

Available online at www.sciencedirect.com



Cognitive Brain Research 19 (2004) 244-258



www.elsevier.com/locate/cogbrainres

Navigating in a virtual three-dimensional maze: how do egocentric and allocentric reference frames interact?

Manuel Vidal^{a,*}, Michel-Ange Amorim^b, Alain Berthoz^a

^aLaboratoire de Physiologie de la Perception et de l'Action, CNRS/Collège de France, 11 place Marcelin Berthelot, 75005 Paris, France ^bResearch Center in Sport Sciences, Université Paris Sud, 91405 Orsay Cedex, France

Accepted 16 December 2003

Abstract

Spatial navigation in the presence of gravity restricts one's displacement to two-dimensional (2D) planes. Therefore, self-motion only includes translations and yaw rotations. In contrast, in weightlessness, one can translate and turn in any direction. In the first experiment, we compared the ability to memorize a virtual three-dimensional (3D) maze after passive exploration in three self-motion conditions, each using a different set of rotations for turning. Subjects indicated which pathway they traversed among four successive corridors presented from an outside perspective. Results showed that exploring in the *terrestrial* condition (including only yaw rotations, the viewer's virtual body remaining upright) allowed better recognition of the corridor than in the *weightless* condition (which included pitch and yaw rotations according to the turns), particularly for more complex 3D structures. The more frequently the viewer-defined (egocentric) and the global environment (allocentric) verticals were aligned during exploration, the more easily subjects could memorize the 3D maze, suggesting that simplifying the relationship between the egocentric and allocentric reference frames facilitates spatial updating. Nevertheless, with practice, performance in the *weightless* condition improved whereas in the natural *terrestrial* condition performance remained at its initial maximum, indicating that the cognitive processes involved were innate for this particular condition. The second experiment revealed that single rotations in the *terrestrial* condition must be performed around the body axis in order to obtain optimal spatial updating performance, and that the latter is independent of the conflict with gravity that might favor this condition when one is actually upright. This suggests that although humans can memorize 3D-structured environments their innate neurocognitive functions appear to be specialized for natural 2D navigation.

Theme: Neural basis of behavior *Topic:* Learning and memory, systems and functions

Keywords: Spatial memory; Reference frames; Human; 3D maze; Virtual reality

1. Introduction

Human navigation relying on spatial knowledge requires the continuous processing of spatial information in order to update this knowledge and execute the planned trajectory. Spatial updating is performed through the integration of one's displacements and through the recognition of environmental landmarks along the way. The former depends principally on the extraction of heading information from optic and acoustic flow [21] and the integration of selfmotion information such as speed and acceleration provided by the vestibular system, proprioception and vision in a process called "path integration" [7-11,14]. It is this latter process that feeds the memorizing process of one's trajectory while exploring an environment, whether known or novel. During walking, terrestrial gravity restricts human displacements to two-dimensional (2D) planes, and the head is most of the time stabilized, in order to keep it continuously upright relative to gravity [16]. Although humans process vertical information (elevation) about their environment, either for altitude variations, as in a town for example [3], or for navigation inside buildings with several floors [15], it has been found that they are not as precise in such processing as they are for horizontal information (azimuth). Astronauts frequently report being disoriented during space flight, especially when they have to go to a specific sector of the space station, or when they have to retrieve a tool they had placed somewhere nearby [6]. Because trajectories in

^{*} Corresponding author. Tel.: +33-1-44-27-14-07; fax: +33-1-44-27-13-82.

E-mail address: manuel.vidal@college-de-france.fr (M. Vidal).

microgravity are no longer restricted to 2D planes, body translations and rotations are possible in any direction of space. Therefore, in flight, astronauts' self-motion can include yaw rotations as in terrestrial navigation, but also pitch and roll rotations. Nevertheless, astronauts tend to avoid adopting unusual body orientations relative to their visual environment [18]. Furthermore, because gravitational cues are suppressed, they can no longer serve as a stable external reference frame for spatial orientation.

In the present study, we tested the ability of human subjects to recognize the geometrical shape of a threedimensional (3D) virtual maze after passive egocentric selfmotion according to different displacement conditions. Based on the abovesummarized literature on navigation, we hypothesized that humans would have difficulty in solving complex spatial problems while navigating inside 3D structures. Two main issues regarding 3D navigation were addressed in this paper. The first issue concerns the capacity to store the 3D structure of an environment: can humans build a mental representation of complex 3D environments where all dimensions have the same probability of occurrence? This question does not concern most natural situations of navigation where processing of vertical information is irrelevant, but rather concerns navigation inside buildings with several levels or inside space stations. The second issue concerns the human capacity for integrating self-motion that includes both yaw and pitch body rotations. In such cases, memorizing the shape of the environment may require a complex coordinate

transformation in order to shift from the egocentric (local) reference frame experienced during navigation to an allocentric (global) reference frame so as to build an object-like representation of the environment. Intuitively, we felt that vertical could be at the core of the allocentric reference frame. Our second question was therefore: what is the effect of tilting the observer's egocentric vertical relative to the environment's allocentric vertical, which occurs in a 3D displacement, on the process of memorizing the trajectory?

Natural human self-motion includes two characteristics that are direct consequences of head stabilization during locomotion and might be helpful to simplify spatial orientation. First, shifting from the egocentric to an allocentric reference frame involves rotations about the body axis only. Second, gravity provides a constant reference that can be used to infer the orientation of the head when it is tilted, allowing the retrieval of spatial information independently of temporary body tilts. For these reasons, we were interested in manipulating the relationship between the egocentric and the allocentric reference frames during navigation, for the memorization of the traveled 3D environment. We designed a first displacement condition called the terrestrial condition, inspired by natural self-motion, in which the observer's vertical orientation is kept constant throughout the exploration, and where going up and down simulates the elevator of a building (see Fig. 1). Based on previous observations, we expected this type of displacement to provide an optimal spatial input for building a mental model of the traveled 3D environment.



Fig. 1. Terrestrial navigation in buildings: spatial inferences are more difficult between levels than within levels [15]. A different cognitive map can be built for each level, with connections to the other levels.

In contrast, we expected that varying the alignment of both the observer's vertical and the environment's vertical usually provided by gravity would disturb the construction of such a mental representation. In order to test this hypothesis, we designed two other displacement conditions. The first preserved this alignment along the horizontal sections of the navigation path only. This displacement mode was called the subaquatic condition by analogy with scuba divers' body orientations during exploration. In the second, called the weightless condition, the observer's vertical could be aligned with any direction, by analogy with the microgravity condition of space flights, where astronauts can adopt any orientation during navigation. We expected that the latter displacement mode would most severely impair the creation of a mental model of the traveled 3D structure. Virtual reality allowed us to visually simulate participants' self-motion in the different displacement modes.

Rodents in their natural environments are faced with different 3D spatial problems than those usually encountered by humans, like digging a tunnel to reach a target at a specific position not directly accessible by simple 2D navigation, or orienting themselves in a big city's sewer network. Furthermore, they can walk on vertical planes, and in such cases do not have to stabilize their heads. For these animals, integrating displacements not restricted to 2D planes and representing 3D routes in memory are cognitive aptitudes necessary for survival. Furthermore, when rats move in 3D mazes, priority is given to processing the vertical dimension rather than the horizontal [5]. These differences of 3D navigation capacities between animal species may stem from differences in the evolutionary pressures on rodents and humans, each species developing cognitive functions adapted to its natural environment. It is of interest to discover whether or not the human cognitive functions involved during natural navigation are also well adapted to 3D navigation. From an evolutionary perspective, we might expect them to be more specifically adapted to 2D navigation, and less appropriate for 3D exploration than in the case of rodents. On the other hand, it is possible that humans, with practice, may learn how to manage complex displacements including yaw and pitch turns (the weightless condition in our experiments). If memorizing a 3D maze with such an unusual displacement condition then becomes possible, we could hypothesize that instead of being innate, as would appear to be the case in rodents, such a cognitive capacity is rather a learned capacity. Accordingly, the spatial learning performance in complex displacements would never attain the same level as in a natural displacement (the terrestrial condition in our experiments).

2. Experiment 1

2.1. Materials and methods

2.1.1. Subjects

Sixteen naive subjects (12 men and 4 women) aged from 20 to 32 years participated in this experiment. Most were students or laboratory staff, and all but one were right-handed. They all gave prior written consent.

2.1.2. Experimental setup

Subjects sat on a chair of adjustable height allowing the line of sight to be centered on a large screen covering a 115°



Fig. 2. A view of the experimental setup. The subject's line of sight was centered on a 115° FOV video projector screen (240 × 180 cm). A PC computer (equipped with a Diamond's FireGL 1 video card) generated the video and the sound for the verbal instructions and recorded the responses entered on the keyboard. The subject's eyes were positioned at a distance of 80 cm from the screen.

horizontal and 100° vertical field of view (FOV) at a distance of 80 cm. The stimuli were projected onto the screen at a resolution of 1024×768 (see Fig. 2). In darkness, subjects responded using a keyboard with keys highlighted by phosphorescent stickers.

2.1.3. Procedure

Each of the 36 trials in the experiment included a visual navigation phase followed by a test phase. During the

navigation phase, subjects were passively driven at constant speed through a virtual cylindrical 3D corridor with stone walls (see Fig 3a). Three different displacement conditions were compared using the same set of 12 different corridors, four with three segments, four with four segments and four with five segments. Four trials were then performed for each level of the independent variable (navigation condition × number of segments). Each segment was aligned with one of the canonical axes defining the allocentric



Fig. 3. (a) An inside static view of one corridor explored by subjects (resolution of 1024×768 at 15 fps). Perspective correction was adjusted to the real FOV experienced by subjects. (b) An outside view of the corridor as seen during the recognition task, the red arrow indicating the point of entry.

reference frame (see Fig. 4d). In order to avoid memorization of a verbal sequence of the directions taken in corridors, subjects performed a dual task consisting of verbal shadowing. They had to repeat out loud random numbers ranging from 10 to 60 that were played through headphones every 2.5 s.

Just after the navigation, subjects were first asked to draw with their finger in 3D space the remembered shape of the corridor (manual reproduction task), and to press a key once they had finished. This first task was used as priming for the following one; therefore, only the reaction time was measured. The second task was to select from among four external views the corridor structure that corresponded to the one previously explored (recognition task). A 'yes' or 'no' key press was recorded for each view presented successively. The principles underlying the construction of the distractors associated with each corridor are given in Appendices A and B (see Table 1). External views were aligned with respect to the observer's orientation in the first segment, which was the same across all displacement conditions. The entrance point to the tunnel was indicated with a red arrow (see Fig. 3b). Each trial lasted approximately 70 s (40-s visual stimulus, 5-s manual reproduction test, 12-s recognition test, 15-s rest).

Both the order of the 36 trials (each corresponding to one of the three displacement conditions in one of the 12



Fig. 4. The orientation of the subjects in the virtual corridor for the (a) *terrestrial*, (b) *subaquatic* and (c) *weightless* conditions. The *z*-axis, *x*-axis and *y*-axis represent the body axis direction, the line of sight, and the left-hand direction, respectively. (d) The allocentric reference frame (X,Y,Z) defined by the egocentric reference frame (x,y,z) at the initial position.

corridors of the set), and the sequential presentation order of the correct external view and of the distractors during the recognition task were randomized. Subjects were never told which condition they were going to be tested in. After each block of 12 trials, the number of correct responses was displayed to the subjects as a score, and was followed by a 5-min pause. This feedback was given in order to keep subjects motivated during the whole experiment. It was not given after each trial, to prevent them from developing unwanted simplifying cognitive strategies. Subjects triggered each trial by pressing a specific key when ready before each exploration. For each trial, the response latency for the manual reproduction, the reaction time for each of the four views of the recognition task and the accuracy of the choice were recorded. Given that the test views were presented successively, the last 'yes' choice among the four views was considered as the subject's definitive choice. This allowed subjects to cancel a former 'yes' that they considered a mistake after seeing a second view that seemed to be the correct one. The experiment was preceded by three practice trials for each of the different visual displacement conditions for subjects to familiarize themselves with the computer interface. The full experiment lasted approximately 1 h.

2.1.4. Displacement conditions

Three visual displacement conditions were studied. The allocentric reference frame corresponding to the initial egocentric reference frame determined the vertical and horizontal references used to describe these displacement modes (see Fig. 4d). In the terrestrial condition (see Fig. 4a), the head was always kept upright, and in vertical segments, the walls scrolled up or down in front of the subject as if inside a transparent elevator. In this condition, before entering a vertical segment, a yaw rotation was made to orient the view in the direction the path followed after the end of the segment. This information was given before vertical translations to permit subjects to know at the same time for all conditions which direction was coming next. In the weightless condition (see Fig. 4c), the viewing direction pointed towards the end of the current segment and at each junction a single yaw or pitch egocentric rotation was performed to reorient the line of sight with the next segment, thereby allowing subjects to experience rotations around all three axes of the allocentric space. The *subaquatic* condition (see Fig. 4b) was similar to the weightless condition except that a second roll rotation could be simultaneously added at turns following a vertical segment, in order to reposition the head upright (as defined by the initial viewing orientation in the first segment). The name of these conditions were inspired by the kind of self-motion one can have in terrestrial, subaquatic and weightless environments. Knowing that in our study only visual motion was simulated, these conditions cannot correspond to a real motion in such environment. Therefore, the naming convention is only a partial analogy with reality.

During displacement through the virtual corridor, the simulated gaze direction rotated in anticipation of the curve, as would occur in natural conditions [4,20]. That is, the virtual viewing direction started rotating 2600 ms before the translation of the viewpoint started to curve. This anticipation delay was estimated empirically; we tested different delays and chose the one that was the most natural and comfortable. Linear speed was kept constant during the whole displacement and was the same for each trial.

Because there were no absolute cues in the virtual visual scene, the actual translation velocity is undefined from the visual flow field alone. However, one can estimate the equivalent displacement velocity supposing that the subject walked on the floor of the tunnel. Consequently, for a subject measuring 1.75 m in height, the virtual speed would be 1.31 m/s (around 4.7 km/h), which corresponds to a normal walking speed for humans.

2.1.5. Data analysis

Analysis of variance (ANOVA) was performed on the different dependent variables (accuracy, latency of correct response and manual reproduction) with displacement condition \times number of segments as within-subject factors. Specific hypothesized effects were tested using contrasted comparisons, and post hoc analyses were performed with the Scheffé test.

2.2. Results

2.2.1. 3D recognition performance

Although subjects reported having difficulty in performing the spatial task at the beginning of the experiment, they said it became easier after the first 12 trials. In a postexperiment debriefing, they all mentioned that the weightless condition was more difficult than the terrestrial condition, but that it was as difficult as the subaquatic condition. The instructions did not inform the subjects why, in the subaquatic condition, a double rotation was sometimes performed simultaneously at turns (to reorient the body's vertical with that of the environment). However, some of the subjects reported that, somehow, they knew before reaching the corner if they were going to have a "strange rotation" or not. This suggests that they were naturally expecting to reorient their virtual body position upright after traveling in a vertical segment with a horizontal position. Because the weightless and the subaquatic conditions only differed when returning to a horizontal corridor segment, this expectation was fulfilled in the subaquatic condition but not in the weightless condition.

The average response accuracy for all subjects is presented in the clustered error bar chart in Fig. 5, for each displacement condition, and for corridors with three, four or five segments, or altogether. Because chance performance was 25%, the results indicate an overall high level of accuracy. The results indicate a significant main effect of the displacement condition on accuracy [F(2,30)=



Fig. 5. Recognition accuracy (means \pm S.E.) according to the displacement conditions and each number of segments, or across all of them. Dashed line indicates the response chance level (25%).

5.16; p < 0.012]. As expected, accuracy decreased proportionally with the number of segments in the corridors [F(2,30)=17.94; p < 0.001], with an average of 93.2%, 83.9% and 69.3% for three-, four- and five-segment trials, respectively.

Because the *weightless* and *subaquatic* conditions only differed on one three-segment corridor among the four used in this experiment, but differed for all four- and fivesegment corridors (see definition of corridors in Table 1 in Appendix A.), we decided to make specific analyses without considering the three-segment trials in order to properly compare performance in the three displacement conditions. We again found a significant main effect of the displacement condition on accuracy [F(2,30) = 10.57; p < 0.001], with an average of 82.8%, 78.9% and 68.0% for the *terrestrial*, *subaquatic* and *weightless* conditions, respectively. A contrasted comparison of the displacement conditions revealed that the *weightless* condition was statistically different from the *subaquatic* condition [F(1,15) = 9.30; p < 0.01] and from the *terrestrial* condition [F(1,15) = 27.21; p < 0.001].

The displacement condition × number of segment (only four and five) interaction with accuracy illustrated in Fig. 5 was significant [F(2,30) = 5.03; p < 0.02]. This interaction was due to the great difference in performance for the five-segment trials between the *weightless* condition and the other conditions; post hoc tests showed that this condition yielded significantly poorer results than the *terrestrial* (p < 0.01) and *subaquatic* (p < 0.02) conditions. Therefore, a clear deterioration of performance was observed in the *weightless* condition when the number of segments reached five, whereas in the other conditions, performance slowly decreased but the level of accuracy remained high.

The average response latency for hits, i.e., when subjects recognized the correct corridor among the views presented, is illustrated in Fig. 6. Misses were replaced by the average value of hit latencies yielded by the subject in the same displacement condition and number of corridor segments. There was a significant effect of the displacement condition on response latencies [F(2,30) = 3.38; p < 0.05]. Latencies for corridors with four and five segments were significantly shorter [F(1,15)=5.16; p < 0.04] for the *terrestrial* condition (2760 ms) than for the weightless and subaquatic conditions (3340 ms in each case). The increase in latency between three and five segments was clearly linear with a slope of about 1000 ms for each additional segment [F(2,30) = 21.38; p < 0.001]. The interaction between the number of segments and the displacement condition was not significant. Thus, the latency of recognition did not reflect the reduction in the accuracy of performance found in the weightless condition.

2.2.2. Learning effects

In order to check if there were distinct spatial learning trends depending on the exploration condition, we examined performance in each condition as a function of the trial order. Successive trials were grouped into subsets of four trials, which reduced the noise introduced by averaging across different numbers of values. The number of values corresponding to each average is indicated by the size of its dot in the plot. This provides additional information about the confidence level of the average.

The learning curves for recognition accuracy according to the navigation condition are detailed in Fig. 7. In the first four trials, the randomization of the order led to very little practice in the *terrestrial* as compared to the *subaquatic* condition; all



Fig. 6. Recognition latencies for correct responses (means ± S.E.) according to the displacement conditions and each number of segments, or across all of them.

subjects contributed in total to 15 *terrestrial* trials against 29 *subaquatic* trials, which explains the initially lower performance in the *terrestrial* condition. Subsequently, performance in this condition reached its learning peak (above 90%) and then stabilized, whereas performance in the *weightless* condition, and to some extent in the *subaquatic* condition, increased gradually over the duration of the experiment to reach approximately the same level as for the *terrestrial* condition. The starting level of performance for the *weightless* condition was below 60% in the first group of trials.

This descriptive analysis of the learning effect suggests that the natural aspect of the *terrestrial* condition required hardly any training to correctly process the spatial information and memorize the corridor. In contrast, because selfmotion with the *weightless* condition does not occur in everyday life, adaptation for this task required practice. Nevertheless, performance in the *weightless* condition clearly improved over time, reaching almost the same level as in the other conditions. Because of the random distribution of trials throughout the experiment, not the same subject individuals and number have contributed to each of the dots in the plot. Therefore, inferential analysis could not be performed to test the statistical significance of these tendencies, and only descriptive conclusions could be drawn.

2.2.3. Manual reproduction and the shadowing task

The mean latencies of manual reproduction were about the same for the three studied modes of navigation (about 4450 ms), and increased significantly [F(2,30) = 52.30;p < 0.001] with the number of corridor segments (from approximately 3000 ms for three segments to 5500 ms for five segments). Subjects said out loud numbers every 2.5 s; moreover, generating random numbers has a high cognitive cost. Thus, although we did not record the verbal responses in the dual task, it can be assumed that the shadowing task was correctly performed. Subjects reported that they paid particular attention to the numbers in the first 12-trial block, but after that they automated the task.

2.3. Discussion

Subjects readily identified the 3D outside view of the shape of the maze explored in the terrestrial condition. The good performance of subjects in this condition shows that it was possible to build a correct mental representation of the path in the corridors. In response to the first issue addressed in this study, this suggests that humans can, to some extent, build a representation of a complex 3D environment in working memory. This is particularly true for environments with segments of constant length and with right-angled turns, as tested in our experiment. If nonhorizontal segments had been oriented at angles other than 90°, the results would probably have been different. Concerning the second issue, as to whether or not humans can integrate self-motion that includes yaw as well as pitch body rotations, the answer is twofold. Overall, recognition accuracy in the weightless condition was considerably impaired when the number of segments of the maze reached five (falling from 83% to 53%). This suggests that the cognitive processes involved in this task for this particular condition were no longer effective (chance level being at 25%). Therefore, processing a 3D displacement that includes yaw and pitch rotations is more difficult (poorer precision and longer latencies) than for a natural 2D displacement such as in the *terrestrial* condition. Although it is possible to build a spatial representation from realistic 3D navigation when exploring a simple structure (with three or four segments), the cognitive functions involved in this task do not appear



Fig. 7. Learning curves of the recognition accuracy for each displacement condition with their respective standard deviation in the bottom.

to be adapted to more complex environments (with more segments). On the other hand, if we consider the training effect, the answer to this issue is different. In fact, we found that subjects' performance in the weightless condition continually improved from the beginning to the end of the experiment, which suggests that after some practice subjects got used to integrating the pitch and yaw rotations and could then memorize their trajectory correctly. Therefore, integrating complex 3D self-motion, as in the weightless condition, is not innate for humans, in contrast to integrating naturalistic self-motion, as in the terrestrial condition. Nevertheless, after exposure to this kind of complex 3D self-motion, it becomes possible to memorize the path traveled, although performance would not reach the same level as with a natural displacement. Indeed, the results from another study revealed that after intensive practice with both conditions in corridors with increased complexity (five and six segments), the performance level reached at the learning plateau of the weightless is still lower than that of the *terrestrial* condition [19].

The relationship between the egocentric and allocentric reference frames in each displacement condition provides a plausible explanation for the observed differences in spatial performance. As mentioned in the introduction, our spatial tasks required, at some point, a shift from an egocentric reference frame—in which subjects had the spatial experience—to an allocentric reference frame—in which the views of the corridors were presented during the recognition task. Indeed, updating the stored spatial information about the corridor requires subjects to extract after each turn their orientation relative to an allocentric reference frame, in order to be able to correctly infer the direction of the following segment. This mental process can be performed

either during the exploration, if subjects adopt the strategy of building a mental image of the corridor during the exploration, or during the recognition task. In the latter case, subjects could for instance adopt the strategy of storing only virtually generated exproprioceptive information during the exploration [13], and then sequentially evaluating, at each turn of the corridors presented, whether the allocentric direction change matches the memorized self-motion. These two modes of processing spatial information correspond to two general strategies that we can find in navigation when subjects are asked to continuously update an object's relative position while walking blindfolded [1].

One important distinction between the terrestrial condition and the other conditions in our study is that subjects only had to integrate yaw rotations to extract their orientation relative to the allocentric reference frame, whereas in the other two conditions, they also needed to integrate pitch and roll rotations. In typical terrestrial navigation, yaw is the only rotation angle one has to integrate in order to infer one's orientation and thus to remember the shape of a path. Besides, in the terrestrial condition, the egocentric reference frame had the particularity of maintaining the body's vertical axis (z) aligned with the allocentric vertical axis (Z) throughout the exploration of the maze. Therefore, shifting from an egocentric to an allocentric frame of reference, based on rotations along the vertical axis (the yaw rotations), would be at advantage. This is consistent with the results in the *terrestrial* condition. In the *subaauatic* condition, due to a double rotation when returning to a horizontal segment, the alignment of the vertical of both the egocentric (z) and allocentric (Z) reference frames was also present during navigation in horizontal segments (along X or Y), so the reference shift was partially facilitated. In contrast, the *weightless* condition showed the poorest performance because during the maze exploration the reference frame shift required rotations about all three axes of the allocentric space. Therefore, the results suggest that the complexity of the relationship between egocentric and allocentric reference frames affects the construction of a 3D spatial mental model. In fact, the spatial processes required by the task are more easily implemented with increasing alignment of verticals of the egocentric (*z*) and allocentric (*Z*) reference frames during navigation, whatever the strategy described previously.

Depending on the navigation conditions, the egocentric visual rotations for any given corridor could be different in number and in nature, which has implications both for the number of turns to integrate, and on the visuo-vestibular conflict introduced. Actually, because subjects remained seated upright during the simulated exploration, virtual body rotations were not sensed by the semicircular canals nor were the gravity orientation changes sensed by the otoliths. The terrestrial condition involved fewer rotations, and only yaw rotations that did not conflict with gravity. In contrast, the weightless and subaquatic conditions involved more rotations, and included pitch rotations that conflicted with gravity (see Fig. 4). This could have affected the integration of the displacement [2], and thus have been responsible for the better results in the terrestrial condition. However, recognition accuracy for corridors with five segments was nearly the same for the subaquatic condition (77%) and the terrestrial condition (78%), whereas it was considerably impaired for the weightless condition (53%). On the other hand, latencies were shorter for the terrestrial (3250 ms) than for the weightless and subaquatic conditions (approximately 3850 ms in each case). We can therefore assume that independent of the number and nature of the rotations involved in the *terrestrial* condition, the spatial updating process required in our task was more accurately performed for naturalistic displacement modes (terrestrial and to some extent subaquatic) than for displacements including yaw and pitch turns (weightless). Nevertheless, the processing time increased with the number of rotations during exploration of the environment and when pitch (or roll) rotations had to be integrated, leading to the increased reaction times observed in both the subaquatic and the weightless conditions.

3. Experiment 2

In order to validate some of the interpretations given in the previous discussion, and to look further into the particularity of the natural *terrestrial* displacement condition, we conducted an additional experiment that addressed two questions. The first question concerned the importance of gravity as an external reference used during navigation, as in the *terrestrial* condition of the previous experiment. As mentioned before, shifting from an egocentric to an allocentric reference frame in order to memorize the 3D path was easier in the *terrestrial* condition because the vertical axis was common to both frames of reference. In normal conditions, gravity provides the vertical axis of the allocentric reference frame used in navigation. Therefore, having an egocentric reference frame consistent with the gravitational vertical during navigation (an upright posture whether virtual or real) possibly facilitates the updating performance because it provides a common stable reference across the different perspectives encountered during navigation. We evaluated this influence on performance when participants performed the task in a nonupright position, thereby removing the possibility of using gravity in the integration process. We compared the effect of observers' actual orientation (upright vs. lying on the side) in two of the three virtual displacement conditions (terrestrial vs. weightless), bringing both conditions to the same level of conflict with regard to gravity during visual motion. Our prediction was that the difference observed between these conditions in the previous experiment would remain when subjects were tilted, despite the modified gravity orientation.

The second question concerned the contribution of having the rotation axis of the *terrestrial* condition aligned with the participant's body axis, which makes virtual orientation changes only through yaw rotations. We wanted to determine whether a *terrestrial*-like condition where orientation depended only on pitch rotations would still result in a better performance than the *weightless* condition in which orientation was a function of yaw and pitch rotations. We therefore added the *pitch terrestrial* condition, in which all simulated rotations along the pathway were performed around a single axis that was this time horizontal while subjects remained upright.

3.1. Materials and methods

3.1.1. Subjects

Twenty-six naive subjects (17 men and 9 women) aged from 19 to 33 years participated in this second experiment. Most of them were university students, and all but two were right-handed. They all gave prior written consent before starting and were remunerated.

3.1.2. Experimental setup

Subjects were either seated, as in the previous experiment, or lay on their right side with the keyboard positioned in the corresponding orientation. From the point of view of the subject, the trials were visually similar to the ones in the first experiment except that the stimuli had a higher resolution (1200×1200 pixels) and refresh rate (85 fps). In order to have comparable stimuli when subjects were seated and when they lay on their side, the vertical and horizontal fields of view were this time equal (107°).

3.1.3. Procedure

The procedure was similar to that of the previous experiment and only the differences and the reasons for them will be described. First, we removed the three-segment corridors from the protocol because they were too easy and only small differences in the results were observed across experimental conditions. We removed the *subaquatic* condition and included the three new conditions referred to above; the lying down *terrestrial* and *weightless* conditions and the *pitch terrestrial* condition.

We replaced the previous recognition task with a 3D reconstruction task. During this task, subjects were asked to draw with the computer the remembered 3D shape of the corridor. In this way, we eliminated the possible influence of choice of distractors in the recognition task. Subjects were first shown an external view of the first segment with an avatar at the entrance point indicating the orientation relative to which the reconstruction had to be made. It was aligned with the subject's body position, such that when subjects were in the upright position the avatar was vertical and when they were in the lying down position the avatar was horizontal, with regard to an upright observer. Four arrows labeled from 1 to 4 indicated the four possible directions of the next segment (see Fig. 8). Segments were added one by one by pressing the key corresponding to the direction chosen. Once the correct number of segments was

entered, a message appeared asking the subject to confirm the drawing by pressing the spacebar key. For every trial, accuracy of the drawn corridor was calculated as the number of segments reconstructed correctly from the beginning minus one (i.e., excluding the first, already drawn segment), divided by the total number of segments minus one. For instance, if the reconstruction of a four-segment corridor had only the first three segments correct, accuracy would be 66.6%. At any time, subjects could cancel their last choice by pressing the backspace key. The chance level for balanced blocks of trials including corridors with four and five segments is 12.4%.

Lastly, we modified the dual task. Because the new reconstruction task was more sequential, subjects would be more inclined to use a verbal strategy. At the beginning of each trial, three random numbers within the range of 20-59 were played through the headphones and subjects had to memorize them in the correct order. Just after the reconstruction task, subjects had to recall this sequence of numbers using the keyboard, and a sound was immediately played if more than one number was incorrect or if the numbers were not in the correct order. Mean accuracy at the



Fig. 8. The reconstruction task in the upright conditions. Segment by segment, subjects had to choose between the four possible directions, each segment direction being parallel to one of the canonical axes. Once the correct number of directions had been entered, a message appeared asking the subject to confirm the 3D drawing by pressing the spacebar key. Subjects could cancel their last choice at any moment by pressing the backspace key.

dual-task was approximately equivalent across conditions (mean: 75.5%; S.E.: 3.1%), and suggested that the dual task was correctly performed.

Five experimental conditions were compared for 10 different corridors, half being randomly selected from a four-segment database and the other half from a fivesegment database. The full experiment for any given subject comprised two sessions, one with 30 trials performed seated upright and the other with 20 trials performed lying down on the right side, each divided into blocks of 10 trials. In the upright position, the terrestrial, weightless and pitch terrestrial navigation conditions were compared, while in the lying down position only the terrestrial and weightless conditions were compared. The order of the sessions was counterbalanced across subjects. Each session started with practice trials: two for each of the navigation conditions corresponding to each of the body positions, for subjects to familiarize themselves with the computer interface. The task being cognitively very demanding, the two sessions for any given subject were done on different days in order to avoid mental saturation. The two sessions lasted about 1 h each.

3.2. Results

The performance accuracy in each experimental condition is presented in Fig. 9.

3.2.1. Body position

A 2 (body position) × 2 (navigation condition) × 2 (number of segments) within-subjects ANOVA design table was used to compare the reconstruction accuracy of the *terrestrial* and the *weightless* exploration conditions according to the body position. Again we found a significant main effect of the number of segments on accuracy [F(1,25)=39.57;



Fig. 9. Reconstruction accuracy (means \pm S.E.) according to the three upright and the two lying on the side exploration conditions. Dashed line indicates the response chance level (12.4%).

p < 0.001], the performance difference between corridors with four and five segments being 11.9% on average. The condition effect on performance was significant for both the upright [F(1,25)=14.05; p < 0.001] and lying on the side position [F(1,25)=7.73; p < 0.01]. Accuracy for the *terrestrial* condition (with 72.5% and 59.1% for the upright and lying on the side positions, respectively) was higher than for the *weightless* condition (with 62.1% and 49.2%, respectively) in each body orientation. This is consistent with the results of the previous experiment, which used a recognition task instead of the current reconstruction task. When subjects lay on their sides, rather than being seated, their performance decreased significantly in both the *terrestrial* [F(1,25)=7.25; p < 0.015] and the *weightless* condition [F(1,25)=7.77; p < 0.01].

3.2.2. Visually pitched terrestrial navigation

A 2 (condition) × 2 (number of segments) within-subjects ANOVA design table was used to compare the reconstruction accuracy of the *terrestrial* and the *pitch terrestrial* exploration conditions. Performance in the *pitch terrestrial* condition was significantly [F(1,25)=17.61; p < 0.001]lower (with 58.2%) than in the *terrestrial* condition (with 72.5%). A post hoc test revealed that this difference was significant for both four-segment trials (p < 0.003) and fivesegment trials (p < 0.05). The *pitch terrestrial* and the upright *weightless* conditions yielded approximately the same level of performance.

3.3. Discussion

The results of the second experiment reinforced the interpretations presented in the discussion of the first experiment. On one hand, by testing the *terrestrial* and *weightless* conditions with subjects lying on their sides, we showed that the previously observed differences in performance were independent of the visio-otolithic conflict. In fact, the same level of conflict with respect to gravity was present in both conditions. Moreover, in the *terrestrial* condition, laying subjects on their sides considerably impaired the reconstruction performance, which shows that in this natural condition, having the body and the rotation axes aligned with gravity facilitates memorization. It suggests that the egocentric to allocentric shift required by the task is easier if it involves rotations around the gravity axis.

On the other hand, we found new evidence in support of the hypothesis that the smaller number of rotations in the *terrestrial* conditions did not contribute to the difference in performance observed between the conditions. Actually, the *pitch terrestrial* condition, which also had fewer rotations to integrate along the displacement, produced approximately the same low level of performance as the *weightless* condition. The difference between the *terrestrial* and *pitch terrestrial* conditions is that, in the latter, only pitch rotations were used instead of yaw rotations. Although the displacements in the 3D maze required the integration of only one rotation type (pitch instead of yaw), it was not sufficient to maintain a high level of performance.

Therefore, the *terrestrial* condition produced a higher performance mainly because shifting from an egocentric to an allocentric reference frame is easier in a natural condition where only yaw turns are required, in which the rotation axis is aligned with gravity. Furthermore, it is not a question of fewer turns to integrate, but rather of the characteristics of the rotations involved. These findings are consistent with the mental rotation literature: Shiffrar and Shepard [17] have shown that performance improved when the axes of the object, the rotation axis and the gravitational vertical were aligned. Tilting one of them resulted in a marked deterioration of speed and accuracy of the mental rotation.

4. General discussion

In summary, we found that it is possible to build a mental representation of a 3D environment, although this representation is probably oriented with respect to the specific direction defined by the vertical of the memorized structure (usually provided by gravity). In other words, cognitive manipulations of this structure might be highly dependent on this vertical axis. The mental model could be a set of superimposed 2D cognitive maps having the vertical segments encoded as junctions between those maps. The processing of vertical and horizontal dimensions would consequently be very distinct and lead to a different spatial performance. If gravity defines this vertical axis, it would have a strong influence on the memorization process as well as on cognitive manipulations of the model such as mental rotations.

Our results show that the ongoing relationship between the egocentric reference frame and the allocentric reference frame has a crucial influence on the spatial updating of the 3D structure being memorized. In particular, humans have difficulty in integrating 3D displacements where any rotation in space can occur. The alignment of gravity with the vertical egocentric axis certainly plays a role in determining spatial performance.

Although we found that, with practice, subjects could learn how to integrate and memorize a displacement that used pitch and yaw rotations, this developed capacity appeared to be rather limited and not innate, in contrast to natural displacements. Based on our results, a new functional explanation for humans trying to keep their heads stabilized during locomotion [16] is that it facilitates the shift from an egocentric to an allocentric reference frame, which is required in order to memorize our trajectory. Indeed, keeping the head stabilized relative to the vertical of the environment, which is probably defined by gravity, reduces the complexity of the change of reference to a simple rotation around the vertical axis, thereby allowing an efficient updating of the cognitive map of the surrounding environment as well as a correct computation of one's orientation in this environment. Evolutionary considerations based on the possibility of stabilizing the head during locomotion could provide an explanation as to why humans and rats have fundamentally different innate navigational abilities. On one hand, humans evolved from monkeys that lived in the rainforest and had to build mental representations of a 3D environment. However, monkeys usually climb trees with their body upright and moving from tree to tree does not include pitch body rotations such as those in our weightless displacement condition. Therefore, human phylogenesis might have led to this head stabilizing process in order to simplify spatial cognition. On the other hand, because rats have a much higher power-to-weight ratio than humans, gravity induces weaker locomotion restrictions and thus they can walk on steeply sloping or even vertical surfaces. In these situations, rats cannot stabilize their heads to the same extent as humans. Spatial orientation in such environments requires them to perform complex referential shifts relying on rotations about any axis, and independently of the orientation of gravity. Therefore, the survival of the species has probably relied on the cognitive capacity to deal with 3D locomotion [12].

Acknowledgements

This research was supported by the Centre National d'Etudes Spatiales (CNES). Manuel Vidal received a grant from the Centre National de Recherche Scientifique (CNRS) for his PhD. The authors would like to thank Joseph McIntyre and Sid Wiener for their helpful comments on the text, as well as France Maloumnian for help with the graphics, and all the subjects who participated in the experiments.

Appendix A. Construction of the virtual mazes

Twelve virtual 3D corridors were used in the first experiment (Expected column of Table 1), each one being explored in a random order using the three displacement conditions (*weightless*, *subaquatic* and *terrestrial*). After each visual stimulus in a specific corridor, four external views of corridors were presented in random order, including the expected corridor and its associated distractors (the three pictures shown to the right of each corridor in Table 1).

Appendix B. Construction of the distractors

The principles that underlie the construction of the distractors associated with each corridor can be described in terms of the number of equal turns starting from the first segment. The distractors are ranked by level of difficulty for rejecting them, as presented in Table 2. A description of the set of distractors associated with a corridor, and the underlying choices of these distractors are summarized in Table 2. The notations used in this table can be explained with the

Table 1											
The sets of corridors	with	three,	four	and	five	segments	used	for	the	first	experiment

Catagory	Expected		Distractors		Category	Expected	Distractors				
Category	Expected	Most similar	Intermediate	Most different	Category	Expected	Most similar	Intermediate	Most different		
	7	`			nued)	٦	ヮ	7	حر		
nents	7	Ļ	7	7	(conti	7	1,	۲	٦		
4 segments	5	7	1			М	ς	Ĺ	مر		
	1	7	1	7	ments		5	Ļ	5		
	ζ	ľ	Ľ	7	5 seg	5		5	N		
	ς		7	<u> </u>		Ļ	ŗ	11	4		

Each of the 12 corridors explored (Expected column) with their three respective distractors ordered by level of difficulty for rejecting them.

following example: let us consider a distractor with the same first turn (two segments) as the explored corridor, but with the second turn leading to a different 3rd segment. This difference characterizes the distractor difficulty: with either a 3rd segment in the symmetrical direction—difficulty noted 3^{sym} , or with a 3rd segment whose direction is rotated by 90° —difficulty noted 3^{rot} . We assumed that the symmetrical difference would be less obvious to detect than the rotated difference.

Table 2

The construction principles of the distractors according to the set of the explored corridors and the level of difficulty of the distractors

Set of corridors	Level of difficulty	Similarity and transformation
Three segments	Most similar	3 ^{sym} (four trials)
	Intermediate	2 ^{sym} (four trials)
	Most different	2^{sym} (one trial), 2^{rot} (three trials)
Four segments	Most similar	3 ^{sym} (four trials)
	Intermediate	2 ^{sym} (four trials)
	Most different	2 ^{sym} (two trials), 2 ^{rot} (two trials)
Five segments	Most similar	5 ^{sym} (one trial), 4 ^{rot} (two trials),
		3 ^{sym} (one trial)
	Intermediate	3^{sym} (one trial), 2^{sym} (three trials)
	Most different	2^{sym} (three trials), 2^{rot} (one trial)

References

- M.A. Amorim, S. Glasauer, K. Corpinot, A. Berthoz, Updating an object's orientation and location during nonvisual navigation: a comparison between two processing modes, Percept. Psychophys. 59 (1997) 404–418.
- [2] S.S. Chance, F. Gaunet, A.C. Beall, J.M. Loomis, Locomotion mode affects the updating of objects encountered during travel: the contribution of vestibular and proprioceptive inputs to path integration, Presence 7 (1998) 168–178.
- [3] T. Gärling, A. Böök, E. Lindberg, C. Arce, Is elevation encoded in cognitive maps? J. Environ. Psychol. 10 (1990) 341–351.
- [4] R. Grasso, S. Glasauer, Y. Takei, A. Berthoz, The predictive brain: anticipatory control of head direction for the steering of locomotion, NeuroReport 7 (1996) 1170–1174.
- [5] M.-C. Grobéty, F. Schenk, Spatial learning in a three-dimensional maze, Anim. Behav. 43 (1992) 1011–1020.
- [6] D.L. Harm, D.E. Parker, Perceived self-orientation and self-motion in microgravity, after landing and during preflight adaptation training, J. Vestib. Res. 3 (1993) 297–305.
- [7] L.R. Harris, M. Jenkin, D.C. Zikovitz, Visual and non-visual cues in the perception of linear self-motion, Exp. Brain Res. 135 (2000) 12–21.
- [8] I. Israël, R. Grasso, P. Georges-Francois, T. Tsuzuku, A. Berthoz, Spatial memory and path integration studied by self-driven passive linear displacement: I. Basic properties, J. Neurophysiol. 77 (1997) 3180–3192.
- [9] Y. Ivanenko, R. Grasso, I. Israël, A. Berthoz, Spatial orientation in humans: perception of angular whole-body displacements in two-dimensional trajectories, Exp. Brain Res. 117 (1997) 419–427.

- [10] R.L. Klatzky, J.M. Loomis, A.C. Beall, S.S. Chance, R.G. Golledge, Spatial updating of self-position and orientation during real, imagined, and virtual locomotion, Psychol. Sci. 9 (1998) 293–298.
- [11] R.L. Klatzky, J.M. Loomis, R.G. Golledge, J.G. Cicinelli, S. Doherty, J.W. Pellegrino, Acquisition of route and survey knowledge in the absence of vision, J. Mot. Behav. 22 (1) (1990) 19-43.
- [12] J.J. Knierim, B.L. McNaughton, G.R. Poe, Three-dimensional spatial selectivity of hippocampal neurons during space flight, Nat. Neurosci. 3 (2000) 209–210.
- [13] J.R. Lishman, D.N. Lee, The autonomy of visual kinaesthesis, Perception 2 (1973) 287–294.
- [14] H. Mittelstaedt, The role of the otoliths in perception of the vertical and in path integration, Ann. N. Y. Acad. Sci. 871 (1999) 334-344.
- [15] D.R. Montello, H.L.J. Pick, Integrating knowledge of vertically aligned large-scale spaces, Environ. Behav. 25 (1993) 457–483.

- [16] T. Pozzo, A. Berthoz, L. Lefort, Head stabilization during various locomotor tasks in humans: I. Normal subjects, Exp. Brain Res. 82 (1990) 97–106.
- [17] M.M. Shiffrar, R.N. Shepard, Comparison of cube rotations around axes inclined relative to the environment or to the cube, J. Exp. Psychol. Hum. Percept. Perform. 17 (1991) 44–54.
- [18] C. Tafforin, R. Campan, Ethological experiments on human orientation behavior within a three-dimensional space-in microgravity, Adv. Space Res. 14 (1994) 415–418.
- [19] M. Vidal, M. Lipshits, J. McIntyre, A. Berthoz, Gravity and spatial orientation in virtual 3D maze. J. Vestib. Res. (in press).
- [20] J.P. Wann, D.K. Swapp, Why you should look where you are going, Nat. Neurosci. 3 (2000) 647–648.
- [21] W.H. Warren Jr., M.W. Morris, M. Kalish, Perception of translational heading from optical flow, J. Exp. Psychol. Hum. Percept. Perform. 14 (1988) 646–660.