, one simple turn can probably be updated rather well, but the sequence of many turns lead to the accumulation of path-integration errors, which is visible in the absolute error data. The ego-orientation error in turning direction consequently showed a large variability, but no overall effect.

Block E with additional but to-be-ignored visual information showed slightly more pronounced differences between UPDATE and CONTROL performance, especially for response times, which were increased by more than 200 ms. This indicates a severe difficulty in ignoring the visual stimulus, even though it was known to be totally irrelevant. However, the configuration error for UPDATE trials was only moderately increased by approximately 2°, indicating that the mental spatial representation was still rather consistent. The other variables show virtually no difference to the blindfolded condition in block F.

In a way, the UPDATE trials in block E (constVis. + vest. cues) can be seen as UPDATE trials for the *vestibular* stimulus and IGNORE trials for the *visual* stimulus. Conversely, the IGNORE condition of block D (HMD just vis. cues) can be seen as well as a UPDATE condition for the (constant) *vestibular* stimulus and an IGNORE condition for the *visual* stimulus. Indeed, the data shows virtually the same impaired performance for the two conditions where the visual cues were to be ignored and the vestibular ones to be trusted (block D IGNORE versus block E UPDATE). Especially the increased response time and configuration error for those conditions indicate a strong visual dominance over the vestibular cues: Even when explicitly intending to trust the vestibular cues more than the visual cues, participants were apparently unable to suppress the visual cues.

2.2.2.3 *Summary and Conclusions.* The data revealed the relative ease and accuracy of automatic spatial updating when one is provided with meaningful visual landmarks arranged into a consistent scene. Blindfolded spatial updating showed response-time differences (UPDATE–CONTROL) similar to those observed in the literature (e.g., Farrell and Robertson [1998]), validating our methodology. Additional conflicting visual information was rather hard to ignore and increased response times further. Maybe the most relevant outcome was the comparability of spatial updating for real and virtual environments. This demonstrates the power and usability of VR for investigating spatial updating. Furthermore, our rapid-pointing paradigm yielded overall response times that were considerably and consistently smaller than all values we could find in the literature.[1] On the one hand, this proves the ease and intuitiveness of our rapid pointing methodology. On the other hand, it allows the investigation of early processes in spatial updating that might not have been accessible before. This might be a critical issue in many spatial updating studies. Participants in the study by Wraga et al. [2004] needed, for example, more than seven times longer for pointing (8–12 s) than for verbal responses (1.1–1.5 s). This might indeed be problematic, since response times of more than 8 s might allow more than enough time for any kind of mental spatial task, like mental rotations, cognitive strategies, etc. It is, consequently, at least debatable whether Wraga et al. [2004] measured, in fact, *automatic* spatial updating performance and not some kind of rather cognitive mental spatial abilities. Furthermore, such long response times might increase their variability to a level where differences in the order of 100 ms (which is the typical difference found between UPDATE and CONTROL trials [Farrell and Robertson 1998; May 2000]) might not be visible any more.

2.2.3 *Obligatory Spatial Updating.* In this subsection, we will analyze the obligatory nature of spatial updating initiated by different combinations of visual and vestibular cues. The reasoning is as follows: If, and only if, spatial updating is obligatory (i.e., largely beyond conscious control) will

[1]Response times for pointings after blindfolded rotations differ considerably, with values ranging from 1.6 [May 2000] and 1.7 s [Farrell and Robertson 1998] over 1.8–3.2 s [Rieser 1989] up to more than 3 s [Creem and Proffitt 2000; Presson and Montello 1994]. A recent study on visually assisted spatial updating in VR reported even response times between 8 and 12 s [Wraga et al. 2004].

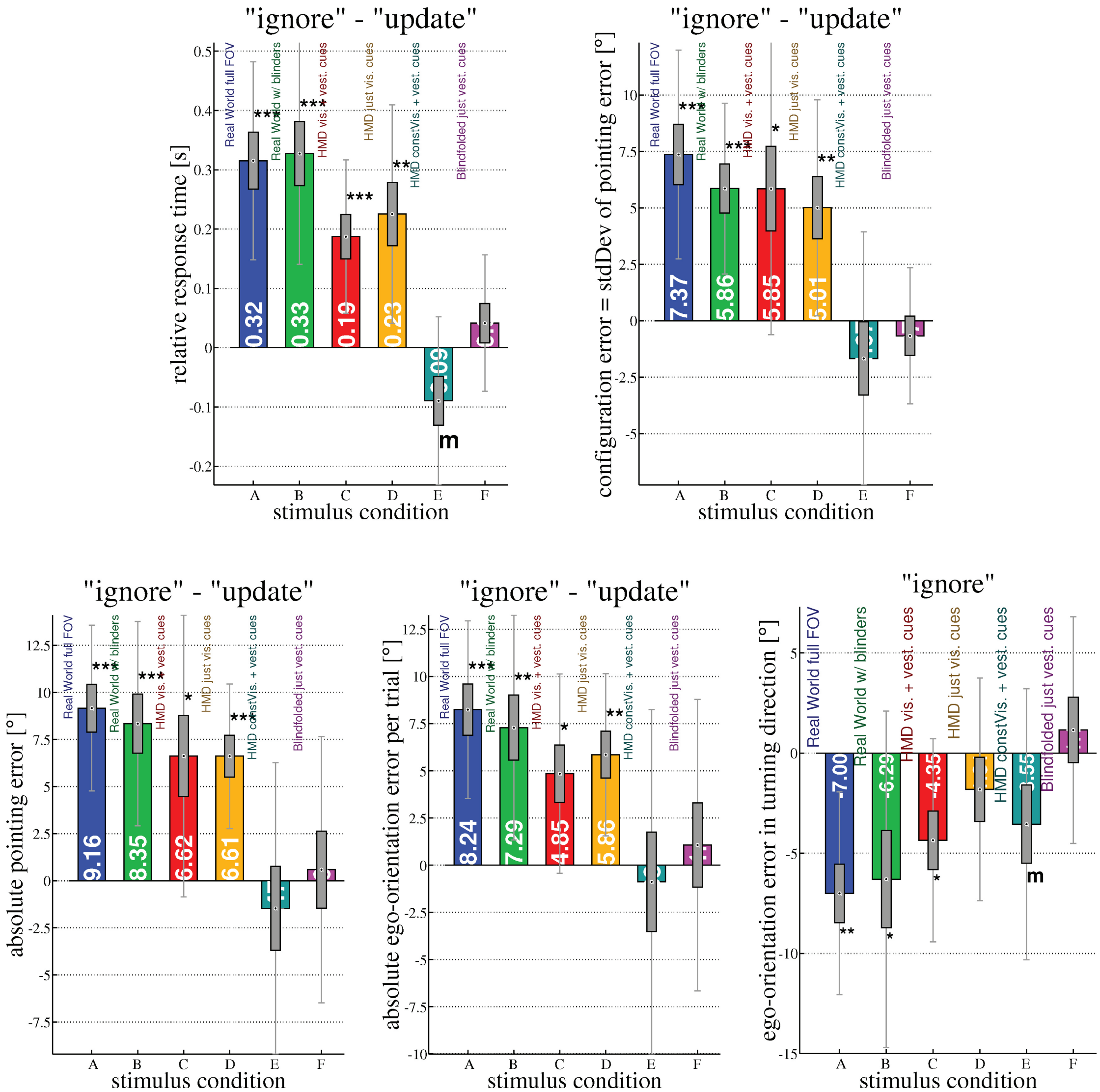


Fig. 6. Obligatory spatial updating performance, quantified as the difference between IGNORE and UPDATE performance (except bottom right plot). Values above zero indicate that ignoring is considerably harder than updating, implying obligatory spatial updating. This was the case for all conditions with useful visual cues (block A–D).

ignoring the turn stimuli be considerably harder than updating them as usual. That is, the difference between IGNORE and UPDATE trials would then be considerably above zero. The corresponding data are summarized in Figure 6 and will be discussed in detail below.

2.2.3.1 *Conditions with Useful Visual Information (Block A–D).* Both real-world conditions (blocks A and B) demonstrate essentially the same obligatory nature of the turn stimuli, without any influence of the FOV: IGNORE response times were increased by more than 300 ms, and all error measures were also largely increased. The considerable increase in configuration error indicates that the mental representation from the previous (to-be-remembered) orientation was less consistent and could not be remembered or accessed properly. The increase in absolute pointing error and absolute ego-orientation error can, for the most part, be explained by a direction-specific misperception of the correct ego-orientation in the opposite direction of the ignore motion [see Figure 6 (*bottom right*)]. That is, participants were apparently unable to correctly recall their previous orientation in the IGNORE trials and pointed as if the former orientation was being rotated in the *opposite* direction, i.e., further away than it actually was. This phenomenon is somewhat counterintuitive and conflicts with a motion capture or representational momentum explanation, which would predict a misperception *in* the direction of the motion, not *against* it. It seems like participants were trying to overcompensate for the actual rotation by pointing as if having turned further than they actually did.

The VR conditions (blocks C and D) demonstrated comparable obligatory spatial updating, even though the difference between IGNORE and UPDATE trials was slightly, but insignificantly, less pronounced. The ego-orientation error against turning direction in the IGNORE condition is greatest for the real-world conditions [block A (7°) and block B (6.3°)], slightly smaller for block C with HMD (4.4°), and negligible for block D without vestibular cues. Even though only the differences between the real-world conditions (blocks A and B) and the purely visual condition (block D) reached significance [$t(11) = 3$, $p = 0.012^*$ and $t(11) = 2.27$, $p = 0.044^*$, respectively], the ego-orientation error against turning direction seems to be an interesting measurand that has, to our knowledge, previously been neglected in spatial updating studies.

2.2.3.2 *Conditions Without Useful Visual Information (Block E and F).* Figure 6 reveals that vestibular turn cues without assisting visual turn cues were essentially as easy or hard to IGNORE as to UPDATE. This was found consistently for all five dependent measures. That is, smooth vestibular turn cues alone were clearly incapable of inducing obligatory spatial updating and turn the world inside our head against our own conscious will. In block E, with the visual stimulus indicating generally the wrong orientation, but no turn, ignoring ego-turns is almost 100 ms faster than updating them, suggesting that ignoring was actually perceived as easier than updating.

The observed ease of ignoring vestibular cues from blindfolded turns in block F was rather surprising, as the literature indicates that blindfolded motions should be much harder to IGNORE than to UPDATE. Farrell and Robertson [1998] found, for example, a response time increase from 1.7 to 3.3 s, accompanied by a moderate increase in absolute pointing error from approximately 24°–31°. Several differences in the experimental paradigms and setups used might be able to explain some of the observed differences.

Our hand-held pointing wand enabled considerably smaller response times than those observed in the literature (see subsection 2.2.2.2). This suggests that our pointing paradigm might be easier and more intuitive to use, which enables us to better investigate the quick process of spatial updating. The overall small response times in our experiments, however, do, by no means, explain the ease in ignoring purely vestibular turn stimuli. The smooth motions used were clearly above detection threshold, but the accelerations and velocities reached might still be considerably below the values typically used in the literature. Furthermore, pointing targets in our study were attached to the walls of the room and, hence, embedded in a consistent, natural scene. In this manner, participants probably did not update or imagine the position of *individual* targets or target arrays, but most likely updated the scene and room geometry as a whole, which is known to be more reliable and less prone to disorientation [Wang and Spelke 2000]. This is different from many spatial updating studies that used arrays of objects that were not well embedded or part of a consistent scene [Easton and Sholl 1995; Farrell and Robertson 1998; Farrell and Thomson 1998; May and Klatzky 2000; May 1996; Presson and Montello 1994; Rieser et al. 1982; Rieser 1989; Wang and Spelke 2000]. Finally, the relatively small turning angles used (<60°) and the repeated turns without in between visibility of the scene in our study might also have contributed to the ease of ignoring the motion.

2.2.3.3 *Summary and Conclusions.* Comparing IGNORE and UPDATE performance revealed obligatory spatial updating for all conditions with useful visual information (blocks A–D). That is, visual cues alone, even without any concurrent vestibular stimuli, can be sufficient for "turning the world inside our head," even against our own conscious will. This clearly indicates reflexlike, cognitively almost impenetrable (i.e., obligatory) spatial updating by visual cues alone. Moreover, a strong visual dominance over the vestibular cues was observed, even in the conditions where participants were explicitly asked to ignore the visual stimulus completely and just trust the vestibular cues (block E). Further experiments are planned to investigate the degree of realism required to enable obligatory spatial updating in virtual environments. So far we can conclude that a highly detailed, photorealistic view onto a consistent, landmark-rich environment is clearly sufficient to enable obligatory as well as automatic spatial updating, whereas presenting only optic flow is not [Klatzky et al. 1998; Riecke et al. 2002b].

Smooth vestibular turn cues without any assisting visual cues, on the other hand, were clearly incapable of triggering reflexlike obligatory spatial updating (block F). This outcome is, to our knowledge, unprecedented and not predicted by the literature. Low accelerations and velocities and a highly consistent scene that is easier to mentally picture might all contribute to this apparent contradiction. Further experiments, however, are needed to understand this fundamental difference and pinpoint the exact condition under which vestibular cues might indeed be sufficient for initiating obligatory spatial updating, as is typically claimed by the literature.

2.2.4 *Additional Analyses*

2.2.4.1 *Learning Effect.* To test whether participants' performance depended on the amount of exposure to the task, a correlation analysis was performed between performance and session number. No performance feedback was given during any of the test phases. That is, potential improvements would be due to practice effects, implicit learning, or improved strategies. The results of the correlation analysis, however, revealed no significant correlation for any of the dependent parameters and spatial updating conditions (all $r^2{}'s \leq 0.032$, all $t's(70) \leq 1.51$, all $p's \geq 0.068$).

2.2.4.2 *Turning-Angle Effect.* If spatial updating was nonautomatic and thus effortful and requiring considerable cognitive effort, like mental spatial rotation, we would expect that smaller turns should be easier and lead to better UPDATE performance than larger turns. This should be reflected in increased errors and especially response times for larger turns. A correlation analysis, however, revealed no significant performance change with turning angle for any of the dependent variables or cue combinations (blocks) ($p > 0.05$). This suggests that spatial updating was performed *during* the motion and not afterward. Furthermore, the lack of a turning-angle effect argues against higher cognitive processes, like mental spatial rotations performed after or during the actual turn. Together with the rather low overall response times, this suggests that spatial updating was, indeed, automatic.

If, on the other hand, participants in the IGNORE condition performed some kind of mental backrotation as is often claimed, response times in the IGNORE condition should be positively correlated with turning angle. Correlation analyses showed such an effect only for the absolute pointing error and absolute ego-orientation error in block B (real-world w/blinders, $r = 0.38$, $r^2 = 0.14$, $t(11) = 2.48$, $p = 0.031^*$ and $r = 0.29$, $r^2 = 0.085$, $t(11) = 2.44$, $p = 0.033^*$, respectively). None of the other correlations attained significance for any of the dependent variables and cue combinations (blocks) ($p > 0.05$). This lack of consistent response-time increase with turning angle does not support the mental rotation hypothesis.

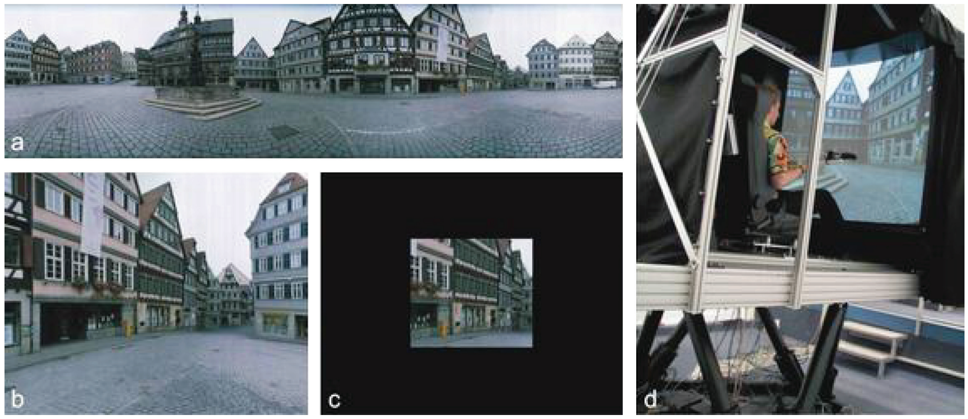


Fig. 7. (a) The virtual replica was created by wrapping a 360° round shot photograph of the Tübingen market place onto a cylinder. This creates an undistorted view for the observer positioned in the center of the cylinder. The 22 pointing targets were marked by little red dots on the image. Pictures (b) and (c) illustrate the full 84° × 63° view and the reduced view (blinders and HMD conditions, 40° × 30° FOV), respectively. (d) Participant seated on the motion platform and facing the curved projection screen, which displays a view of the Tübingen market place. The physical field of view is 84° × 63° and matches the simulated FOV.

3. CONTROL EXPERIMENT

3.1 Introduction

The previous experiment examined extreme cases of either full or no useful information in the visual and vestibular domain, and showed the typical response pattern for obligatory spatial updating whenever useful visual information was available. The purpose of this control experiment was to extend the previous findings and address questions that remained unanswered. This includes, on the one hand, questions about the importance of visual display parameters like FOV and projection screen versus HMD usage for spatial updating. On the other hand, questions about the influence of various visuovestibular motion parameters, like the relation between visual and vestibular motion and the turning amplitude and velocity, were addressed. We even tested the extreme case of no motion cues at all in a "teleport"-like condition; there, participants saw views from novel orientations in a slide-show type presentations, without any visual or vestibular motion cues whatsoever. Due to the large amount of potentially relevant parameters, the nature of this control experiment is rather exploratory. The main purpose was, consequently, to scan the realm of parameters to identify critical parameters that are worth investigating in more depth in later studies.

3.2 Methods

The overall methodology of this control experiment is similar to the first experiment, with a few changes described below that addressed open questions and potential difficulties in the first experiment.

3.2.1 *Participants.* A group of 6 female and 2 male naive participants took part in the control experiment, with ages ranging from 17 to 38 years (mean: 27.9). All participants had been living in Tübingen for several years and were familiar with the Tübingen market place used as the visual scene.

3.2.2 *Stimuli and Apparatus.* Three different visualization conditions were used in this experiment:

1. An **HMD** condition, which was comparable to block C of the previous experiment.
2. A **projection screen** condition, where participants were seated in front of a projection screen mounted on top of the platform, as described below [see Figure 7(d)].

3. A **blinders** condition, where participants were again seated in front of the projection screen, but wore cardboard goggles limiting the FOV to that of the HMD ($40^\circ \times 30^\circ$). The blinders were the same as in the previous experiment, but used to limit the view onto the computer-rendered image on the projection screen, not the real surround.

The projection setup was designed by the authors mainly to allow for a larger FOV and to avoid drawbacks often associated with HMDs: HMDs are known to lead to a number of problems, including discomfort, fatigue, headaches, eye strain, and blurred vision [Cobb et al. 1999; Hettinger et al. 1996; Howarth and Costello 1997; Mon-Williams and Wann 1998; Stanney et al. 1998]. Further drawbacks associated with HMDs include distortions of the perceived space, as well as impaired spatial orientation performance [Arthur 2000; Bakker et al. 1999, 2001; Creem-Regehr et al. 2003; Hettinger et al. 1996; Kearns et al. 2002; Nelson et al. 1998; Riecke et al. 2005]. The whole projection setup is mounted on top of the motion platform [see Figure 7(d)] in order to allow for optimal viewing conditions and immersion. This projection screen is 1.68 m wide and has a curvature radius of 2 m. This yields a physical FOV of $84^\circ \times 63^\circ$ for the participant seated at a distance of about 1.14 m. A modified wide-angle LCD video projector (Sony VPL-PX 21 with wide-angle lens VPL-FM 21) is mounted above and behind the head of the participant and projects a computer-rendered 1024×768 pixel image non-stereoscopically onto the screen. To enhance immersion and to reduce the influence of the external reference frame of the physical surround (Motion-Lab) as much as possible, the projection screen is surrounded by a black frame and the whole projection setup is surrounded by light-proof black curtains on all sides.

Instead of using 12 regularly arranged target objects attached to the wall of the rectangular Motion-Lab as in the main experiment, we used a considerably larger, more complex, and less regular environment with 22 target landmarks arranged irregularly (the Tübingen market place, see Figure 7). This should render abstract cognitive strategies (like using symmetries and counting targets) virtually impossible, thus forcing participants to resort to automatized spatial updating whenever possible. Furthermore, this procedure pushed participants closer to their performance limit, which might yield even clearer results.

In most spatial updating studies in the literature, however, target configuration and/or room geometry were quite simple: Typically, a rather limited number of four to eight targets is used, which are often arranged regularly or along the cardinal directions (front–back/left–right), and are typically embedded in just a simple and featureless symmetrical room [Carpenter and Proffit 2001; Creem et al. 2001; Creem and Proffitt 2000; Easton and Sholl 1995; Farrell and Robertson 1998, 2000; Klatzky et al. 1998; May and Klatzky 2000; May 1996, 2004; Presson and Montello 1994; Rieser 1989; Shelton and McNamara 2001; Simons and Wang 1998; Wang and Simons 1999; Wang and Spelke 2000; Wraga et al. 1999a, 1999b, 2000, 2004; Yardley and Higgins 1998]. Hence, it seems quite possible that participants did not always update the room or target configuration properly. Instead, they might, for example, have used simpler strategies like reversing left and right directions instead of updating 180° turns, updating individual targets analytically instead of using normal, holistic spatial updating, or counting objects instead of updating turns (e.g., take the third target to the left instead of updating a 135° turn).

The asymmetric scene and target configuration in this experiment made all of these cognitive strategies impracticable. Furthermore, the abundance of landmarks ensured that participants could not update all targets individually or rote-learn all relative angular distances between all targets (which would add up to $\sum_{n=1}^{N-1=21} n = 231$). Instead, participants had to update the whole scenery to deduct individual target locations, which is exactly what we intended and what "normal" spatial updating is about.

3.2.3 *Procedure.* An extended training phase preceded the test phase in order to familiarize participants with the rapid pointing procedure, the different spatial updating conditions, the VR setup, and

Table III. Summary of the 8 + 3 Different Cue Combinations (Blocks) Used in the Control Experiment

Block	Visualization Setup	Field of View (FOV)	Gain vest/vis	Angular Range (visual)	Mean vis. Turn vel.	Cue Conflict
A	HMD	40° × 30°	1	[−57°, +57°]	20°/s	no
B	blinders + screen	40° × 30°	1	[−57°, +57°]	20°/s	no
C	proj. screen	84° × 63°	1	[−57°, +57°]	20°/s	no
D	proj. screen	84° × 63°	0.5	[−114°, +114°]	20°/s	gain
E	blinders + screen	40° × 30°	0.25	[−228°, +228°]	20°/s	gain
F	proj. screen	84° × 63°	0.25	[−228°, +228°]	20°/s	gain
G	blinders + screen	40° × 30°	0	[−228°, +228°]	20°/s	yes
H	proj. screen	84° × 63°	0	[−228°, +228°]	20°/s	yes
I	proj. screen	84° × 63°	0	[−228°, +228°]	jump	yes
J	proj. screen	84° × 63°	0.25	[−57°, +57°]	20°/s	gain
K	proj. screen	84° × 63°	0.25	[−228°, +228°]	80°/s	gain

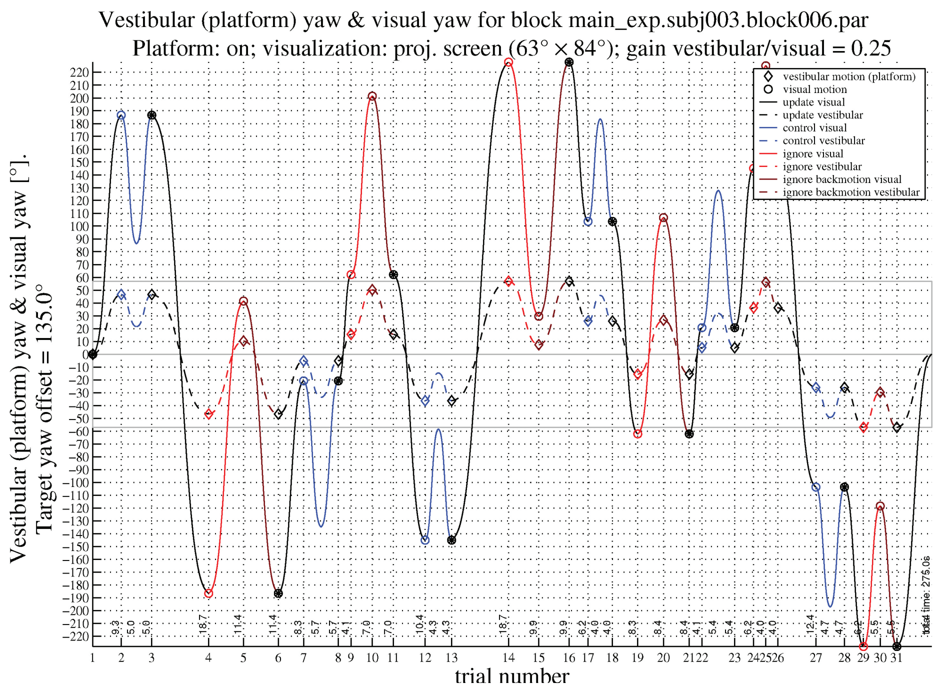


Fig. 8. Vestibular (platform) motion and visual motion for one representative block (D) of the control experiment. Depicted are the vestibular (platform) and visual yaw angle, demonstrating the sequence of the spatial updating conditions (Control → Update → Ignore → Ignore Backmotion → Update, etc.) and the effect of the gain factor. In this block, the vestibulo-visual gain factor was $g = 1/4$, indicating that the visual motions were four times as large as the physical (vestibular) ones. Pointings occurred at all circles and diamonds of the trajectory.

the landmarks used. As before, a repeated-measures, within-subject design was used in the test phase. The experimental design and the cue combinations used for each block are summarized in Table III and described in more detail in subsection 3.2.4. Each participant completed eight blocks of different cue combinations in pseudo-balanced order. Each block consisted of 32 trials and lasted about 18 minutes. The 32 trials were split into 12 Update trails, and 6 trials each for the other spatial updating conditions (Control, Ignore, and Ignore Backmotion), as illustrated in Figure 8. The blocks were performed in two or three sessions on different days to avoid fatigue effects and obviate the influence of declining alertness.

In a follow-up study, the same participants performed three more control conditions, first block I (jump condition), followed by two more blocks (J and K) in balanced order (see subsection 3.2.4 and Table III). For the sake of comparability, the results of the follow-up study are presented together with the original eight blocks. This comparison might seem critical due to potential learning or practice effects. The learning-effect analysis in subsection 3.3.4, however, revealed only a significant learning effect in terms of response time. This response time decrease was found for both the UPDATE and IGNORE conditions, but not the CONTROL condition. This suggests that careful comparisons between the first eight blocks and the three subsequent blocks might be legitimate for the other four dependent variables and for the CONTROL condition, in general.

The pointing paradigm was similar to the previous experiment and pointing targets were selected to be outside of the FOV of the projection screen and within a comfortable pointing range ($|\alpha_{pointer} - \alpha_{straight\ ahead}| \in [42^\circ, 110^\circ]$). As the target names in this experiment had different lengths, but were unique after the second syllable, the response-time computation was adapted accordingly by defining $t = 0$ to be at the mean pronunciation time of the first two syllables (1.43 s). Furthermore, the summed response time between the first target announcement and the fourth and last pointing was announced acoustically just before the next trial. Pilot experiments had shown that this performance feedback effectively motivates participants to perform as good as they could. It was further found to decrease boredom and keep participants alert by enhancing the gamelike character of the experiment.

3.2.4 *Stimulus Conditions and Motivation.* The different stimulus conditions for the eleven blocks were chosen to allow for comparisons between different visual display parameters like FOV and projection screen versus HMD usage, on the one hand, and visuo vestibular motion parameters, like gain factors and turning amplitude and velocity, on the other. The parameter combinations are compiled in Table III and motivated below in more detail.

The HMD condition (block A) was chosen to allow for comparisons to the previous experiment, which used the same HMD and the same yaw range, but a different scene. Block B and C used the projection screen with and without blinders and were apart from this identical to block A. This allows for comparisons between the different visualization systems (HMD versus blinders) and FOV (projection screen with and without blinders).

Blocks D–H were designed to investigate larger turning angles, which are often assumed to be more difficult to update. To do this, gain factors $g_{vestibular/visual} < 1$ had to be introduced, as the motion platform used has a limited turning range of $\pm 57^\circ$. Preexperiments had shown that vestibular/visual gain factors down to 1/4 are accepted and typically pass unnoticed by participants when they are involved in an engaging task like rapid pointing. Block D used an intermediate gain factor of $g = 1/2$ and block E and F reduced the gain factor to $g = 1/4$. In block G and H, the gain factors were set to zero, to investigate spatial updating performance if no concurrent vestibular turn stimuli were presented at all. To investigate the influence of the FOV, participants wore blinders in blocks E and G; the results were compared to blocks F and H, respectively, with unrestricted vision.

The three follow-up conditions (blocks I–K) were run in one separate session after the first eight blocks were completed. They were designed to answer some of the questions not sufficiently addressed by blocks A–H. It seemed as if stimulus parameters and movement specifics had little, if any, influence on spatial updating performance in blocks A–H and the previous experiment as long as sufficient visual cues were provided. This raised the question whether a smooth, continuous spatial updating, similar to processes known as mental spatial rotation, can fully explain the observed results. To address this issue, a "jump" condition was introduced in block I. In this condition, participants were presented with new views *without* any continuous motion in between, similar to a slide-show type presentation, as if they were teleported directly to the new orientation. If motion parameters have any influence on visually assisted spatial updating, this jump to new orientation should disrupt spatial updating performance considerably and improve ignore performance. Similar results for continuous and discontinuous (jumplike) spatial updating, on the other hand, would suggest that the mere view of the new orientation at a given instant of time is sufficient to somehow teleport the mental spatial representation to the new orientation and reanchor the mental representation almost instantaneously, thus challenging the prevailing notion of spatial updating.

Block J used a gain factor of $g = 1/4$ and a visual yaw range of $\pm 57^\circ$, resulting in a vestibular (physical) yaw range of only $\pm 14.25^\circ$. This allowed the disambiguation of effects of gain factor and turning angle by comparing the results to block B and F. Block K was designed to elucidate the effect of movement velocity by using visual rotational velocities that were four times higher than in the other conditions. Apart from that, block K was identical to block F.

3.3 Results and Discussion

As this control experiment is exploratory in nature, the line of argument will be rather qualitative, focusing on the interesting effects and tendencies. For reference, the corresponding statistical analyzes of interest are summarized in Table IV. The fifth dependent variable (ego-orientation error in turning direction) did not show any significant differences between the stimulus conditions and will, hence, only be reported in the context of the IGNORE condition (see subsection 3.3.3.6). Due to the small number of participants, the power of the tests is, of course, limited and only strong effects might be clearly visible. Nevertheless, this procedure seems sufficient to scan the realm of potentially relevant parameters and identify the ones that are worth being investigated in more detail in future studies.

To provide a first impression of the results, the data for the HMD condition (block A) are summarized in Figure 9. As in the previous experiment, Figure 9 demonstrates the typical response pattern for obligatory spatial updating: UPDATE and IGNORE BACKMOTION performance are almost as good as baseline CONTROL performance, whereas IGNORE performance is considerably decreased.

Compared to the results in the similar HMD condition (block C: HMD vis. + vest. cues) of the previous experiment, however, overall performance is slightly decreased: Configuration error, absolute pointing error, as well as absolute ego-orientation error are all decreased by 4–5°. Furthermore, mean response times in the CONTROL condition were increased by 250 ms, indicating that the task was considerably harder, even for the supposedly simple baseline task.

Apart from possible overall differences in the participant populations used, differences in the experimental procedures might account for the observed performance differences. Most prominently, the scenery used in the control experiment was considerably more complex and without any potentially helpful symmetry or regularity. Furthermore, the number of targets was almost doubled; targets were arranged irregularly without any symmetry. Nevertheless, participants were able to successfully update to new orientations and had a hard time ignoring the turn, indicating that spatial updating was still working and the rapid pointing approach was still a successful paradigm.

3.3.1 *Baseline* (CONTROL) *Performance.* The results for the baseline (CONTROL) condition are summarized in Figure 10, with the corresponding statistical analyses in Table IV. In general, CONTROL performance showed a considerable between-subject variability, indicated by the often, rather large, standard deviations, but only few systematic differences between blocks. That is, neither the difference between HMD and blinder usage, nor the gain factor, turning angle, or movement velocity produced any clear performance differences for the simple forth-and-back motions (cf. Table IV). Even the jump condition (block I), where participants saw an intervening view of a new orientation before seeing the same

Table IV. Tabular Overview of the Paired Two-Tailed t-Test for the Different Comparisons

		Response Time		Configuration Error		Absolute Pointing Error		Absolute Ego-Orientation Error	
		$t(7)$	p	$t(7)$	p	$t(7)$	p	$t(7)$	p
CONTROL									
HMD vs. blinders	A vs. B	−1.18	.28	.0362	.97	−.949	.37	−.891	.4
	B vs. C	.691	.51	.747	.48	1.15	.29	1.38	.21
FOV effect	E vs. F	−1.94	*.093m*	1.88	.1	3.29	**.013***	3.61	**.0086***
	G vs. H	−.11	.92	.615	.56	1.39	.21	1.12	.3
gain factor &	B vs. E	1.49	.18	−.251	.81	−.873	.41	−1.46	.19
turn angle (blinders)	B vs. G	.901	.4	−.678	.52	−.406	.7	−1.02	.34
gain factor &	C vs. F	−.0483	.96	1.17	.28	1.2	.27	.141	.89
turn angle (proj. scr.)	C vs. H	−.396	.7	−.266	.8	.0216	.98	−1.13	.29
	E vs. G	−.829	.43	−.583	.58	.229	.83	.245	.81
just gain factor	F vs. H	−.289	.78	−.856	.42	−1.19	.27	−2.98	**.02***
	B vs. J	2.28	*.057m*	1.13	.29	1.73	.13	1.71	.13
just turn angle	F vs. J	2.79	**.027***	−.306	.77	−.119	.91	.383	.71
turn velocity	F vs. K	1.19	.27	.0571	.96	.187	.86	.468	.65
turn vel. & turn angle	J vs. K	−1.16	.28	1.21	.27	.731	.49	.0684	.95
turn angle & jump	I vs. J	1.11	.31	.42	.69	.302	.77	.687	.51
jump vs. cont.	I vs. K	−.383	.71	.564	.59	.543	.6	.937	.38
UPDATE-CONTROL									
HMD vs. blinders	A vs. B	.491	.64	−.775	.46	.952	.37	2.06	*.079m*
	B vs. C	.109	.92	−.0125	.99	−.523	.62	−1	.35
FOV effect	E vs. F	−.247	.81	.409	.69	−.754	.48	−.865	.42
	G vs. H	.432	.68	.367	.72	−.0324	.98	.76	.47
gain factor &	B vs. E	1.5	.18	−.634	.55	−.396	.7	.135	.9
turn angle (blinders)	B vs. G	.0674	.95	−.432	.68	−.914	.39	−.811	.44
gain factor &	C vs. F	1.69	.14	−.667	.53	−.777	.46	.243	.81
turn angle (proj. scr.)	C vs. H	.376	.72	.182	.86	−.0176	.99	1.31	.23
	E vs. G	−1.66	.14	.318	.76	−.313	.76	−.85	.42
just gain factor	F vs. H	−.484	.64	.983	.36	.802	.45	1.85	.11
	B vs. J$	.655	.53	−.743	.48	−.873	.41	−.541	.61
just turn angle	F vs. J$	−.351	.74	−.215	.84	.761	.47	.651	.54
turn velocity	F vs. K$	1.34	.22	.0548	.96	−.439	.67	−.569	.59
turn vel. & turn angle	J vs. K	1.26	.25	.326	.75	−2.24	*.06m*	−1.3	.23
turn angle & jump	I vs. J	−.517	.62	−.626	.55	1.07	.32	.786	.46
jump vs. cont.	I vs. K	1.34	.22	−.602	.57	−.945	.38	−.621	.55
IGNORE-UPDATE									
HMD vs. blinders	A vs. B	−2.28	*.057m*	−.852	.42	−2.01	*.084m*	−3.03	**.019***
	B vs. C	.672	.52	−.87	.41	.978	.36	1.5	.18
FOV effect	E vs. F	−3.28	**.014***	−3.44	**.011***	−1.32	.23	−.912	.39
	G vs. H	−1.8	.12	−.247	.81	−.652	.54	−.384	.71
gain factor &	B vs. E	1.8	.12	−.658	.53	−.85	.42	−.0831	.94
turn angle (blinders)	B vs. G	1.09	.31	−1.45	.19	−1.27	.25	−1.11	.31
gain factor &	C vs. F	−2.05	*.08m*	−2	*.086m*	−2.8	**.026***	−1.72	.13
turn angle (proj. scr.)	C vs. H	−.856	.42	−1.56	.16	−4.12	**.0045****	−3.19	**.015***
	E vs. G	−2.77	**.028***	−1.52	.17	−.653	.53	−.752	.48
just gain factor	F vs. H	.655	.53	.0553	.96	−.658	.53	−.492	.64
	B vs. J$	1.4	.2	.663	.53	1.07	.32	1.88	.1
just turn angle	F vs. J$	1.89	.1	2.7	**.031***	2.2	*.064m*	1.48	.18
turn velocity	F vs. K$	2.08	*.076m*	−.285	.78	.171	.87	−.0101	.99
turn vel. & turn angle	J vs. K	−1.21	.26	−2.88	**.024***	−3.05	**.019***	−1.95	*.093m*
turn angle & jump	I vs. J	1.39	.21	2.19	*.065m*	2.4	**.047***	1.81	.11
jump vs. cont.	I vs. K	.352	.74	−1.17	.28	.978	.36	1.09	.31

Block comparisons marked with a "$" are comparisons between one of the first eight conditions and one of the three follow-up conditions performed afterward. Hence, potential order or learning effects cannot be fully excluded, even though those effects did not reach statistical significance (see subsection 3.3.4). Note that the only consistent effect for the CONTROL condition (*top*) was a small benefit for an increased FOV (E versus F). Furthermore, there were no significant differences in terms of automatic spatial updating whatsoever (UPDATE–CONTROL, *middle*). The effects in terms of obligatory spatial updating (IGNORE–UPDATE, *bottom*) reached significance for the parameters presentation device (HMD versus blinders), FOV, and turning angle.

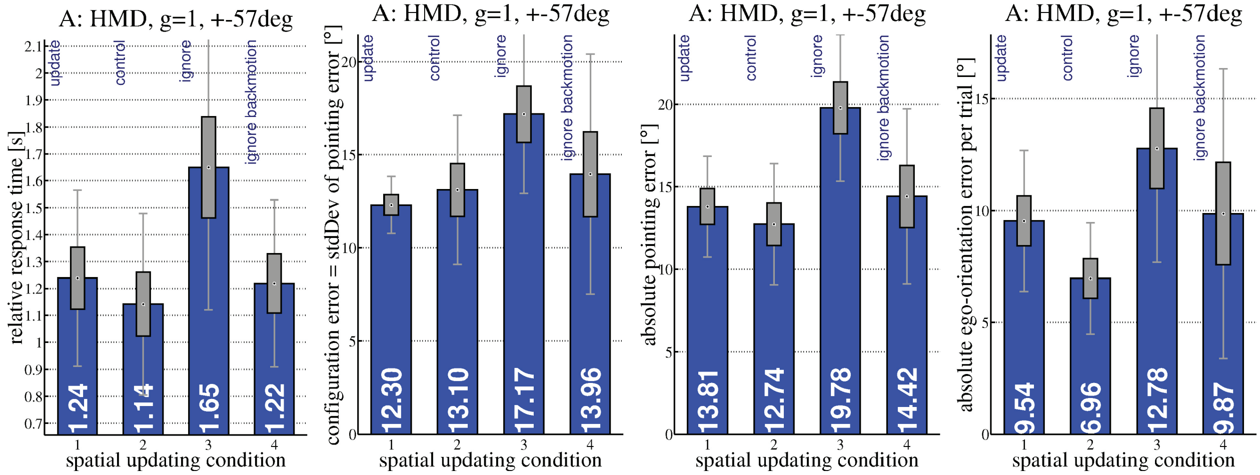
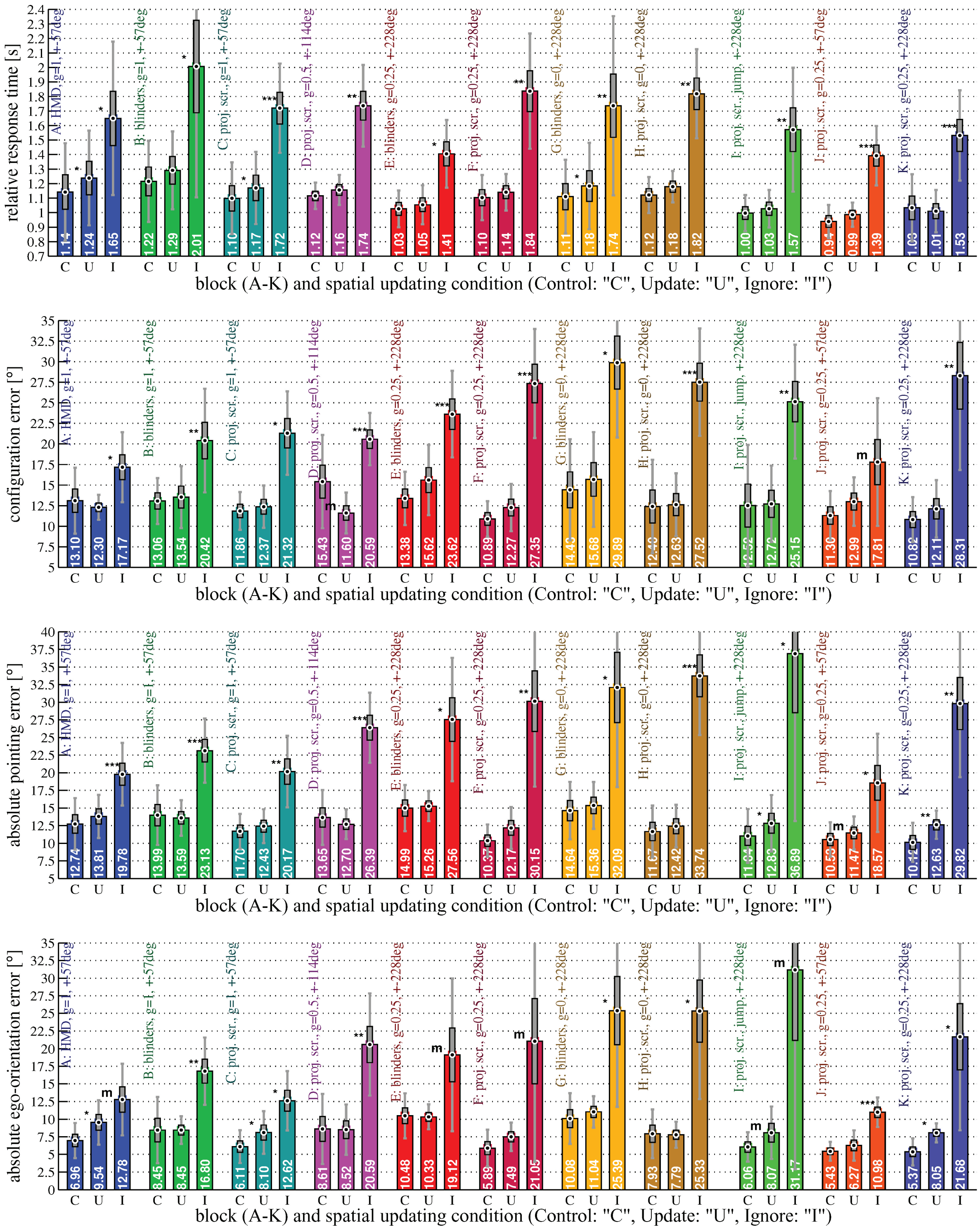


Fig. 9. Pointing performance in the HMD condition (block A) of the control experiment showing the typical response pattern for spatial updating. See Figure 3 for further explanation.

view again, did not decrease performance. Only the FOV had a consistent effect on pointing accuracy (blocks B versus C, E versus F, and G versus H). Reducing the FOV via blinders increased configuration error, absolute pointing error, and absolute ego-orientation error by up to 5°. This reduction in pointing accuracy in the baseline trials is even more pronounced than the reduction by 2–3° in the previous experiment, where the FOV was reduced from an *unlimited* FOV to the same blinders-limited FOV of $40^\circ \times 30^\circ$. This suggests that the more complex target configuration used in the current experiment did, indeed, lead to clearer differences between the different cue combinations. Compared to the previous experiment, however, we did not observe any clear response-time differences for the different FOVs. This might indicate that the response-time advantage for an unlimited FOV in the previous experiment was caused by the peripheral vision of the targets to point to. It could, however, also be caused by performance trade-offs between response time and pointing accuracy in the control experiment, as only the response time was fed back after each trial and participants might have focused more on responding quickly than on pointing accurately and consistently.

3.3.2 *Automatic Spatial Updating.* As in the previous experiment, automatic spatial updating was investigated by comparing UPDATE and CONTROL performance for the different cue combinations. The data are compiled in Figure 10, with the corresponding statistical analyses being summarized in Table IV. Figure 10 reveals a small response-time increase especially for the HMD condition (block A). All other dependent measures showed only marginal differences of typically less than 2.5° between UPDATE and CONTROL trials. Furthermore, there were apparently no clear differences between the different cue combinations whatsoever, which is corroborated by the lack of any significant differences in the pairwise t-test presented in Table IV. Taken together, this indicates that for all cue combinations tested, automatic spatial updating was almost as easy and accurate as baseline performance. Hence, we can conclude that photorealistic visual cues from a consistent, landmark-rich environment are under a wide range of simulation parameters sufficient to enable quick and accurate spatial updating. That is, neither the difference between HMD and blinder usage, nor the absence of any vestibular turn cues or any of the parameters FOV, gain factor, turning angle, movement velocity, or discontinuous (jumplike) updating (block I) reduced automatic spatial updating performance consistently. This is in accordance with the previous experiment, which also showed decent automatic spatial updating performance for


relative response time [s]
configuration error [°]
absolute pointing error [°]
absolute ego-orientation error [°]
block (A-K) and spatial updating condition (Control: "C", Update: "U", Ignore: "I")
A: HMD, g=1, +-57deg
B: blinders, g=1, +-57deg
C: proj. scr., g=1, +-57deg
D: proj. scr., g=0.5, +-114deg
E: blinders, g=0.25, +-228deg
F: proj. scr., g=0.25, +-228deg
G: blinders, g=0, +-228deg
H: proj. scr., g=0, +-228deg
I: proj. scr., jump, +-228deg
J: proj. scr., g=0.25, +-57deg
K: proj. scr., g=0.25, +-228deg

all conditions with useful visual information (blocks A–D) and no clear effect of the stimulus conditions for the difference between UPDATE and CONTROL trials. This again, indicates, the power and usability of good visual stimuli for spatial updating, and points toward a visual dominance.

In UPDATE trials of the jump condition (block I), the mental representation of the new orientation had to be instantiated before being able to perform the pointing task. Hence, the response-time difference of approximately 30 ms between UPDATE and CONTROL trials in the jump condition (block I) might be interpreted as the time needed to establish or reanchor a new egocentric representation. If this interpretation was true, 30 ms would indeed be rather quick for such a complex task like changing the orientation of the egocentric representation of the surrounding scene, even though the scene is well known and highly trained. This would call for a highly automatized updating of the egocentric representation. To address this effect directly, however, one might want to compare UPDATE trials to conditions with constant static visual information, i.e., without the in-between flashing of a different orientation as was done in the CONTROL condition.

3.3.3 *Obligatory Spatial Updating.* As in the previous experiment, the power of the available turn cues for turning our mental spatial representation was investigated by asking participants to ignore all turn stimuli and respond as if still being at the previous orientation. Only if spatial updating was indeed obligatory in the sense of being reflexlike and hard to suppress will IGNORE performance be considerably worse than UPDATE performance. The differences between IGNORE and UPDATE trials are graphically presented in Figure 10; the corresponding statistical analyses are compiled in Table IV. A first look at the figures shows that under all conditions, ignoring the turn stimuli was considerably harder than using them to UPDATE, as usual. That is, spatial updating was always obligatory and reflexlike. There were, however, systematic differences between the different cue combinations (blocks), which will be elaborated upon as follows. Each subsection is concerned with answering one main question, with the order of the questions being the same as the order of the statistical tests in Table IV.

3.3.3.1 *HMD Versus Blinders.* The visual cues were easier to ignore when using the HMD than the blinders (block B versus A). All four dependent variables point toward this direction, suggesting that visual cues presented via HMD are less convincing in triggering obligatory spatial updating and turning the mental spatial representation against our own conscious decision. To the best of our knowledge, this is the first study that directly compared spatial updating performance between an HMD and a projection screen. Hence, we can only speculate about the origin of the reduced performance for the HMD condition. Most studies seem to agree, however, that HMDs can cause a number of perceptual distortions. Most prominently, perceived distances are typically largely underestimated and compressed when presented through an HMD (see Creem-Regehr et al. [2003] for an excellent overview). A recent study found a similar compression and underestimation for rotations [Riecke et al. 2005]. Furthermore, using the same VR setup as in the current study, Riecke et al. [2005] showed that the FOV of the HMD was overestimated by more than 100%, where as the FOV of the blinders was estimated almost veridically. Even though we cannot pinpoint the exact reasons for the reduced spatial updating performance for the HMD in the current study, it seems evident that HMDs produce a number of undesirable and hardly understood perceptual distortions. Hence, care should be taken when using HMDs in tasks related to

←

Fig. 10. Spatial updating performance for the control experiment, plotted for the eleven different cue combinations (blocks A–K, color-coded) and the spatial updating conditions UPDATE ("U"), CONTROL ("C"), and IGNORE ("I"). The asterisks "*" between bars indicate whether the mean of these two bars differs significantly (on a 5, 0.5, or 0.05% significance level, using a two-tailed t-test; an "m" indicates a marginal (10%) significance). Note that spatial updating was automatic and obligatory for all conditions and dependent variables: UPDATE performance was comparable to CONTROL performance, whereas IGNORE performance was considerably impaired.

spatial orientation. Together with the investigation by Riecke et al. [2005], the current study provides some initial evidence that curved projection screens, even when viewed through a vision delimiting device, might be more suitable for presenting ego-motions and enabling good spatial orientation in virtual environments than HMDs.

3.3.3.2 *Influence of FOV.* Comparing block E and F reveals that turns presented *without* blinders were considerably harder to ignore (increased response times) and resulted in an increased configuration error, compared to turns with blinders-restricted FOV. That is, the increased FOV seems to render spatial updating more obligatory and reflexlike. This effect was less pronounced for purely visual turns (blocks G versus H). Smaller turns and a gain factor of 1, on the other hand, did not result in any consistent FOV effect (blocks B versus C), potentially due to the simpler task.

3.3.3.3 *Influence of Visuovestibular Gain Factors and Vestibular Cues.* Comparing spatial updating for different gain factors did not show any clear effect, suggesting that the presence of concurrent vestibular turn stimuli is not required as long as a consistent, landmark-rich visual scene is presented (blocks B versus E, B versus G, E versus G, F versus H, and B versus J, see Table IV). This might be interpreted in the direction that "good" visual cues alone *can* be fully sufficient for initiating obligatory spatial updating and, hence, enabling good and robust spatial orientation.

3.3.3.4 *Influence of Turning Angle and Movement Velocity.* Comparing blocks with different yaw ranges showed increased response times, configuration errors, absolute pointing errors, and absolute ego-orientation errors for larger turns (blocks C versus F, C versus H, and J versus F). Not all of these differences reached statistical significance (see Table IV), which might, of course, be due to the limited number of participants. It nevertheless provides a first hint that the turning angle might be a critical variable, as larger turns seem to be considerably harder to ignore than smaller turns. Comparing block J and K reveals considerable performance deficits in all four measurands for block K, where both turning angles and response times were four times larger than in block J, but the average movement duration was the same. Movement velocity is most likely not the cause of the difference, as block F and K differed only in movement velocity and showed virtually the same performance.[2] This corroborates the above argument that larger turning angles are, indeed, harder to ignore and, hence, lead to more obligatory spatial updating.

3.3.3.5 *Continuous Versus Discontinuous (Jumplike) Motions.* Block I investigated spatial updating performance without any explicit or apparent motion cues whatsoever. We have already seen that this slide-show type presentation of new orientations was as efficient in enabling automatic spatial updating as comparable conditions (e.g., block K). However, will it be sufficient to trigger obligatory spatial updating as well?

Comparing IGNORE and UPDATE performance for block I and K reveals the same response-time increase of more than 500 ms for the IGNORE task, indicating that the IGNORE task is quite hard and rather time-consuming. The configuration error was slightly less pronounced in the discontinuous (jump) condition, whereas the absolute pointing error as well as the absolute ego-orientation error were more pronounced. However, none of these differences reached statistical significance (see Table IV). Taken together, this implies that discontinuous, slide-show type presentations of new orientations are virtually as successful in triggering obligatory spatial updating as are continuous motions to the new orientation. This finding was rather unexpected and awaits further investigations to allow for a comprehensive explanation (cf. section 5).

[2]The observed difference in response time is most likely due to the previously mentioned overall improvement in response time over the course of the experiment.

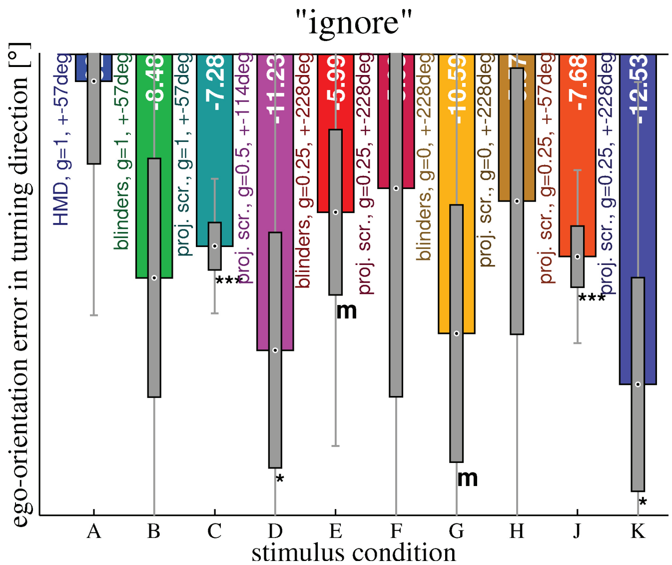


Fig. 11. Ego-orientation error in turing direction for the Ignore trials of the control experiment. Note the considerable systematic error in all but the HMD condition. Participants responded as if the to-be-ignored rotation was larger than it actually was, which might be interpreted as an overcompensation of the to-be-ignored motion.

3.3.3.6 *Ego-Orientation Error in Turning Direction.* For the continuous motions in block J and K, where participants were highly trained on the task, the ego-orientation error against turning direction (see Figure 11) explains a considerable part of the absolute pointing errors and absolute ego-orientation errors. This implies that even though participants might have improved their overall pointing accuracy somewhat, they were nevertheless unable to correctly remember their previous orientation, especially when exposed to rapid and large turns. This occurred even though they had already been exposed to the Ignore task for several hours and were highly trained with explicit feedback, as they were always moved back to the previous orientation in the following Ignore Backmotion trial.

The most obvious hypothesis about the underlying processes would be that participants tried to remember the previous orientation by remembering the landmark they were facing. Many participants reported, in fact, using this strategy. The observed systematic ego-orientation error of about 13° (Block K) was, however, almost as large as the mean angular separation of the targets (about 16°), indicating that participants' judged previous ego-orientation was closer to the wrong landmark (i.e., the landmark *adjacent* to the one faced) than to the correct one (i.e., the landmark previously faced). Hence, the underlying processes remain unclear. No matter what the underlying reasons are, the Ignore trials did show significant ego-orientation errors against turning direction in both this experiment and the main experiment. As the representational momentum literature has to our knowledge so far only been concerned with object or object array motions, the observed self-motion effect might be an interesting extension that could provide further insights into the underlying mechanisms and awaits further investigation.

3.3.4 *Learning Effect.* To quantify potential performance improvements over time within the first eight blocks of the experiment, correlation analyses were performed between all four dependent variables and the block order for the Update, Control, and Ignore conditions. Only the response time was significantly negatively correlated with block number. This was found for both the Update and Ignore conditions, but not the CONTROL condition.[3] As the only consistent overall learning effect was a decrease in response time in the UPDATE and IGNORE conditions, comparisons between the first eight blocks and the three subsequent blocks seem legitimate for the other four dependent variables and for the CONTROL condition in general. To double-check the comparability, the last two blocks of the control study [i.e., the tenth and eleventh block (either J or K)] were also compared using two-sided t-tests. The statistical analysis showed no significant effect of block order for any of the dependent variables ($p > 0.05$), suggesting that all order effects and learning curves had settled by that time.

4. GENERAL DISCUSSION AND CONCLUSIONS

4.1 Main Experiment

The rapid pointing paradigm introduced in the main experiment allowed response times and accuracies below the values typically observed in the literature, indicating the ease and intuitive usability of our pointing device and the usefulness for spatial updating research. For all conditions with useful visual cues (Block A–D), the typical response pattern for obligatory spatial updating was observed: UPDATE performance was almost as good as CONTROL performance, whereas IGNORE performance was considerably worse. This confirms that spatial updating can be reliably quantified with our rapid-pointing paradigm.

This response pattern was found irrespective of concurrent vestibular motion cues, indicating that visual cues alone were sufficient to elicit reflexlike obligatory spatial updating. Furthermore, performance in VR was about as good as performance in its real-world counterpart (as long as the FOV was the same). That is, a simulated, photorealistic view onto a consistent, landmark-rich environment was almost as powerful in turning our mental spatial representation against our own conscious will as a corresponding view onto the real-world. This highlights the power and flexibility of using highly photorealistic VR for investigating human spatial orientation and spatial cognition. Last, but not least, it validates our VR-based experimental paradigm and demonstrates the generalizability of results obtained in this VR setup to comparable real-world tasks.

4.2 Control Experiment

The control experiment was designed to cross-validate as well as extend findings from the main experiment by exploring the influence of further motion and simulation parameters on spatial updating performance. The baseline (CONTROL) condition revealed a performance decrease for reducing the FOV, thus confirming results from the main experiment.

Automatic spatial updating, quantified as the difference between UPDATE and CONTROL performance, however, was completely independent of all parameters varied in the control experiment. That is, as long as participants were presented with a photorealistic view onto the well-known Tübingen market place in VR, they could readily adopt the new orientation and knew immediately where the currently nonvisible landmarks were. It did not matter how the scene was presented, how large the FOV was, or how far they were moved. Even the complete absence of any concurrent vestibular turn stimuli or the discontinuous, jumplike presentation of a new orientation did not prevent participants from being able to indicate quickly and accurately where the surrounding objects of interest were.

Moreover, when participants were asked to ignore all visual cues and point as if not having turned, they were virtually unable to do so successfully: Whenever visual cues were provided, spatial updating proved to be obligatory to the extent that the simulated movement was considerably harder to IGNORE than to UPDATE, corroborating the findings from the main experiment.

[3] $r = -0.22, r^2 = 0.047, t(62) = 1.75, p = 0.042^*$ for the UPDATE condition and $r = -0.23, r^2 = 0.051, t(62) = 1.82, p = 0.037^*$ for the IGNORE condition.

This difficulty of ignoring the visual scene increased with larger turning angles and for the larger FOV. Conversely, this suggests that a reduced FOV, especially when using an HMD, seems to render the visual motion stimulus less convincing. Actually, all blocks with the same movement range and the same, large FOV (i.e., blocks F, G, H, and K of the control experiment) showed virtually the same performance. That is, if the FOV is large enough, other factors like movement velocity, gain factor, and even discontinuous (jumplike) versus continuous motions do not seem to influence spatial updating performance. This argues against the usage of HMDs and any display with rather limited FOV for applications involving simulated movements of the observer. This supports results from other studies that found similar drawbacks and considerable systematic errors when using HMDs for executing simulated self-rotations [Bakker et al. 1999, 2001; Riecke et al. 2005] as well as for simple navigation tasks [Kearns et al. 2002]. This might also explain the small performance differences observed between the real-world and HMD condition of the main experiment.

Combining the results from the control experiment with the findings from the main experiment, we can conclude that photorealistic visual stimuli of consistent, landmark-rich scenes are clearly sufficient for enabling excellent automatic spatial updating as well as triggering obligatory spatial updating, especially when presented through a large FOV. This was found irrespective of concurrent vestibular turn cues. Even discontinuous, slide-show like presentation of new orientations was sufficient and yielded virtually the same, excellent performance. On the other hand, vestibular cues from smooth turns alone were clearly incapable of initiating obligatory spatial updating, as they could as easily be ignored as they could be used to deliberately update to the new orientations.

This result conflicts with the prevailing opinion that vestibular cues are required or even sufficient for proper updating of ego-turns. Several factors might explain this difference: primarily the immersiveness of our visualization setup and the abundance of natural landmarks embedded in a well-known scene. Furthermore, in the blocks without useful visual cues, the consecutive motions without in-between visibility of the scene might have led to some disorientation, and it is conceivable that participants' mental representation of the scene somehow faded or was no longer strong and well-anchored.

5. CONTINUOUS VERSUS INSTANT-BASED SPATIAL UPDATING

Apart form the well-known smooth spatial updating induced by *continuous* movement information, the control experiment also demonstrated a *discontinuous*, jumplike spatial updating that allowed participants to quickly adopt a new orientation at a given instant of time, without any explicit motion cues and independent of the previous orientation. These slide-show type presentations of new orientations were even sufficient to trigger obligatory, reflexlike spatial updating. Even though further experiments are required to further investigate and corroborate this effect, it already provides a considerable challenge to the prevailing understanding of spatial updating, which cannot convincingly explain the observed data.

Hence, we propose a refinement to our concept of spatial updating by introducing a distinction between the classical *continuous spatial updating* known from the blindfolded spatial updating literature and a second, complementary spatial updating process that we termed "*instant-based spatial updating*" that is able to explain the jump condition in the control experiment. In the following, we will elaborate upon this distinction and will try to argue why it might actually make sense to have two spatial updating mechanisms running in parallel.

In the literature, spatial updating is typically investigated in situations that have no reliable landmarks usable for position-fixing, e.g., by blindfolding participants or displaying optic flow only [Farrell and Robertson 1998; Klatzky et al. 1998; May and Klatzky 2000; May 2000; Presson and Montello 1994; Rieser et al. 1982; Rieser 1989; Simons and Wang 1998; Wang and Simons 1999]. Only recently were visual landmark cues integrated in human spatial updating research (e.g., Christou et al. [1999];

Wang and Spelke [2000]; Wraga et al. [2004]). Without any available landmarks, only relative or motion information can be used for spatial updating. In blindfolded navigation studies, for example, velocity and acceleration information from the vestibular systems can be used to continuously update the mental spatial representation using path integration. Even proprioception generally provides only relative movement information (e.g., the number of steps traveled). That is, the body senses provide primarily relative information and are thus prone to accumulation errors during path integration. Nevertheless, the literature shows clearly that vestibular and proprioceptive cues are, under many conditions, sufficient to enable automatic spatial updating. As this process is essentially based on path integration, any interruption or impairment due to, e.g., high cognitive load or distractions could potentially yield a completely misaligned mental representation, which would then be useless. Thus, it seems natural to propose a spatial updating process that operates continuously and autonomously and, thus, needs to be highly automated. This is what we refer to as **continuous spatial updating**. Any discontinuity in a spatial updating process based on relative spatial information or motion cues would critically disrupt its usability and reliability. Hence, one might argue that this continuous spatial updating should also be obligatory in the sense of reflexlike and hard-to-suppress.

As continuous spatial updating is based on path integration and leads to increasing alignment errors over time, it seems sensible to propose a second process that can automatically reanchor the potentially misaligned mental representation to the physical surround. We would like to introduce the term *instant-based spatial updating* to refer to this process.

Returning to our initial example, what happens when you are at home at night when the main fuse blows? When walking around in complete darkness trying to find the fuse box or some light source, we become increasingly uncertain about our current ego-position. That is, we still have some intuitive feeling of where we are, but we would not bet a lot on the exact location. The situation changes as soon as we can perceive the location of a known landmark. This instantaneous position-fixing could occur via different sensory modalities: Auditorily, for example the phone could be ringing; haptically, we might touch or run into the kitchen table; visually, somebody else might already have replaced the fuse, or lightning might have lit the room for a fraction of a second. That is, any clearly identifiable spatial cue (landmark) could reanchor our mental representation in an instant, without much cognitive effort or time needed.

This process of automatically realigning or reanchoring our egocentric mental representation to the surround is what we refer to as **instant-based spatial updating**. We use the term *spatial updating* to specify the automatized nature of the transformation process and to distinguish it from more general (nonautomatized) reorientation or self-alignment processes. The term *instant-based* spatial updating was chosen because the process is based on what is perceived at a given instant of time. That is, instant-based spatial updating does not require knowledge of our previous position or orientation with respect to our surround and thus operates, in a sense, in a history-free manner. This is a fundamental difference to continuous spatial updating, which is based on knowledge of the previous position/orientation and a relative motion signal (typically velocity or acceleration) that can be continuously integrated over time to yield a current estimate of the self-to-surround relationships. This distinction between continuous and instant-based spatial updating has recently been incorporated into a theoretical spatial orientation framework [Riecke and von der Heyde 2002].

When locomoting under full-cue conditions, this instant-based spatial updating probably occurs at any instant in time and is thus indistinguishable from continuous spatial updating, as both processes are in close agreement and complement each other. Moreover, they can be considered as a mutual back-up system for the case that one of them fails or does not receive sufficient information. As pointed out earlier, blind navigation is a prototypical example where continuous spatial updating needs to bridge the potentially large gap between possible realignments using instant-based spatial updating.

Conversely, being a passenger in a bus driving over cobblestones, for example, the constant jerks and vibrations render continuous spatial updating by vestibular and proprioceptive cues utterly useless for navigation. Hence, the visual cues will most likely be used for repeated position-fixing via instant-based spatial updating. This redundancy of having two spatial updating processes running in parallel whenever possible is thus quite useful and improves the overall robustness and reliability of human spatial orientation.

Our distinction between continuous and instant-based spatial updating bears some resemblance to Kosslyn's distinction between "shift transformations" and "blink transformations," respectively [Kosslyn 1994]. Shift transformations are responsible for smooth and seemingly continuous transformations of mental images, like object translations and rotations. If, however, "an image object must be transformed a large amount, the image may be allowed to fade and a new one is generated" [Kosslyn 1994, p. 402], which Kosslyn refers to as "blink transformation." Note that shift and blink transformations refer to mental object image transformations that are continuous and discontinuous, respectively. Continuous and instant-based spatial updating, on the other hand, refer to the transformation of the complete mental egocentric spatial representation (i.e., a "scene"), which involves a change in the observer's position and/or orientation. Furthermore, spatial updating is normally automated and reflexlike (obligatory), whereas Kosslyn's image transformations are typically deliberate cognitive processes (i.e., neither automatic nor obligatory). These fundamental differences might explain the often found advantage of self-motions over object motions for the updating of physical as well as imagined rotations [Simons and Wang 1998; Wang and Simons 1999; Simons et al. 2002; Wraga et al. 1999a, 2000, 2004].

We hope that the results presented in this paper and our distinction between continuous and instant-based spatial updating, in particular, will enable us to further elucidate the amazing capabilities of human spatial orientation. Ultimately, knowing what sensory cues are required and/or sufficient for enabling robust spatial orientation and spatial updating, we can apply this knowledge by focusing on simulating those cues in ego-motion simulations with high fidelity, thus devising a perceptually oriented ego-motion simulation paradigm that is truly lean and elegant.

ACKNOWLEDGMENTS

This research was funded by the Max-Planck Society, the EU grant POEMS-IST-2001-39223 (Perceptually Oriented Ego-Motion Simulation, see www.poems-project.info), and the Deutsche Forschungsgemeinschaft (SFB 550 "Erkennen, Lokalisieren, Handeln: neurokognitive Mechanismen und ihre Flexibilität").

REFERENCES

AMORIM, M. A. AND STUCCHI, N. 1997. Viewer- and object-centered mental explorations of an imagined environment are not equivalent. *Cognit. Brain Res. 5*, 3 (Mar.), 229–239.

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. Available: ftp://ftp.cs.unc.edu/pub/publications/techreports/00-019.pdf.

BAKKER, N. H., WERKHOVEN, P. J., AND PASSENIER, P. O. 1999. The effects of proprioceptive and visual feedback on geographical orientation in virtual environments. *Presence—Teleoperators and Virtual Environments 8*, 1 (Feb.), 36–53.

BAKKER, N. H., WERKHOVEN, P. J., AND PASSENIER, P. O. 2001. Calibrating Visual Path Integration in VEs. *Presence—Teleoperators and Virtual Environments 10*, 2 (Apr.), 216–224.

BATSCHELET, E. 1981. *Circular statistics in biology*. Academic Press, London.

CARPENTER, M. AND PROFFIT, D. R. 2001. Comparing viewer and array mental rotations in different planes. *Memory & Cognition 29*, 3 (Apr.), 441–448.

CHANCE, S. S., GAUNET, F., BEALL, A. C., AND 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 and Virtual Environments* 7, 2 (Apr.), 168–178.

Christou, C., Tjan, B., and Bülthoff, H. 1999. Viewpoint information provided by familiar environment facilitates object identification. Tech. Rep. 68, Max-Planck Institut für biologische Kybernetik. Available: www.kyb.mpg.de /publication.html?publ=1536.

Cobb, S. V. G., Nichols, S., Ramsey, A., and Wilson, J. R. 1999. Virtual reality-induced symptoms and effects (VRISE). *Presence—Teleoperators and Virtual Environments 8*, 2 (Apr.), 169–186.

Creem, S. H. and Proffitt, D. R. 2000. Egocentric measures of spatial updating: Is there an advantage for action? Poster presented at the *41st Meeting of the Psychonomic Society, New Orleans, LA*.

Creem, S. H., Wraga, M., and Proffitt, D. R. 2001. Imagining physically impossible self-rotations: Geometry is more important than gravity. *Cognition 81*, 1 (Aug.), 41–64.

Creem-Regehr, S. H., Willemsen, P., Gooch, A. A., and Thompson, W. B. 2003. The influence of restricted viewing conditions on egocentric distance perception: Implications for real and virtual environments. Tech. Rep. UUCS-03-016, School of Computing, University of Utah, Salt Lake City, USA. Available: www.cs.utah.edu/techreports/2003/pdf/UUCS-03-016.pdf.

Easton, R. D. and Sholl, M. J. 1995. Object-array structure, frames of reference, and retrieval of spatial knowledge. *J. Exp. Psychol.-Learn. Mem. Cogn. 21*, 2 (Mar.), 483–500.

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

Farrell, M. J. and Thomson, J. A. 1998. Automatic spatial updating during locomotion without vision. *Q. J. Exp. Psychol. Sect A-Hum. Exp. Psychol. 51*, 3 (Aug.), 637–654.

Farrell, M. J. and Robertson, I. H. 2000. The automatic updating of egocentric spatial relationships and its impairment due to right posterior cortical lesions. *Neuropsychologia 38*, 5, 585–595.

Hettinger, L. J., Nelson, W. T., and Haas, M. W. 1996. Target detection performance in helmet-mounted and conventional dome displays. *International Journal of Aviation Psychology 6*, 4, 321–334.

Hollins, M. and Kelley, E. K. 1988. Spatial updating in blind and sighted people. *Percept. Psychophys. 43*, 4 (Apr.), 380–388.

Howarth, P. A. and Costello, P. J. 1997. The occurrence of virtual simulation sickness symptoms when an HMD was used as a personal viewing system. *Displays 18*, 2 (Dec.), 107–116.

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

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

Kosslyn, S. M. 1994. *Image and brain: The resolution of the imagery debate*. MIT Press, Cambridge, MA.

Loomis, J. M., Da Silva, J. A., Philbeck, J. W., and Fukusima, S. S. 1996. Visual perception of location and distance. *Current Directions in Psychological Science 5*, 3 (Jan.), 72–77.

May, M. 1996. Cognitive and embodied modes of spatial imagery. *Psychologische Beiträge 38*, 418–434.

May, M. 2000. *Kognition im Umraum*. Studien zur Kognitionswissenschaft. DUV: Kognitionswissenschaft, Wiesbaden.

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

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

Mon-Williams, M. and Wann, J. P. 1998. Binocular virtual reality displays: When problems do and don't occur. *Human Factors 40*, 1 (Mar.), 42–49.

Nelson, W. T., Hettinger, L. J., Cunningham, J. A., Brickman, B. J., Haas, M. W., and McKinley, R. L. 1998. Effects of localized auditory information on visual target detection performance using a helmet-mounted display. *Human Factors 40*, 3 (Sept.), 452–460.

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

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

Riecke, B. E. and von der Heyde, M. 2002. Qualitative modeling of spatial orientation processes using logical propositions: Interconnecting spatial presence, spatial updating, piloting, and spatial cognition. Tech. Rep. 100, MPI for Biological Cybernetics. Avaliable: www.kyb.mpg.de/publication.html?publ=2021.

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

RIECKE, B. E., VON DER HEYDE, M., AND BÜLTHOFF, H. H. 2002b. Spatial updating in virtual environments: What are vestibular cues good for? *Journal of Vision 2*, 7 (May), 421a. http://journalofvision.org/2/7/421/ Available: www.kyb.mpg.de /publication.html?publ=628.

RIECKE, B. E., SCHULTE-PELKUM, J., AND BÜLTHOFF, H. H. 2005. Perceiving simulated ego-motions in virtual reality—comparing large screen displays with HMDs. In *SPIE—Invited paper on VALVE: Vision, Action, and Locomotion in Virtual (and Real) Environments* (San Jose CA). Available: www.kyb.mpg.de/publication.html?publ=3233.

RIESER, JOHN, J., GUTH, DAVID, A., AND HILL, EVERETT, W. 1982. Mental processes mediating independent travel: Implications for orientation and mobility. *Journal of Visual Impairment and Blindness 76*, 6 (Jun.), 213–218.

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

SHELTON, A. L. AND MCNAMARA, T. P. 2001. Systems of spatial reference in human memory. *Cognitive Psychology 43*, 4 (Dec.), 274–310.

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

SIMONS, D. J. AND WANG, R. F. 1998. Perceiving real-world viewpoint changes. *Psychol. Sci. 9*, 4 (Jul.), 315–320.

SIMONS, D. J., WANG, R. X. F., AND RODDENBERRY, D. 2002. Object recognition is mediated by extraretinal information. *Perception & Psychophysics 64*, 4 (May), 521–530.

STANNEY, K. M., MOURANT, R. R., AND KENNEDY, R. S. 1998. Human factors issues in virtual environments: A review of the literature. *Presence—Teleoperators and Virtual Environments* 7, 4 (Aug.), 327–351.

VON DER HEYDE, M. 2000. The Motion-Lab—A Virtual Reality Laboratory for Spatial Updating Experiments. Tech. Rep. 86, Max Planck Institute for Biological Cybernetics, Tübingen, Germany. Dec. Available: www.kyb.mpg.de/publication.html?publ=633.

WANG, R. X. F. AND SIMONS, D. J. 1999. Active and passive scene recognition across views. *Cognition 70*, 2 (Mar.), 191–210.

WANG, R. X. F. AND SPELKE, E. S. 2000. Updating egocentric representations in human navigation. *Cognition 77*, 3 (Dec.), 215–250.

WRAGA, M., CREEM, S. H., AND PROFFITT, D. R. 1999a. The influence of spatial reference frames on imagined object- and viewer rotations. *Acta Psychol. 102*, 2–3 (Sept.), 247–264.

WRAGA, M., CREEM, S. H., AND PROFFITT, D. R. 1999b. Spatial updating of an irregularly shaped virtual array during self- and display rotations. *Invest. Ophthalmol. Vis. Sci. 40*, 4 (Mar.), 1.

WRAGA, M., CREEM, S. H., AND PROFFITT, D. R. 2000. Updating displays after imagined object and viewer rotations. *J. Exp. Psychol.-Learn. Mem. Cogn. 26*, 1 (Jan.), 151–168.

WRAGA, M., CREEM-REGEHR, S. H., AND PROFFITT, D. R. 2004. Spatial updating of virtual displays during self- and display rotation. *Memory & Cognition 32*, 3 (Apr.), 399–415.

YARDLEY, L. AND HIGGINS, M. 1998. Spatial updating during rotation: The role of vestibular information and mental activity. *J. Vestib. Res.-Equilib. Orientat. 8*, 6 (Nov.-Dec.), 435–442.

Received November 2004; revised ????; accepted February 2005