Original Articles

Qualitative differences in memory for vista and environmental spaces are caused by opaque borders, not movement or successive presentation

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

Max Planck Institute for Biological Cybernetics, Tübingen, Germany

A B S T R A C T

Two classes of space define our everyday experience within our surrounding environment: vista spaces, such as rooms or streets which can be perceived from one vantage point, and environmental spaces, for example, buildings and towns which are grasped from multiple views acquired during locomotion. However, theories of spatial representations often treat both spaces as equal. The present experiments show that this assumption cannot be upheld. Participants learned exactly the same layout of objects either within a single room or spread across multiple corridors. By utilizing a pointing and a placement task we tested the acquired configurational memory. In Experiment 1 retrieving memory of the object layout acquired in environmental space was affected by the distance of the traveled path and the order in which the objects were learned. In contrast, memory retrieval of objects learned in vista space was not bound to distance and relied on different ordering schemes (e.g., along the layout structure). Furthermore, spatial memory of both spaces differed with respect to the employed reference frame orientation. Environmental space memory was organized along the learning experience rather than layout intrinsic structure. In Experiment 2 participants memorized the object layout presented within the vista space room of Experiment 1 while the learning procedure emulated environmental space learning (movement, successive object presentation). Neither factor rendered similar results as found in environmental space learning. This shows that memory differences between vista and environmental space originated mainly from the spatial compartmentalization which was unique to environmental space learning. Our results suggest that transferring conclusions from findings obtained in vista space to environmental spaces and vice versa should be made with caution.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

The ability to remember the location of non-visible targets is essential for a multitude of everyday life tasks, such as communicating the direction to the train station to a non-local person or pointing to a certain cupboard in the kitchen to guide your cooking mate. In order to solve such problems, target locations have to be represented in memory. People have the ability to remember locations in their immediate visible surrounding, i.e., vista space, such as rooms, corridors or open spaces (Montello, 1993). In vista spaces, properties of the surroundings and configuration of objects in space can be perceived from one vantage point by taking a look around. Yet, people are also capable of combining information from several interconnected vista spaces, i.e., an environmental

space, such as in buildings or cities (Montello, 1993). Information, in this case, has to be gathered by traversing through and experiencing multiple spaces. Object-to-object relations have to be established mentally, for example, by integrating them into a single reference frame.

Prior studies have already indicated differences between spatial representations acquired in vista and environmental spaces. Firstly, it was found that borders of visibility often determine mental updating of object locations. Namely, locations beyond the currently visible vista space (e.g., locations on a campus) are less likely to be updated compared to locations within the same vista space (e.g., objects in a room) (Avraamides & Kelly, 2010; Kelly, Avraamides, & Loomis, 2007; Wang & Brockmole, 2003a, 2003b). Such results suggest that the self-to-object updating process concentrates more on the immediate environment and less on distant targets exceeding the current vista space. Secondly, locations within one vista space unit seem to have a greater degree of "mental closeness" than locations separated by spatial borders. Despite having the same Euclidean distance, the distances between objects

* Corresponding author at: Max Planck Institute for Biological Cybernetics, Spemannstr. 38, 72076 Tübingen, Germany.
E-mail address: tobias.meilinger@tuebingen.mpg.de (T. Meilinger).
[1] Equal contributions.

is judged as being shorter within a single unit (e.g., room) compared to across units (e.g., to the next room) (Kosslyn, Pick, & Fariello, 1974; McNamara, 1986; Newcombe & Liben, 1982). Thirdly, switching between distinct environmental representations is costly, which manifests in increased response times (Brockmole & Wang, 2002, 2003; Wang & Brockmole, 2003a, 2003b). Also, memory of environmental spaces can be comprised of multiple, local reference frames, one for each single vista unit of the environmental space (e.g., for each traveled passage of a route) (Meilinger, Riecke, & Bülthoff, 2014; Werner & Schmidt, 1999). In general, these results suggest that entering a new vista space by passing a visual border strongly affects how we represent the space and that an environmental space is potentially represented segmentally, comprising multiple vista space units.

Importantly, most of these experiments did not control for the amount of information that is needed to be processed within a vista or an environmental space. The number of objects that had to be taken into account and the area that needed to be covered mentally was always larger for the environmental space compared to the vista space, for example when retrieving memory of object location within and beyond the current test room, thus increasing memory load for the environmental space compared to the vista space (e.g., Brockmole & Wang, 2002, 2003; Kosslyn et al., 1974; McNamara, 1986; Newcombe & Liben, 1982). Therefore, effects might at least partially be explained by these differences. In order to match information quantity, we examined participants' configurational memory after learning *exactly the same* object layout (keeping distances and angles constant) either within a vista space or in an environmental space.[2] In the following we will derive three hypotheses about potential differences that may arise in the spatial representations of the layout. In a second step, we will examine how distinct learning characteristics within vista and environmental spaces may underlie these differences.

1.1. Order effects

Learning an environmental space is inevitably temporal. One needs to pass through a discrete vista unit to perceive the next one. Thus, objects are encountered successively in a specific order. Several studies have examined the effects of order during spatial tasks. Results by Strickrodt, O'Malley, and Wiener (2015) suggest that when learning a route, people memorize the sequence of encountered landmarks along the way in combination with the corresponding turning direction. Landmark and turning information of the preceding intersection were used to infer the correct direction of turn at the following decision point. Object order is also used to identify overall route direction, i.e., forward direction or return path (Wiener, Kmecova, & de Condappa, 2012). How engrained object order is in spatial memory was demonstrated in a priming experiment by Janzen (2006). After learning a route in a large-scale virtual environment containing a range of landmarks, subsequent recognition was faster when participants were primed with a former predecessor landmark, compared to a former successor landmark (see also Schweizer, Herrmann, Janzen, & Katz, 1998). These results are in line with the assumption that the representation of a route is highly integrated, following a stimulus-response-stimulus pattern that allows memorizing route landmarks as a sequence (e.g., O'Keefe & Nadel, 1978).

These studies all target characteristics of the acquired route knowledge. Interestingly, in addition to the above-mentioned results, route direction was also shown to influence performance in tasks designed to address configurational memory (survey tasks), even though, typically configurational knowledge is thought to be uncoupled from the order of learning. In a study by Moar and Carleton (1982), participants were more accurate in directional and distance judgements to targets along a route when probed in the direction they had previously learned the route than in the opposite direction. For example, performance was better while standing at the location of the first object along the route and pointing to the third object encountered during learning than pointing from the third object to the first object. These results suggest that route direction is preserved within configurational memory and used not only for route tasks, but also for survey tasks. This result only represents an indirect examination of whether object order is incorporated in participants' configurational knowledge when learning takes place in an environmental space. In the current study, however, we aimed for a direct measure by letting participants perform a configurational placement task, where the layout of environmental objects had to be reproduced from memory. We predict that, when learned in environmental space, the reconstruction of objects follows the order in which they were first encountered. This order should be easiest to retrieve and, as a result, most preferred. In contrast, presentation of an object layout in a vista space does not impose a predetermined learning order. All objects are visible at once. Access of configurational memory could be flexible following random order. Alternatively, scanning patterns during learning might influence retrieval. These scanning paths might be random as well, thus, being unique for every participant. There is also evidence for systematic scanning paths of grid layouts along horizontal paths (Gilchrist & Harvey, 2006; Hardiess, Gillner, & Mallot, 2008). In sum, whereas environmental space learning should predetermine one specific order, the order of retrieving configurational memory from vista space should be much more varied.

1.2. Distance effects

Following the abovementioned results (Avraamides & Kelly, 2010; Brockmole & Wang, 2002, 2003; Kelly et al., 2007; Kosslyn et al., 1974; McNamara, 1986; Meilinger et al., 2014; Newcombe & Liben, 1982; Wang & Brockmole, 2003a, 2003b; Werner & Schmidt, 1999), a compartmentalized space might cause the mental representation to be compartmentalized as well. Learning an environmental space is highly restricted compared to vista space learning. Vision of the entire space is obstructed, the order connected vista spaces are successively entered is predefined as well as the walking distance between locations along the route. We assume that retrieving spatial information will depend on this predefined structure of space.

There is evidence suggesting that distance information from the learning experience might still be preserved within configurational memory. Thorndyke and Hayes-Roth (1982) reported an increased error in directional and distance judgements dependent on the number of corridors between the participant's current position and target location. One possible explanation for this increasing error with distance could indeed be that during task execution (retrieval process), memory of the environmental space is retrieved successively, along the route from which the environment was experienced from. This might be realized, for example, by mentally walking down the memorized route starting from the current location and approaching the target (Byrne, Becker, & Burgess, 2007; Sanders, Rennó-Costa, Idiart, & Lisman, 2015) or by constructing a mental model of the non-visible parts of the environment corridor-by-corridor from one's

---

[2] Studies utilizing vista space learning usually test what they call *object-to-object relations* (e.g., Avraamides & Kelly, 2010; Mou & McNamara, 2002; Yamamoto & Shelton, 2009). Studies exploring navigation and wayfinding in environmental space typically examine object-to-object relations as well, but subsume it under the term *survey knowledge* (i.e., knowing where a target is located in terms of direction and distance without necessary knowing a route leading there; e.g., Siegel & White, 1975).

current location (Meilinger, 2008). Both theories predict an increase in computational effort for larger distances (route distance or amount of corridors) between current and target location, since spatial information must be activated successively following the encoding procedure. However, providing evidence for the increase of pointing error with traveled distance and not for an increase in pointing latency (Thorndyke & Hayes-Roth, 1982) legitimates an alternative explanation: the accumulation of error during encoding (compare to path integration model of Fujita, Klatzky, Loomis, & Golledge, 1993). For example, corridors might be assumed shorter than they actually are, the angle of turn at an intersection might be encoded as a regular 90° turn, whereas in fact being 80°. On average, this error will be larger the more distance traveled.

We assume that pointing accuracy and latency are indicative of distinct processes of spatial learning. Whereas accuracy might be associated with the encoding process, that is, the precision of memory, latency during pointing relates to the process of retrieval, that is, accessing the memory content (see also Pantelides, Kelly, & Avraamides, 2016). The assumption that error and latency do reflect distinct aspects of cognition is used in other literature as well (Prinzmetal, McCool, & Park, 2005; Sternberg, 1969). We expect error accumulation during encoding to be independent of the time needed to retrieve the distorted memory. Therefore, even if a complete, integrated representation of the environment was built, where no additional processing is needed regardless of inter-object distance, the representation itself could be distorted, leading to an error increase with distance. By observing latency we investigate the retrieval process, which in turn should be bound to the structure of spatial representation. Studies examining path integration already demonstrated that both, error as well as latency, increase in a multi segment path completion task with increasing overall path length and/or number of legs (Klatzky et al., 1990; Loomis et al., 1993; but see Wan, Wang, & Crowell, 2013; Wiener & Mallot, 2006). In contrast to these studies, retrieving configurational memory of an object layout for executing survey tasks strongly relies on a long-term representation of space. Evidence regarding latency for retrieving survey knowledge is still missing.

In a single vista space, learning is comparably unrestricted. Relations and distances between to-be-learned objects can be directly perceived in a commonly visible reference space. Typically, no walking path or encoding order is prescribed. When examining direction and distance judgement between targets learned as spread across a fully visible space, McNamara (1986) demonstrated that the accuracy of judgements was sensitive to the Euclidean distance (i.e., the straight-line distance) between two object locations. Whether this is due to an unprecise layout memory or the retrieval process again can only be assessed when analyzing latency. Indeed, Kosslyn, Ball, and Reiser (1978) found that when learning the positions of landmarks from a map the time to mentally scan from one to another landmark depends on the straight-line distances between them.

In our study, we set out to examine whether and how the structure of environmental and vista space influences the structure of the corresponding representation. We instructed participants to perform a pointing task after learning an object layout either in vista or in environmental space. In case of a segmented, non-integrated representation of the environmental space, retrieval of spatial memory is expected to be successive, following the corridors. This would lead to an increase of pointing latency with increasing corridor distance to the target. Again, in contrast to changes in error size, alteration in latency would explicitly imply the need to adjust processing time in order to solve the task. Access of vista space memory might be affected by the Euclidean distance, facilitating retrieval of objects nearby.

1.3. Reference frame orientation

It has been shown that the representation of space is orientation-dependent (for a review see McNamara, 2003). Here, orientation refers to the alignment of the body or visual field with respect to the environment, thus, the perspective onto the environment (independent of the target bearing). After learning the locations of objects within an environment and being subsequently tested for configurational memory, pointing between objects from certain perspectives leads to better performance compared to other perspectives. When learning took place in a vista space often the best pointing performance is shown from the originally learned orientation, i.e., the initial view upon the object layout (often referred to as 0°), compared to novel orientations. Additionally, contra-aligned (180°) and orthogonal orientations (±90°) seem to be retrieved better than oblique orientation (e.g., 45°). Hence, pointing performance usually yields a w-shape, or saw tooth, performance pattern in error and latency along the range of tested body orientations (e.g., Kelly & McNamara, 2008; Meilinger & Bülthoff, 2013; Mou & McNamara, 2002; Shelton & McNamara, 2001). One explanation for this pattern is the encoding of object locations relative to one or two orthogonal reference axes which are retrieved rather effortlessly. Testing from other perspectives requires additional inferential processes (McNamara, Sluzenski, & Rump, 2008; Mou, McNamara, Valiquette, & Rump, 2004; for an alternative explanation of this pattern see Street & Wang, 2014).

The alignment of the spatial reference frame (i.e., orientations on which maximum pointing performance is centered) was found to be influenced by multiple factors. Not only the perspective during encoding (experienced views) is thought to be used, but also environmental geometry (extra-layout cues), such as the shape of the room or the mat on which objects were placed, and the intrinsic configuration of the object layout itself (intra-layout cues) (Mou & McNamara, 2002; Shelton & McNamara, 2001). We assume that for environmental spaces the initial view and the global layout-intrinsic orientation are less determining for setting the reference frame alignment. Deriving the global layout is effortful and cannot be done until the last unit is reached. Instead, each room, corridor and street, constitutes a separate entity, which itself entails discrete intra and extra-layout cues. When walking down, for example, a corridor, the observers' view will naturally become aligned with the geometric axis of the corridor. Such a viewer-space-alignment experienced later during learning was found to be more important for determining reference frame orientation than initial views on a room (Kelly & McNamara, 2008; Shelton & McNamara, 2001; Valiquette & McNamara, 2007). A relatively simple environmental space (e.g., few orthogonally interlinked corridors) may still be represented along a single main axis that spans the entire environmental space (axis presumably aligned with the first vista space encountered; Meilinger, Frankenstein, Watanabe, Bülthoff, & Hölscher, 2015; Tlauka, Carter, Mahlberg, & Wilson, 2011; Wilson, Wilson, Griffiths, & Fox, 2007). However, sufficiently complex environmental spaces seem to be represented within multiple local reference frames, with each local corridor or street occupying a distinct reference frame aligned with the respective corridor (Meilinger et al., 2014; Werner & Schmidt, 1999).

In the current study we kept both the initial view within the environments and the orientation of the global object layout constant while setting the geometric axes of both learning spaces in conflict; this contrasts the reference frame alignment of vista and environmental space learning. We predicted that the reference frame in both environments should be aligned with the visible context. That is the room orientation in vista space and the corridor orientation in environmental space.

The current study is concerned with the acquisition of object-to-object relations under different learning conditions. Experiment

1 examined whether the acquired memory is different depending on whether exactly the same object layout is either learned in vista or in environmental space. We predicted that knowledge acquired from environmental space preserves features of the spatio-temporal learning process, resulting in higher latency when pointing to targets with increasing corridor distance, and recall in the order objects were encountered in. We expected recall latency and order of vista space memory to be influenced by the layout structure instead. Furthermore, we predicted reference frame orientation in environmental space to be aligned with the visible context of the corridor and the learning experience rather than the initial view or the intrinsic layout orientation. In a second step (Experiment 2) we mimicked characteristics of environmental space learning (i.e., movement through space, successive learning experience) within a vista space to isolate the distinguishing factors between spaces.

2. Experiment 1

Performance in visual pointing and object placement was ascertained after learning an object layout either in an environmental space (ES) or in a vista space (VS).

2.1. Method

2.1.1. Participants

26 naïve participants were recruited from a subject database, gave written informed consent and participated in exchange for monetary compensation. Participants were randomly assigned to either of two conditions (ES or VS). Two participants had to be excluded. One participant did not perform significantly better than chance level (90°) in the pointing task. The other participant had to be excluded due to a lack of comprehension of task instructions. The remaining sample of 24 participants (12 for each condition) had a mean age of $M = 26.09$ ($SD = 6.94$, [19;52]) and included twelve females (seven randomly assigned to ES condition, five to VS condition). The experimental procedure was approved by the ethical committee of the University Hospital Tübingen.

2.1.2. Material

We used Virtools® 5.0 (Dassault Systemes) for programming the virtual environment and the experimental procedure. The experiment took place in a 12×12 m tracking hall, enabling free movement in real space while wearing a head mounted display (HMD) visualizing the virtual space. Participants' head coordinates were tracked by 16 high-speed motion capture cameras with 120 Hz (Vicon® MX 13) to render a real-time egocentric view of the virtual environment. We used a NVIDIA Quadro FX 3700M graphics card with 1024 MB RAM and a nVisor SX111 HMD with a field of view of 102° (horizontal) × 64° (vertical), a resolution of 1280 × 1024 pixels for each eye, and 66% overlap. The interpupillary distance was fixed at 6 cm. We adjusted the HMD fit and screen position for each participant. This virtual reality setup provided important depth cues such as stereo vision, texture gradients, and motion parallax and enabled participant to physically walk through a virtual world.

The object layout participants were asked to learn consisted of seven target objects lying on the floor of the virtual environment arranged within an incomplete 3×3 grid with bilateral symmetry (Fig. 1). From left to right a teapot, a hammer and a banana were located in the closest row, and the middle row held a horse, a telephone and a tennis racket. A trumpet was located in the center of the furthest row. It is assumed that the linear relation between hammer, hair dryer and trumpet determines the global main axis of the object layout, namely 0° (compare to Greenauer & Waller,

2010; Kelly & McNamara, 2008). Additional objects by or on the walls, such as a vase or a fireplace, served as aids for orientation within the environment and remained visible throughout learning and testing phase, whereas the target objects were absent in the testing phase. Note that both ES and VS contained the same target objects and objects aiding orientation.

Both, in ES and VS condition, exactly the same object layout was arranged on the floor, thus, distance and relations between the objects were identical. In the VS condition a rectangular room was presented to the participants from a constant point of view located in front of the object layout (X in Fig. 1, right). Body location within the environment was kept constant during the whole experiment. However, participants were allowed to look around. In the ES condition the environment consisted of walls placed to arrange four parallel, interlinked corridors, offset by 45° to the main axes of object layout and room orientation in the VS condition (Fig. 1, left). To see all objects participants had to walk from the start point X through all corridors sequentially passing each object. Initial view of both environments was set along the main axis of the object layout (grey arrows above the X's in Fig. 1). This view also defined the zero-point of body orientation, as illustrated by the arrow flanked by 0° below the sketches of the environments in Fig. 1, top. Thus, for example, turning 45° to the left in ES in order to be aligned with the corridor would correspond to a body orientation of −45° with respect to the reference orientation of 0°.

2.1.3. Procedure

After participants were familiarized with the equipment the *learning phase* started. Participants were positioned standing in the corner of the experiment room facing the opposite corner, shortly before being equipped with the HMD. We instructed participants to learn in depth where the virtual objects on the floor were located within the virtual environment and gave no time restrictions or details about later performance tests.

Participants in ES condition moved through the environmental space, following the corridors. They moved twice from their start point, to the end of the last corridor and back to the start to ensure sufficient learning. Since no movement was required in VS condition, participants were instructed to inspect the whole room, including the walls and corners at their back, for later orientation. At no point they were allowed to leave their current position. To ensure correct object identification each object was tagged (in English) by the participant in both conditions. The experimenter corrected misidentifications (i.e., object names not used later in the experiment). After traversing through space twice in ES condition or indicating sufficient learning in VS condition, we removed the target objects and participants proceeded to the learning test. We successively presented blue spheres at former object positions. Participants then had to recollect and name the object located at this position from memory. In ES participants were obliged to walk through the environment again to encounter all blue spheres. The order of testing was the following: Teapot, hammer, hair dryer, racket, trumpet, banana, and horse. This order was neither along rows and columns of the object layout nor along the order of first contact in ES condition. One or more errors resulted in a new learning trial (walking there and back in ES condition and self-paced learning in VS condition). This procedure was repeated until all locations were associated with the correct object. Subsequently, after a short break, the *test phase* started. The procedure of the test phase was identical for ES and VS condition. Participants stood in front of a table, mounted by a joystick and, first, had to conduct a visual pointing task within their previously learned environment (in the absence of the objects) and, subsequently, perform an object placement task.

In the *visual pointing task* all target objects were removed from the environment. Participants were teleported to a former object
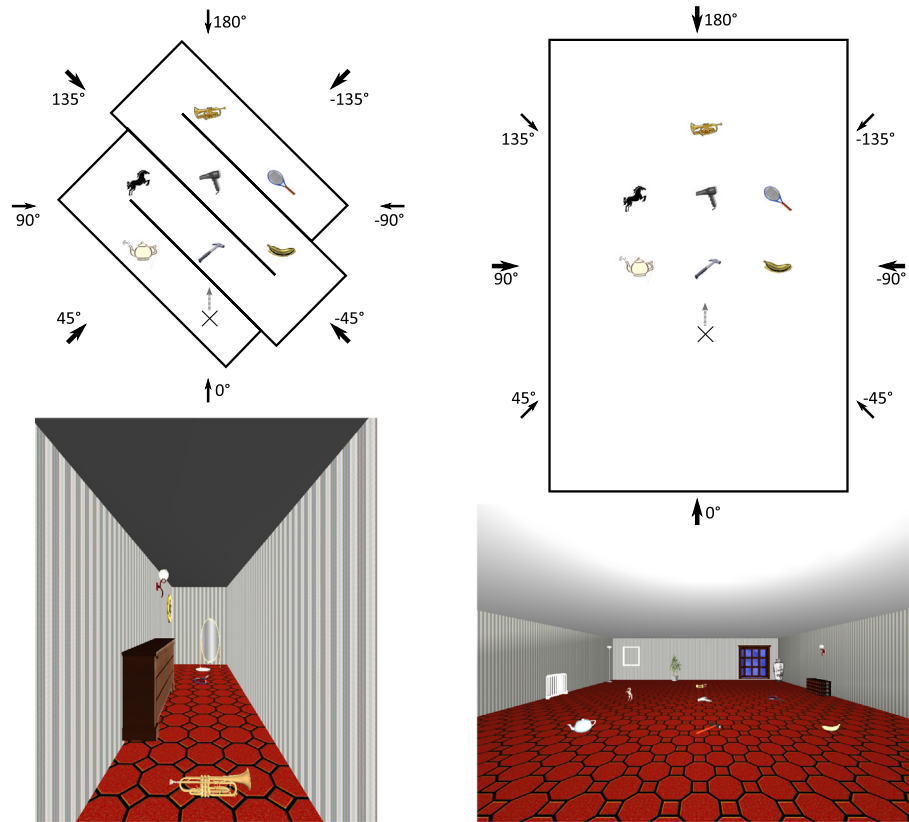
Fig. 1. Left: The layout of ES condition from a birds-eye perspective and participants view from within the environment. Right: Layout and participants view from within the VS condition. Xs indicate the starting position, grey arrows above the Xs the initial view upon the environment. Layout orientation is similar in both environments. Alignment of the visible geometry differs, as indicated by the black, bold arrows.

location in each trial, being randomly aligned with 1 out of 8 body orientations. Note that the physical orientation of the participant in the real tracking hall did not change (aligned with the table mounted with the joystick). Rather, for each trial the virtual reality was adjusted in position and orientation to render the desired trial characteristics. The possible orientations participants bodies were then aligned with in the virtual reality are illustrated by the arrows encircling the outline of the environments in Fig. 1, top. They are spaced around a full circle in steps of 45°. The current location (e.g., "You are at the hair dryer") and the pointing target (e.g., "Point to the banana") was indicated during each trial on the HMD screen. This example illustrated in Fig. 2, top, emulated a body orientation of −45° with respect to the reference orientation of 0°. Participants had to identify their orientation based on the visual input from looking around. They were not allowed to walk through the environment during the test phase.

Each of the eight body orientations was tested nine times from different object locations resulting in 72 pointing trials. For 8 of the 9 pointing trials of each body orientation the correct bearings of the target (the correct pointing direction) were spaced around a full circle in steps of 45°. In each body orientation participants had to point to the front, right-front, right, right-back, back, etc. To analyze distance effects, four of the remaining eight trials were set up to cover the largest distance in terms of corridors, i.e., pointing from the teapot to the racket/trumpet and vice versa. This led to correct pointing directions of either 18°, 64°, −108° and −63° relative to the current body orientation. For the remaining four trials targets were set to be located in the same corridor (minimum distance in terms of corridors) with the correct pointing direction of 0°, 45°, −45° and −90°. In summary, 20 trials covered the minimum distance in terms of corridors, 29 trials covered a short
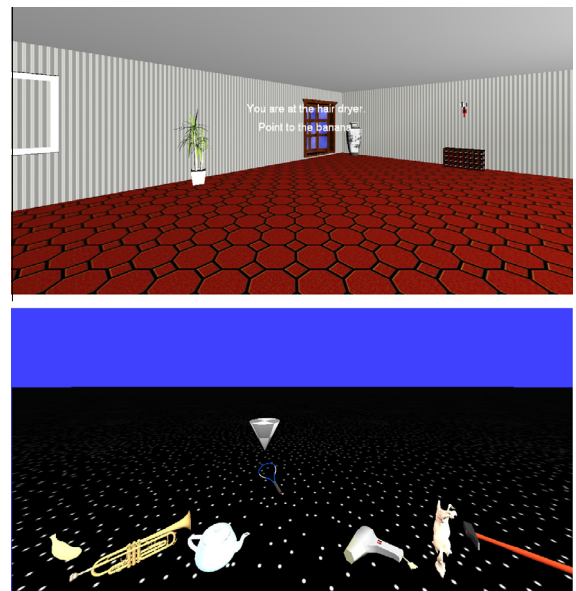

Fig. 2. Top: View upon the environment of one exemplary trial in the pointing task. Bottom: Reconstruction of the layout in the placement task. The currently selected object for rearrangement is marked with an inverted cone. Note that when presenting the environments with the correct field of view no visual distortion was present.

distance (next corridor), 19 trials covered a medium distance (second next corridor) and 4 trials covered maximum distance. Trials were presented in random order to every participant. Participants

executed pointing by moving a joystick handle, enabling pointing measurements across a 360° circle. For example, assuming the target to be located in front of one's current position, the joystick handle had to be pushed straight forward; assuming the target to be located 135° to the left of one's current position, the joystick handle had to be pulled backwards and to the left. The current pointing direction was accepted with a button press.

Participants were allowed to examine the environment by looking around. Consequently, during this time head orientation deviated from the body orientation. In order to point to the target, however, participants were instructed to look straight again, thus, realigning with the body orientation in this trial (also realigning with the joystick in front of them). Participants had to align properly in order to have their pointing response recorded and to continue with the next trial. Head orientations deviating more than 10° were not accepted by the program. Latency consisted of the duration between trial onset (appearance within the environment) and button press. Pointing error consisted of the absolute deviation (in °) between pointing direction and correct direction.

For the *object placement task*, participants were placed in a new virtual surrounding, containing a horizontal plane only (Fig. 2, bottom). Objects were located in a horizontal row in front of the participant, ordered in random sequence determined for each participant. We instructed participants to arrange the objects in the layout they had previously seen in the virtual environment. Objects could be placed in any preferred order and could be (re)arranged until the participant was satisfied. Participants used buttons to switch between objects and the joystick to move the currently selected object. We recorded the order of replacement, i.e., which object was moved away from the start location first, which object was moved second, etc.

2.1.4. Data analysis

From the 24 participants approximately 4% of the pointing performance data was deleted due to deviation of more than two *SD* from a participant's overall mean. All tests conducted were corrected for nonsphericity or inequality of variance when appropriate. As adding participants' gender to the analysis did not change any of the reported effects, we only report the pooled data.

2.2. Results

2.2.1. Distance to target

We examined the influence of distance to a target on pointing latency in order to estimate whether this spatial information structured spatial memory. Two potentially meaningful distances between the current position and the target object were analyzed separately. Firstly, the distance in terms of corridors, and secondly, the Euclidean distance. Corridor distance distinguishes trials representing pointing within the same corridor (0), to the next corridor (1), across two (2) or three corridors (3). This classification is based on the ES layout, but was likewise applied to VS trials for control. Euclidean distance represents the straight-line distance between two objects. We differentiate the relative Euclidean distance value of 1, representing the smallest possible distance between two objects (e.g., from teapot to horse) and ascending from there Euclidean distance $\sqrt{2}$ (e.g., from teapot to hair dryer), 2 (e.g., from teapot to banana) and $\sqrt{5}$ (e.g., from teapot to trumpet). A relative Euclidean distance value of 1 represents an absolute physically distance of ca. 2.83 m. For analysis it is not meaningful whether absolute or relative values are used. It should be noted here that the two distance dimensions tended to be associated with one another in our study by $r = 0.22$. Also we did not analyze absolute error as

this was not the goal of the study. The following analyses were conducted for both distance dimensions separately.

Starting with the dimension of corridor distance and its influence on pointing latency, we conducted an ANOVA with the between-participant factor *environment* (ES vs. VS) and the within-participant factor *corridor distance*. A main effect of *corridor distance*, $F(3,66) = 5.47$, $p = 0.008$, $\eta_p^2 = 0.20$, and an interaction of *corridor distance × environment*, $F(3,66) = 4.58$, $p = 0.017$, $\eta_p^2 = 0.17$, but no main effect of *environment*, $F < 1$, $p > 0.800$, were found. Thus, the influence of corridor distance on latency differed between environments. To further examine this interaction we regressed pointing latency onto corridor distance separately for each participant. From these regressions b was extracted. This standardized slope describes the linear change of latency with increasing corridor distance for every participant. Fig. 3, top left, depicts the individual and mean slopes for the two environmental conditions. *T*-tests were used to analyze slopes. As expected, slopes derived from ES did exceed 0, mean $b = 0.22$, $t(11) = 8.15$, $p < 0.001$, $d = 2.35$ (see Fig. 3, top right), indicating an increase in pointing latency the more corridors are residing between current and target location. With each additional corridor pointing took on average 1.02 s ($SD = 0.53$) longer. In VS, which worked as the control condition where no latency increase with ascending corridor distance was expected, the mean slope did not differ from 0, mean $b = 0.02$, $t(11) = 0.63$, $p = 0.542$, $d = 0.18$. Thus, pointing did not take longer the further away targets were located with respect to corridors. Additionally, a comparison of mean slopes between ES and VS revealed that the linear increase of latency across ascending corridor distance was higher for ES compared to VS, $t(22) = -4.28$, $p < 0.001$, $d_s = 1.75$. This pattern was also evident in individual slopes. Whereas in ES data of each single participant rendered a positive slope, in VS this was only the case for 7 out of 12 participants. Results suggest that in ES memory, the spatio-temporal pattern of learning was preserved and this was not due to the structure of the layout itself.

As corridors were not meaningful (since nonexistent) in VS we conducted a control analysis with Euclidean distance. The ANOVA revealed a significant interaction of *Euclidean distance × environment*, $F(3,66) = 5.19$, $p = 0.020$, $\eta_p^2 = 0.19$. Even though not significant, *Euclidean distance* tended to influence pointing latency, $F(3,66) = 3.13$, $p = 0.073$, $\eta_p^2 = 0.13$. No main effect of *environment*, $F < 1$, $p > 0.660$ was found. Following this, the effect of Euclidean distance seems to differ between the two environments. Fig. 3, bottom left, depicts the mean and individual slopes derived from the regression of pointing latency onto Euclidean distance. The level of standardized slopes derived from ES did exceed 0, mean $b = 0.08$, $t(11) = 2.57$, $p = 0.026$, $d = 0.74$ (Fig. 3, bottom right), although smaller in size compared to the analysis of corridor distance. In the VS condition, the mean slope did not differ from a 0 slope, mean $b = 0.05$, $t(11) = 1.46$, $p = 0.172$, $d = 0.42$. Thus, only in ES pointing latency increased with increasing Euclidean distance between current and target location. When directly comparing ES and VS slopes did not differ between the environments, $t(22) = -0.75$, $p = 0.470$, $d_s = 0.30$. Looking at individual slopes 8 out of 12 participants (ca. 66%) had a positive slope in the VS condition, 9 out of 12 (75%) in the ES condition. Straight-line distances did not seem to play a prominent role when VS memory was retrieved.

2.2.2. Reference frame orientation

To evaluate the reference frame orientation in memory, data derived from the visual pointing task was submitted to a repeated measure ANOVA with the factors *environment* (ES vs. VS) and *body orientation* (−135° to 180° in steps of 45°). We expected an interaction. Second, pointing error and latency were fitted to a w-contrast, which describes the primary pattern of performance found in prior research. Using a contrast avoids execution of multiple pairwise
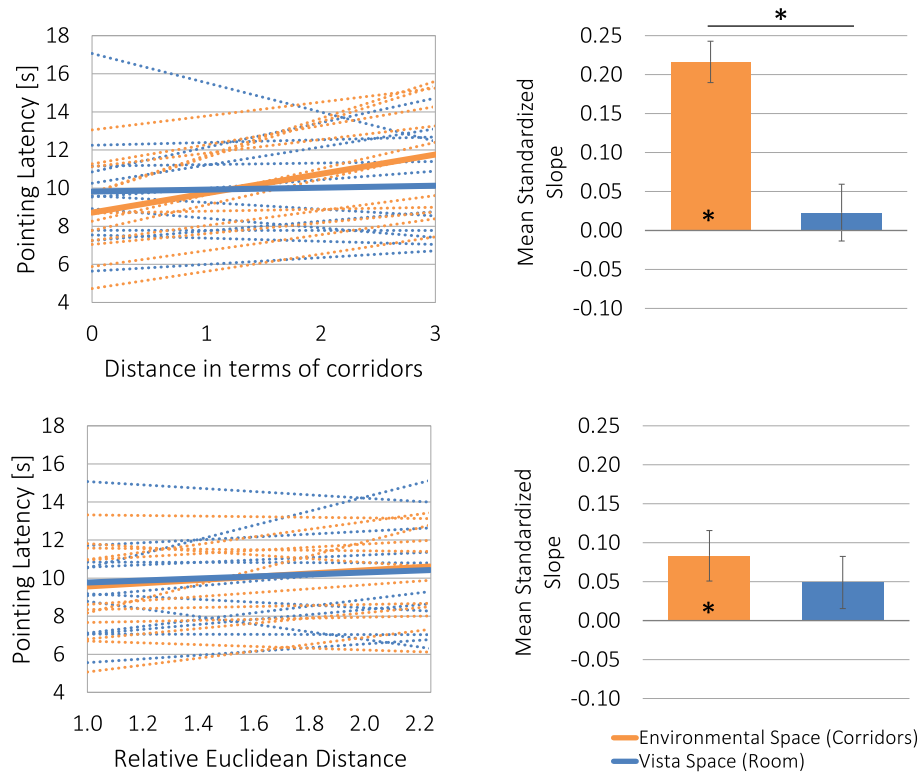
Fig. 3. Top: Linear increase of pointing latency with corridor distance between current position and target. Bottom: Linear effect of Euclidean distance on latency. Left: Individual (dashed line) and averaged (solid line) slope of pointing latency as a function of distance to a target. Right: Mean standardized slopes. Error bars depict *SEM*. * p<0.05. Asterisks within bars indicate deviation from 0.

comparisons between levels of conditions, but instead makes it possible to describe the fit of a curve to a predefined shape with one single parameter (Greenauer & Waller, 2010; Levin & Neumann, 1999). This predefined shape or pattern has to be specified in advance reflecting your hypothesis. As learning perspective, visible intrinsic object layout, and room orientation in VS condition all entail the same main axis (0° perspective), this value was selected as baseline for the contrast. The utilized w-contrast describes a saw tooth pattern centered on this orientation of 0° (see the legend of Figs. 4 and 5, middle, for illustration). The contrasts weights were defined to be lowest (−1) for the 0° orientation and all orientation which are orthogonal (±90°) or opposed to it (180°), thus, predicting lowest error rate and fastest responses at these orientations. Highest error rate and slowest responses (worst performance) was ascribed to oblique orientations (±45° and ±135°) by setting higher contrasts weights (1). Note that a contrast weight of 0 would predict average performance. To calculate contrast fit to the data, for each participant contrast weights were multiplied with the average performance in the respective perspective and added up (e.g., −1 × average in −180° + 1 × average in −135°, etc.). Contrast fits were inspected using *t*-tests. We predicted a high, therefore, positive contrast fit for learning an object layout in VS. Here the body orientations leading to best performance should be the initial view upon the environment (0°), and ±90° and 180° deviation from it. For learning in ES setup, however, even though the global layout and the initial view were aligned with 0°, we predicted a reference frame alignment with ±45° and ±135°, according to the visual input when walking through a corridor. A negative w-contrast fit is expected here. A contrast fit of 0 would indicate, that data can't be described by a w-shape.

The underlying assumption of the w-contrast is that the space is represented along two orthogonal axes (four body orientations rendering highest performance, four orientations rendering lowest performance). Alternatively, the space could be represented along a single axis, which can be expressed by an m-contrast centered on either one axis of the corresponding w-contrast. The pattern of an m-contrast implies that pointing performance is best when aligned with this specific axis. Thus, being aligned or directly opposed to one view should then yield best performance whereas decline occurs when deviating from these views, e.g., best performance with 0° and 180° body orientation, worst performance with ±90°. In order to examine which pattern represents our data best, in the last step we tested whether a m-contrast centered on 0° (along long room axis) or a m-contrast centered on ±90° (along short room axis) renders a better data fit than the previously examined w-contrast centered on 0°/±90° for the VS data. Similarly, for ES condition we tested whether the m-contrast centered on 45°/−135° (facing the corridor wall) or the m-contrast centered on −45°/135° (along corridor axis) renders a better data fit than a w-contrast centered on ±45°/±135°. In short, we tested whether the assumption of a single axis fits the data structure better than the assumption of two orthogonal axes. In Figs. 4 and 5, right, w- and m-contrasts of the respective learning condition are depicted. Note that the m-contrast centered on 45°/−135° is the inverse of the m-contrast centered on −45/135°. The same holds for m-contrasts centered on ±90° and 0°/180°. Thus, a positive fit of either corresponds to a negative fit of equal size for the respective other. Regarding the question of best fit, consequently, only positive fits are of interest and were compared via *t*-tests to the corresponding w-contrast.

2.2.2.1. Pointing latency. Fig. 4, left, depicts individual (dashed line) and averaged (solid line) pointing latencies of both conditions. The ANOVA yielded a significant main effect of *body orientation* on pointing latency, $F(7, 154) = 4.50$, $p = 0.002$, $\eta_p^2 = 0.17$, and an interaction of *body orientation* × *environment*, $F(7, 154) = 4.93$, $p = 0.001$,
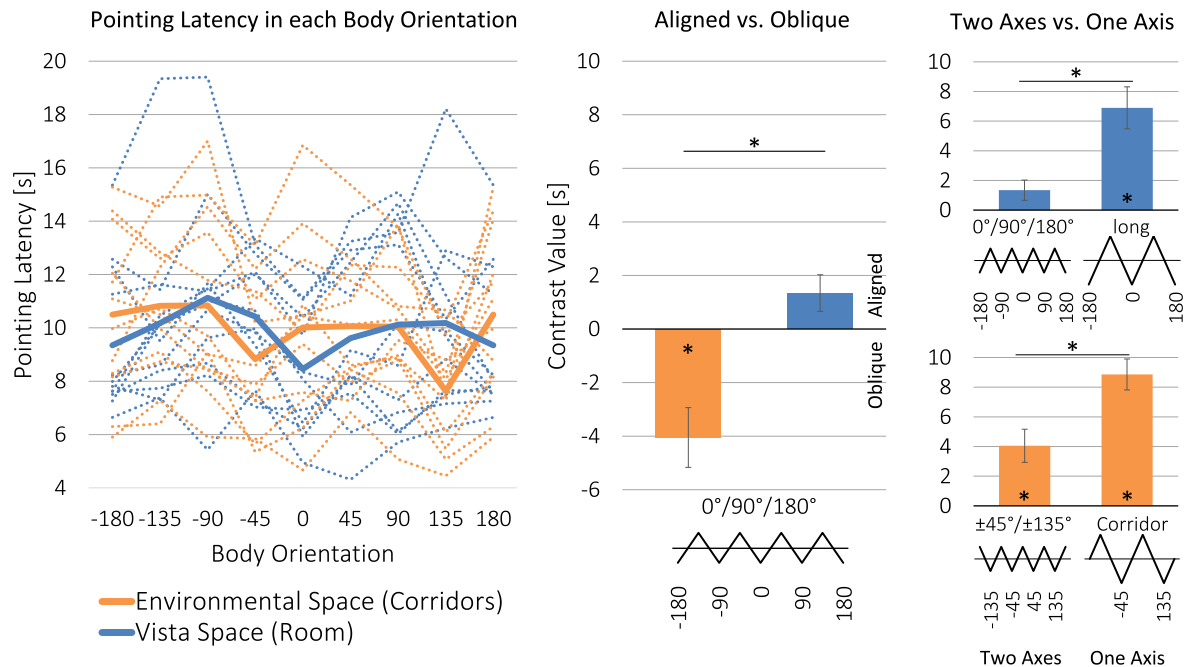
Fig. 4. Left: Individual (dashed line) and averaged (solid line) pointing latency as a function of body orientation. 180° is displayed twice for symmetry. Middle: Values for w-contrast fit centered on 0°/±90°/180° for pointing latency data. Right: Absolute values of contrast fit for w- and m-contrasts, separately for VS (upper) and ES (lower) condition. Pictograms define the used contrasts. Long = along long axis of room. Corridor = along corridor axis.
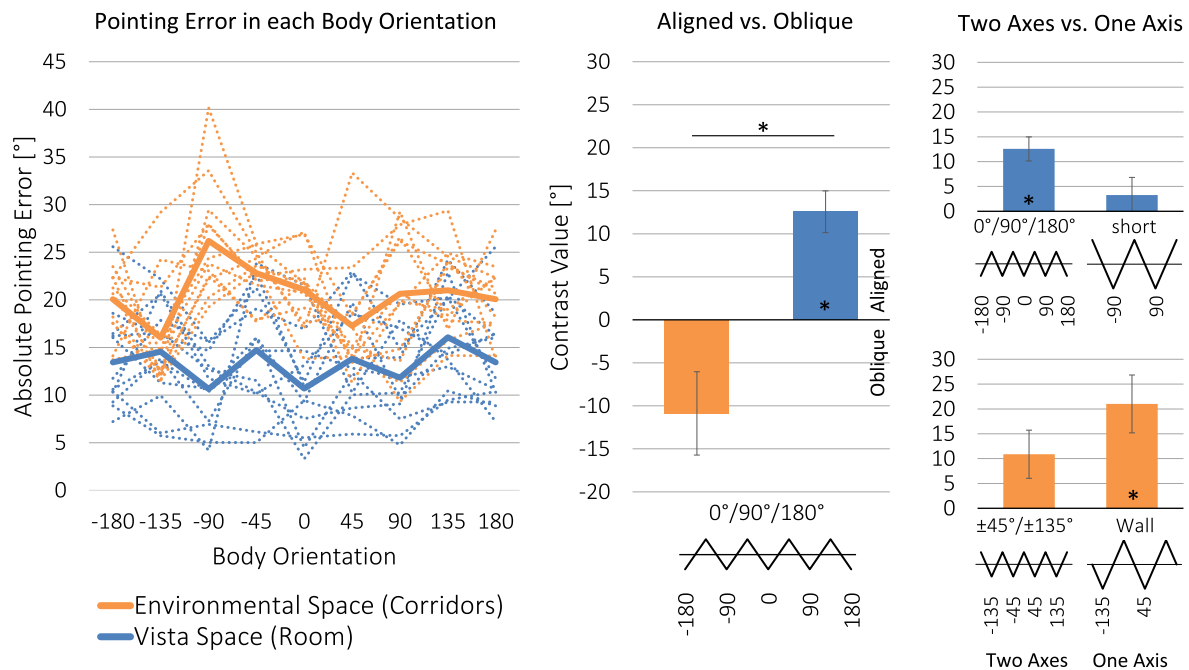

Fig. 5. Left: Individual (dashed line) and averaged (solid line) absolute pointing error. Middle: Mean values for w-contrast fit centered on 0°/±90°/180° for pointing error. Right: Contrast fit for w- and m-contrast, separately for VS (upper) and ES (lower) condition. Pictograms define the center of the contrasts. Short = along short axis of room. Wall = facing corridor wall.

$\eta_p^2 = 0.18$, but no main effect of *environment*, F < 1, *p* > 0.900. As predicted, the interaction demonstrates that pointing latency at specific body orientations differed depending on the learning environment. W-contrast fits further identified the nature of these differences (Fig. 4, middle). In line with our hypotheses, w-contrast fit of pointing latency data was higher in VS than in ES, *t*(22) = 3.95, *p* = 0.001, d_s = 1.61, suggesting different reference frame orientations. Average values for contrast fit (Fig. 4, middle) were

significantly below 0 for ES, *t*(11) = −3.48, *p* = 0.005, *d* = 1.00. This indicates that pointing latency in ES can be well explained by an inverted w-contrast, centered on the oblique ±45° orientations. Even though not significant, a trend was found for the contrast fit for VS to be larger than 0, *t*(11) = 1.88, *p* = 0.087, *d* = 0.54.

Descriptive data of single participants support the pattern of w-contrast fits. In VS 9 out of 12 participants (75%) pointed faster in trials of aligned body orientation (0°, ±90° and 180°) compared to

trials of oblique body orientation (±45°, ±135°). In contrast, in ES only 1 out of 12 participants showed this pattern. Thus, the remaining participants (ca. 92%) pointed faster from oblique body orientations.

To test whether the assumption of a single axis fits the data structure better than the assumption of two orthogonal axes, data fit to the corresponding m-contrasts was compared to the w-contrast fit for each condition. In Fig. 4, right, w- and m-contrast fits of the respective learning condition are depicted. For VS the positive m-contrast fit centered on 0°/180° (long room axis) significantly exceeded the w-contrast fit centered on 0°/±90°/180°, $t(11)$ = −3.22, $p = 0.008$, $d_z = 0.93$. For ES the m-contrast centered on −45/°135° (along corridor axis) produced a positive fit that significantly exceeded the w-contrast fit centered on ±45°/±135°, $t(11)$ = −4.78, $p = 0.001$, $d_z = 1.38$. Furthermore, both described m-contrasts exceeded 0 significantly, t's > 4.67, p's < 0.002. Thus, regarding pointing latency assuming a single reference axis aligned with the longest axis of the visible space fits the data better compared to two orthogonal axes.

2.2.2.2. Absolute pointing error. Fig. 5, left, depicts individual (dashed line) and averaged (solid line) absolute pointing error of both conditions. The ANOVA revealed a main effect of *environment*, $F(1,22) = 27.6$, $p < 0.001$, $\eta_p^2 = 0.556$. Participants pointed more accurately within the VS, indicating an advantage for environmental learning when the object layout was fully visible from one point of view. We also found a main effect of *body orientation*, $F(7,154)$ = 2.75, $p = 0.010$, $\eta_p^2 = 0.111$, qualified by a significant interaction of *body orientation × environment*, $F(7,154) = 6.47$, $p < 0.001$, $\eta_p^2 = 0.23$. Hence, as predicted, error size at specific body orientations differed depending on the learning environment (environmental vs. vista). According to our predictions, w-contrast fit was higher in VS than in ES, $t(22) = 4.14$, $p < 0.001$, $d_s = 1.69$, suggesting differently oriented reference frame orientations (Fig. 5, middle). Moreover, average values for w-contrast fit were above 0 for VS, $t(11) = 4.96$, $p < 0.001$, $d = 1.43$. This indicates that reference frames in VS were oriented along orthogonal directions of room walls and the intrinsic orientation of the object layout (i.e., 0°, ±90° and 180°). In contrast, reference frame orientations in ES tended to be oriented along oblique orientations, $t(11) = −2.15$, $p = 0.055$, $d = 0.62$. Clearly, learning in ES determined a reference frame perspective different from VS learning, even though the object layout was exactly the same.

Descriptive data of single participants mirror these effects. In VS 11 out of 12 participants (ca. 92%) showed better pointing performance in aligned trials (0°, ±90° and 180°) compared to oblique trials (±45°, ±135°). In ES only 3 out of 12 participants showed this pattern. 75% of the participants in this condition pointed more accurate from oblique body orientations.

As for latency, we tested employment of one vs. two reference axes. Fig. 5, right, depicts the indicative w- and m-contrast fits of the respective learning condition. Again, only positive m-contrast fits were of interest and compared to the corresponding w-contrast. In contrast to latency now positive values were produced for m-contrast centered on ±90° for VS condition (along short room axis) and m-contrast centered on 45°/−135° for ES condition (facing the wall of a corridor). The latter was larger than 0, $t(11) = 3.45$, $p = 0.005$, d = 1.00. However, no difference between m- and w-contrast fit was observed, ES: $t(11) = −1.75$, $p = 0.107$, $d_s = 0.51$, VS: $t(11) = 1.64$, $p = 0.130$, $d_s = 0.47$.

2.2.3. Further pointing results

Our results suggest that pointing performance varied as a function of corridor distance and body orientation and differently so for VS and ES learning. Did these factors cover most for the variability in the data or are there important communalities between the

learning situations remaining, originating from the common object layout? For example, it might be easier to memorize and recall the position of the teapot compared to the position of the trumpet. In order to explore this, we calculated the residuals for latency and error that express unexplained variance after both, corridor distance as well as orientation during pointing, were accounted for. For each target location error and latency residuals were separately averaged over participants. We then calculated the correlation of these residuals between the two conditions. The same was done for each location participants were currently pointing from. Correlating the residuals of latency in VS and ES across the positions one is currently pointing from rendered a medium but non-significant result, $n = 7$, $r = 0.598$, $p = 0.156$. This correlation was mainly driven by the objects *hammer* and *hair dryer*. Both were located in the middle of the layout. Pointing from objects at the edge of the layout allows for a rather fast and rough estimation of the correct pointing direction since all remaining objects lie in somewhat similar direction. For example, being teleported to the trumpet, facing 0° (short wall in VS) any target object will lie behind you. In contrast, being positioned at hammer or hair dryer one is surrounded by targets. Pointing from there may naturally lead to longer decision times. We want to emphasize, however, that this correlation is not significant, thus, diminishing its importance. The remaining correlations for error residuals across current position and for error and latency residuals across target locations did not render significant results, $n = 7$, r's < 0.37, p's > 0.42. While some idiosyncrasies of the layout might be present in both conditions it seems that corridor distance and body orientation explain large parts of participants' performance.

Considering either latency or error in isolation can lead to huge misunderstandings of the data and effects reported above. Therefore, we looked into possible interrelations of latency and error in our data. Not a single participants showed a significant negative correlation between pointing error and latency, all $r > −0.16$, $p > 0.097$. Considering this analysis one can assume that our pointing data is not suffering from a speed-accuracy trade-off.

2.2.4. Order of object placement

We analyzed the order in which objects were relocated in the placement task. For each single participant we computed the Kendal-tau ordinal correlation as a measure of agreement between the relocation order in the placement task and the order of first encounter as predefined in ES learning (i.e., teapot, horse, hammer, banana, hair dryer, trumpet and racket). This correlation criterion reflects how much participants revert to this one possible learning order. The agreement of the placement order in the VS condition with ES learning order, indeed, is expected to be non-existent. It rather functions as an important baseline value ES correlation is compared against. Fig. 6, left and middle, visualizes the degree to which participants preserved the order of learning in ES and used it to relocate the objects in the placement task. When learning took place in ES a clear relationship emerged with an average $r = 0.82$. The correlation is larger than 0, $t(11) = 8.62$, $p < 0.001$, $d = 2.49$, and exceeds the correlation found in the VS condition, $r = 0.11$, $t(22) = −4.34$, $p < 0.001$, $d_s = 1.77$, which itself, as expected, does not differ from 0, $t(11) = 0.79$, $p = 0.446$, $d = 0.23$. The strong correlation found in ES is also supported by the individual data: 8 out of 12 participants provided a perfect match of $r = 1$, remaining participants uniformly showed a positive correlation. Hence, order induced by the learning process in the ES was still present in the context of conducting a configurational judgment task for which order was virtually irrelevant.

As VS memory might be bound to a specific order as well, we examined multiple potential alternatives. The examined orders were clustered in three groups: (1) Recall along rows and columns of the object layout (including 16 plausible orders), (2) following
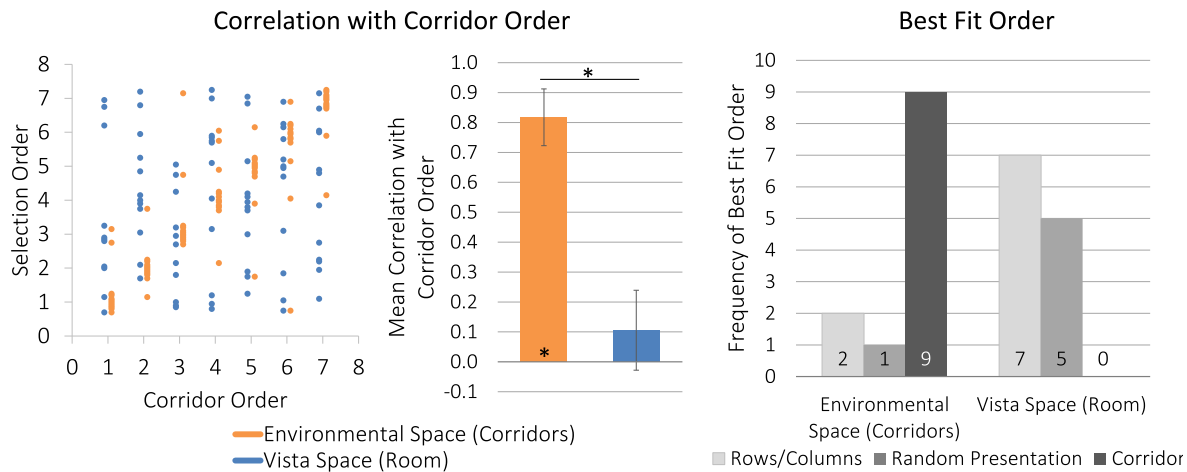
Fig. 6. Left: Placement order plotted as a function of the order objects were encountered when moving through the corridors. Middle: Mean correlation with corridor order. Right: Number of participants revealing best fit to either of three order clusters (along rows/columns, along object presentation during the placement task and along corridor based encounter) when rearranging the object layout from memory.

the random object presentation during the placement task (including eight plausible orders)[3] and (3) along corridor based encounter (single ES learning order as described above). The order in which the objects were presented to the participants in the placement task (i.e., a row) was randomized for each participant. However, orders along rows/columns of the uniform object layout are partly overlapping with the corridor order (mean correlation $r = 0.50$). For example, the potential order of learning and recalling the layout starting on the left and nearest object (teapot) and from there column by column to the right (each column starting with the nearest object) is highly associated with the corridor order, $r = 0.81$. A participant showing a high selection agreement with either will, thus, automatically show a high agreement with the other as well. Due to this non-independency a direct comparison between the strength of the different order correlations was not considered appropriate. Rather, we clustered participants with regard to the order that rendered the highest agreement with their individual placement order.

Fig. 6, right, depicts the amount of people scoring highest in any of the three order clusters.[4] The frequencies people were assigned to the three order groups differed significantly between the two learning environments, $p = 0.001$, two-tailed Fisher's exact test. More precisely, whereas in VS recall along rows/columns (7 participants; 3× best fit with order along rows, 4× best fit with order along columns) and along presentation of the objects in the placement task (5 participants) seems to be common, in ES most participants (9 out of 12) showed highest agreement with the corridor order.

2.2.5. Learning time and repetitions

Participants learning in ES needed significantly more time to meet the learning criterion (100% correct identification of objects at their corresponding location) compared to participants learning in VS, ES: $M = 8.13$ min, $SD = 1.82$, VS: $M = 4.38$ min, $SD = 2.08$, $t(22) = 4.71$, $p < 0.001$. This result is hardly surprising, since in ES

---

[3] Orders along rows/columns included, for example, from nearest to farthest row, each row from left to right; or from rightmost to leftmost column, across all columns alternating between nearest to farthest object and vice versa; et cetera. Orders following object presentation in the placement task when objects are arranged in random order along a single row included, for example, simply from left to right; or starting in the middle of the single row with the already active object to the leftmost object and from there the remaining object up until the rightmost object; et cetera.

[4] For one participant of the VS condition the correlation of placement order with presentation order and row/column order was equally high. We considered presentation order as the more conservative, thus, more appropriate fit, as it does not assume a specific order preserved in memory.

movement was required to explore the environment. The number of repetitions needed to meet the learning criterion did not differ significantly between learning in ES, $M = 1.25$, $SD = 0.13$, and VS, $M = 1.17$, $SD = 0.11$, $t(22) = -0.48$, $p = 0.633$, $d_s = 0.19$.

2.3. Discussion

Many studies have examined spatial learning either in VS (e.g., Kelly & McNamara, 2008; Meilinger & Bülthoff, 2013; Mou & McNamara, 2002; Shelton & McNamara, 2001) or in ES (e.g., Avraamides & Kelly, 2010; Brockmole & Wang, 2002, 2003; Cohen, Baldwin, & Sherman, 1978; Kelly et al., 2007; Kosslyn et al., 1974; McNamara, 1986; Meilinger et al., 2014; Newcombe & Liben, 1982; Wang & Brockmole, 2003a, 2003b; Werner & Schmidt, 1999). The present experiment aimed at answering the question of how spatial memory is different for VS and ES by comparing memory for exactly the same object layout. In summary, the observed distance, order, and alignment effects indicate that spatial memory for ES and VS differ in terms of preserved spatio-temporal learning experience, and employed reference frame orientation.

We examined whether the process of estimating object-to-object relations depended on corridor or Euclidean distance to the target. To date, better performance for closer targets with regards to traveled distances in ES was mainly found in the inspection of errors (Thorndyke & Hayes-Roth, 1982), but not latency. However, since an increase in error over corridor distance could result also from error accumulation during learning (thus, memory precision), we assume latency to be a more suitable criterion for indicating processing time (thus, retrieval) (see also Pantelides et al., 2016). We hypothesized an incremental process when recalling memory acquired in ES. Indeed, when learning took place in ES results from the visual pointing task suggest that pointing latency increased with increasing corridor distance. With each additional corridor separating participant and target, more time was needed to point to the target. To our knowledge, this was the first time a traveled distance effect on retrieval latency of survey memory has been demonstrated. It points to the fact that in ES the process of recalling locations beyond the current VS cannot be a simple and direct read-out from a single memory unit. Estimation of object-to-object relations in ES rather happens incrementally during pointing, successively activating spatially and/or temporally related information from memory. Whether this process consists of

mentally walking down the route (Byrne et al., 2007; Sanders et al., 2015) or of constructing a mental model of the non-visible environment parts from one's current location (Meilinger, 2008), however, remains an issue to exam further.

The layout itself was not responsible for the corridor distance effect, since the effect was not evoked by VS learning. Similarly, pointing latency did not increase with increasing Euclidean distance between object pairs. As McNamara (1986) showed that Euclidean distance affects the accuracy of survey judgements, these results again support the disparity of error and latency measurements. Assuming that pointing latency is a manifestation of the recall process rather than the precision of memory, all objects learned within the VS seemed to be retrieved equally fast irrespective of both distance metrics. Likely, no further mental constructions were needed here; rather, configurational memory already comprised all object locations in a single reference frame as suggested by prior research (McNamara et al., 2008; Mou et al., 2004). Additionally, this result might indicate that the positive association between pointing latency and Euclidean distance in ES was merely driven by the correlation of corridor and Euclidean distance. Taken together, learning differs across ES and VS environments which manifests in the absence of a distance effect in the VS environment, and an effect of corridor distance in the ES environment. Specifically, even though both groups of participants were instructed to learn the same object layout, in the ES environment the temporal and spatial learning procedure was preserved in memory retrieval, and thus, the structure of the layout representation.

We predicted that retrieving memory for ES would be bound to the learning order due to the successive entering of corridors in ES learning. Indeed, a high correlation between the order of first object encounter in ES and the order of object relocation during the placement task was found for ES. It is important to note that this order effect was observed in the retrieval procedure of a survey task, a task that did not require placing the objects along a specific order. In fact, in order to select objects according to the encoding order in ES learning, participants even had to deselect items, disregard the random objects order presented in front of them and deliberately select other items. The strong relation between encoding order and placement order in ES is in contradiction to a VS process of reading-out from a single reference frame representation (McNamara et al., 2008; Mou et al., 2004), which itself does not predict a preferred order, as mere inter-object relations are memorized. Interestingly, from twelve participants in the ES condition, eight provided a perfect order match and all others showed a positive correlation as well. Hence, never was an object layout reconstructed from the endpoint of the ES to the starting position, even though this sequence was encountered just as often as walking from start to end. Previous research showed how explicitly landmark sequence is preserved and affects memory retrieval when learning a route (Janzen, 2006; Schweizer et al., 1998; Strickrodt et al., 2015; Wiener et al., 2012) and when learning the configuration of an ES (Moar & Carleton, 1982). The present experiment extends these findings by demonstrating that order effects are determined by the very first experience within the environment (original forwards direction) even when walking in both directions (forwards and return path).

Similar to the effect of corridor distance, we did not expect the learning order of ES to be reflected in spatial memory acquired in a VS. VS learning functioned as a control condition and the absence of a meaningful correlation with corridor order demonstrates that effects in ES were not due to the layout structure itself. Regarding the best order fit analysis, placement order was much less consistent in VS. Participants showed placement order patterns along rows and columns (with no systematic preference for either following rows or columns) or merely along the random presentation order in the placement task. From this result we cannot determine whether layout memory of VS is generally structured along rows and columns but easily overwritten by presenting alternative sequences or whether just some participant preserve a row-/columns structure in memory. It should be noted that we consider a statistical comparison within a learning condition as problematic as the amount of potential orders that were covered in our analysis differed between rows/columns, presentation and corridor order. We examined 16 plausible orders representing selection along rows and columns, but only eight plausible orders of following the random object presentation in the placement task and one order representing corridor learning. Thus, the chance to be clustered in either order group is unequal. This is problematic for the distributions of orders within a learning group, but less so for comparisons between learning groups. It should also be noted that the analysis of a best fit sequence depends on the selected orders that are taken into account. We clearly did not cover all feasible orders, but only a subset of reasonable orders. As a result, most participants have been clustered into order groups because all remaining correlation rendered lower values – even though their maximum correlation was not significant. In sum, we showed a clear difference between VS and ES learning. ES learning clearly preserved the distinctive initial learning encounter.

As expected, we found evidence for differently oriented reference frames in VS and ES indicated both by error and latency. W-contrast fits were positive for learning in VS, hence, reference frames were centered on the aligned orientation of initial view, room geometry and object layout (0°, ±90°, 180°) as in prior studies (e.g., Kelly & McNamara, 2008; Shelton & McNamara, 2001; Valiquette & McNamara, 2007). In contrast, in ES negative w-contrasts were found, i.e., overall lowest error rate and fastest responses were shown when aligned with or orthogonal to the corridors (±45° and ±135°). Such an alignment with vista units of an ES has also been observed before (Meilinger et al., 2014; Werner & Schmidt, 1999). Our study demonstrated in addition that the initial view and the global layout-intrinsic orientation are less important for setting the reference frame in an ES. In VS studies it has already been shown that the initial view can be dominated by another, if this new view is aligned with a geometric feature (e.g., global room, mat, object layout) (Kelly & McNamara, 2008; Shelton & McNamara, 2001; Valiquette & McNamara, 2007). Similarly, the same seems to account for ES learning. During walking participants are aligned with the corridor walls as well as with the locally visible objects. These factors seem to determine the alignment of spatial memory.

We further examined whether the assumption of two orthogonal axes constituting a reference frame, i.e., performance following a w-shaped pattern, holds. Alternatively, space could be represented along a single axis, i.e., performance following an m-shaped pattern. Considering both, pointing error and latency, our results are inconclusive about whether a single or two orthogonally aligned reference axes were involved. Pointing latency was better described by an m-contrast centered on the long axis of the room (0°/180°) in VS, and by an m-contrast centered on the corridor axis (−45°/135°) in ES compared to the corresponding w-contrasts. In contrast, when analyzing pointing error in ES the m-contrast centered on the orientations when facing a corridor wall (45°/−135°) produced a positive fit. Also, VS now evidenced two reference axes rather than one. From this no clear conclusions about the number of reference frame axes can be made. Nonetheless, our results demonstrated that performance pattern in ES are in clear opposition to the performance pattern in VS learning, which we assume to be the crucial point here. The overall orientation of the reference frame seems to be well captured in the w-contrast fit rendering this measurement a reliable, even rather conservative mean to detect differences in reference frame alignment between ES and VS.

In addition to the alignment effect, VS learning also resulted in higher pointing accuracy than ES learning. This difference likely originated from the specific differences of VS and ES learning such as successive vs. instant visibility of objects, required movement and common visible anchor (i.e., the room) for VS, but not ES learning. In ES, participants need to relate locations that were never encountered together. The mental effort to construct a mental representation of the object layout is likely to be higher and the process more error-prone in ES compared to VS learning.

This is the first work showing that retrieving configurational memory for ES is bound to the traveled distance and order of learning. Observed effects cannot be accounted for by a simple read-out process from a single reference frame, which typically explains memory retrieval for VS. Albeit these results clearly show where memory for VS and ES differ, they do not answer the question of the underlying reasons, which is the subject of Experiment 2.

3. Experiment 2

Experiment 2 was concerned with what aspects of the learning situations may cause differences in the memory structure. Most importantly, the separation of ES in multiple VS units (compartmentalization), the movement through space, and the successive encounter of objects should be treated as potentially relevant factors for a divergence.

3.1. Compartmentalization

The nature of ES is that the environment is separated into units by spatial borders. Opaque barriers were found to elicit overestimation of physical distance between targets (Kosslyn et al., 1974). The effects of distance, order, and reference frame alignment found in Experiment 1 identify additional characteristics on which ES memory differs from VS memory which, indeed, might have been caused by opaque borders. There is, however, evidence that not just opaque borders, but also non-opaque borders elicit distinct distortions in spatial judgements. McNamara (1986), for example, reported a bias in distance estimation when learning an environment in which spatial borders were merely set by strings on the floor (i.e., no opaque border). As the compartmentalization of space in the ES condition was inevitably linked to the need of movement and to successive object encounter it is important to identify the cause of effects found in Experiment 1. In the VS condition of Experiment 1 participants were restricted to learn the fully visible object layout standing at one location. To account for the potential influence of movement and successive object presentation we emulated both in a VS setting in Experiment 2. Indeed, in a real-world scenario one can easily move around in VS as well, successively passing the objects within. However, exploring a VS is not subject to restrictions comparable to restrictions imposed by an ES structure. Most importantly, a VS provides a common reference space objects within the space can jointly be related to. This common reference space might facilitate the integration of object locations into a single reference frame compared to learning in a compartmentalized space (ES), regardless of movement and successive objects presentation.

3.2. Movement

The translation through space when learning an ES makes it possible to experience a multitude of visual and proprioceptive information. In contrast, many studies concerned with the learning of object layouts in VS typically exclude walking from the learning procedure. Often, visual information are presented from one up to a few predefined vantage points. Indeed, learning in VS does per definition not require movement, since it involves all spatial information that can be gathered from a single vantage point.

To examine whether the effects of order, distance and alignment originate from movement, in Experiment 2 we now had participants walk through the room along a path matching exactly the movement through ES. Now the path determined, for example, that in order to travel from the teapot to the hammer the horse has to be passed by making a detour. The prevention of a direct path between teapot and hammer might be interpreted as a boundary, which in turn might influence memory structure. Furthermore, now a walking distance between pairs of objects was provided. Both spatio-temporal information, the impression of a non-visible boundary and the experienced walking distance, might promote order and distance effects. At the same time, these effects might be diminished since the VS itself allows for a global observation of all environmental features.

Finally, walking across the object layout might also induce a different reference frame orientation. By introducing the path traveled in ES in a single room we created a conflict between multiple inputs. The visible context of the room (room geometry, global object layout) and the initial view now have to compete against varying viewpoints, perspectives and body orientations during movement with the main learning orientation being oblique to the room axes. Previous papers have demonstrated the importance of self-to-geometry alignment that is experienced later during learning, after the initial view for setting the reference frame orientation (Kelly & McNamara, 2008). Also bodily cues were found to be of importance. Yamamoto and Shelton (2005, 2007) showed that proprioceptive learning (blindfolded walking) by itself can yield a reference frame orientation seemingly independent of and comparable in strength to visual learning. Hence, multiple encoded views and different body orientations during learning might counterbalance the visible context and influence reference frame usage. Varying whether participants walk along the route or exhibit the object layout from a constant view will help to understand how these factors influence the structure of spatial memory.

3.3. Object presentation

Another aspect distinguishing learning in ES from learning in VS is that the visibility of objects is not simultaneous. The environmental borders and transition points from one spatial entity to the next determine the sequence in which objects are encountered. While objects of the previous VS will be out of sight, objects in the currently visited VS will now be attended. As an important aspect of the learning procedure in ES, we wanted to examine whether successive presentation cause or contribute to the maintenance of spatio-temporal encoding information in memory, i.e., distance and order effects, and to the alignment of the reference frame. Therefore, in Experiment 2 some of the participants were confronted with the target objects step-by-step, adopting the object encounter of the ES condition of Experiment 1 within a VS. Objects located within the same corridor in the ES condition of Experiment 1 were now, within the VS room, visible at the same time, alternating with the next object pair and so forth. Such a learning procedure will set the spotlight to discrete object pairs while preventing the view upon another proportion of the layout objects. Thus, similar to movement through space, successive object presentation determines a specific spatio-temporal learning experience that might as well induce order and distance effects. Furthermore, now the global object layout as a potential cue influencing reference frame alignment, will not be apparent anymore. Rather, pairs of objects aligned with orientations oblique to the room geometry constitute another visual cue, which might affect reference frame alignment.

In Experiment 2 we set out to examine which aspects of the learning procedure that distinguish ES from VS learning lead to divergence in the spatial representation of the same object layout. As spatial separation along opaque barriers cannot be varied independently of movement and successive object presentation we eliminate the compartmentalization of space, but varied the other two factors. Learning conditions in VS were step-by-step adapted to ES learning. We had three learning groups: Participants viewed objects successively from a static position (*stat-succ*), viewed the objects simultaneously but followed the path executed in ES (*move-simult*), or viewed the objects successively while following the path (*move-succ*). The last condition differed from ES learning only by the absence of walls, i.e., the absence of multiple VS units. In combination with VS learning in Experiment 1 (*stat-simult*) this yielded a 2 (object presentation: simultaneous vs. successive) × 2 (movement: static position vs. movement through space) plan to examine how far any of these conditions leads to results matching findings of ES. Finding that movement or successive object presentation (or their combination) in VS elicit similar effects to learning in ES would assign them to be determining factors for configurational learning of ES. However, the absence of order effects, distance effects, or reference frame alignment along oblique orientations in Experiment 2 would identify the separation of space as the remaining, determining factor for the distinct construction of configurational knowledge.

3.4. Method

Methods were identical to Experiment 1 except for the alternations described.

3.4.1. Participants

36 participants (19 females) with a mean age of 26.97 years participated (*SD* = 7.57, [16;48]) and were randomly assigned to one of the three groups (12 per group). One participant withdrew from participation after completing the visual pointing task; hence, placement data of this participant was not recorded. From the original sample of 39, two participants were excluded since they did not perform significantly better than chance level of 90° absolute pointing error. Another participant was excluded due to computer problems during the pointing task.

3.4.2. Materials and procedure

All participants learned the object layout within the rectangular room of the VS condition of Experiment 1 and the same initial view (Fig. 1, right). The remaining procedure was adjusted according to the conditions. In condition *stat-succ* no movement was required. Participants were not allowed to leave their current position, but they were obliged to look around. Objects were presented in successive order matching the presentation of objects of ES learning. Objects formerly presented within the same corridor in ES were now presented at the same time (object pairs). Common visibility of objects, henceforth, was as follows: teapot – horse and hammer – banana and hair dryer – trumpet and racket. To enable self-paced learning participants pressed the button of a controller. Following a duration of continued button press a switch from one object (pair) to the next took place. This duration matched average walking time through a corridor in ES as determined in pre-experiments. Importantly, when the former object (pair) disappeared, the next object (pair) appeared. Since in ES learning of Experiment 1 participants were allowed to stop at any point during their movement through space and, hence, determined encoding time themselves, also participants in the *stat-succ* condition could prolong the view upon the current object (pair) by pausing the button press of the controller. When reaching the last object pair (trumpet and racket), object pairs were presented in backwards

order again. This procedure was repeated one more time (similar to walking twice from start to the end point in ES).

Participants in the *move-simult* condition moved through VS, matching the path through the ES corridors of Experiment 1. Grey discs on the floor led participants on a specified path. When reaching a disc, the next target disc lit up. Disc locations corresponded to the location of direction change in ES, i.e., the end and start point of each corridor. Only the disc at the current and the next position was displayed. Participants had to walk on a straight line towards the next disc without detours. However, they were allowed to stop at any point and to look around. Participants in the movement conditions had to follow the corridor route four times, alternating between forward and return path. In the *move-simult* condition the complete object layout was visible simultaneously during learning, analogous to VS learning in Experiment 1.

The *move-succ* condition matched both the movement and the visibility of objects to ES learning. Objects were presented successively, as in the *stat-succ* condition, while participants had to follow the path mirroring movement in ES similar to the *move-simult* condition. The presentation of object pairs occurred automatically when participants reached positions that corresponded to the passage between two ES corridors.

After translation between start and endpoint (movement conditions) or being confronted with all object pairs four times (*stat-succ*), the objects were removed and acquired object knowledge was tested. For the static learning condition, the learning test was identical to VS learning in Experiment 1. For the movement conditions, the learning test was identical to the environmental condition of Experiment 1, except that participants walked through the room, not the corridors and their walking was again guided by the discs. After reaching the learning criterion (100% correct identifications) the *test phase* started.

3.4.3. Data analysis

From the 36 participants, ca. 4% of the pointing performance data was deleted due to deviation of more than 2 *SD* from a participant's overall mean. One participant in the *move-simult* condition withdrew from continuing the experiment after the pointing task, rendering only eleven participants in this condition for the analysis of placement order. The aim of Experiment 2 was to understand the impact of movement, successive object presentation and compartmentalization of space on the acquired memory. We focused on quantitative effects on which VS and ES condition clearly differed in Experiment 1. Thus, we confined our analysis to latency increase with movement/presentation distance (i.e., corridor distance) as represented by individual regression slopes, pointing performance fits to a w-pattern centered on 0°/±90°/180° and correlations of placement order with ES learning order (i.e., corridor order). These analyses enabled us to judge how much movement, object presentation and their combination render equal values to those obtained from ES learning. We analyzed these data with a 2 (movement) × 2 (object presentation) ANOVA (including VS condition of Experiment 1) to estimate the influence of each factor separately, as well as of their combination. Subsequently, we compared the level of each parameter with the respective value in the ES condition of Experiment 1. Please note that comparisons across experiments were valid as Experiment 1 and 2 were conducted together and participants were assigned randomly to all conditions.

3.5. Results

3.5.1. Distance to target

For the distance analysis of Experiment 2 we regressed pointing latency of each single participant onto the two distance metrics and extracted the standardized slopes *b*. Fig. 7 depicts the mean slopes for the linear increase in pointing latency as a function of

corridor distance (left panel) or Euclidean distance (right panel) to the target. The three bars on the right in each panel mark data collected in Experiment 2. Data from Experiment 1 is depicted again in the two bars on the left for ease of comparison. Note that corridor distance in Experiment 2 was not induced by actual corridors but by walking and successive layout experience. The ANOVA on the linear increase of latency across ascending corridor distances revealed that neither *movement* nor *object presentation* exert a main effect on the level of slope. Further, no significant interaction could be found, F's < 2.1, *p*'s > 0.159. Neither learning condition led to a slope larger than 0, *t*'s < 0.95, *p*'s > 0.363, similar to results of the VS condition in Experiment 1 (*stat-simult*). Thus, the manipulation of movement and object presentation in VS did not lead to an increase of pointing latency across ascending corridor distances. When considering individual slopes, in each new VS condition 6 to 7 out of 12 participants revealed a positive slope. This further supports the aforementioned results. Consistently, all VS slopes differed from the slope attained in ES, *t*'s > 4.01, *p*'s < 0.002. Since neither movement, nor object presentation or the interaction rendered effects this suggests that the increase in reaction time across corridor distance found in Experiment 1 was due to the visual separation of the space.

A similar analysis was conducted for the effect of Euclidean distance. Standardized slopes representing the potential linear increase of latency across Euclidean distance were submitted to an ANOVA. Results mainly resembled those attained for corridor distance: No main effects or interaction of *movement* and *object presentation* could be found, F's < 1.71, *p*'s > 0.197. Also, neither VS condition that mirrored aspects of ES learning rendered slopes that exceed 0, *t*'s < 0.96, *p*'s > 0.359. Individual slope distribution was again near chance: 5 of 12 participants (ca. 42%) in each new VS condition yielded a positive slope. When comparing the linear increase of latency across ascending Euclidean distance obtained in this experiment to the ES condition from Experiment 1, only conditions *stat-succ* and *move-succ* differed significantly from ES, *t*'s > 2.13, *p*'s < 0.045. This supports results found in Experiment 1. Euclidean distance again was of no significant importance when learning took place in a single room (VS).

3.5.2. Reference frame orientation
3.5.2.1. Pointing latency.
Pointing latency as a function of body orientation is depicted in Fig. 8, top left. Contrast fits to a w-pattern centered on 0/±90°/180° are displayed on the top right, the three bars on the right of the panel marking data collected in Experiment

2. No main effects of *movement* or *object* presentation on w-contrast fit were found, F's < 0.17, *p*'s > 0.676. There was, however, a trend for an interaction of *movement ∗ object presentation*, $F(1,44)$ = 3.256, $p = 0.078$, $\eta_p^2 = 0.069$. Condition *stat-succ* rendered highest contrast fits. Average contrast fits for pointing latency in all VS conditions were positive, for *stat-succ* significantly above 0, $t(11)$ = 2.37, $p = 0.037$, $d = 0.68$. In line with the results of Experiment 1, fits were clearly different from ES learning, *t*'s > 2.30, *p*'s < 0.032. Thus, neither movement nor successive presentation (or their combination) yielded a similar shift in reference frame orientation towards oblique directions (±45°, ±135°) as found in ES learning. The same pattern was found when looking at single participants: In the *stat-succ*, *move-simult* and *move-succ* condition 9 (75%), 7 (58%) and 6 (50%) out of 12 participants showed numerically faster pointing performance in trials of aligned body orientation (0°/±90°/180°) compared to trials of oblique body orientation (±45°/±135°), respectively.

3.5.2.2. Absolute pointing error.
Fig. 8, bottom, displays results for the absolute pointing error which parallel those of pointing latency. No main effect of *movement* or *object presentation* or an interaction between the two could be found, F's < 0.51, *p*'s > 0.482. Again, presenting objects successively or allowing translation through space, did not seem to influence w-contrast fits, i.e., the selection of reference frame orientation. Contrast fit for *move-simult* exceeded 0, $t(11) = 2.96$, $p = 0.013$, $d = 0.85$, and for *stat-succ* by trend also, $t(11) = 2.08$, $p = 0.062$, $d = 0.60$. Just as for pointing latency, w-fits in all VS conditions were numerically larger than 0 and significantly different from the contrast fit of ES learning of Experiment 1, *t*'s > 2.10, *p*'s < 0.048. This pattern again was mirrored in individual data: In the *stat-succ*, *move-simult* and *move-succ* condition 8 (67%), 10 (83%) and 9 (75%) out of 12 participants pointed more accurate in trials of room aligned body orientation (0°/±90°/180°) compared to trials of oblique body orientation, respectively.

Pointing latency and error were negatively correlated in 1 out of 36 participants, $r = -0.31$, $p = 0.006$, remaining correlations $r < 0.21$. The average correlation of error and latency across the three VS condition of Experiment 2 was $r = 0.01$. As in Experiment 1, data does not seem to exhibit a speed-accuracy trade-off.

3.5.3. Order of object placement
Fig. 9 shows the mean correlations between experienced order and placement order. Considering the different VS conditions nei-
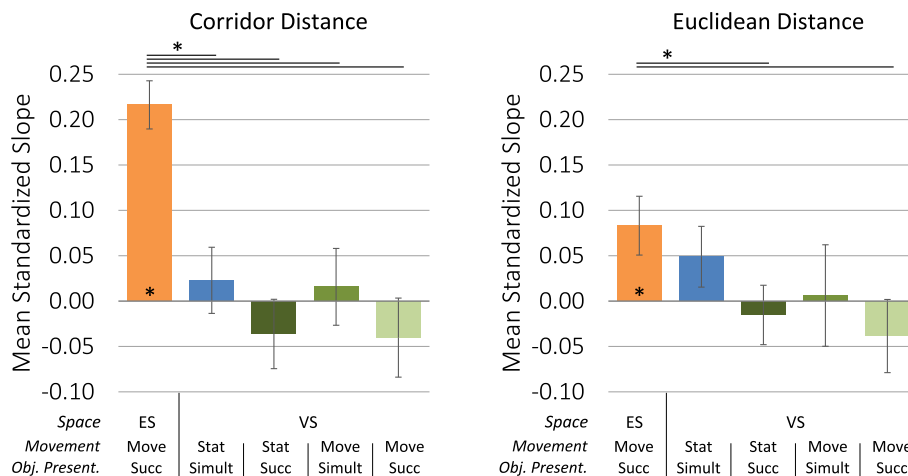

Fig. 7. Mean standardized slopes of the effect of corridor distance (left) and Euclidean distance (right) on pointing latency. The three bars on the right of each panel represent the new learning conditions of Experiment 2, mirroring ES learning in a single room. Move = Movement, Stat = stationary without movement, Succ = successive presentation of objects, Simult = simultaneous presentation of objects.
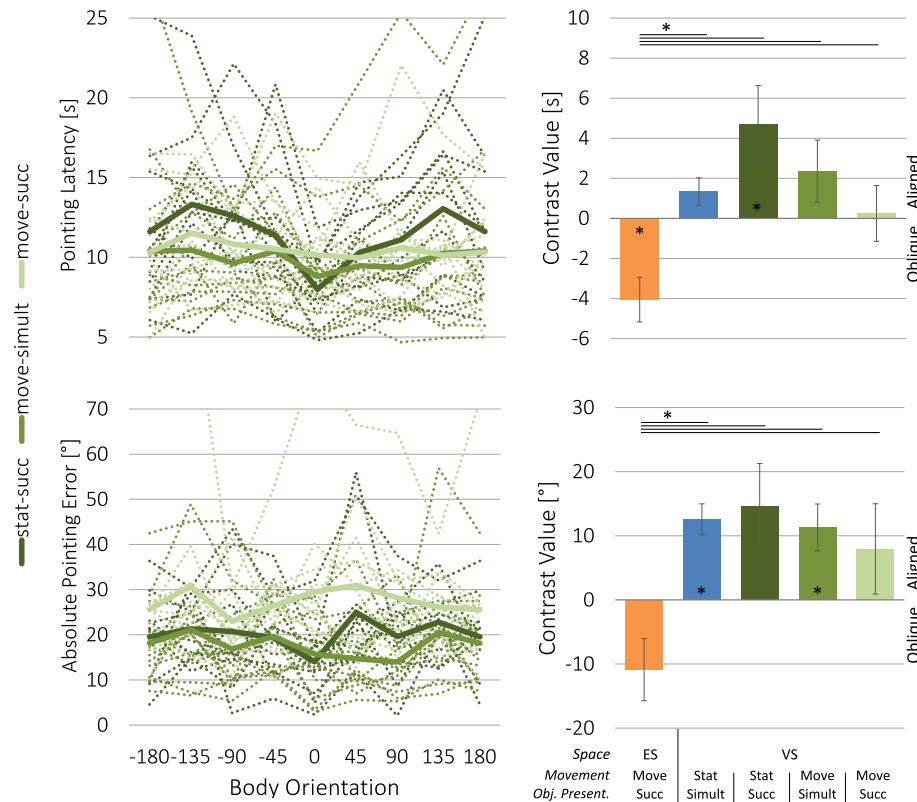
Fig. 8. Pointing latency (top) and absolute pointing error (bottom) as a function of body orientation. Left: Individual (dashed line) and averaged (solid line) pointing performance of the new VS learning conditions of Experiment 2. Right: Data fit to the w-contrast centered on 0°/±90°/180°.
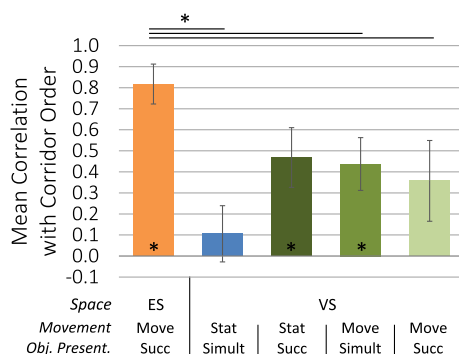

Fig. 9. Mean correlations of corridor order and placement order for each learning condition.

ther a main effect of *movement* or *object presentation*, nor a significant interaction was present, *F*'s < 2.14, *p*'s > 0.151. Thus, varying movement and object presentation in VS did not seem to have influenced the corridor order effect. However, in contrast to VS from Experiment 1 the mean correlations of *stat-succ*, *t*(11) = 3.30, *p* = 0.007, *d* = 0.95, and *move-simult*, *t*(11) = 3.49, *p* = 0.006, *d* = 1.05, differed from 0, and by trend also *move-succ*, *t*(11) = 1.861, *p* = 0.090, *d* = 0.54. Although lager than 0, the correlations were still smaller than in ES for *move-simult*, *t*(21) = −2.42, *p* = 0.024, *d_s* = 1.02, *move-succ*, *t*(16.1) = −2.15, *p* = 0.047, *d_s* = 0.88, and by trend also *stat-succ*, *t*(22) = −2.05, *p* = 0.053, *d_s* = 0.84. Consequently, it can be inferred that movement and successive object presentation either alone or in combination led to a mediocre relocation preference along the order of learning. However, the order effect still differed from the effect found in ES learning. When

controlling for gender as a covariate, a main effect of gender was found, *F*(1,44) = 6.583, *p* = 0.014, *η_p^2* = 0.141, which, however, did not yield any major changes in abovementioned results. Overall, female participants exhibited a larger correlation between learning and placing order, *r* = 0.54, than males, *r* = 0.13, but both seem to be similarly affected by movement and object presentation.

3.5.4. Learning time and repetitions

The time needed to learn the environment varied across the different VS conditions. We observed a significant main effect of *movement*, *F*(1,44) = 19.06, *p* < 0.001, *η_p^2* = 0.30, as well as a trend for *object presentation*, *F*(1,44) = 3.47, *p* = 0.069, *η_p^2* = 0.07, on learning time. Learning took more time when movement and successive object presentation was introduced. The interaction of *movement* and *object presentation* was not significant, *F*(1,44) = 0.29, *p* = 0.598, *η_p^2* = 0.006. Participants spent on average 6.77 min (*SD = 3.20*) learning in the *stat-succ* condition, 11.26 min (*SD = 4.28*) in the *move-simult* condition, and 15.55 min (*SD = 11.02*) in the *move-succ* condition.

Average number of learning repetitions required to pass the learning criterion (100% correct identification of objects at their corresponding location) were *M* = 1.42, *SD* = 0.15 for *stat-succ*, *M* = 1.75, *SD* = 0.18 for *move-simult* and *M* = 1.83, *SD* = 0.27 for *move-succ*. Movement conditions evoked more learning repetitions than learning from a static position, as indicated by the main effect of *movement*, *F*(1,44) = 7.14, *p* = 0.011, *η_p^2* = 0.14. No main effect of *object presentation* or interaction of *object presentation* × *movement* could be found, *F*'s < 0.80, *p*'s > 0.377. Findings suggest that movement makes it harder to learn the locations of objects in space. Nevertheless, individual adaptation of learning time and number of repetitions ensured that acquired object location knowledge was sufficiently comparable between groups.

3.6. Discussion

Experiment 2 investigated possible underlying mechanisms differentiating VS and ES learning. Learning experience in ES differed from VS learning. Specifically, in ES navigators were confronted with a compartmentalized space, had to walk through the environment and experienced successive presentation of objects. This study manipulated movement and successive object presentation and their combination within the visual context of a VS to estimate their respective and combined influence. Our results across all three parameters revealed a clear picture: neither changes in movement, object presentation or their combination influenced the acquired layout memory in a way equal to learning in ES. In fact, basically each single VS condition differed on each parameter from the ES condition (only learning order for *stat-succ* did – with $p = 0.053$ – not reach significance). We conclude that the separation via opaque borders must be the main source of differentiation.

No evidence could be found, that either the successive visibility of objects, movement through space or the combination of both had a specific effect on pointing latency with increasing movement/presentation distance (i.e., corridor distance). In Experiment 2 no additional time was needed to activate memory for objects which were passed and/or perceived later during learning, regardless of the strong spatio-temporal characteristics of movement and successive object presentation. Movement, providing additional proprioceptive input, has previously been found to shape spatial knowledge (Chrastil & Warren, 2013; Waller, Loomis, & Haun, 2004; Yamamoto & Shelton, 2005, 2007). However, bodily walking cues did not yield an ES-like memory structure when learning in VS. Similar pointing latencies for both corridor and Euclidean distance also strengthen the conclusion that object locations in each VS condition were memorized within a single, integrated representation, irrespective of the learning procedure and straight-line distances between objects. This implies that the presence of borders between corridors in ES learning is responsible for the incremental processing during retrieval that was found in Experiment 1.

In the placement task we observed medium size order effects. The correlation between learning order and placement order were larger than 0 in every VS condition incorporating movement and/or successive object presentation (for *move-succ* at least by trend). This dissociates VS conditions of Experiment 2 from the original VS condition of Experiment 1 (*stat-simult*), where no order effect was present. Creating a spatio-temporal contingency by guiding movement and/or restricting the attentional focus led to the incorporation of learning order in configurational memory for VS as well, guiding memory retrieval, however, to a lesser extent compared to ES learning. Sensitivity for route direction or order was mainly shown in studies utilizing ES (e.g., Janzen, 2006; Moar & Carleton, 1982; Schweizer et al., 1998; Wiener et al., 2012). Our results imply that this sensitivity seems to be – at least partly – independent of whether there is a common reference space (VS) or not (ES). Most importantly, none of the VS conditions induced similar learning order effects as ES learning. This illustrates that guidance of attention can only partly explain the order effect found in ES and it reveals the impact of opaque borders on shaping order dependency. Relying on order when learning in ES or uncoupling from the learning order in VS can each for itself yield advantages. In ES order might be particularly advantageous to not confuse the sequence of single corridors. This is exceedingly important if ES memory consists of multiple subunits. In contrast, VS memory that is not bound to a specific order might be retrieved more flexibly.

Why did movement and successive object presentation in VS yield middle sized order effects, but no effects of movement/presentation distance (i.e., corridor distance)? We speculate that placement order in the layout reproduction task (irrespective of where exactly participants place the objects) is associated with the *temporal* aspect of spatial knowledge, whereas pointing latency across varying distances captures how *spatial* aspects (direct relations between pairs of objects) are retrieved from memory. Such a dissociation of memory systems specialized in spatial locations vs. behavioral responses, which incorporate also the temporal order, have been proposed before (Packard & McGaugh, 1996; Restle, 1957). In a case study van der Ham et al. (2010) demonstrated how temporal and spatial aspects of navigation are dissociated in humans. Impairment in a route ordering task did not similarly lead to impairment in route continuation task, or vice versa. Likewise, addressing different aspects of survey knowledge might be prone to an analog dissociation between temporal and spatial aspects. Furthermore, predefining an order by movement or successive object presentation might generate an additional verbal memory trace constructed along the learning order. Verbal memory was shown to be involved within route learning (Meilinger, Knauff, & Bülthoff, 2008; Wen, Ishikawa, & Sato, 2011) as well as learning of an object layout (Meilinger & Bülthoff, 2013). Memory retrieval in the subsequent placement task might be initialized following this verbal code. Female participants exhibited larger order correlations than males. This was the only effect of gender observed in both experiments. We speculate that this effect might originate from a stronger reliance on a verbal coding strategy for spatial material in women (Coluccia & Louse, 2004). Thus, the dissociation between spatial and temporal aspects of spatial memory may explain the emergence of mediocre order effects in the absence of distance effects, and/or verbal coding along the learning experience might be responsible for part of the order effect observed.

In Experiment 2, we induced a conflict of available reference axes evoked by movement, successive object presentation and a common reference space. Participants moved along paths and/or were confronted with pairs of objects which were aligned with an axis (main axis of $-45°$ to $135°$) that is oblique to the initial view, room geometry and global object layout ($0°/\pm90°/180°$). W-contrasts of all VS conditions differed clearly from ES learning. This does not imply that movement and successive object presentation have no effect on the alignment of the reference frame. For example, the pattern of orientation dependency of the pointing performance in *move-succ* (most similar to ES learning) seems to become more leveled, not showing a clear trend in either direction. Here the maximum conflict of available geometric axes and views is experienced. However, as the pattern even in this condition induced no conversion of the dependency pattern of body orientation and clearly differed from ES learning, we conclude that on their own movement and successive object presentation are not sufficient to assimilate the clear reference frame alignment along oblique orientations that was found in ES. Following this, we conclude that compartmentalization through opaque barriers – the remaining factor differentiating ES and VS learning of Experiment 1 – was responsible for the clear shift to oblique orientations in ES. Within a single corridor of the ES the visible objects, corridor walls and experienced views through movement were jointly aligned, supporting a corresponding reference frame alignment (Kelly & McNamara, 2008; Shelton & McNamara, 2001; Valiquette & McNamara, 2007). We assume that the potentially conflicting cue of the initial view was easily overwritten by the viewer-space-alignment when walking through the corridors (compare to Kelly & McNamara, 2008). Furthermore, the opaque borders literally cut off the perception of the potential conflicting cue of the global layout orientation. This could only be inferred at the moment the last object was discovered and indeed only a mental, probably distorted representation could have been used. As our results demonstrated, no effortful restructuring and realignment of layout memory on the basis of an inferred global layout orientation – a cue that extends beyond the current corridor unit

– was carried out in ES. Both in VS and ES the visible surrounding geometry seems to serve as the main cue organizing a reference frame for remembering locations in space.

In conclusion, our results clearly show that *movement* and *object presentation* introduced in a VS do not render the performance pattern observed in ES learning. This leaves the *compartmentalization of space* as a main factor causing the memory structure of an ES to differ from the memory structure of a VS. Having a common, continuously visible reference within the VS allowed participants to (1) integrate all target locations into one representation without successively activating spatially distant information from memory, (2) rely less on the order of learning although full decoupling was not observed, and (3) to employ reference frames different from the ones used in ES learning.

4. General discussion

We examined memory for an object configuration learned within a VS (a single room) or within an ES that is spread across multiple corridors. Experiment 1 showed that configurational memory differed qualitatively: Contrary to VS learning, retrieving memory of the ES was bound to the distance experienced and to the order in which the objects were learned. Also, ES learning employed different reference frames whose orientation followed the orientation of corridors rather than the initial view of the environment or the layout-intrinsic orientation. Experiment 2 revealed that neither the movement trajectory, nor the successive presentation of objects, or the combination of both could fully account for the qualitative differences. Having examined these factors we conclude that compartmentalization into multiple sub-spaces is the main factor responsible for the dissociation of memory between these two classes of space.

Our results blend nicely into existing findings. Spatial borders were found to affect updating (Avraamides & Kelly, 2010; Kelly et al., 2007; Wang & Brockmole, 2003a, 2003b), distance estimation (Cohen et al., 1978; Kosslyn et al., 1974; McNamara, 1986; Newcombe & Liben, 1982), reference frame selection (Meilinger et al., 2014; Werner & Schmidt, 1999) and switching costs between spatial units (Brockmole & Wang, 2002, 2003). Also interpretation of the current results clearly supports the theoretical distinction between VS and ES proposed by Montello (1993). Our results extend prior findings in that they demonstrate clear differences in the memory structure of different spaces on three different aspects by directly comparing VS and ES learning with the same material, thus, excluding additional differences.

A distinction between ES and VS based on opaque borders is found in neuroscience as well (for recent overviews on navigation see Spiers & Barry, 2015; Wolbers & Wiener, 2014). Visual borders were shown to influence the organization of spatial representations on the level of single neurons. Specialized cells fire whenever a rat is close to an enclosing wall (Solstad, Boccara, Kropff, Moser, & Moser, 2008) and opaque borders strongly influence the firing patterns of hippocampal place cells (O'Keefe & Burgess, 1996) as well as entorhinal grid cells (Stensola, Stensola, Moser, & Moser, 2015). A place cell fires at – and therefore identifies – a specific location within an environment (e.g., the south-west corner of a room). In a single room the same cell will show reactivation (in addition to base rate activity) if the same location is visited again. Importantly, across multiple interconnected spaces (i.e., within ES) often cells are not firing at a unique location only. Rather, the same cell may fire again (is reused) within different vista spaces (Grieves, Jenkins, Harland, Wood, & Dudchenko, 2016; Skaggs & McNaughton, 1998; Spiers, Hayman, Jovalekic, Marozzi, & Jeffery, 2015). Transferred to the present experimentation, a single place cell might fire in multiple corridors, but it will not do so at corre-

sponding locations within a single VS room. Not only place cells, but also grid cells are sensitive to compartmentalization along opaque VS borders (Carpenter, Manson, Jeffery, Burgess, & Barry, 2015). An entorhinal grid cell fires at repeated locations arranged along a regular grid covering the whole space. Interestingly, rats were found to use different grids for different corridors (ES), but a single grid pattern when walking similar trajectories within a single VS (Derdikman et al., 2009). These findings indicate that compartmentalization of ES into multiple VS along opaque borders is also reflected in the neuronal response.

When navigating towards a goal location, hippocampal place cells are activated consecutively along the route to that goal, even before physical movement (Pfeiffer & Foster, 2013). This successive activation has been proposed to correspond to mind (i.e., non-physical) travel – or mental walk – along a route (Byrne et al., 2007; Sanders et al., 2015). Path integration along mind travel may then be used to estimate a vector towards the goal. Indeed, such survey estimations were associated with hippocampal activity in humans before (Wolbers & Büchel, 2005). One specific prediction for path integration via mind travel is that longer distances towards a goal will result in more place cell activity and therefore larger overall hippocampal activation. Indeed changes in blood flow associated with higher summed activity at longer paths to a goal location was observed in humans as well, while watching a video of a travel through a familiar city part (Howard et al., 2014) and when sequentially presenting pictures of close-by and distant city locations (Morgan, Macevoy, Aguirre, & Epstein, 2011). It should, however, be noted that this increase of hippocampal activity can also be explained by an alternative process, namely, by mentally adding blocks of vista spaces to form a mental model of the non-visible surrounding (Meilinger, 2008). Importantly, this positive correlation of distance and hippocampal activity reversed when distances were introduced within a VS, i.e., a virtual room (Viard, Doeller, Hartley, Bird, & Burgess, 2011), or an endless plane (Sherrill et al., 2013). Thus, the human hippocampus presumably has a share in both spaces, but the processing involved differs qualitatively. In summary, the VS-ES distinction brought forward in the current study corresponds to some recent distinctions obtained from single cell activity in rodents as well as summed activity within humans. We do think that future experimentation along these lines will be fruitful.

Consistent with the literature our results show that memory for VS and ES differ due to visual borders. But how is that memory organized? A VS clearly seems to be treated as a unique unit. As in other studies where learning took place in a single room (e.g., Mou & McNamara, 2002; Shelton & McNamara, 2001) our results suggest a common reference frame for all locations in the VS conditions. Close-by and distant pointing targets were processed equally fast indicating similar access from within a common memory unit. While we do find reminiscence of the experienced order, this order effect still differs from ES learning and might be based on a memory system independent of the organization of spatial relations (Packard & McGaugh, 1996; Restle, 1957; van der Ham et al., 2010). Findings showing that place and grid cells in rats do not remap within a single constant VS (Derdikman et al., 2009; Skaggs & McNaughton, 1998) do further support the assumption of a single VS unit in memory. Similarly, this accounts for results demonstrating an advantage of mentally switching between object locations within a VS (Brockmole & Wang, 2002, 2003) and the preferred updating of object locations within a VS as compared to locations in neighbouring spaces (Avraamides & Kelly, 2010; Kelly et al., 2007; Wang & Brockmole, 2003a, 2003b).

Contrary to VS, data from our and other studies suggest that ES memory is split into multiple units. Each unit, in our case, each individual corridor of the ES is assumed to operate as a VS. Thus, conclusions about underlying learning mechanisms drawn from

the VS room should likewise be effective in a single corridor. In line with this, reference frame orientation followed the immediate visible input, both in a VS room as well as in an individual ES corridor. The observed distance effect in pointing latency suggests that memory access is fastest within the pre-activated memory unit, i.e., within the corridor one is currently located in. Beyond the visible unit the distance effect indicates a process of successive activation corridor per corridor, unit by unit, not a onetime recall of a single unit. Similar to switching costs, that are interpreted as an effortful retrieval of a new memory unit (Brockmole & Wang, 2002, 2003), latency increase with traveled distance can be interpreted as a successive activation (and integration) of the individually represented VS. This might be reflected in higher hippocampal activity with increasing path distance to the target location (Howard et al., 2014; Morgan et al., 2011). Our order effect in ES, which is exceeding the effect found under conditions of movement and successive presentation in VS, is also consistent with a structure of multiple connected subspaces which are accessed in the order of connection. Such a structure has already been proposed in the literature (Chrastil & Warren, 2014; Mallot & Basten, 2009; Meilinger, 2008; Trullier, Wiener, Berthoz, & Meyer, 1997). Updating of object locations across subspaces might not naturally and easily emerge, as would be expected from single VS spatial unit. This is exactly what prior results showed (Avraamides & Kelly, 2010; Kelly et al., 2007; Wang & Brockmole, 2003a, 2003b). In sum, results from the present study as well as from the literature are consistent with the idea that ES are represented within multiple spatial units.

The representation of subspaces may rely on distinct reference frames, and the orientation of each reference frame might depend on the spatial cues available in each vista unit (Meilinger et al., 2014; Werner & Schmidt, 1999). These units might be further embedded within a hierarchical structure (Mallot & Basten, 2009; McNamara et al., 2008) with a common top-level reference frame encompassing multiple subunits. Such a top level reference frame might play a stronger role for individuals with high spatial abilities (e.g., Meilinger et al., 2014) or when familiarity with an environment increases. Importantly, extending the assumption of a single, common reference unit (e.g., Gallistel, 1990; O'Keefe & Nadel, 1978; Sholl, 2001) to ES without postulating a hierarchy with sublevels needs additional specification of the processes that generated the observed ES-specific distance and order effects and an explanation why these processes were not evoked in a VS.

The main conclusion from the present study is, that memory for VS and ES is structurally different – even if the same spatial information was learned. Accessing memory for ES was constrained by the distance and order in which objects were learned. We demonstrated that these effects cannot be fully explained by movement through the environment and successive object presentation, rather spatial separation is needed for that. The visible geometry of corridor and room determined the reference frame orientation in memory, and also likely the units ES memory was subdivided into. These results ultimately emphasize that transferring conclusions of findings obtained in VS studies to the more complex learning of ES (and vice versa) should be made cautiously.

Acknowledgements

Appendix A. Supplementary material

All data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cognition.2016.06.003.

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://dx.doi.org/10.1037/a0017040.

Brockmole, J. R., & Wang, R. F. (2002). Switching between environmental representations in memory. *Cognition, 83*(3), 295–316. http://dx.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://dx.doi.org/10.1016/S0.

Byrne, P., Becker, S., & Burgess, N. (2007). Remembering the past and imagining the future: A neural model of spatial memory and imagery. *Psychological Review, 114*(2), 340–375. http://dx.doi.org/10.1037/0033-295X.114.2.340.

Carpenter, F., Manson, D., Jeffery, K., Burgess, N., & Barry, C. (2015). Grid cells form a global representation of connected environments. *Current Biology, 25*(9), 1176–1182. http://dx.doi.org/10.1016/j.cub.2015.02.037.

Chrastil, E. R., & Warren, W. H. (2013). Active and passive spatial learning in human navigation: Acquisition of survey knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(5), 1520–1537. http://dx.doi.org/10.1037/a0032382.

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

Cohen, R., Baldwin, L. M., & Sherman, R. C. (1978). Cognitive maps of a naturalistic setting. *Child Development, 49*(4), 1216–1218. http://dx.doi.org/10.2307/1128763.

Coluccia, E., & Louse, G. (2004). Gender differences in spatial orientation: A review. *Journal of Environmental Psychology, 24*(3), 329–340. http://dx.doi.org/10.1016/j.jenvp.2004.08.006.

Derdikman, D., Whitlock, J. R., Tsao, A., Fyhn, M., Hafting, T., Moser, M.-B., & Moser, E. I. (2009). Fragmentation of grid cell maps in a multicompartment environment. *Nature Neuroscience, 12*(10), 1325–1332. http://dx.doi.org/10.1038/nn.2396.

Fujita, N., Klatzky, R. L., Loomis, J. M., & Golledge, R. G. (1993). The encoding-error model of pathway completion without vision. *Geographical Analysis, 25*(4), 295–314. http://dx.doi.org/10.1111/j.1538-4632.1993.tb00300.x.

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

Gilchrist, I. D., & Harvey, M. (2006). Evidence for a systematic component within scan paths in visual search. *Visual Cognition, 14*(4–8), 704–715. http://dx.doi.org/10.1080/13506280500193719.

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), 938–957. http://dx.doi.org/10.1037/a0019647.

Grieves, R. M., Jenkins, B. W., Harland, B. C., Wood, E. R., & Dudchenko, P. A. (2016). Place field repetition and spatial learning in a multicompartment environment. *Hippocampus, 26*(1), 118–134. http://dx.doi.org/10.1002/hipo.22496.

Hardiess, G., Gillner, S., & Mallot, H. A. (2008). Head and eye movements and the role of memory limitations in a visual search paradigm. *Journal of Vision, 8*(1), 7.1–13. http://dx.doi.org/10.1167/8.1.7.

Howard, L. R., Javadi, A. H., Yu, Y., Mill, R. D., Morrison, L. C., Knight, R., ... Spiers, H. J. (2014). The hippocampus and entorhinal cortex encode the path and Euclidean distances to goals during navigation. *Current Biology, 24*(12), 1331–1340. http://dx.doi.org/10.1016/j.cub.2014.05.001.

Janzen, G. (2006). Memory for object location and route direction in virtual large-scale space. *The Quarterly Journal of Experimental Psychology, 59*(3), 493–508. http://dx.doi.org/10.1080/02724980443000746.

Kelly, J. W., Avraamides, M. N., & Loomis, J. M. (2007). Sensorimotor alignment effects in the learning environment and in novel environments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(6), 1092–1107. http://dx.doi.org/10.1037/0278-7393.33.6.1092.

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://dx.doi.org/10.3758/PBR.15.2.322.

Klatzky, R. L., Loomis, J. M., Golledge, R. G., Cicinelli, J. G., Doherty, S., & Pellegrino, J. W. (1990). Acquisition of route and survey knowledge in the absence of vision. *Journal of Motor Behavior, 22*(1), 19–43. http://dx.doi.org/10.1080/00222895.1990.10735500.

Kosslyn, S. M., Ball, T. M., & Reiser, B. J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance, 4*(1), 47–60. http://dx.doi.org/10.1037/0096-1523.4.1.47.

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

Levin, J. R., & Neumann, E. (1999). Testing for predicted patterns: When interest in the whole is greater than in some of its parts. *Psychological Methods, 4*(1), 44–57. http://dx.doi.org/10.1037/1082-989X.4.1.44.

Loomis, J. M., Klatzky, R. L., Golledge, R. G., Cicinelli, J. G., Pellegrino, J. W., & Fry, P. a. (1993). Nonvisual navigation by blind and sighted: Assessment of path integration ability. *Journal of Experimental Psychology: General, 122*(1), 73–91. http://dx.doi.org/10.1037/0096-3445.122.1.73.

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

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

McNamara, T. P. (2003). How are the locations of objects in the environment represented in memory? In *Spatial cognition III* (pp. 174–191). Berlin, Heidelberg: Springer Berlin Heidelberg. http://dx.doi.org/10.1007/3-540-45004-1_11.

McNamara, T. P., Sluzenski, J., & Rump, M. (2008). Human spatial memory and navigation. In H. L. Roediger, III (Ed.). *Cognitive psychology of memory. Learning and memory: A comprehensive reference* (Vol. 2, pp. 157–178). Oxford: Elsevier.

Meilinger, T. (2008). The network of reference frames theory: A synthesis of graphs and cognitive maps. *Spatial cognition VI. Learning, reasoning, and talking about space* (Vol. 5248, pp. 344–360). Berlin, Heidelberg: Springer. http://dx.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://dx.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://dx.doi.org/10.1007/s00426-014-0629-6.

Meilinger, T., Knauff, M., & Bülthoff, H. (2008). Working memory in wayfinding—A dual task experiment in a virtual city. *Cognitive Science: A Multidisciplinary Journal, 32*(4), 755–770. http://dx.doi.org/10.1080/03640210802067004.

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), 542–569. http://dx.doi.org/10.1080/17470218.2013.821145.

Moar, I., & Carleton, L. R. (1982). Memory for routes. *The Quarterly Journal of Experimental Psychology Section A, 34*(3), 381–394. http://dx.doi.org/10.1080/14640748208400850.

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

Morgan, L. K., Macevoy, S. P., Aguirre, G. K., & Epstein, R. A. (2011). Distances between real-world locations are represented in the human hippocampus. *The Journal of Neuroscience, 31*(4), 1238–1245. http://dx.doi.org/10.1523/JNEUROSCI.4667-10.2011.

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://dx.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://dx.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://dx.doi.org/10.1016/0022-0965(82)90030-3.

O'Keefe, J., & Burgess, N. (1996). Geometric determinants of the place fields of hippocampal neurons. *Nature, 382*(6581), 425–428.

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

Packard, M. G., & McGaugh, J. L. (1996). Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. *Neurobiology of Learning and Memory, 65*(1), 65–72. http://dx.doi.org/10.1006/nlme.1996.0007.

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://dx.doi.org/10.1080/20445911.2015.1102144.

Pfeiffer, B. E., & Foster, D. J. (2013). Hippocampal place-cell sequences depict future paths to remembered goals. *Nature, 497*(7447), 1–8. http://dx.doi.org/10.1038/nature12112.

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://dx.doi.org/10.1037/0096-3445.134.1.73.

Restle, F. (1957). Discrimination of cues in mazes: A resolution of the "place-vs.-response" question. *Psychological Review, 64*(4), 217–228. http://dx.doi.org/10.1037/h0040678.

Sanders, H., Rennó-Costa, C., Idiart, M., & Lisman, J. E. (2015). Grid cells and place cells: An integrated view of their navigational/memory function. *Trends in Neurosciences, 38*(12), 763–775. http://dx.doi.org/10.1016/j.tins.2015.10.004.

Schweizer, K., Herrmann, T., Janzen, G., & Katz, S. (1998). The route direction effect and its constraints. *Spatial cognition* (Vol. 15, pp. 19–38). . http://dx.doi.org/10.1007/3-540-69342-4_2.

Shelton, A. L., & McNamara, T. P. (2001). Systems of spatial reference in human memory. *Cognitive Psychology, 43*(4), 274–310. http://dx.doi.org/10.1006/cogp.2001.0758.

Sherrill, K. R., Erdem, U. M., Ross, R. S., Brown, T. I., Hasselmo, M. E., & Stern, C. E. (2013). Hippocampus and retrosplenial cortex combine path integration signals for successful navigation. *Journal of Neuroscience, 33*(49), 19304–19313. http://dx.doi.org/10.1523/JNEUROSCI.1825-13.2013.

Sholl, M. J. (2001). The role of a self-reference system in spatial navigation. In D. R. Montello (Ed.), *Cosit 2001* (pp. 217–232). Berlin, Heidelberg: Springer. http://dx.doi.org/10.1007/3-540-45424-1_15.

Siegel, A. W., & White, S. H. (1975). The development of spatial representations of large-scale environments. *Advances in Child Development and Behavior, 10*, 9–55. http://dx.doi.org/10.1016/S0065-2407(08)60007-5.

Skaggs, W. E., & McNaughton, B. L. (1998). Spatial firing properties of hippocampal CA1 populations in an environment containing two visually identical regions. *The Journal of Neuroscience, 18*(20), 8455–8466.

Solstad, T., Boccara, C. N., Kropff, E., Moser, M.-B., & Moser, E. I. (2008). Representation of geometric borders in the entorhinal cortex. *Science, 322* (5909), 1865–1868. http://dx.doi.org/10.1126/science.1166466.

Spiers, H. J., & Barry, C. (2015). Neural systems supporting navigation. *Current Opinion in Behavioral Sciences, 1*, 47–55. http://dx.doi.org/10.1016/j.cobeha.2014.08.005.

Spiers, H. J., Hayman, R. M. a., Jovalekic, A., Marozzi, E., & Jeffery, K. J. (2015). Place field repetition and purely local remapping in a multicompartment environment. *Cerebral Cortex, 25*(1), 10–25. http://dx.doi.org/10.1093/cercor/bht198.

Stensola, T., Stensola, H., Moser, M.-B., & Moser, E. I. (2015). Shearing-induced asymmetry in entorhinal grid cells. *Nature, 518*(7538), 207–212. http://dx.doi.org/10.1038/nature14151.

Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist, 57*(4), 421–457.

Street, W. N., & Wang, R. F. (2014). Differentiating spatial memory from spatial transformations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(2), 602–608. http://dx.doi.org/10.1037/a0035279.

Strickrodt, M., O'Malley, M., & Wiener, J. M. (2015). This place looks familiar—How navigators distinguish places with ambiguous landmark objects when learning novel routes. *Frontiers in Psychology, 6*(1936), 1–12. http://dx.doi.org/10.3389/fpsyg.2015.01936.

Thorndyke, P. W., & Hayes-Roth, B. (1982). Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology, 14*(4), 560–589. http://dx.doi.org/10.1016/0010-0285(82)90019-6.

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://dx.doi.org/10.1080/17470218.2011.586710.

Trullier, O., Wiener, S. I., Berthoz, A., & Meyer, J. A. (1997). Biologically based artificial navigation systems: Review and prospects. *Progress in Neurobiology, 51* (5), 483–544. http://dx.doi.org/10.1016/S0301-0082(96)00060-3.

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

van der Ham, I. J. M., van Zandvoort, M. J. E., Meilinger, T., Bosch, S. E., Kant, N., & Postma, A. (2010). Spatial and temporal aspects of navigation in two neurological patients. *NeuroReport, 21*(10), 685–689. http://dx.doi.org/10.1097/WNR.0b013e32833aea78.

Viard, A., Doeller, C. F., Hartley, T., Bird, C. M., & Burgess, N. (2011). Anterior hippocampus and goal-directed spatial decision making. *Journal of Neuroscience, 31*(12), 4613–4621. http://dx.doi.org/10.1523/JNEUROSCI.4640-10.2011.

Waller, D., Loomis, J. M., & Haun, D. B. M. (2004). Body-based senses enhance knowledge of directions in large-scale environments. *Psychonomic Bulletin & Review, 11*(1), 157–163. http://dx.doi.org/10.3758/BF03206476.

Wan, X., Wang, R. F., & Crowell, J. a. (2013). Effects of basic path properties on human path integration. *Spatial Cognition & Computation, 13*(1), 79–101. http://dx.doi.org/10.1080/13875868.2012.678521.

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

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

Wen, W., Ishikawa, T., & Sato, T. (2011). Working memory in spatial knowledge acquisition: Differences in encoding processes and sense of direction. *Applied Cognitive Psychology, 25*(4), 654–662. http://dx.doi.org/10.1002/acp.1737.

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

Wiener, J. M., Kmecova, H., & de Condappa, O. (2012). Route repetition and route retracing: Effects of cognitive aging. *Frontiers in Aging Neuroscience, 4*, 1–7. http://dx.doi.org/10.3389/fnagi.2012.00007.

Wiener, J. M., & Mallot, H. A. (2006). Path complexity does not impair visual path integration. *Spatial Cognition & Computation, 6*(4), 333–346. http://dx.doi.org/10.1207/s15427633scc0604_3.

Wilson, P. N., Wilson, D. a., Griffiths, L., & Fox, S. (2007). First-perspective spatial alignment effects from real-world exploration. *Memory & Cognition, 35*(6), 1432–1444. http://dx.doi.org/10.3758/BF03193613.

Wolbers, T., & Büchel, C. (2005). Dissociable retrosplenial and hippocampal contributions to successful formation of survey representations. *Journal of Neuroscience, 25*(13), 3333–3340. http://dx.doi.org/10.1523/JNEUROSCI.4705-04.2005.

Wolbers, T., & Wiener, J. M. (2014). Challenges for identifying the neural mechanisms that support spatial navigation: The impact of spatial scale. *Frontiers in Human Neuroscience, 8*, 1–12. http://dx.doi.org/10.3389/fnhum.2014.00571.

Yamamoto, N., & Shelton, A. (2009). Sequential versus simultaneous viewing of an environment: Effects of focal attention to individual object locations on visual spatial learning. *Visual Cognition, 17*(4), 457–483. http://dx.doi.org/10.1080/13506280701653644.

Yamamoto, N., & Shelton, A. L. (2005). Visual and proprioceptive representations in spatial memory. *Memory & Cognition, 33*(1), 140–150. http://dx.doi.org/10.3758/BF03195304.

Yamamoto, N., & Shelton, A. L. (2007). Path information effects in visual and proprioceptive spatial learning. *Acta Psychologica, 125*(3), 346–360. http://dx.doi.org/10.1016/j.actpsy.2006.09.001.