Memory for navigable space is flexible and not restricted to exclusive local or global memory units

Marianne Strickrodt[1,2], Heinrich H. Bülthoff[1], Tobias Meilinger[1]

[1]Max-Planck-Institute for Biological Cybernetics, Tübingen

[2]Graduate Training Centre of Neuroscience, Tübingen University

Author Note

Correspondence concerning this article should be addressed to Marianne Strickrodt, Max Planck Institute for Biological Cybernetics, Max-Planck-Ring 8, D-72076 Tübingen, Germany. E-mail: marianne.strickrodt@tuebingen.mpg.de. Phone: +49 (0) 7071 601 615. Fax: +49 (0) 7071 601 616

Abstract

Objects learned within single enclosed spaces (e.g., rooms) can be represented within a single reference frame. Contrarily, the representation of navigable spaces (multiple interconnected enclosed spaces) is less well understood. In this study we examined different levels of integration within memory (local, regional, global), when learning object locations in navigable space. Participants consecutively learned two distinctive regions of a virtual environment that eventually converged at a common transition point and subsequently solved a pointing task. In Experiment 1 pointing latency increased with increasing corridor distance to the target and additionally when pointing into the other region. Further, when pointing within a region alignment with local and regional reference frames, when pointing across regional boundaries alignment with a global reference frame was found to accelerate pointing. Thus, participants memorized local corridors, clustered corridors into regions and formed global representations of the entire environment. Introducing the transition point at the beginning of learning each region in Experiment 2 caused previous region effects to vanish. Our findings emphasize the importance of locally confined spaces for structuring spatial memory and suggest that the opportunity to integrate novel into existing spatial information early during learning may influence unit formation on the regional level. Further, global representations seem to be consulted only when accessing spatial information beyond regional borders. Our results are inconsistent with conceptions of spatial memory for large scale environments based either exclusively on local reference frames or upon a single reference frame encompassing the whole environment, but rather support hierarchical representation of space.

Keywords: spatial memory; reference frames; hierarchical models of spatial memory; levels of spatial integration

Introduction

When learning a new environment, for example, after moving to a new city, spatial memory of our surrounding is gathered successively. Each time we leave the house we visit multiple places sequentially, thus, we are confronted with chunks of spatial information over time. One day we leave our home and walk to the market place, thereby passing the bank, the pharmacy and the bakery. Yet, on another day, we take the bus or subway to the remote book shop in another borough of the city. From there we stroll through the city, passing a café, then a butcher, until we, unexpectedly, end up at the market place again. Not only our everyday experience but also scientific evidence shows that we are quite able to learn the relationship between two (or more) separately learned neighborhoods. Schinazi, Nardi, Newcombe, Shipley, and Epstein (2013), for example, who had people learn two separate routes first and introduced the connecting route in a second learning session, found a rapid learning of between route constellation of landmarks (see also Ishikawa & Montello, 2006; Weisberg, Schinazi, Newcombe, Shipley, & Epstein, 2014). What remains unclear: How do we deal with those chunks of information? How do we relate the locations of pharmacy and café without having experienced them coincidentally? These questions are addressed within the present work.

There are theories postulating that we rely on a network of locally confined places when representing navigable space, such as buildings or cities (Chrastil & Warren, 2014; Meilinger, 2008; Warren, Rothman, Schnapp, & Ericson, 2017). Chrastil and Warren (2014), for example, postulate the use of a labelled graph. Places are represented as nodes and the connectivity between places are represented as edges. This graph is labelled with additional, metric information, such as distances between the places and angles between the edges starting from the same node. In the graph model of Meilinger (2008) nodes represent reference frames limited to vista spaces, i.e., places one can see from a single vantage point (e.g., a street, a room) (Montello, 1993). Object locations and surrounding geometry within a vista space are stored relative to this local reference frame. The edges connecting local reference frames are specified perspective shifts (distance and angles). In both theories, a complex navigable space is represented without the need for a coarser, global integration that is stored in long-term memory, but with the help of local information only. To point to distant targets from memory local information are integrated during the recall process online within working memory to, for example, form a transient reference frame incorporating one's current position and the target

(Meilinger, 2008). Following these theories new connected routes are simply added to the existent graph representation by adding new nodes and edges.

There is ample evidence that immediately visible surroundings (i.e., vista spaces) are core units when memorizing and dealing with navigable space. Updating of objects located in a room is disrupted when leaving this room (Wang & Brockmole, 2003). Visual borders clustering an object array seem to likewise cluster the representation of the array (e.g., Kosslyn, Pick, & Fariello, 1974; Marchette, Ryan, & Epstein, 2017; Meilinger, Strickrodt, & Bülthoff, 2016). Also, when participants learn the layout of objects that are spread across multiple streets or corridors (i.e., multiple interconnected vista spaces) evidence was found that multiple reference frames are memorized, which are defined by and confined to the immediately visible surrounding of a street and/or corridor (Meilinger, Riecke, & Bülthoff, 2014; Werner & Schmidt, 1999). Spatial memory based on reference frames is usually detected by utilizing one key characteristic of reference frames, namely, the orientation dependency of spatial memory. Accessing relative directions between object locations that are stored in memory is faster and more accurate when one is (physically or mentally) aligned with certain orientations compared to other orientations. This is explained by encoding objects relative to one or more orthogonal reference axes. Alignment with an axis allows for a rather effortless retrieval, while being misaligned requires costly additional transformations (McNamara, Sluzenski, & Rump, 2008; Mou, McNamara, Valiquette, & Rump, 2004). Meilinger, Riecke and Bülthoff (2014) showed that participants that learned a rather complex virtual environment consisting of seven corridors performed best in a subsequent pointing task when they were bodily aligned with the initial view within each single corridor. This indicates that local memory units were stored, more precisely, in the form of one reference frame for each corridor.

An alternative approach to spatial memory structures based on mere local units are hierarchical conceptions of spatial memory (e.g., Mallot & Basten, 2009; Stevens & Coupe, 1978; Wiener & Mallot, 2003). They do not oppose concepts based on mere local units but extend them. They assume multiple levels of representation and start from the premise that local spatial entities (places) can be subsumed under another memory unit on a superordinate hierarchical level to form a region, for instance. There is evidence that we indeed represent regions and use this information for spatial tasks. Route decisions were shown to be made according to the least number of transitions between predefined regions (e.g., Schick, Halfmann,

& Mallot, 2015; Wiener & Mallot, 2003) and judgements about the relative position of two remote cities (e.g., San Diego in California and Reno in Nevada) was found to be based on the relative position of the two federal states the cities are in (e.g., California is west of Nevada), rather than the actual location of the cities (e.g., San Diego, California is east of Reno, Nevada) (Stevens & Coupe, 1978). Such studies, however, remain uninformative about the format of this superordinate memory unit. What do we store on the higher level of hierarchy? Do we store conceptual or topological knowledge (Hirtle & Mascolo, 1986; Meilinger, 2008; Wang & Brockmole, 2003, 2003; Wiener & Mallot, 2003), or do we embed multiple vista spaces that form a region into a new metric relational scheme, another superordinate spatial reference frame (e.g., McNamara, Sluzenski, & Rump, 2008; Meilinger & Vosgerau, 2010)?

In fact, it was shown that reference frames can spread across multiple adjacent vista spaces. For example, Wilson, Wilson, Griffiths and Fox (2007) had people learn object locations within a simple three-corridor environment, containing four target objects. A subsequent pointing task provided evidence for a global reference frame covering all three corridors. Irrespective of the corridor participants were in, they performed best when aligned with the perspective of the very first corridor compared to other orientations. Thus, there seem to be circumstances under which spatial information from multiple enclosed vista spaces are aggregated and integrated into one reference frame (see also Meilinger, Frankenstein, Watanabe, Bülthoff, & Hölscher, 2015; Richardson, Montello, & Hegarty, 1999; Tlauka, Carter, Mahlberg, & Wilson, 2011).

An interesting study by Greenauer and Waller (2010) showed that people seem to build not only micro- but also macroreference frames when learning two arrays of objects within the same room (vista space). Depending on whether participants pointed within one object array or within the other, orthogonally-aligned object array or whether they pointed across arrays, different body perspectives were identified to elicit best pointing performance. In the case of within-array pointing best pointing performance was found when participants were aligned with the geometry (i.e., main axis) of the respective object array pointing was currently performed in. This indicates that two orthogonally aligned microreference frames were formed, one reference frame for each object array, and used when the current position and the target object both are located within the same microreference frame. In contrast, when the participants imagined themselves to be standing within one object array, but the target was located in the other object array another, single orientation was found to elicit best pointing performance. In these across-array pointing trials evidence for a macroreference frame was found that aligned with the salient axis of the between-array geometry.

This indicates that a hierarchy of reference frames can be formed for chunks of spatial information presented in a vista space. For spaces extending a vista space a hierarchy of reference frames was discussed in the literature (McNamara et al., 2008; Meilinger & Vosgerau, 2010) assuming that reference frames for vista spaces can be grouped together to form distinct regions which are integrated to form distinct regional reference frames on a higher level of the hierarchy. However, it was not yet experimentally tested for.

Altogether, past research produced results that indicate that local vista spaces serve as memory units for navigable space (e.g., Marchette, Ryan, & Epstein, 2017; Meilinger, Riecke, & Bülthoff, 2014; Meilinger, Strickrodt, & Bülthoff, 2016), but likewise can be aggregated and integrated into a single reference frame (e.g., Meilinger, Frankenstein, Watanabe, Bülthoff, & Hölscher, 2015; Wilson, Wilson, Griffiths, & Fox, 2007). Additionally, regions that are clustering navigable space are stored and used for route planning (e.g., Wiener & Mallot, 2003) and relational judgements (e.g., Stevens & Coupe, 1978). Therefore, in this study we investigate whether spatial memory acquired in navigable space is stored on multiple levels, similar to the micro- and macroreference frames found in vista spaces (Greenauer & Waller, 2010). To examine this, we built a clustered virtual environment consisting of two obliquely aligned regions. Corridors within a region were similar in many attributes (e.g., color, category of landmarks, distance, see Materials) while being maximally different from the attributes of the other region. Assuming a hierarchy of reference frames opens two questions: First, how many hierarchical levels are formed, and which areas do units on each level comprise? For example, multiple *local* units might be stored on a single level where each reference frame is limited to a single corridor (vista space). Corridors belonging to the same region might be aggregated and integrated to form a *regional* memory unit on a superordinate level. It is also possible that a *global* memory unit is formed that spans across the entire environment, thus, comprising all spatial information from all encountered vista spaces of the environment within a single reference frame. Different combinations of local corridor units, regional units or a global unit are conceivable, for example, local memory units on the lower level and a single, global reference frame on a superordinate level of hierarchy, while there is no hierarchical level for regional units.

The second question is: What sets the main orientation of the reference frames on the different levels of hierarchy? We addressed these questions by assessing participants performance in a spatial memory task where they pointed from selected locations in the environment to memorized landmarks.

The first question—expansion of reference frames and number of levels—is addressed based on two main assumptions: (a) Each new memory unit that is accessed during recall will lead to an increase in pointing latency (see also Meilinger, Strickrodt, & Bülthoff, 2016) and (b) being aligned with the orientation of a stored reference frame during recall will lead to fastest pointing latency compared to being aligned otherwise (e.g., McNamara, Sluzenski, & Rump, 2008; Mou, McNamara, Valiquette, & Rump, 2004). We focus on latency rather than accuracy, based on the assumption that both are indicative of distinct processes of learning and memory (e.g., Prinzmetal, McCool, & Park, 2005; Sternberg, 1969). We more strongly associate pointing accuracy with the precision of memory set mainly during the encoding process, whereas latency more strongly relates to the retrieval of the memory content, may it be more or less precise (see also Meilinger, Strickrodt, & Bülthoff, 2016; Pantelides, Kelly, & Avraamides, 2016).

Based on the two above stated assumptions, clear performance patterns were specified that would be in favor of the distinct hierarchical levels. If local corridor units and the successive connections between them are stored, then the relative position of objects will be retrieved via the successive activation of all local memory units that are represented between one’s current position in the environment and the location of the target. This means, pointing to a target that is three corridors away requires the consecutive activation of one’s current position, the two corridors in-between and the memory unit of the target corridor. Thus, pointing latency should increase with increased corridor distance between current and target corridor, as the number of costly transitions between connected local reference frames increases (corridor distance effect). More precisely, pointing to a target in an adjacent corridor should be faster than pointing to a target two corridors away, and so on. A similar pattern can be predicted for the regional level: If regional units and their connection are memorized, having to point from one’s current location to the target in the other region should lead to performance loss as a new memory unit (the other region) must be activated (effect of target region). Both distance effects—corridor distance and target region effect—can occur in parallel. If all corridors are exclusively integrated into a global reference frame covering the whole environment pointing latency should be independent of corridor distance or target region because the memory unit of the environment is activated no matter where people are located. This approach has some similarity with the concept of spatial priming. For example, McNamara (1986) had participants learn the arrangement of multiple objects within a large hall (vista space). The hall was divided into sections by long bands on the floor. He used a primed object recognition task, where the previous trial either contained an object from within the same region (section) or another region, and either far or close with respect to Euclidean distance. One assumption was that recognition time is dependent on the level of activation that was induced by the prime. Another assumption was that recognition priming mirrors an automatic process of retrieving spatial memory, not influenced by recall strategies. Therefore, speed of recognition should tell about the structure of spatial memory. McNamara (1986) found both distance and region effects for this clustered vista space. Our approach is similar to spatial priming as both regard the time needed for recalling spatial memory to be highly dependent and thus informative about the underlying memory structure. Yet it is different as we concentrate on a more complex task in a complex navigable space, the precise recall of target locations during pointing.

The effect of body orientation during pointing should be telling with regards to deployed reference frames. If pointing is based on local reference frames, pointing should be fastest when aligned with the initial view within every corridor. In case of regional reference frames, aggregating multiple corridors, we should be able to identify one main orientation within each region that leads to best performance when aligned with it. Yet again, if a global reference frame is deployed, there should be one facing orientation across the whole environment yielding best pointing performance. All three alignment effects might be observed in parallel.

There are many studies examining reference frame orientation for object locations in vista spaces, but less that examine navigable space. Therefore, as a second question, we ask what sets the reference frame orientation beyond an enclosed vista space, thus, on higher levels of the spatial hierarchy? Vista space studies suggest that there are multiple factors influencing the orientation of a reference system, namely, egocentric experience (e.g., first perspective) (e.g., Kelly & McNamara, 2008; Rieser, 1989), salient layout intrinsic cues (formed rows and columns) (e.g., Kelly & McNamara, 2008), salient layout extrinsic cues (room geometry) (e.g., Shelton & McNamara, 2001; Valiquette & McNamara, 2007) and even instructions (e.g., Mou & McNamara, 2002). Whether all these factors play a role when integrating multiple vista spaces into a reference frame on a higher level is not yet known. However, it seems plausible to assume that the notion of space that is interpreted in terms of a conceptual north (e.g., Shelton & McNamara, 2001) is not only accounting for vista, but also for navigable spaces (e.g., McNamara & Valiquette, 2004).

Some studies examining navigable space emphasize the importance of the first perspective experienced within the environment (e.g., Tlauka, Carter, Mahlberg, & Wilson, 2011; Wilson, Wilson, Griffiths, & Fox, 2007) and report, what they call, the First Perspective Alignment effect. Yet again, as has been shown in vista space studies already, the selection of a reference frame direction is not constrained to the assimilation of later information into an already existing reference frame, but can likewise succeed via accommodation (Piaget & Inhelder, 1969), thus, adapting a representation by integrating the cumulated input to form a reference frame with a main orientation distinct from the initial perspective. When learning an object array in vista space the saliency of an initial view can easily be overwritten by emphasizing specific perspectives without directly experiencing it (Greenauer & Waller, 2010; Mou & McNamara, 2002) or by an alignment with a layout intrinsic or extrinsic cue when examining an object array from a different perspective at a later time (Kelly & McNamara, 2010; Shelton & McNamara, 2001). The possibility to update earlier experienced spatial memory through movement can contribute to that (Wang, 2016). When examining the environmental layouts used in studies evidencing a global reference frame in navigable space (e.g., Meilinger, Frankenstein, Watanabe, Bülthoff, & Hölscher, 2015; Meilinger, Riecke, & Bülthoff, 2014; Richardson, Montello, & Hegarty, 1999; Tlauka, Carter, Mahlberg, & Wilson, 2011; Wilson, Wilson, Griffiths, & Fox, 2007) we identified two alternative factors besides the driving force of the first perspective, that might (partly) explain the global alignment. Often the first corridor/street of the environments used in those studies was also the longest. Thus, the first perspective was confounded with the most frequent perspective experienced during the initial walk through the environment. Additionally, often environments with a simple geometric structure were used, for example, three corridors in an upside-down U-shape. Thus, starting from the initial segment participants turned 90° towards the second segment and 90° thereafter following the same turning direction as before, ending up in a corridor parallel to the first one. Already Wertheimer (1923) identified parallelism (among others) as a key principle of perceptual grouping of 2D elements (Gestalt laws). Indeed, recognizing parallelism on 2D planes seems to be a core component of geometrical understanding in humans (see for example Dehaene, Izard, Pica, & Spelke, 2006; Dillon, Huang, & Spelke, 2013). Also, studies investigating human spatial memory in orientation tasks after disorientation within rooms of various shapes showed sensitivity to environmental geometry (e.g., Hermer & Spelke, 1996, 1994). Humans even have the tendency to remember irregular environments as more parallel and regular than they are (e.g., Byrne, 1979; Moar & Bower, 1983; Tversky, 1981). Thus, when confronted with natural, navigable space, we seem to be prone to search for, even superimpose easy structures to conceptualize space and to anchor our knowledge. We therefore reckon that the overall geometry of such a U-shaped environment—two parallel leg segments which converge to the same orthogonal segment, thereby forming a weak bilateral symmetry—is easy to infer and might serve as a salient cue for setting the reference frame orientation along this U-shape. Of course, when first entering the U-shape its structure is not apparent. Only after traversing through the connector segment and leaving the U-shape from the second parallel leg one might be able to understand the configuration. When reentering the U-shape later in time the shape is recognized (i.e., "I am back at the U-shape") and anticipation of the subsequent corridors is facilitated, therefore, boosting saliency of the orientation when entering the U-shape. Using such a cue requires relating multiple corridors and, therefore, might be identified only after sufficient experience with the environment.

We addressed this issue of multiple salient cues for reference frame orientation by examining three potential reference frame alignments for the two superordinate levels: (1) Alignment with the first perspective within each region and across the overall environment, (2) alignment with the most frequently encountered perspective during the first walk through each region and across the overall environment, and (3) alignment with the salient geometric cue of a U-shape within each region or across the environment. To our knowledge this is the first study testing for multiple potential reference frames on three levels of hierarchy. Predictions are described and visualized in the result section of Experiment 1 (Deployed Reference Frames).

In summary, our study aimed to explore the architecture of spatial memory acquired in navigable space. Participants were asked to explore and learn the layout of a virtual environment consisting of two regions by walking. Two main questions were addressed: Do people memorize the environment in the form of local, regional or global memory units or within multiple forms, thus, revealing hierarchical structure of spatial memory? And what sets the main orientation of the reference frames on the different levels of hierarchy, the first perspective, the most frequently experienced perspective, or the geometric salience of a U-shape? Between the two experiments we manipulated the regional start point of learning to vary ease of global integration. In Experiment 1 participants started at the outer ends of both regions and reached the regional transition point only after walking through the entire region. Conceivably, this makes it harder to integrate both regions into a global memory unit, while facilitating regional clustering. In Experiment 2 participants started learning from the transition point between the two regions, enabling them to encode the connection of the two regions from the start and integrate each additional corridor into their existing knowledge. Compared to Experiment 1 this learning procedure might facilitate global integration. After learning the environment, we tested participants spatial memory with a pointing task. Their performance was analyzed for distance and alignment effects on the local, regional and global level to shed light on how their memory was structured.

Experiment 1

Method

Participants

After providing informed consent, 23 participants partook in the experiment, all naïve to the research question and the experiment. They received monetary compensation for participating. The experiment was approved by the local ethics committee. Participants were required to have normal or corrected to normal visual acuity. For three participants, the experiment was terminated before completion: one was unable to reach the learning criteria (see Procedure), two did not finish the experiment in the maximally available time of three and a half hours. We excluded another two because of bad performance (see Results). The remaining 18 subjects were included in the analysis, 6 of which were male. Their average age was 29.72 years (*SD* = 10.61).

Material

During learning participants walked freely in a large tracking space of 12×12 m, while their head position was tracked by 16 high-speed infrared cameras at 150 Hz (Vicon® MX 13) and transmitted to a computer (with NVIDIA GTX1080 graphics card). The computer rendered the respective egocentric view within the virtual environment, which was displayed to the participants via an Oculus Rift Development Kit 2 that provided a field of view of ca. 100°×100° at a resolution of 960×1080 pixels for each eye. Head-mounted display (HMD) and computer were connected via a long cable. The experimenter followed participants closely during learning, carrying and repositioning the cable to ensure safe movement through the tracking space and to eliminate direction cues through the cable origin. This immersive VR setup allowed for a realistic learning experience, including proprioceptive cues and stereopsis as well as self-determined exploration of the virtual world.

The virtual environment (Figure 1) consisted of eight interconnected corridors, each containing one distinct virtual 3D object standing on a pedestal in the middle of the corridor on either side. The environment was constructed to induce maximal separation between the two regions of the environment. Besides visual cues (colored walls and pedestals: blue and red) additional factors were held constant within one half while being maximally different from the other region: semantic similarity (categories: animals used as landmark objects in the blue region, red region contained only tools), distance between regions (longest corridors in the middle of maze at regional transition point), and complexity of turning angles between corridors (90° angles within a region, 45° at transition point). Also, the two halves of the environment were explored one after another (spatio-temporal learning contingency), with an overlap only at the regional transition point (mandala). Each region comprised the same number of corridors and objects, ensuring similar memory load.

The objects served as landmarks for self-localization and orientation and as target objects in a subsequent pointing task. For each participant, a sample of eight landmarks (four animals, four tools) were randomly selected from a pool of 16 objects (eight animals, eight tools) and distributed across corridors (animals in blue, tools in red corridors). The remaining eight objects served as distractors objects in a primed recognition task (not described here). The pool of 16 objects had previously been validated in a pre-test and selected from a set of 24 objects as being well recognizable. In the virtual environment landmarks were either standing on the right or on the left corridor wall and facing either direction along the corridor. For each participant, the four possible combinations of side and facing (four possible positions of a landmark) were assigned randomly to the four corridors of a region.

Procedure

Participants were instructed that they were about to learn a virtual environment, starting with one half of the environment, then followed by the other half of the environment. Participants either started in the blue or in the red region (pseudo-randomly assigned) to account for saliency effects of individual regions.

Learning phase Before exploring the environment participants were briefed about the experimental procedure. They were informed that they were about to learn a virtual environment, first one then the other half of the environment and that they need to memorize the whole environment and how the objects and the corridors relate to each other. The instructions made clear that they will be teleported to different positions within the environment later, where they must be able to orient themselves and to relate to any other object from there. Participants then were equipped with the HMD. With the display turned off, they were disoriented, led through a few left and right turns to a predefined starting position.

In both halves participants started at the very end of a region, standing with their back to the dead-end of the last corridor, facing along the corridor (end-to-transition condition). They then walked through the corridors towards the transition point (mandala). Standing at the transition point the view into the other, yet unexplored region was blocked by a curtain. Only parts of the corridor of the other region was visible, however, neither the object and the pedestal nor the turn to the next corridor of the other region was visible. Standing at the mandala they had clear visual information about how the two corridors of the two regions are connected to each other. Here participants were explicitly informed: “You are now at the point where the two halves of the environment come together. Only currently the view into the next corridor is blocked. You will learn this part of the environment after you have learned the first half’. The experimenter emphasized again that participants would be required to relate objects from both halves of the environment later in the testing phase and therefore would need to know how the halves relate to each other.

Participants walked from the dead-end to the mandala and back again to the dead-end corridor three times. They had to name the objects during learning to ensure they could identify each object by name. Apart from that, participant could walk in their own speed and stop anytime they liked to look around the environment. However, they were not allowed to walk back before fully translating the region.

After learning the environment three times a learn-check was carried out. Objects and pedestals were removed from the scene and fog occluded the turning direction to the next corridor. While walking through the deprived environment, participants had to recall from memory name, side, and facing direction of the object in the respective corridor as well as the turning direction to get into the next corridor. All four corridors were queried, two from each learning direction (dead-end to transition point and vice versa). If participants did not reach 100% accuracy (learning criteria), they had to learn the environment again. This time they walked to the mandala and back only once and then did the learn-check again. This procedure was repeated until participants reached the learning criteria, or a maximum of six times. Participants were excluded from the experiment in case they were not able to learn the environment within a maximum of six learning repetitions. After reaching the learning criteria participants took a short break and continued learning the second half of the environment, following the same end-to-transition point procedure as in the first region. When reaching the transition point for the first time in the second region they were explicitly informed that “This is where the two parts of the environment come together. You are now at the exact same position as before when you were at the end of the first half of the environment.”, and the experimenter emphasized again that it is important to know how the two halves relate to each other to solve the subsequent memory test. Like this, we ensured participant would pay enough attention to this part of the environment, a part that they would never fully explore by walking across this point. The learning phase ended after reaching the learning criteria for the second region as well. After a short break, the testing phase started.

Testing phase The first task was a primed recognition task, which is not reported here. Participants recognized target objects from the environment among distractors after being primed by other target objects or distractors. The task took approximately 20 minutes. Subsequently, after a short break, the pointing task started.

During pointing participants sat on a high chair. A joystick was used for recording responses, standing right in front of the participant at belly height. In each trial of the pointing task, participants were teleported to the middle of a corridor of the environment next to a landmark. Their head and body was aligned with one of eight possible orientations, evenly spread around 360° in steps of 45° (Figure 1, black arrows). First, they looked around, self-localizing themselves with the help of the object in the corridor. On both sides of the corridor arms the view was blocked by a white fog, thus, the turning direction to the next corridor was not visible. After confirming self-localization and orientation by clicking a button on the joystick, the name of one of the remaining seven objects was projected over the rendered environment and participants were required to point to this object. This means, they had to indicate the direction towards the target with respect to their current position and orientation. The direction had to be indicated by moving the joystick handle (allowing for the full range of 360°) and confirming the target direction with a button press. No visual feedback about the joystick direction was given. Participants had to look straight ahead while giving their response (both for self-localization and pointing) otherwise their response was not registered. Afterwards, the experiment continued with the next trial.

From each of the eight corridors participants pointed eight times, covering all eight body orientations, to a target within the same region, and again eight times, covering all body orientations, while pointing to a target of the other region. This adds up to 128 trials. Target objects were selected randomly for every trial under the constraint that each object within a region was used equally often as a target. Trial order was random. Each task started with eight practice trials, randomly chosen from the pool of 128 trials, to familiarize participants with the task. Those were not included in the analysis. Including the practice trials there were 136 trials with four blocks of 34 trials each. Each block was followed by a pause of self-determined length. Participants were instructed to point as fast and as accurate as possible. No feedback was given about the accuracy of the response and task execution was not constrained by any time limit. At the end of the experiment we asked participants to judge their general sense of direction (SOD) on a five-point Likert scale.

To summarize, we varied the within-subject factors *body orientation* (while pointing being aligned with 0°, 45°, 90°, 135°, 180°, 225°, 270°, 315° relative to the environment, see Figure 1), *target region* (target within same region vs. across region with respect to current position) and corridor *distance* (current position and target one to seven corridors apart). We recorded self-localization time (not reported), pointing latency (i.e., the time between displaying the target name and registering a response), and pointing error (i.e., absolute deviation between correct and indicated target direction).

Results

As mentioned in the sample description we excluded two participants based on their pointing performance. Either pointing accuracy did not significantly exceed chance level or latency was 2 SD slower than the sample mean. We further excluded data points deviating more than ±2 SD from a participant's overall mean pointing latency and pointing error[1]. On average, we excluded 4.12% (*SD* = 2.57) values from pointing error and 4.64% (*SD* = 1.33) values from pointing latency per participant. Participants needed on average 2.00 learning repetitions (*SD* = 1.19) to learn the blue region (reflecting about 8 walks through the environment, exclusive of walks during learn-check) and 2.11 learning repetitions (*SD* = 1.13) to learn the red region (reflecting about 8.22 walks). They spend on average 35.79 (*SD* = 12.07) minutes in the environment.

Distance and cluster effects

We conducted an ANOVA with the within-subject factors *distance* and *target region*. Across region pointing trials with a distance to the targets of more than three corridors were excluded from the analysis to match corridor distance for both target locations (within and across region). Thus, 88 out of the 128 trials were used for this analysis, 64 within-region and 24 across region trials.

Pointing latency (Figure 2, left) was indicative of mental separation between corridors and regions. We found main effects of *distance*, $F(2, 34) = 9.74$, $p < .001$, $\eta_p^2 = .36$, and *target region*, $F(1, 17) = 4.88$, $p = .041$, $\eta_p^2 = .22$. Pointing latency significantly differed between corridor distance 2 and 3, $t(34) = -2.856$, $p = .019$, as well as 1 and 3, $t(34) = -4.342$, $p < .001$, but not between corridor distance 1 and 2, $t(34) = -1.487$, $p = .310$ (p-value adjusted by Tukey method for multiple comparisons). The response was on average 1.98 sec faster when pointing within the same region compared to pointing to a target in the opposite region. No interaction between *target region x distance* was found, $F(1.41, 23.92) = 2.19$, $p = .146$, $\eta_p^2 = .11$ (Greenhouse-Geisser-adjusted). Thus, there is no evidence that the effect of target region varies across distances.

---

[1] Exclusion of latency values did not coincide with exclusion of error values of the same trial. Rather outlier exclusion was done separately for both dependent variables. As stated in the instruction we reckon latency and error to mirror different aspects of spatial memory (precision of memory vs. retrieval of the memory content) (Meilinger, Strickrodt, & Bülthoff, 2016; Pantelides, Kelly, & Avraamides, 2016). Mutual exclusion of latency and error values (i.e., excluding the whole trial) did lead to higher percentage of excluded trials per participant, but did not lead to any significant changes to the results of separate exclusion, which are reported in this paper.

For pointing error (Figure 2, right) we found a main effect of *distance*, $F(2, 34) = 19.57$, $p < .001$, $\eta_p^2 = .54$. Pointing error was significantly lower for distance 1 compared to distance 2, $t(34) = -3.978$, $p = .001$, as well as compared to distance 3, $t(34) = -6.171$, $p < .001$, but not significantly different between distance 2 and 3, $t(34) = -2.193$, $p = .087$ (p-value adjusted by Tukey method for multiple comparisons). We did not observe a main effect of *target region*, $F(1, 17) = 0.35$, $p = .561$, $\eta_p^2 = .02$, or an interaction of *target region x distance*, $F(2, 34) = 2.44$, $p = .102$, $\eta_p^2 = .13$. Thus, the analysis does not provide evidence that pointing error is higher when pointing across regional boundaries as compared to pointing within ones' current region.

Adding gender, first learned region (blue or red), number of learning trials, and self-reported sense of direction (SOD) as covariates to the analyses of latency and accuracy did not change the inference statistical values and interpretation of the abovementioned effects in a meaningful way. The results described here are therefore limited to reports of statistics without the consideration of covariates. Further, no main effect of gender, first learned region (blue or red), and SOD on latency or accuracy were found. Only the number of learning trials was associated with pointing error, $r = 0.69$, $p = .002$. Participants who quickly learned the environment were also comparatively better pointers.

Deployed reference frames

The further analysis concentrated on the pointing latency to assess the processes underlying spatial recall (compare to Meilinger, Strickrodt, & Bülthoff, 2016; Pantelides, Kelly, & Avraamides, 2016, and see Introduction). In contrast to the distance analysis, where trials with distances beyond three corridors were excluded, we now used the full data set of 128 trials (excl. outlier values) per participant. To account for the strong distance effect we first conducted an ANOVA with distance as the within-subject factor on the full dataset. The computed residual values instead of the raw values of pointing latency were now used for the following analysis of reference frames. By controlling for the known effect of distance on latency, we hereby removed unwanted variance, and in the following, were aiming to explain variance in the data not yet explained by distance. Interpreting residual values is not as intuitive as interpreting raw pointing latency values (e.g., they center on 0), nevertheless they do preserve the relation and magnitude of difference between data points that are manifest in the raw data (i.e., faster or slower pointing performance). In the following description of the results we will use the more intuitive term of "pointing latency" and "reaction time", but we mean to refer to "latency residuals" after statistically accounting for the effect of distance.

The environment used in the current study consisted of two obliquely aligned regions. Literature suggests that one's current position and target location affect which reference frames are selected to base spatial recall upon. For example, results from Greenauer and Waller (2010) suggest that macroreference frames are used when pointing from one to another object array in a vista space, but not when pointing within one array (see also Zhang, Mou, McNamara, & Wang, 2014, for nested spaces). Therefore, and based on the region effect found in the previous analysis we analyzed the effect of body orientation on pointing performance separately for within and across region trials.

At each position in the environment participants were tested while being aligned with one of eight body orientations.[2] Depending on the reference frame(s) deployed clear predictions can be made about which body orientation should elicit fastest memory access compared to the other body orientations. Figure 3 illustrates our predictions. For each prediction at each position in the environment one orientation can be identified as being superior. If participants use local reference frames, pointing latency should be fastest when being aligned with the first perspective experienced in each corridor compared to being aligned with one of the seven remaining orientations (e.g., facing a corridor wall). If participants use a reference frame covering a whole region and the reference frame direction was set by the first perspective experienced in this region, latency should be fastest when being aligned with the perspective experienced in the very first corridor, independent of the current position within this region. Thus, for example, facing the left wall in the second corridor of the blue region, should yield faster response times than being aligned otherwise (e.g., facing straight along the corridor). Likewise, if participants form and use a global reference frame covering the whole environment and the reference frame direction was set by the first perspective experienced in the first region, pointing latency should be fastest when being aligned with this first perspective at each single position tested. This means, having started off learning the blue region followed by the red region, should yield faster pointing when being aligned with the first experienced orientation in the blue region – even if

---

[2] Supplementary material, Figure S1 and Figure S2, left, give an overview of the pattern of pointing latency before further aggregation to test for the specific reference frame predictions. The figures provide a visualization of pointing latency as a function of body orientation on corridor scale (Figure S1, left) and on global scale (Figure S2, left) (global scale corresponds to compass rose in Figure 3).

located in the red region during recall. In this latter case the superior orientation is obliquely aligned with one's current visual scene in a red corridor. Both on the regional and global scale alternative cues could be used to set the reference frame orientation. Therefore, the most experienced perspective during the first walk through the environment and the salient geometry of the environment, in our case the U-shape of a region are explored as well[3]. For each participant and each prediction, we contrasted the mean pointing latency in the superior orientation trials with the remaining, non-aligned trials. This reaction time differences were then further assessed. A mean positive reaction time difference is indicative of a fit to the respective prediction. Zero corresponds to no difference in pointing latency.

For within region pointing Figure 4, left, side shows the fit for the seven considered predictions, Table 1, left side, the results of the analysis. For each prediction, we analyzed with a t-test whether the reaction time difference is significantly larger than zero. When subjects pointing within their current region we found a significant fit for the prediction of local reference frames and of regional reference frames, both for the first perspective experienced within a region as well as the U-shape. No evidence could be found that participants used a regional reference frame following the most frequently experienced perspective or that they use global reference frames. As mentioned in footnote 2, our reference frame predictions are not uncorrelated. Furthermore, their correlation varies. Local corridor alignment correlates with the prediction of regional reference frame by $r = 0.29$ for first perspective alignment, $r = 0.43$ for most frequently experienced perspective and $r = 0.14$ for U-shape. We therefore reanalyzed the fit to our regional and global reference frame predictions after correcting for the effect of local alignment. Similar results were obtained. We still find the fit to the predictions of regional reference frames following the first perspective, $t(17) = 1.785$, $p = .046$, $d = 0.42$, as well as the regional reference frames following the U-shape, $t(17) = 2.309$, $p = .017$, $d = 0.54$, while no other predictions fit well, t's < 1.043, p's > .155. Similarly, we also corrected the data for the two

---

[3] The predictions tested for are by no means comprehensive, but based on observations of environmental layouts of previous studies. Further, it should be noted that the predictions are not orthogonal. For example, the prediction of a regional reference frame following the first perspective alignment is identical with the prediction of a regional reference frame following the U-shape when considering the blue region, but not when considering the red region (compare to Figure 3). When combining both regions, as done in our analysis, a correlation of first perspective and U-shape prediction remains with $r = .43$. Same holds for the global level, when considering all subjects, half of which learned in the blue, half in the red region. Likewise, the prediction of a regional reference frame following the U-shape is correlated with the prediction of a global reference frame following the U-shape of the first region by $r = .43$. Predictions of U-shape and most frequent alignment can be considered orthogonal with a correlation of $r = -.14$.

meaningful regional reference frame predictions separately to see whether we still find a significant fit to the local prediction. The fit to the local prediction remains significant after correcting for the regional U-shape prediction, $t(17) = 1.761$, $p = .048$, $d = 0.42$, and remains as a trend after correcting for the regional first perspective prediction, $t(17) = 1.419$, $p = .087$, $d = 0.33$.

A similar analysis as for within region pointing was conducted for trials in which participants pointed across the regional boundaries. This time, results are inconclusive. Across region data fitted to none of our predictions, *t's* < 1.14, *p's* > 0.136, thus, being aligned with the local corridor or with one of the three possible regional or global reference frame does not seem to facilitate spatial memory access compared to being otherwise aligned. Reckoning that this might be due to greater noise in the across region data, we compared individual standard deviations for within and across region pointing. Indeed, standard deviations were significantly higher in across region trials, $t(17) = -2.25$, $p = .038$, $d = -0.75$, thus, potentially obstruct detection of used reference frames. Therefore, we decided to re-run the same analysis of the fit to the seven predicted reference frames with only a subset of the across region trials, namely, only for trials querying targets that are lying one, two or three corridors away, excluding trials with corridor distance four and higher. With this trial subset (24 trials per participant) the same corridor distances as for the evaluation of the within-region pointing trials are covered, similar as for the analysis of distance and cluster effects. Indeed, individual standard deviation of this subset was comparable to variance for within region trials, $t(17) = -1.50$, $p = .151$, $d = -0.50$. Table 1, right, shows the results of the analysis of across region trials with corridor distance one to three, in Figure 4, right side, latency differences are depicted. As before, data did not show significant fit to neither local nor regional reference frame predictions. However, now significant fits to the predictions of global reference frames were found, more precisely, for the prediction of the first experienced perspective of the first learned region as well as the U-shape experienced in the first learned region. The prediction of a global reference frame based upon the most frequently experienced perspective in the first learned region did not reach significance.

Continuing with the previously used data set we also examined whether pointing across region led to fastest pointing when aligned with the main reference direction of the target rather than the main reference direction of one's current position – a result found by Zhang, Mou, McNamara and Wang (2014) for pointing between two nested spaces. In our study no difference in pointing latency on across region trials could be found between being aligned with the four main orientations of one's current region and being aligned with the four main orientations of the neighboring region containing the target, $t(17) = 0.43$, $p = .674$, $d = .14$, thus, we did not find evidence for the alignment with the target reference frame.

Discussion

In Experiment 1 participants memorized a regionalized environment. They explored both halves of the environment separately by physical walking, starting from the outer ends and walking towards the transition point, where they had a glance into part of the other region. We were interested in how their memory might be structured and what spatial units they form and use to solve a subsequent pointing task. During this pointing task participants were randomly teleported to different locations within the environment, being aligned with one of eight possible body orientations, while pointing to targets within the same region as their current standpoint or to targets located in the other region. The first experiment revealed evidence that participant formed local, regional as well as global spatial units that are accessed during recall.

Pointing latency increased with increasing corridor distance to the target, independent of whether participants pointed within their current region or across regional boundaries. One possible explanation is, that local corridor units were stored in memory which then had to be successively activated in a fixed, learned order, to recall the location of a target object multiple corridors away (e.g., Meilinger, Strickrodt, & Bülthoff, 2016). For every new unit activated the mental effort increases reaction time (and is accumulating error). Besides this latency increase across spatial corridor units, we also found a facilitative effect of local alignment with the corridor. Participants were significantly faster when aligned with the first perspective experienced within each corridor, compared to the seven remaining orientations.

Pointing to targets within the other region, i.e., standing in the red region pointing to a target in the blue region or standing in the blue region pointing to a target in the red region, led to higher pointing latency compared to pointing to a target within one's current region. This effect was not dependent on the corridor distance to the target (no interaction of target region and distance). Additionally, pointing performance in within-region trials was significantly faster when subjects were aligned with the first perspective experienced within a region (first corridor orientation). For example, facing the right wall in the second corridor of the red region, should yield faster response times than being aligned otherwise (except when being aligned with the first experienced perspective in the corridor). Also, pointing performance was significantly faster when subjects were aligned with the orientation when entering the U-shape that is made up by three consecutive corridors. For example, facing the left wall in the third corridor of the red region, should yield faster response times than being aligned otherwise. No fit (even a negative value) was found to the prediction of a regional reference frame following the most frequently experienced perspective within a region. The fit to a regional reference frame following the first perspective and the U-shape remains significant even after statistically accounting for the effect of local corridor alignment. Taken together, participants seem to have formed obliquely aligned, regional reference frames of the blue and red region.

The fit to regional reference frame predictions cannot be explained by a facilitative contra- or orthogonal alignment with the reference frames of the local corridors. This would allow for a speeded access as soon as participants are aligned with either of the four main orientations of the local reference frames (0°, ±90°, 180°). In this case we would expect a fit to all three regional predictions. However, we only find a fit to two predictions.

We further found support for the formation of a global reference frame encompassing both regions. Data of within region pointing did not fit to either of the three predictions for global reference frames. However, pointing performance of across region trials querying targets one to three corridors away revealed a fit to the global predictions following the U-shape as well as the first perspective experienced in the first learned region. This support is somewhat weakened by the fact that it is only found for a subset of the across region trials. Analyzing latency patterns of the full data set of across region trials (corrected for the effect of distance) was non-descriptive. Neither prediction for a specific reference frame alignment fit our data, potentially due to the higher variance in pointing performance for higher corridor distances. Nevertheless, taken equal corridor distances, the selection of remembered reference frames seems to be depending on the relationship between one’s current position and the target location. This will be elaborated upon more in the General Discussion. An alternative approach based on results obtained by Zhang, Mou, McNamara and Wang (2014), that showed that the reference frame direction of the target is used, did not describe our data well.

Our results imply that participants embedded multiple vista spaces, formed distinct memory units, one for each region, and memorized the whole environment in a global format.

Importantly, as indicated by the alignment effect, those memory units seem to be metric relational schemes, linking multiple vista spaces to one superordinate reference frame direction, thus, they are spatial in nature (compare to McNamara, Sluzenski, & Rump, 2008; Meilinger & Vosgerau, 2010) (see also General Discussion).

Being able to identify a fit to the U-shape prediction besides a first perspective alignment on the regional level is a first indication that—also for navigable space—reference frames are formed that do not merely assimilate newly incoming spatial information into an already existing reference system (i.e., assimilate new corridor information into the direction of the first perspective), but that information of multiple corridors are accumulated and accommodated (geometry of three corridors) to form a reference frame with a new, distinct reference direction (Piaget & Inhelder, 1969). This is in line with studies examining object arrays in vista space, which show that egocentric alignment with a salient geometric cue at a later time point during learning can shape the reference frame accordingly (Kelly & McNamara, 2010; Shelton & McNamara, 2001).

The prediction for a reference frame following the U-shape structure was tailored to the orientation experienced when entering the U-shape, walking towards the connector segment of the two parallel corridors. We decided for a single main orientation to equalize predictions for all potential reference frames (one main orientation vs. the remaining orientations) and because we assumed that the moment one enters the simple shape is particularly prominent as (after enough exposure to the environment) anticipation of the two subsequent corridors is facilitated. During learning each participant entered the U-shape of each region at least six times (3x end-to-transition, 3x transition-to-end). Alternatively, the orientation when leaving the U-shape might be of significant saliency as well, as this orientation is accompanied with the emerging realization of the easy structure of the just travelled corridors. To test for this possibility, we analyzed fit of the within region pointing latency to the prediction of a regional reference following the U-shape with the main orientation centered on the orientation when exiting the shape, thus 180° opposite to the originally tested orientation when entering the U-shape. Interestingly, also here a significant fit was found[4]. Based on this we cannot conclude whether the main orientation of a superordinate reference frame is set when entering or exiting the U-

---

[4] This analysis was not mentioned in the result section. Significant fit of pointing latency to the prediction of a U-shape but with the main orientation centered on the orientation experienced when exiting the U-shape, $t(17) = 2.389$, $p = .014$, $d = 0.56$. Significant fit remains even after correcting for local alignment, $t(17) = 2.116$, $p = .025$, $d = 0.50$.

shape. Yet, it remains that the simple structure of two parallel leg segments connected to the same orthogonal segment, thereby forming a weak bilateral symmetry, might serve as an important cue for setting the orientation of superordinate reference frames.

Despite taking longer to recall spatial memory for the other region, pointing accuracy was not affected by the regional belonging of the target. Pointing to targets within the other region did not lead to higher error in pointing. This shows that participants were capable to learn the two halves of the environment separately with overlap at only one common position, the transition point. And this, even though the turning angle at the transition point was more complex (45°) than within region angles (90°)—complexity was previously shown to enlarge pointing error (e.g., Moar & Bower, 1983)—and even though participants were only allowed to examine this transition point visually, but never walked across it. This indicates that the regional transition point was memorized just as accurately as the corridor angles within a region.

Results of Experiment 1 indicate that corridors triggered a clustering of spatial memory into vista space units and that we could successfully trigger clustering into regional memory units. Additionally, it seems that integration into a global unit spanning the whole environment occurred, although this was only detectable for smaller corridor distances. All three levels manifested in spatial metric schemata, the reference frames on the local, regional and global level. In Experiment 2 we set out to replicate findings of Experiment 1 and elucidate the indeterminate findings regarding a global memory unit. Participants in Experiment 1 learned both halves separately, always starting at the outer end of each region, walking towards the transition point. This learning procedure makes it particularly difficult to integrate both regions into a global memory unit. After the first half participants memorized the layout of four corridors and the relative position of the four objects within. When starting to learn the second half, relating each new corridor to the previously formed memory unit was not immediately possible. Participants first had to walk through the four corridors of the second region before ending up at the transition point again. Only then were they able to understand the layout of the whole environment. By reaching the regional transition point again they might already have formed a first “draft” of the second region that now must be connected to the representation of the first learned region. Integrating both representations into one spatial unit might entail considerable effort. In a second experiment, we wanted to see whether lowering this effort might lead to clearer results regarding global integration. Regions were still learned separately. However, now participants started learning from the transition point, walking towards the outer ends of each region. Like this, global integration, corridor by corridor, should be possible right away.

Experiment 2

Method

Participants

Twenty-one naïve participants partook in the experiment, receiving monetary compensation. The experiment was approved by the local ethics committee. Participants were required to have normal or corrected to normal visual acuity. Two participants had to stop the experiment before completion: one was unable to reach the learning criteria, one experienced motion sickness during testing. One additional participant was excluded due to chance level pointing performance. The remaining 18 subjects (10 males) were included in the analysis. Their average age was 28.22 years (*SD* = 8.59).

Material

The material and equipment used was identical to Experiment 1.

Procedure

The procedure was identical to Experiment 1, except of the learning direction. While participants in Experiment 1 started to learn each half of the environment at the outer end of the succession of corridors, walking towards the transition point, participants in Experiment 2 started to learn each half with their initial position at the transition point. Thus, from the beginning they were aware of how both regions are connected to each other. Also, starting with the second region it was immediately possible to relate each newly explored corridor to the previously learned corridors.

Results

We excluded data points deviating more than ±2 SD from a participant's overall mean pointing latency and pointing error. This resulted in exclusion of on average 5.38% (*SD* = 2.14) values from pointing error and, separately, 4.39% (*SD* = 1.20) values from pointing latency per participant. Participants needed on average 1.61 learning repetitions (*SD* = 0.98) to learn the blue region and 1.61 learning repetitions (*SD* = 0.85) to learn the red region (reflecting about 7.22 walks through the environment for each region, exclusive of walks during learn-check). They spend about 35.21 (*SD* = 7.14) minutes in the environment. We rerun the same steps of analysis as ran in Experiment 1.

Distance and cluster effects

We conducted an ANOVA with the within-subject factors *distance* and *target region*. Figure 5 depicts pointing latency and error across the full range of distances. But again, across region pointing trials with a distance to the targets of more than three corridors were excluded from the analysis to match corridor distance between conditions (64 within-region and 24 across-region trials remaining).

We found a main effect of *distance*, $F(1.26, 21.34) = 19.01$, $p < .001$, $\eta_p^2 = .53$ (Greenhouse-Geisser-adjusted) on pointing latency. Pointing latency differed significantly between corridor distance 1 and 2, $t(34) = -3.762$, $p = .002$, as well as 1 and 3, $t(34) = -6.111$, $p < .001$, but only by trend between corridor distance 2 and 3, $t(34) = -2.349$, $p = .062$ (p-values adjusted by Tukey method for multiple comparisons). In contrast to Experiment 1 no significant effect of *target region*, $F(1, 17) = 2.06$, $p = .169$, $\eta_p^2 = .11$, was found. Thus, now participants didn't need more time to access spatial memory from the other region compared to recalling targets located within the same region. We observed no interaction between *target region x distance*, $F(2, 34) = 0.15$, $p = .864$, $\eta_p^2 = .01$. To further evaluate the non-significant result of target region we run a Bayesian repeated measure ANOVA with the factors *distance* and *target region* (r scale for fixed effects = 0.5) to evaluate the likelihood of a null-effect of target region. The highest Bayes factors value was found for the model assuming a single main effect of *distance*, $BF_{10} = 69{,}530{,}000$. This is the model that outperformed the null model the most. The model assuming a single main effect of *target region* revealed a particularly small value, $BF_{10} = 0.371$. This is, according to the classification defined by Lee and Wagenmakers (2013) (adjusted from Jeffreys, 1961), within the range of $BF_{10} = 1/3 – 1$, which is interpreted as anecdotal evidence for H0 (no effect of target region) compared to H1. Both additionally tested models of two main effects (*distance* and *target region*, $BF_{10} = 34{,}830{,}000$) and the model assuming two main effects and an interaction ($BF_{10} = 4{,}779{,}000$) only reach BFs smaller than the main effect model of distance. More precisely, the *distance* model is 2.00 times more likely than the model assuming two main effects and 14.55 times more likely than the full model (main effects and interaction).

For pointing error (Figure 5, right) we again found a main effect of *distance*, $F(2, 34) = 10.29$, $p < .001$, $\eta_p^2 = .38$. Pointing error was significantly lower for distance 1 compared to distance 2, t(34) = -2.522, $p = .042$, as well as compared to distance 3, t(34) = -4.526, $p < .001$, but not significantly different between distance 2 and 3, t(34) = -2.004, $p = .127$ (p-values adjusted by Tukey method for multiple comparisons). We did not observe a main effect of *target region*, $F(1, 17) = 1.29$, $p = .272$, $\eta_p^2 = .07$, or an interaction of *target region x distance*, $F(1.47, 25.00) = 0.49$, $p = .562$, $\eta_p^2 = .03$. Like Experiment 1 participants pointing error was not significantly larger when pointing across regional boundaries compared to pointing within one's current region.

Similar to Experiment 1, adding gender, first learned region (blue or red), number of learning trials, and SOD as covariates to the analyses of latency and accuracy did not change the results for accuracy and latency. We therefore concentrated on reporting statistics without modelling covariates. Besides the number of learning trials, $r = 0.68$, $p = .002$ (more error the more learning trials), no other covariate was associated with latency and error.

Deployed reference frames

After statistically accounting for the effect of distance on pointing latency (see Experiment 1 for a more detailed explanation) we continued with the analysis of deployed reference frames, again concentrating on trials participants pointed within their current region and across regional boundaries separately.[5]

Figure 6, left, shows the fit for the seven considered predictions for within-region pointing, Table 2, left, the results of the analysis. When participants pointed within their current region we observed a significant fit for the prediction of local reference frames and of regional reference frames, this time for the most frequently experienced perspective. As mentioned in footnote 2 our reference frame predictions are correlated. The strongest correlation can be found for the predictions of local corridor alignment with the regional reference frame following the most frequently experienced perspective ($r = 0.43$). Indeed, when correcting data of within

---

[5] Supplementary material, Figure S1 and Figure S2, right, give an overview of the pattern of pointing latency before further aggregation to test for the specific reference frame predictions. The figures provide a visualization of pointing latency as a function of body orientation on corridor scale (Figure S1, right) and on global scale (Figure S2, right) (global scale corresponds to compass rose in Figure 3).

region pointing further for the facilitative effect of local alignment the significant fit to the prediction of regional reference frames following the most frequently exposed perspective disappears, $t(17) = 1.381$, $p = .093$, $d = 0.33$. The fits to the other predictions remain non-significant, t's < 0.605, p's > .275. Likewise, correcting for the effect of regional alignment with the most frequently experienced perspective, causes the fit to the local reference frame prediction to disappear as well, $t(17) = 0.590$, $p = .281$, $d = 0.14$. The contribution of each single factor hence remains unclear. However, considering the analysis of distance and cluster effects which is supporting local memory units only it is quite likely that local reference frames are producing the observed fits. Data of Experiment 2 did not fit the predictions of regional reference frame following the first perspective or the U-shape. Similar to Experiment 1, no evidence could be found that participants used global reference frames in within region pointing trials.

Finally, trials were analyzed in which participants pointed across the regional boundaries. Again, results for the full set of across region trials are inconclusive. No fit to either prediction of local, regional or global reference frames was found, *t's* < 1.21, *p's* > .122. As again variance in pointing was higher for across compared to within region pointing trials, $t(17) = -2.35$, $p = .031$, $d = -0.78$, we re-ran the reference frame analysis for across region trials with corridor distance one, two and three only (comparable standard deviation for within and across trials, $t(17) = -1.23$, $p = .236$, $d = -0.41$). A figure and results of the analysis can be found in Figure 6, right side, and Table 2, right side. Again, a significant fit to a global reference frame prediction was found, namely, for a global reference frame following the U-shape experienced in the first learned region. Neither of the remaining predictions fitted well. Similar to Experiment 1 participants were not faster when aligned with the reference direction of the target region (four main orientations of the obliquely aligned other region), $t(17) = -1.25$, $p = .230$, $d = -0.42$ (based on dataset with all distances), speaking against an alignment with the target reference frame.

Discussion

In the second experiment, a new sample of participants memorized the same regionalized environment as in Experiment 1. Again, both halves of the environment were explored separately. However, this time, participants started learning from the transition point between the two regions, walking towards the dead end of either region. Thus, in contrast to Experiment 1, participants now had the chance to encode the connection of the two regions from the beginning and integrate each additional corridor into their existing knowledge. We again observed effects in favor of local corridor units and global integration. Results regarding regional memory units are less clear.

Pointing latency increased with increasing corridor distance to the target. Similar to Experiment 1, this pattern can be explained by local corridor units that are stored in memory and connected along the learning order. Accessing the location of a target is bound to successively activating each memory unit until reaching the unit that contains the target. Each transition costs time and accuracy. In line with the distance effect, the examination of orientation dependency showed that participants were significantly faster when aligned with the first perspective experienced in each corridor, compared to the seven remaining orientations. This as well speaks for local memory units.

As in Experiment 1 participants learned the connection of both regions well, even though they never experienced the entire environment at once. Pointing accuracy depended on corridor distance, but not on whether participants pointed within or across region. In Experiment 1 we found a facilitative effect on pointing latency when recalling spatial information within one’s current region compared to a more time-consuming access of targets positioned in the other region. This effect disappeared in Experiment 2. Time needed for pointing within one’s current region and across regional boundaries now was of comparable length. The analysis of reference frames for within region pointing revealed—besides the local corridor alignment—a fit to a regional reference frame following the most frequently experienced perspective. This, however, could simply be an artifact of the collinearity of the consulted predictions. Accounting for the facilitative effect of local corridor alignment dissolves the facilitative effect of the most frequently experienced perspective within a region and vice versa. Since the regional reference frame analysis is ambiguous and we find no region distance effect in latency we are reluctant to argue in favor of regional memory units. We cannot exclude them with absolute certainty, but if they are present they seem less strong than in Experiment 1.

Just as in Experiment 1 results suggest that no global reference frames were used for pointing to target location within one’s current region. However, again evidence for global embedding was found in the reference frame analysis of across region trials when limiting target distance to maximum three corridors. Like Experiment 1 the full set of across region trials (target distance 1–7) was indecisive with regards to any reference direction tested.

In sum, our results of Experiment 2 suggest that subjects relied on local memory units and that integration into a global memory unit might have occurred. Surprisingly, changing the learning procedure between Experiment 1 (end-to-transition point) and Experiment 2 (transition-to-end point) appears to mitigate the clustering effects of regions. And this, even though there are still other cues triggering a regionalization. Like Experiment 1, regions in Experiment 2 were still dissociated by color, semantic membership of landmarks, complexity of the angle of turn and spatio-temporal learning experience. Although the effort for global embedding was presumably lowered in Experiment 2 by allowing for a continuous integration of new spatial information, specifically when starting to learn the second region at the transition point participants are already familiar with, this did not elucidate further the findings regarding a global memory unit that were found in Experiment 1. This will be elaborated upon in the General Discussion.

General Discussion

In two experiments participants memorized an environment consisting of two connected regions by active exploration of a virtual environment. Participants either learned each region starting at the dead-end, exploring the four corridors of the region until ending up at the inter-regional transition point (Experiment 1) or they learned each region starting at the transition point, walking ‘outwards’ to the dead-end of each region (Experiment 2). Subsequently, participants pointed from different locations within the environment to targets within the same region or the other region. The aim of this study was to identify how the memory for this clustered space might be structured. Specifically, we aimed to establish whether participants stored the environment on a local corridor level or (also) formed memory units comprising individual regions or the environment as a whole, thus, revealing hierarchical structure of spatial memory. Additionally, we addressed different cues that might set the orientation of potentially formed reference frames on the different levels of hierarchy. In sum, we found evidence for local, regional as well as global memory units across two experiments, and different cues driving reference frame orientation.

To examine the existence of different hierarchical levels we decided for a two-fold approach when analyzing participants pointing performance. First, an analysis of pointing latency with varying corridor distance to and regional belonging of the target was conducted. The speed of memory retrieval should be telling about the memory structure, as pointing latency should increase with every new memory unit that is accessed (see also Meilinger, Strickrodt, & Bülthoff, 2016). Second, we tested how well participants' performance in different body orientations fit to a number of potential spatial reference systems, either limited to local corridors or spread across regions or the whole environment. Bodily alignment with the main orientation of the formed reference frame should facilitate memory recall (e.g., McNamara, Sluzenski, & Rump, 2008; Mou, McNamara, Valiquette, & Rump, 2004). Since measured within the same spatial task (i.e., pointing to non-visible targets located beyond ones' current vista space) both approaches should draw from the same memory source, thus, jointly add to the picture of how the spatial memory for the regionalized space is structured. Whereas the first can give insights into the number and expansion of memory unit(s), the second allows to make conclusions about the nature of those memory units, that is whether they possess the spatial feature of an oriented reference system.

We found clear evidence for the formation of *local* memory units in both experiments and both in the analysis of distance and cluster effects and the analysis of reference frames. Pointing latency increased with increasing corridor distance to the target. This pattern can be explained by the presence of local, interconnected corridor units in memory and by a time-consuming process of successively retrieving all local units that lie between a participant's current location and the target location (see also Meilinger, Strickrodt, & Bülthoff, 2016; Pantelides, Kelly, & Avraamides, 2016). We also found quicker pointing when participants were aligned with the first experienced orientation within each local corridor, compared to being aligned with the remaining body orientations, which indicates that local reference frames were employed, one for each corridor. This was found irrespective of the learning direction (Experiment 1 and 2). Other studies observed evidence for local corridor units as well, in the form of distance effects (Meilinger et al., 2016), local alignment effects (Meilinger et al., 2014; Werner & Schmidt, 1999), and confusion errors based on vista space information (Marchette et al., 2017). The importance of local, bounded places as a core unit for the representation of large-scale space is further substantiated by our results.

Experiment 1 indicates that participants formed *regional* memory units as well. In addition to the local corridor effects, pointing latency also increased when pointing to targets across regions compared to pointing the same distance within a region. Accordingly, the analysis of reference frames revealed evidence for distinct regional reference frames that go beyond the facilitative effect of local corridor alignment. We found a fit to predictions of regional reference frames that follow the first experienced perspective (first corridor of region) and the salient geometry of a U-shape when concentrating on within region pointing trials. Thus, we established clustering into regional units. So far, literature concerned with hierarchical spatial representation and the formation of regions is inconsistent in its understanding of how exactly regions are represented. The format of representing regions or clusters is argued to be either rather non-spatial, such as conceptual or semantic labels (Hirtle & Jonides, 1985; Hirtle & Mascolo, 1986), as well as spatial, in the form of a topological understanding of connectivity and containment (Stevens & Coupe, 1978; Wang & Brockmole, 2003, 2003; Wiener & Mallot, 2003) or in the form of a metrical relational representation (Greenauer & Waller, 2010; McNamara et al., 2008; Meilinger & Vosgerau, 2010). Our study is the first (at least to our knowledge) to provide experimental particulars permitting the conclusion that two distinct reference frames accumulating multiple vista spaces can be formed in regionalized space. Jointly considering both analyses and in particular detecting regional memory units not only by latency increase beyond regional borders, but also by facilitative effects when aligned with regional reference frame levels highlights the spatial character of the regional memory unit. Thus, our results support the concept of regional memory units that are stored in the form of metric relational schemes. For example, multiple vista spaces are linked to one superordinate regional reference frame direction (McNamara et al., 2008; Meilinger & Vosgerau, 2010). Hereby, they extend results by Greenauer and Waller (2010) of micro- and macroreference frames formed for object arrays within a single vista space onto navigable space.

Experiment 2 was less conclusive about the structure of spatial memory on the regional level. In contrast to Experiment 1, the fact that the reference frame alignment effect with the most frequently experienced perspective vanishes when controlling for local corridor alignment does not allow for the conclusion that regional memory units were formed. This is consistent with the absence of a region effect on pointing latency. Corridor distance effects remain across both experiments, supporting the finding of local reference frames. Thus, following the most conservative approach we reckon that the most solid interpretation is that no or only very weak regional units have been formed in Experiment 2. This contrasts studies that succeeded in clustering a navigable space into regions by far less regional cues than those used in Experiment 2 of the current study. Utilizing semantic membership of landmarks and color alone was sufficient to affect subsequent route decisions (Schick et al., 2015; Wiener & Mallot, 2003). One tentative explanation for this inconsistency across studies might be that indeed different formats of regional memory co-occur—semantic (Hirtle & Jonides, 1985; Hirtle & Mascolo, 1986), topological (Stevens & Coupe, 1978; Wang & Brockmole, 2003, 2003; Wiener & Mallot, 2003), and metrical (Greenauer & Waller, 2010; McNamara et al., 2008; Meilinger & Vosgerau, 2010)—, but do not necessarily synchronize. Depending on the used task (i.e., pointing, route planning, etc.) different formats could be targeted. Building on this rationale it seems plausible to assume that the region effects found in Experiment 1 do not originate from effects of categorical belonging to semantic groups (animals vs. tools) or color (blue vs. red). In this case we would have expected to find similar effects on latency when pointing across regional boundaries in Experiment 2 as well. Nevertheless, it is beyond the scope of the current study to try and clarify this discrepancy.

Importantly, all we changed between Experiment 1 and 2 was the learning direction, and thus, participants awareness of how both regions are connected to each other at the very start of learning both regions. Although our data does not allow for strong claims, one possible explanation for the absence of regional memory units in Experiment 2 could be the opportunity to immediately relate each newly explored corridor to the existing memory structure of the first region when starting to learn the second region at the transition point again. In contrast, in Experiment 1 (end-to-transition) the immediate connection of unfamiliar corridors with existing memory structures was not possible. A novel, independent spatial unit might have been formed at the start of the second region with the new, following corridors added immediately. Restructuring the already existing two separate regions when reaching the familiar transition point into one common unit might be associated with higher mental effort compared to simply learn how the two regions are connected to each other, leading to the observed region effects. Indeed, comparable results to our study were found by Han and Becker (2014), who had participants learn landmarks located in two connected regions while varying the point of time the connecting route was introduced (e.g., immediately/very early in the learning phase vs. only after a few blocks of learning regions separately and being tested throughout) and the extend and means of encounter with it (e.g., watching a video of the connecting route vs. seeing the connecting route and parts of the other region constantly vs. actively navigating across regions). Slower pointing latencies for across region pointing where only found when the connecting route was introduced in later blocks, while enabling immediate and very early encounter always led to comparable pointing latencies for within and across region pointing. From studies investigating film and narrative comprehension it is known that, among others, (unexplained) changes in spatial locations of the protagonist negatively effects reading time of a narrative (e.g., Rinck & Weber, 2003, Who when where; Scott Rich & Taylor, 2000), film cuts alter comprehension and memory of the movie (Schwan, Garsoffsky & Hesse, 2000, Do film cuts facilitate the perceptual and cognitive organization of activity sequences), and temporal shifts in a narrative can weaken memory binding of pre- and post-shift content (Ezzyat & Davachi, 2011; Zwaan, 1996; Zwaan, Langston, & Graesser, 1995). Hence, strong discontinuities in an episode alter how such content is perceived and remembered. This is in line with, for example some associative memory models that suggest that (un)available shared context affects associative binding between incoming information (Howard & Kahana, 2002; Polyn, Norman, & Kahana, 2009a, 2009b). Translated to the context of this study, not having access to an immediate reference to the previously learned region when stating to learn the second region in Experiment 1 may both lead to the detection of a spatial discontinuity ("I am somewhere else") as well as stronger temporal discontinuity compared to Experiment 2 because the transition point is discovered later in time. In sum, the availability of a reference point early on during learning connected spaces that at best allows for a continuous flow of incoming information across regional boundaries might be a very crucial factor for the integration of the two spaces and, vice versa, the unavailability thereof determine clustering in memory. Although this explanation fits to the pattern observed across the two experiments, further research is necessary to understand the processes of regional clustering. For example, alternatively the walking direction itself, namely whether regions were first explored on a convergent (end-to-transition point, Experiment 1) or on a divergent path (transition-to-end, Experiment 2), irrespective of the time the transition point was learned, could alter memory formation.

In both Experiments we observed evidence for the integration of all eight corridors into a global memory unit. In Experiment 1 global reference frames following the first experienced perspective and the U-shape of the first learned region were indicated by the reference frame analysis of across region trials, in Experiment 2 for the U-shape. However, this was only the case when reverting to trials querying corridor distance one to three, leaving out trials with higher corridor distances. Individual standard deviation suggests that this might be due to the increased noise in the across region data of all distances compared to within region trials were only corridor distance one, two and three were tested. It should be noted here that the analysis of distance and cluster effects is tailored to detect either local and regional memory units or, in contrast, the absence of local and regional memory units and the exclusive presence of one global memory unit. In the latter case, no increase in latency across corridors or the regional boundary would have been expected, since all spatial information was accessible with similar ease from within one common memory unit as when learning from a city map (Frankenstein, Mohler, Bülthoff, & Meilinger, 2012). However, we observed distance and regional cluster effects that are in favor of local and regional memory units; and based on these results alone no conclusions about the formation of an additional global memory unit can be made. Therefore, consulting results from the reference frame analysis is essential. Interestingly, jointly considering results from the distance and cluster analysis and from the reference frame alignment for across region pointing covering distance one to three reveals successive activation of local memory units and the formation of a global reference frame. This indicates that although a global reference frame is used for recalling across region information no simple all-at-once readout from this memory unit occurred. Rather, even a retrieval process that is based on a global embedding seem to be bound to successive activation of local memory units. A similar interpretation can be made based on the observation of local and regional memory units for within region pointing in Experiment 1. Having access to and using a regional reference frame does still involve successive activation of local memory units, leading to both local and regional effects. If successive activation of local memory units is essential even in the presence of and during utilization of higher order memory units it, first of all, raises the question of what exactly is stored in a global (and regional) reference frame if this information cannot be used for an easy and fast read-out. Secondly, it might explain why including higher corridor distances led to non-conclusive results for across region trials. The successive integration of local memory units—even though stored additionally within a global reference system—at the moment of testing might be subject to limitations of working memory capacity. Utilizing both representations from local and global memory units might functions exclusively as long as this can be done within working memory capacity. The process, however, cannot be upheld when going beyond the limit at a specific corridor distance, therefore, our effects vanish with corridor distances higher than three. Considering Figure 2 and 5 this might well be the case at around corridor distance 4 or 5, where the distance effect on latency seems to flatten out. It remains a question for further investigation what exactly is happening to the spatial recall process if working memory limit is reached.

We cannot completely rule out an alternative explanation for the fit to a global reference frame prediction for smaller distances in across region pointing, namely, that no global but an additional memory unit on the regional level was formed which is not covering the whole environment, but only the corridors around the transition point. Indeed, across region trials covering maximum corridor distance of three are limited to the area around the transition point, namely from there three corridors of the blue and three of the red region, leaving out the outermost corridors. However, if another unit on the regional level was formed we would expect to find a fit to local reference frame predictions as well, just as for pointing within region. As this was not the case, formation of a global reference frame seems more likely.

The use of global reference frames for pointing across regional boundaries only is in line with findings on spatial layouts learned in vista space by Greenauer and Waller (XXX) and extend them to environmental space. Greenauer and Waller (2010) found evidence for a macroreference frame when pointing across two separate object layouts that were learned in a single room, but not when pointing within one object array. Particularly, the study implies that superordinate reference frames might only be used when pointing to a target which is located beyond the scope of a smaller microreference frame. Likewise, in our study global reference frames seem to be accessed flexibly only when required, namely, when recalling the relative direction of a target in the other region.

As pointed out before, the reported effects indicating local and regional reference frames are restricted to trials of within region pointing, while pointing trials that target objects within the other region show recall behavior independent of local and regional main orientations. Importantly, even though we also see corridor distance effects in across region trials, here no evidence for local reference frames can be detected. In other words, even though the distance effect supports the use of local memory units when pointing across regional boundaries, the facilitative effect of local alignment disappears as soon as participant point to a target within the other region. An attempted explanation could be that, when pointing across region, the corridor orientation visible from one’s current location might conflict with the obliquely oriented geometry of the other region, if visualized in the same reference frame. A study by Meilinger and

Bülthoff (2013) suggests that visual pointing can lead to interference between the surrounding visual geometry and the geometry of the memorized environment. However, this interpretation is speculative. Whether and how visual input can interfere with selected reference frames, however, cannot be resolved based on our results.

What is setting the direction of a regional reference frame? Former studies concerned with the formation of a reference frame covering multiple vista spaces focused on the importance of the first perspective taken in the navigable space that must be memorized (e.g., Richardson, Montello, & Hegarty, 1999; Tlauka, Carter, Mahlberg, & Wilson, 2011; Wilson, Wilson, Griffiths, & Fox, 2007). In these studies, the first corridor walked was typically also the longest and/or parallel to the last corridor, forming a salient geometric U-shape. Thus, the salience of the first perspective was partially associated with the most frequently exposed perspective and the parallel environmental structure. In Experiment 1 we found support for the influence of the first perspective as well as the U-shape for setting the regional and global reference frame direction and Experiment 2 replicated the fit to a global reference frame following the U-shape of the first learned region. This indicates that the representation of navigable space might not only be shaped by initial views (first corridor) that form the basis of a reference frame into which subsequent spatial information (following corridors) are integrated into (assimilation), but by spatial information gathered across a sequence of corridors which, in combination, can form a new distinct reference direction (accommodation) (Piaget & Inhelder, 1969). This is in line with an outlook given by McNamara and Valiquette (2004) after outlining their theoretical framework of spatial reference systems, where they state “the first segment is a strong candidate because of the salience conferred by novelty”, but furthermore “It is also possible that one or more of the other segments of the path might be used to establish a reference system” (pp. 21-22). Many studies examined the factors influencing the alignment of the mental reference frame in vista space (e.g., Kelly & McNamara, 2008; Mou & McNamara, 2002; Shelton & McNamara, 2001; Valiquette & McNamara, 2007). Our study highlights, that comparable efforts should now be taken to shed light on the factors influencing reference frame selection in navigable space as well.

We realize that one must remain cautious in making too strong interpretation based on the analysis of reference frames alone, and in particular, about the evidence for regional and global reference frames. Significant fits reported are based on one-sided t-tests, without correction for multiple tests evaluating the three theoretical predictions (first perspective, most frequent, U-shape) on regional and global level for within and across pointing. Nevertheless, both the separation of within and across trials and the three potential cues for reference frame alignment are theoretically motivated and specified as directional hypotheses. Furthermore, results of local and global reference frames could be replicated across two experiments. And most importantly, the reference frame results are in high accordance with the analysis of distance and region effects.

Participants were well able to learn the environment. Pointing accuracy did not depend on whether participants pointed to targets within one's current region or to targets within the other region. And this, despite separate learning of both regions and despite a single, comparatively complex (45°) common reference point at the transition between the two halves of the environment. Previous studies already showed the rapid learning of spatial relations between two areas immediately after being exposed to a connecting route (e.g., Ishikawa & Montello, 2006; Schinazi, Nardi, Newcombe, Shipley, & Epstein, 2013) and our study is in line with those results. In addition, we present evidence that global metric embedding can take place without ever walking the entire environment at once.

Substantiating the presence of local, regional and global reference frames, our study promotes hierarchical concepts of spatial memory (e.g., Mallot & Basten, 2009; Stevens & Coupe, 1978; Wiener & Mallot, 2003). It should not be left unsaid that the current analysis is indecisive of whether multiple hierarchical levels occur within a single participant or only across participants. The analysis is based on the average pointing performance of the whole sample. Efforts to identify single level or hierarchical memory structures within single participants failed to produce interpretable results. Also, it was not possible to separate participants in local-only, region-only or global-only groups based on their performance. It remains an issue of further examination to ascertain whether multiple reference frames are used by single individuals (i.e., hierarchical representation), or whether our results mirror the average of individual strategies across participants, where each participant relies exclusively on the local, regional or global level.

In past research going beyond the immediately visible surrounding of vista space uncovered multiple aspects of spatial learning. Opaque borders seem to distort distance judgements (e.g., Kosslyn et al., 1974; Newcombe & Liben, 1982), affect online updating of landmark locations (e.g., Avraamides & Kelly, 2010; Wang & Brockmole, 2003, 2003) and lead to latency costs when switching between spatial units (e.g., Brockmole & Wang, 2002, 2003) or when pointing to targets located in increasingly distant corridors (Meilinger et al., 2016). Also, exact spatial knowledge of a target within a vista space does not coincide with a similarly good knowledge about where this vista space itself is located (Marchette et al., 2017). Additionally, individual reference frames were found to be formed for individual vista spaces (Meilinger et al., 2014; Werner & Schmidt, 1999). Considered jointly these findings suggest that visual boundaries enclosing vista spaces seem to serve as molds for local spatial memory units. Results of these studies can be well explained by non-hierarchical theories postulating the formation of local memory units and their successive connection (e.g., Chrastil & Warren, 2014; Meilinger, 2008; Warren, Rothman, Schnapp, & Ericson, 2017). Non-hierarchical theories assuming not multiple local, but a single global representation of all spatial information cannot account for these results.

Notwithstanding, there is a large number of studies indicating hierarchical structures in human spatial memory. For example, Stevens and Coupe (1978) showed that the relative position between remote cities seemed to be judged based on the relative position of the federal states the cities are in rather than their actual location. Both for vista (McNamara, Hardy, & Hirtle, 1989) and environmental space (Hirtle & Jonides, 1985) hierarchical clusters were unveiled based on participants landmark recall protocols. Landmark proximity in these hierarchical clusters were associated with distorted distance estimations between landmarks (Hirtle & Jonides, 1985; McNamara et al., 1989) and priming effects in a recognition task (McNamara et al., 1989). Evidencing clustering of spatial information is a prerequisite for hierarchical organization of space. However, a common concern regarding these studies is, whether the results could likewise be explained by non-hierarchical theories when assuming that object locations are indeed clustered, but they are not organized in a hierarchical fashion. In this case, judgements of relation and distance as well as priming effects might only reflect a distorted memory misrepresenting physical space, but not necessarily a memory that does possess another level of hierarchy for regions and clusters.

We suspect that to make profound interpretations regarding hierarchical spatial representations often additional dependent variables and/or approaches need to be consulted, as was done in a number of other studies. For example, McNamara (1986) contrasted participants distance judgements between landmarks (reflecting possible distortions in memory) and spatial priming effects in a recognition task after participants learned a clustered vista space containing four regions. While non-hierarchical theories would predict an exponential decay of priming effects with psychological distances this was not reflected in the data. Thus, priming effects seemed to not simply mirror an erroneous representation distorted according to formed clusters, but a hierarchical representation. In a route choice task Wiener and Mallot (2003) showed that participants tended to approach the region containing the target object directly rather than choosing a path with a longer dwelling time in the non-target region but equivalent in length and complexity. Importantly, for one of the environments they used it seems reasonable to assume that the two alternative routes would also lead to equally long psychological route distances. The environment was a grid field of two rectangles arranged opposite each other that formed two regions. One would expect a similar distortion in spatial memory within both regions, potentially towards the centroids of each region (e.g., Huttenlocher, Hedges, & Duncan, 1991), thus, distortions in memory could not account for a bias in route selection. Still detecting region effects on route planning suggests an additional memory layer for representing regions. Our results line up nicely with these studies by highlighting local, regional and global effects with two complementary analyses that jointly indicate hierarchical structures in human spatial memory. Latency increase across corridor distance and regional boundaries indicate that new memory units must have been activated. The reference frame analysis supports the effects further and acknowledges the spatial character of the units, while complementing them with evidence for global reference frames. We reckon that non-hierarchical theories have difficulties explaining these results.

Following an alternative approach, the superordinate level, which manifests in regional boundary and regional and global alignment effects, might not consist of a coordinate system that is yielding metric embedding of subordinate units within a region or across the environment, but rather of a common reference direction or vector. This does not touch the assumption that memory units for vista spaces are stored in the form of locally confined, spatial reference frames, leading to the observed distance and local alignment effects. However, the encountered regional and global effects could also be caused by additional regional or one additional global reference vector stored in memory on a single superordinate level, a main orientation extrapolated across a limited or the entire number of vista spaces. Like the spatial reference system proposed by Shelton and McNamara (2001) the direction itself can be described as a conceptual “north”, a privileged direction in the environment. It might be set, for example, by the first segment walked (first perspective alignment), or by salient inner-regional structures such as a U-shape. However, in contrast to Shelton and McNamara (2001), the superordinate vector is different in a sense that spatial relations are not explicitly specified with respect to a spatial reference system as coordinates, but rather it reflects an anchor orientation propagated across multiple corridors, for example, via a global sense of direction system (Sholl, Kenny, & DellaPorta, 2006). Bodily alignment with this superordinate vector does allow for an easy access of the remaining object locations *not* due to the availability of and alignment with a spatial reference system that contains the relative position of objects located beyond one's current vista space, but rather because the superordinate vector facilitates the coordination and alignment of the local memory units stored on the subordinate level. Such a vector approach could explain the fact that distance effects prevail—suggesting successive activation of local memory unit, rather than an all-at-once readout—in the presence of regional and global orientation dependencies, as found in our study. However, based on a global sense of direction system (Sholl, Kenny, & DellaPorta, 2006) it also limits the hierarchical representation to a maximum of two levels, a local level and an additional vector either encompassing the corridors of a single region or spanning both regions - with sufficient accuracy of the common global direction around the transition point. This would indeed imply that the detection of regional and global memory units in Experiment 1 are due to individual differences, meaning that some participants formed two separate regional vectors, others formed a single global vector. This would be an additional conjecture to explain our results. At the same time this approach is more economic than assuming an additional third level to be represented as well as in the case of the hierarchical representation of local, regional and global reference frames. Such a superordinate vector could account for alignment effects found across multiple vista spaces and still allow for globally inconsistent spatial memory and biases (e.g., Warren, Rothman, Schnapp, & Ericson, 2017) as only orientation, but not location is specified on the superordinate level. Forming such a vector across multiple corridors during the learning phase should involve accumulation of error during updating, thus, accuracy and conformity of this vector should decrease with increasing corridor distance. Besides the aspect stated earlier, namely, that limited working memory capacity affects successive recall of local units, also error accumulation during learning could explain why higher corridor distances for across region pointing led to inconclusive results about a uniform global reference direction. In that sense a superordinate spatial reference system might consist of a full-blown coordinate system or a single reference direction. Although such an approach might be uncommon and in need of additional theoretical polishing, we regard it worth further investigation in future experiments.

Taken together, we find that participants encoded navigation spaces on the local (Experiment 1 and 2), regional (Experiment 1) and global level (Experiment 1 and 2). The presence of regional reference frames indicated in the alignment effects suggest that multiple vista spaces that form clearly circumscribed areas may serve as molds for spatial memory units on the regional level, and that the representation of regions is not limited to conceptual or topological knowledge (Hirtle & Mascolo, 1986; Meilinger, 2008; Wang & Brockmole, 2003, 2003; Wiener & Mallot, 2003), but can indeed rely on distinct mental reference systems subsuming multiple places, thus, possessing a spatial character (e.g., McNamara et al., 2008; Meilinger & Vosgerau, 2010). The segmentation into distinct spatial reference systems on the regional level might be dependent on the learning procedure. Experiment 2, which allowed for an immediate relation of new spatial information to existing knowledge by starting to learn the second region at a familiar reference point, does not seem to trigger the formation of regionally confined spatial reference systems as strongly as being introduced to a common reference point between two areas at a later time during learning (Experiment 1). Thus, the point in time when the connectivity between separately learned spaces is introduced might play a role for the emerging structure of spatial memory. Spatial information of the entire environment was found to be integrated on a global level. The use of this level, however, was dependent on the relationship between one’s current position and the target location (i.e., only for across region pointing), indicating that the stored mental structures representing navigable space are not consulted exhaustively every time spatial information are recalled but used in customized fashion instead. Importantly, our findings demonstrate that the representation of navigable space is not limited to local memory units encompassing single corridors or streets (e.g., Chrastil & Warren, 2014; Meilinger, 2008; Warren, Rothman, Schnapp, & Ericson, 2017). Similarly, the evidence for local and regional memory units besides support for a global embedding speaks against a purely global integration into a single, common reference frame (e.g., Gallistel, 1990; Ishikawa & Montello, 2006; O’Keefe & Nadel, 1978; Sholl, 2001).

Ultimately, our results show that object locations in navigable space seem to be represented flexibly on different levels (maybe not necessarily within a single individual) and thereby support hierarchical theories of spatial memory (e.g., Mallot & Basten, 2009; Stevens & Coupe, 1978; Wiener & Mallot, 2003). We found strong support that local memory units seem to be a key component for memorizing navigable space, as they are pervasive across both experiments. Potentially, they form the basic units in spatial memory, that subsequently can be consolidated to form memory units on one or more superordinate levels, for example, in the form of regional and global reference frames. The processes involved in the generation of such hierarchical representation and how the memory units interact within and across hierarchical levels will have to be clarified in future research.

Acknowledgments

We would like to thank Joachim Tesch for help with the virtual environment setup.

References

Avraamides, M. N., & Kelly, J. W. (2010). Multiple systems of spatial memory: evidence from described scenes. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *36*(3), 635–645. http://doi.org/10.1037/a0017040

Brockmole, J. R., & Wang, R. F. (2002). Switching between environmental representations in memory. *Cognition*, *83*(3), 295–316. http://doi.org/10.1016/S0010-0277(02)00012-4

Brockmole, J. R., & Wang, R. F. (2003). Changing perspective within and across environments. *Brain & Development*, *25*(4), 291–293. http://doi.org/10.1016/S0

Byrne, R. W. (1979). Memory for urban geography. *Quarterly Journal of Experimental Psychology*, *31*(1), 147–154. http://doi.org/10.1080/14640747908400714

Chrastil, E. R., & Warren, W. H. (2014). From Cognitive Maps to Cognitive Graphs. *PLoS ONE*, *9*(11), e112544. http://doi.org/10.1371/journal.pone.0112544

Dehaene, S., Izard, V., Pica, P., & Spelke, E. (2006). Core knowledge of geometry in an Amazonian indigene group. *Science*, *311*(5759), 381–384. http://doi.org/10.1126/science.1121739

Dillon, M. R., Huang, Y., & Spelke, E. S. (2013). Core foundations of abstract geometry. *Proceedings of the National Academy of Sciences*, *110*(35), 14191–14195. http://doi.org/10.1073/pnas.1312640110

Frankenstein, J., Mohler, B. J., Bülthoff, H. H., & Meilinger, T. (2012). Is the map in our head oriented north? *Psychological Science*, *23*(2), 120–5. http://doi.org/10.1177/0956797611429467

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

Greenauer, N., & Waller, D. (2010). Micro- and macroreference frames: Specifying the relations between spatial categories in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(4), 938–57. http://doi.org/10.1037/a0019647

Hermer, L., & Spelke, E. (1996). Modularity and development: the case of spatial reorientation. *Cognition*, *61*, 195–232. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.73.3709&rep=rep1&type=pdf

Hermer, L., & Spelke, E. S. (1994). A geometric process for spatial reorientation in young children. *Nature*, *370*(6484), 57–59. http://doi.org/10.1038/370057a0

Hirtle, S. C., & Jonides, J. (1985). Evidence of hierarchies in cognitive maps. *Memory &*

Cognition, *13*(3), 208–217. http://doi.org/10.3758/BF03197683

Hirtle, S. C., & Mascolo, M. F. (1986). Effect of semantic clustering on the memory of spatial locations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(2), 182–189. http://doi.org/10.1037/0278-7393.12.2.182

Huttenlocher, J., Hedges, L. V, & Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review*, *98*(3), 352–376. Retrieved from http://psycnet.apa.org/journals/rev/98/3/352/

Ishikawa, T., & Montello, D. R. (2006). Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places. *Cognitive Psychology*, *52*(2), 93–129. http://doi.org/10.1016/j.cogpsych.2005.08.003

Jeanne Sholl, M. (2001). The Role of a Self-Reference System in Spatial Navigation. In D. R. Montello (Ed.), *Cosit 2001* (pp. 217–232). Springer Berlin Heidelberg. http://doi.org/10.1007/3-540-45424-1_15

Jeffreys, H. (1961). *Theory of Probability* (3rd ed.). Oxford: Oxford University Press.

Kelly, J. W., & McNamara, T. P. (2008). Spatial memories of virtual environments: How egocentric experience, intrinsic structure, and extrinsic structure interact. *Psychonomic Bulletin & Review*, *15*(2), 322–327. http://doi.org/10.3758/PBR.15.2.322

Kelly, J. W., & McNamara, T. P. (2010). Reference frames during the acquisition and development of spatial memories. *Cognition*, *116*(3), 409–420. http://doi.org/10.1016/j.cognition.2010.06.002

Kosslyn, S. M., Pick, H. L., & Fariello, G. R. (1974). Cognitive maps in children and men. *Child Development*, *45*(3), 707–716. http://doi.org/10.1111/1467-8624.ep12147012

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian Cognitive Modeling: A Practical Course*. Cambridge University Press. Retrieved from https://faculty.washington.edu/jmiyamot/p548/leemd bayesian cog modeling - practical crs.pdf

Mallot, H. A., & Basten, K. (2009). Embodied spatial cognition: Biological and artificial systems. *Image and Vision Computing*, *27*(11), 1658–1670. http://doi.org/10.1016/j.imavis.2008.09.001

Marchette, S. A., Ryan, J., & Epstein, R. A. (2017). Schematic representations of local environmental space guide goal-directed navigation. *Cognition*, *158*, 68–80. http://doi.org/10.1016/j.cognition.2016.10.005

McNamara, T. P. (1986). Mental representations of spatial relations. *Cognitive Psychology*, *18*(1), 87–121. http://doi.org/10.1016/0010-0285(86)90016-2

McNamara, T. P., Hardy, J. K., & Hirtle, S. C. (1989). Subjective hierarchies in spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(2), 211–227. http://doi.org/10.1037/0278-7393.15.2.211

McNamara, T. P., Sluzenski, J., & Rump, B. (2008). Human Spatial Memory and Navigation. In H. L. Roediger III (Ed.), *Learning and Memory: A Comprehensive Reference* (Vol. 2, pp. 157–178). Oxford: Elsevier. http://doi.org/10.1016/B978-012370509-9.00176-5

McNamara, T. P., & Valiquette, C. M. (2004). Remembering where things are. In G. L. Allen (Ed.), *Human Spatial Memory: Remembering Where* (pp. 3–24). Mahwah, NJ: Lawrence Erlbaum Associates. http://doi.org/10.4324/9781410609984

Meilinger, T. (2008). The Network of Reference Frames Theory: A Synthesis of Graphs and Cognitive Maps. In *Spatial Cognition VI. Learning, Reasoning, and Talking about Space* (Vol. 5248 LNAI, pp. 344–360). Berlin, Heidelberg: Springer Berlin Heidelberg. http://doi.org/10.1007/978-3-540-87601-4_25

Meilinger, T., & Bülthoff, H. H. (2013). Verbal shadowing and visual interference in spatial memory. *PloS One*, *8*(9), e74177. http://doi.org/10.1371/journal.pone.0074177

Meilinger, T., Frankenstein, J., Watanabe, K., Bülthoff, H. H., & Hölscher, C. (2015). Reference frames in learning from maps and navigation. *Psychological Research*, *79*(6), 1000–1008. http://doi.org/10.1007/s00426-014-0629-6

Meilinger, T., Riecke, B. E., & Bülthoff, H. H. (2014). Local and global reference frames for environmental spaces. *Quarterly Journal of Experimental Psychology (2006)*, *67*(3), 542–69. http://doi.org/10.1080/17470218.2013.821145

Meilinger, T., Strickrodt, M., & Bülthoff, H. H. (2016). Qualitative differences in memory for vista and environmental spaces are caused by opaque borders, not movement or successive presentation. *Cognition*, *155*, 77–95. http://doi.org/10.1016/j.cognition.2016.06.003

Meilinger, T., & Vosgerau, G. (2010). Putting Egocentric and Allocentric into Perspective. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 6222 LNAI, pp. 207–221).

http://doi.org/10.1007/978-3-642-14749-4_19

Moar, I., & Bower, G. H. (1983). Inconsistency in spatial knowledge. *Memory & Cognition*, *11*(2), 107–113. http://doi.org/10.3758/BF03213464

Montello, D. R. (1993). Scale and multiple psychologies of space. In *Spatial Information Theory* (pp. 312–321). http://doi.org/10.1007/3-540-57207-4_21

Mou, W., & McNamara, T. P. (2002). Intrinsic frames of reference in spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(1), 162–170. http://doi.org/10.1037//0278-7393.28.1.162

Mou, W., McNamara, T. P., Valiquette, C. M., & Rump, B. (2004). Allocentric and egocentric updating of spatial memories. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *30*(1), 142–157. http://doi.org/10.1037/0278-7393.30.1.142

Newcombe, N., & Liben, L. S. (1982). Barrier effects in the cognitive maps of children and adults. *Journal of Experimental Child Psychology*, *34*(1), 46–58. http://doi.org/10.1016/0022-0965(82)90030-3

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

Pantelides, S. N., Kelly, J. W., & Avraamides, M. N. (2016). Integration of spatial information across vision and language. *Journal of Cognitive Psychology*, *28*(2), 171–185. http://doi.org/10.1080/20445911.2015.1102144

Piaget, J., & Inhelder, B. (1969). *The psychology of the child*. New York: Basic Books.

Prinzmetal, W., McCool, C., & Park, S. (2005). Attention: reaction time and accuracy reveal different mechanisms. *Journal of Experimental Psychology. General*, *134*(1), 73–92. http://doi.org/10.1037/0096-3445.134.1.73

Richardson, A. E., Montello, D. R., & Hegarty, M. (1999). Spatial knowledge acquisition from maps and from navigation in real and virtual environments. *Memory & Cognition*, *27*(4), 741–750. http://doi.org/10.3758/BF03211566

Rieser, J. J. (1989). Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(6), 1157–1165. http://doi.org/10.1037/0278-7393.15.6.1157

Schick, W., Halfmann, M., & Mallot, H. A. (2015). How to construct a linguistic landmark: language cues in the formation of hierarchical representations of space. *Cognitive*

Processing, *16*(1), 383–388. http://doi.org/10.1007/s10339-015-0722-9

Schinazi, V. R., Nardi, D., Newcombe, N. S., Shipley, T. F., & Epstein, R. A. (2013). Hippocampal size predicts rapid learning of a cognitive map in humans. *Hippocampus*, *23*(6), 515–528. http://doi.org/10.1002/hipo.22111

Shelton, A. L., & McNamara, T. P. (2001). Systems of Spatial Reference in Human Memory. *Cognitive Psychology*, *43*(4), 274–310. http://doi.org/10.1006/cogp.2001.0758

Sholl, M. J., Kenny, R. J., & DellaPorta, K. A. (2006). Allocentric-heading recall and its relation to self-reported sense-of-direction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(3), 516–533. Retrieved from http://psycnet.apa.org/fulltext/2006-06644-006.html

Sternberg, S. (1969). Memory-scanning: mental processes revealed by reaction-time experiments. *American Scientist*, *57*(4), 421–57. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5360276

Stevens, A., & Coupe, P. (1978). Distortions in judged spatial relations. *Cognitive Psychology*, *10*(4), 422–437. http://doi.org/10.1016/0010-0285(78)90006-3

Tlauka, M., Carter, P., Mahlberg, T., & Wilson, P. N. (2011). The first-perspective alignment effect: The role of environmental complexity and familiarity with surroundings. *The Quarterly Journal of Experimental Psychology*, *64*(11), 2236–2250. http://doi.org/10.1080/17470218.2011.586710

Tversky, B. (1981). Distortions in Memory for Maps. *COGNITIVE PSYCHOLOGY*, *13*, 407–433. Retrieved from https://www.tc.columbia.edu/faculty/bt2158/faculty-profile/files/1981_Tversky_Distortionsinmemoryformaps.pdf

Valiquette, C., & McNamara, T. P. (2007). Different mental representations for place recognition and goal localization. *Psychonomic Bulletin & Review*, *14*(4), 676–680. http://doi.org/10.3758/BF03196820

Wang, R. F. (2016). Building a cognitive map by assembling multiple path integration systems. *Psychonomic Bulletin & Review*, *23*(3), 692–702. http://doi.org/10.3758/s13423-015-0952-y

Wang, R. F., & Brockmole, J. R. (2003). Human navigation in nested environments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(3), 398–404. http://doi.org/10.1037/0278-7393.29.3.398

Wang, R. F., & Brockmole, J. R. (2003). Simultaneous spatial updating in nested environments. *Psychonomic Bulletin & Review*, *10*(4), 981–986. http://doi.org/10.3758/BF03196562

Warren, W. H., Rothman, D. B., Schnapp, B. H., & Ericson, J. D. (2017). Wormholes in virtual space: From cognitive maps to cognitive graphs. *Cognition*, *166*, 152–163. http://doi.org/10.1016/j.cognition.2017.05.020

Weisberg, S. M., Schinazi, V. R., Newcombe, N. S., Shipley, T. F., & Epstein, R. a. (2014). Variations in cognitive maps: Understanding individual differences in navigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(3), 669–682. http://doi.org/10.1037/a0035261

Werner, S., & Schmidt, K. (1999). Environmental reference systems for large-scale spaces. *Spatial Cognition and Computation*, *1*, 447–473. http://doi.org/10.1023/A:1010095831166

Wertheimer, M. (1923). *Laws of organization in perceptual forms. A source book of Gestalt Psychology*. *psychclassics.yorku.ca*. Retrieved from http://psychclassics.yorku.ca/Wertheimer/Forms/forms.htm

Wiener, J. M., & Mallot, H. A. (2003). “Fine-to-Coarse” Route Planning and Navigation in Regionalized Environments. *Spatial Cognition & Computation*, *3*(4), 331–358. http://doi.org/10.1207/s15427633scc0304_5

Wilson, P. N., Wilson, D. a, Griffiths, L., & Fox, S. (2007). First-perspective spatial alignment effects from real-world exploration. *Memory & Cognition*, *35*(6), 1432–44. http://doi.org/10.3758/BF03193613

Zhang, H., Mou, W., McNamara, T. P., & Wang, L. (2014). Connecting spatial memories of two nested spaces. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *40*(1), 191–202. http://doi.org/10.1037/a0033776

Appendix

Figures

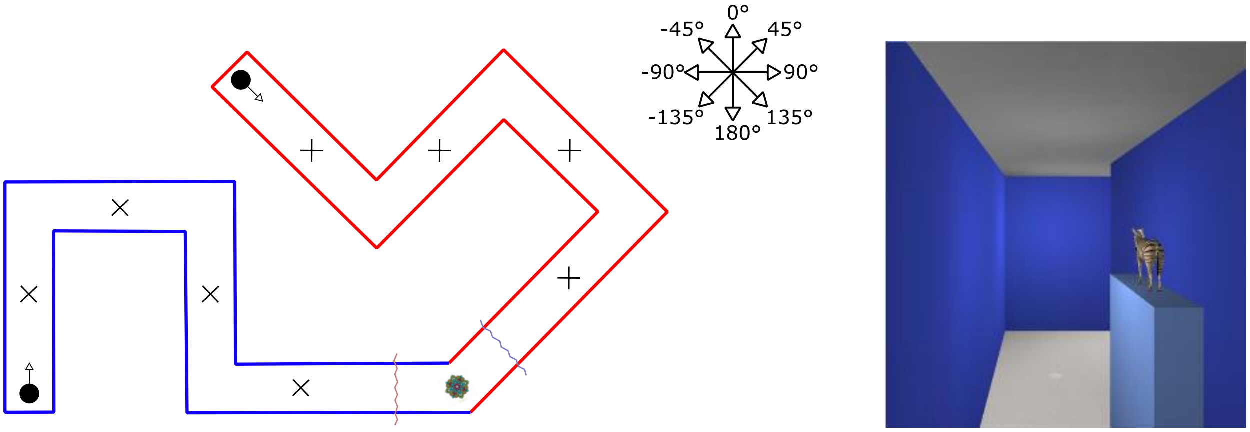

Figure 1. Left: Schematic, aerial view of the environment. Each corridor contained one object (x). A mandala marks the transition point between the two regions. Black dots with arrows indicate initial positions and orientation within both regions (Experiment 1). Wave lines indicate the approximate position of the curtain obstructing the view into the other region when standing at the mandala. Star of black arrows: Eight body orientations tested during pointing while standing at different predefined positions within the environment (at the x's within each corridor). Right: Example of egocentric view in the environment as experienced by participants.

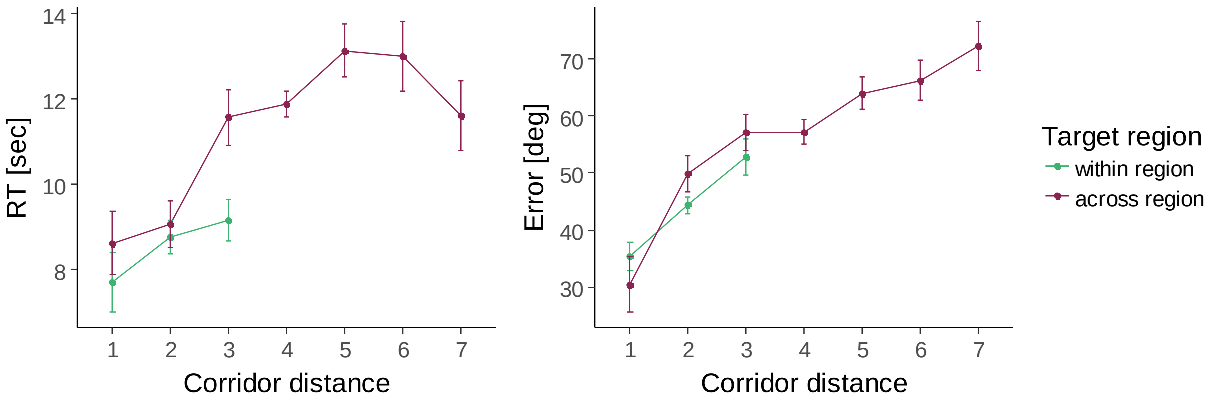

Figure 2. Distance and cluster effects in Experiment 1. Pointing latency (left) and error (right) as a function of corridor distance to the target, pointing to targets within the same region vs. across region. Means and SEMs are depicted. Only data points from corridor distance 1, 2 and 3 were analysed.

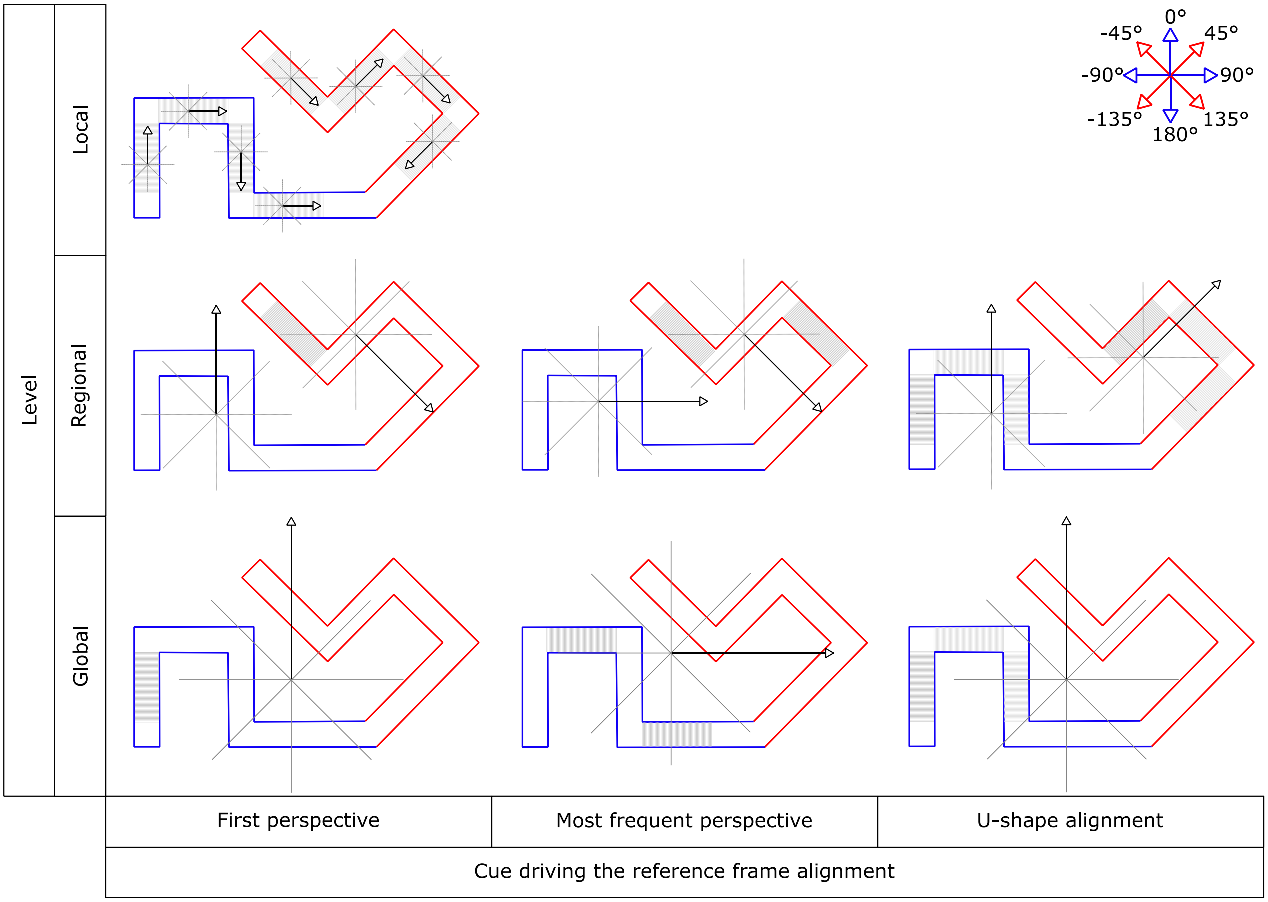

Figure 3. Visualization of the predictions for formed reference frames. We tested for reference frames on the local, regional or global level (rows) that are either following the first perspective, the most frequent perspective or the alignment of the u-shape of three consecutive corridors (columns). Cues driving the reference frames are marked with grey lines in the respective corridor(s). Following each prediction one orientation can be identified as being superior, marked by a black, solid arrow, compared to the remaining orientations, marked by grey, dashed lines. Being aligned with this orientation should yield fastest spatial recall compared to being aligned otherwise. The predictions depicted here are exemplary – they represent predictions for Experiment 1 (learning from dead-end to transition point). Further, the global prediction accounts for subjects that start off learning the blue region and later continue with the red, therefore, global reference frame predictions follow the blue (first learned) region. Learning from transition point to dead-end (Experiment 2) and starting off with region red (half the sample of Experiment 1 and 2) are not depicted here but can be inferred following the same logic.

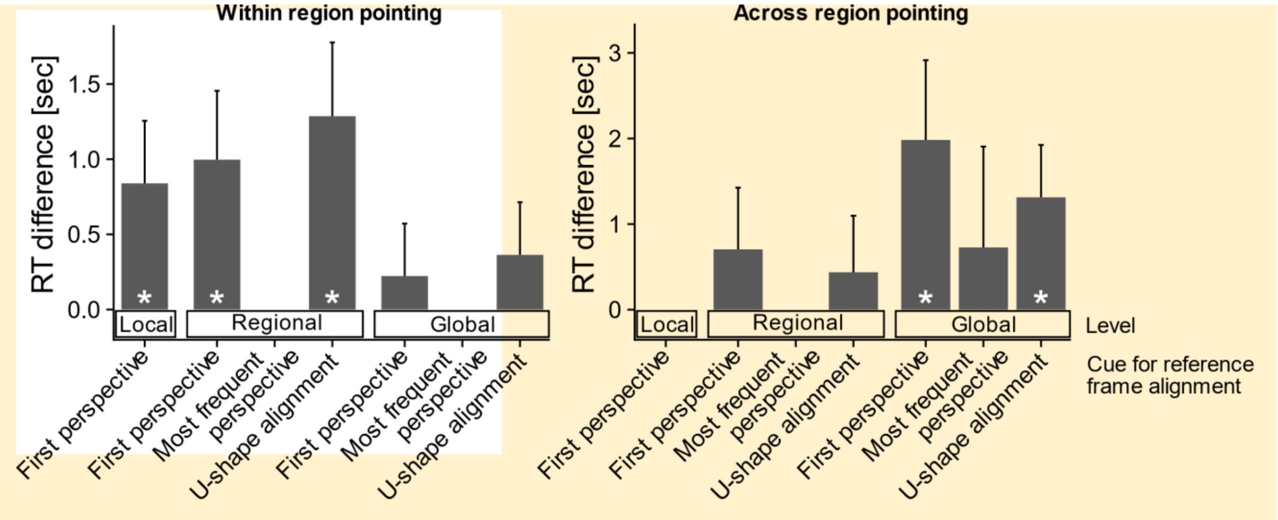

Figure 4. Reference frame alignment in Experiment 1. Difference in pointing latency when being aligned with the superior orientation identified for each prediction compared to the remaining seven orientations, separately for within (left) and across region trials (right) covering only corridor distance one, two and three. Positive values are in favor of the prediction, showing faster pointing when aligned with the superior orientation. We do not show negative values in this figure. We tested for seven potential outcomes: Local reference frames, and then a reference frame following the first experienced perspective, the most frequently experienced perspective or the salient geometry of a U-shaped environment, either on a regional or a global level.

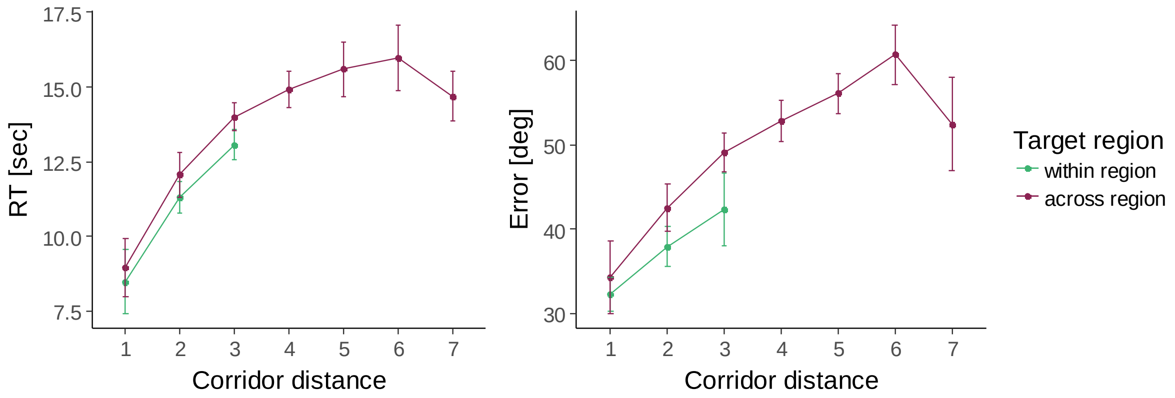

Figure 5. Distance effects in Experiment 2. Pointing latency (left) and error (right) as a function of corridor distance to the target, pointing to targets within the same region vs. across region. Means and SEMs are depicted. Only data points from corridor distance 1, 2 and 3 were analysed.

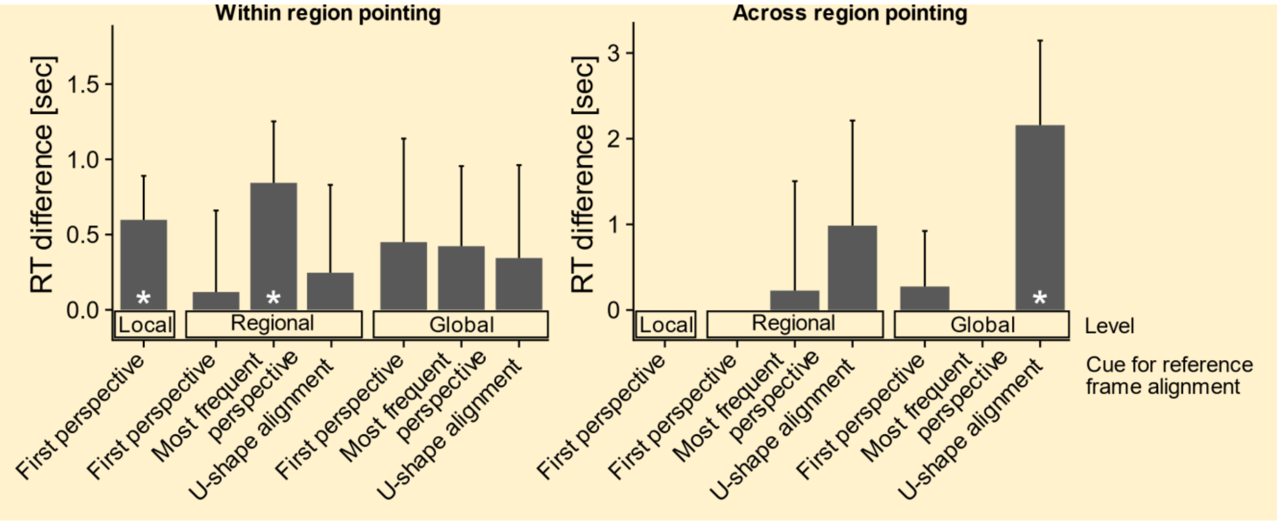

Figure 6. Reference frame alignment in Experiment 2. Difference in pointing latency when being aligned with the superior orientation identified for each prediction compared to the remaining seven orientations, separately for within (left) and across region trials (right) covering only corridor distance one, two and three. Positive values are in favor of the prediction, showing faster pointing when aligned with the superior orientation. We do not show negative values in this figure.

Tables

Table 1. T-tests exploring whether the difference in pointing latency of Experiment 1 is significantly larger than zero, which indicates a fit to the respective prediction. For pointing across region (right) only a subset of trials with a maximum corridor distance of three are included in the analysis reported in this table.

		Pointing within region			Pointing across region		
Prediction		*t*(17)	*p*	*d*	*t*(17)	*p*	*d*
Local	First perspective in corridor	2.019	.030 *	0.48	-0.511	.692	-0.12
Region	First perspective	2.176	.022 *	0.51	0.977	.171	0.23
	Most frequent perspective	-0.718	.759	-0.17	-0.170	.566	-0.04
	U-shape geometry	2.622	.009 *	0.62	0.643	.265	0.16
Global	First perspective	0.639	.266	0.15	2.128	.024 *	0.51
	Most frequent perspective	-1.554	.931	-0.37	0.600	.278	0.15
	U-shape geometry	1.039	.157	0.24	2.147	.023 *	0.51

* <.05, one-sided, larger than zero, no correction for multiple comparisons.

Table 2. T-tests exploring whether the difference in pointing latency of Experiment 2 is significantly larger than zero, which indicates a fit to the respective prediction. For pointing across region (right) only a subset of trials with a maximum corridor distance of three are included in the analysis reported in this table.

		Pointing within region			Pointing across region		
Prediction		*t*(17)	*p*	*d*	*t*(17)	*p*	*d*
Local	First perspective in corridor	2.045	.028 *	0.48	-0.580	.715	-0.14
Region	First perspective	0.218	.415	0.05	-0.191	.575	-0.05
	Most frequent perspective	2.064	.027 *	0.49	0.177	.431	0.04
	U-shape geometry	0.422	.339	0.10	0.803	.217	0.19
Global	First perspective	0.654	.261	0.15	0.423	.339	0.10
	Most frequent perspective	0.794	.219	0.19	-1.743	.950	-0.42
	U-shape geometry	0.558	.292	0.13	2.185	.022 *	0.52

* <.05, one-sided, larger than zero, no correction for multiple comparisons.