

As the world turns: Short-term human spatial memory in egocentric and allocentric coordinates

Pamela Banta Lavenex^{a,*}, Sandro Lecci^a, Vincent Prêtre^a, Catherine Brandner^b, Christian Mazza^c, Jérôme Pasquier^c, Pierre Lavenex^a

^a Laboratory of Brain and Cognitive Development, Department of Medicine, Unit of Physiology, University of Fribourg, Chemin du Musée 5, 1700 Fribourg, Switzerland

^b Department of Psychology, University of Lausanne, Quartier UNIL-Dorigny, Bâtiment Anthropole 5189, CH-1015 Lausanne, Switzerland

^c Department of Mathematics, University of Fribourg, Chemin du Musée 23, 1700 Fribourg, Switzerland

We aimed to determine whether human subjects' reliance on different sources of spatial information encoded in different frames of reference (i.e., egocentric versus allocentric) affects their performance, decision time and memory capacity in a short-term spatial memory task performed in the real world. Subjects were asked to play the Memory game (a.k.a. the Concentration game) without an opponent, in four different conditions that controlled for the subjects' reliance on egocentric and/or allocentric frames of reference for the elaboration of a spatial representation of the image locations enabling maximal efficiency. We report experimental data from young adult men and women, and describe a mathematical model to estimate human short-term spatial memory capacity. We found that short-term spatial memory capacity was greatest when an egocentric spatial frame of reference enabled subjects to encode and remember the image locations. However, when egocentric information was not reliable, short-term spatial memory capacity was greater and decision time shorter when an allocentric representation of the image locations with respect to distant objects in the surrounding environment was available, as compared to when only a spatial representation encoding the relationships between the individual images, independent of the surrounding environment, was available. Our findings thus further demonstrate that changes in viewpoint produced by the movement of images placed in front of a stationary subject is not equivalent to the movement of the subject around stationary images. We discuss possible limitations of classical neuropsychological and virtual reality experiments of spatial memory, which typically restrict the sensory information normally available to human subjects in the real world.

1. Introduction

Allocentric, spatial relational memory is the memory of our surroundings that encodes the interrelationships between individual elements that compose the environment in which we live. As such, spatial relational memory is a fundamental component of each of our autobiographical memories, also known as episodic memories (i.e., the memory of our personal life events that occur in unique spatio-temporal contexts). In everyday life, the construction of a spatial relational representation of the environment is associated with the movement of an individual through its environment [1,2]. Such a spatial representation therefore integrates coherent information derived from all sensory modalities, including visual, vestibular, proprioceptive, somesthetic, auditory and olfactory information, to create a multimodal memory of the envi-

ronment [1–6]. In rodents and non-human primates, the ability to elaborate and remember an allocentric, spatial relational representation of the environment has been shown to be dependent on the integrity of the hippocampal formation [7–9]. Importantly, the tasks used to demonstrate the fundamental role of the hippocampal formation in spatial relational memory processes are paradigms that allow animals to navigate freely in the experimental environment. In contrast, tasks in which the animals' ability to move about in the environment is constrained have led to inconsistent conclusions regarding the role of the hippocampal formation in spatial memory. Such discrepant findings were the result of experimental designs that failed to recognize, and control for, animals' ability to rely on both egocentric (centered on the individual, a.k.a., viewer-centered and viewpoint-dependent) and allocentric (centered on the environment, a.k.a., environment-centered, viewpoint-independent) spatial frames of reference in order to solve these tasks (see [8] for a detailed discussion).

Recently, a great deal of research has focused on describing the deficits in human spatial memory following unilateral and/or

* Corresponding author. Tel.: +41 26 300 87 67; fax: +41 26 300 97 34.
E-mail address: pamelabantalavenex@unifr.ch (P. Banta Lavenex).

bilateral medial temporal lobe damage (damage that is more or less confined to the hippocampal formation; e.g., [10–12]). Recognizing the limitations of standard neuropsychological tests of spatial capacities in humans, which are generally administered on a tabletop located in front of the subject, researchers have begun to take advantage of the development of 3-D graphics and computer technology to design spatial tasks which take place in virtual environments [10–12]. Because it is generally considered that navigation in these virtual environments mimics navigation in the real world, researchers have used the performance of human subjects in these tasks as a measure for their allocentric spatial relational memory capacities. Importantly, however, even though the subject is able to explore the environment by using a joystick to virtually move throughout the environment displayed on a computer screen or a virtual reality personal display system (goggles), the subject's actual position within the real world remains stationary throughout the test. Thus, the subject must selectively attend to the sensory information relevant to solve the task in the virtual-environment (i.e., central vision) and disregard all other sensory information normally consistent with visual inputs (i.e., vestibular, proprioceptive and somesthetic information, as well as optical flow derived from peripheral vision). When considered in this manner, it is not necessarily clear how similar virtual navigation is to navigation in the real world [13]. Specifically, it is not known whether or how the absence, or incoherence, of spatial information derived from different sensory modalities might influence task performance and the elaboration of a spatial relational memory of the environment, or more precisely the memory for the part of the environment that the subject is experiencing.

Previous studies have partially addressed this question by studying how well humans remember the location of objects in the real world, when either the view of an array of objects (placed on a tabletop) is changed by the rotation of the array while the subject remains stationary, or when the view of the objects is changed by the movement of the subject around the array which remains stationary [14–16]. Specifically, subjects were presented with an array of five objects on a circular tabletop for 3 s. After a 7–13-s delay during which the display was hidden and one of the objects was moved by the experimenter, subjects were asked to indicate which object had moved. Overall, these studies demonstrated that task performance is maximal when the spatial information that is derived from different sensory modalities and encoded in distinct egocentric and allocentric frames of references is coherent, whereas performance declines when spatial information derived from different sensory modalities and encoded in distinct egocentric and allocentric frames of reference is incoherent. Most interestingly, these studies suggest that viewpoint changes caused by the movement of the subject around the apparatus have little or no effect on subjects' ability to detect changes in object locations, whereas viewpoint changes caused by the rotation of the array produce a greater disruption of task performance [14–16]. These results have important implications regarding the role and relative importance of egocentric and allocentric information in the construction of spatial representations of the environment. Moreover, with the advancement of virtual reality technology and its increasing use in fMRI studies of normal human spatial memory and the study of patients with specific brain lesions, it is particularly important to determine whether the elaboration of an allocentric spatial relational representation of the environment is differentially influenced by the ability of the subjects to actively move about in a static environment (e.g., as in animal experiments and real world situations with human participants) as compared to when part of the environment moves around subjects who remain stationary (e.g., neuropsychological tabletop experiments and virtual reality environments). Despite fundamental differences in the sensory information that is available and possibly integrated into the spatial

representation of the part of the world that the subject is experiencing under real world versus virtual environments, researchers have generally considered that when subjects effectuate a viewpoint change in virtual reality (either by pushing arrow buttons on a keyboard or by manipulating a joystick), they actually imagine themselves moving, rather than imagining an array of objects and landmarks rotating around them [12]. Few experiments, however, have directly compared subject performance in order to determine whether the brain performs equivalently in these two differing conditions (see Moffat [17] for a review).

Here, we aimed to determine whether restricting human subjects' ability to rely on either allocentric or egocentric frames of reference, which, by consequence, restricts their access to specific types of sensory information, affects their performance, decision time and memory capacity in a short-term spatial memory task performed in the real world. To this end, we developed a task based on the popular game Memory (a.k.a. the Concentration game). In brief, when playing the Memory game, participants are presented with an array of an even number of overturned cards that are randomly arranged on a table top. The cards have a variety of images on their face side, and each image has an identical pair on the tray. It is the goal of the participant to find the matching pairs by overturning two cards at a time (either simultaneously or sequentially). Matching pairs are removed from the game. The identity and location of cards that are overturned but not matched are placed into memory for future use in the game. The Memory game has been used previously to compare the spatial memory capacities of adult male and female subjects [18,19]. However, our paradigm is unique as subjects were tested in four different conditions that allowed us to distinguish and control for the subjects' reliance on different frames of reference (i.e., egocentric versus allocentric) in the elaboration of a spatial representation of the image locations in order to enable maximal efficiency when playing the game (Fig. 1). Our experimental design is thus partially similar to that used previously by [14–16]. Importantly, our task evaluates human spatial memory capacity in a real-world environment, in which moving subjects are forced to rely, alternately, on egocentric and allocentric representations of the image locations that are defined by either coherent or incoherent visual, vestibular, proprioceptive, somesthetic and auditory information.

We first report experimental data from control subjects (men and women) in the different experimental conditions, which varied with respect to the types of spatial representations available to solve the task and the number of image pairs to remember. We then describe a mathematical model that estimates the short-term spatial memory capacity of these subjects in the various conditions. Our findings, using a new experimental design in a real-world environment, reveal no gender differences but fundamental differences in task performance, decision time and short-term spatial memory capacity that are critically dependent on whether spatial information derived from different sensory modalities is coherent, like in the real world, or incoherent, like in virtual reality. In particular, they demonstrate that solving a spatial memory task in which the changes in viewpoint are produced by the movement of an array of images placed in front of the experimental subject is not equivalent to when the subject moves around the array of images.

2. Methods

2.1. Apparatus

The experimental apparatus consisted of a wooden tray on which pairs of identical images could be randomly arranged in a regular, symmetrical array (Fig. 1).

Images were depictions of animals painted on wooden squares (55 mm × 55 mm × 5 mm), hereafter referred to as cards. The English name of the animal also appeared on the card in locations that varied from card to card. The location and orientation of each individual image on the tray was randomized at

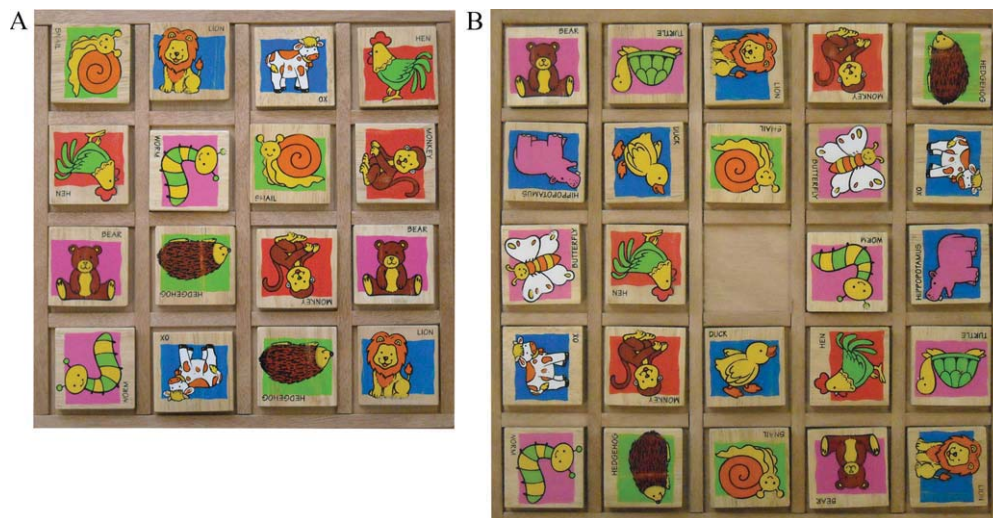


Fig. 1. Experimental apparatus used in the study. A. 8 image pairs used in Experiment 1. B. Twelve image pairs used in Experiment 2. Blocks are presented face-up to illustrate the images used and their random distribution at the beginning of each trial. Experimental subjects never saw the apparatus as it is illustrated here; images were always shuffled face-down on a table and then randomly placed face down in the tray prior to beginning each repetition of the game.

the beginning of each trial. The tray was placed on a platform that could be rotated freely by the subject and was situated about 120 cm above the ground (platform height could be adjusted for individual subjects). In order to assess the effect of memory load on performance, we tested different groups of subjects under different memory load conditions. In Experiment 1, the wooden tray was 29 cm × 29 cm and 8 pairs of images were randomly distributed in a 4 × 4 array (Fig. 1a; 16 locations). In Experiment 2, the wooden tray was 36 cm × 36 cm and twelve pairs of images were randomly arranged in a 5 × 5 array (Fig. 1b; 24 locations, the array's central location always remained empty).

2.2. Procedure

Subjects were asked to play the game memory without an opponent. Subjects were instructed to turn two cards sequentially (i.e., one card and then a second card, hereafter defined as a 'move'), and remove these cards from the game if they matched. Matching pairs were placed in the pocket of an apron that was fastened around the subject's waist. If the images did not match, subjects had to return the cards to their face down position in the exact location where they found them. The test was complete when all pairs were found. Each session consisted of four different experimental trial conditions organized in a pre-determined, pseudo-random order, and all subjects were tested for ten experimental sessions (thus, each subject played the game a total of 40 times).

In the first condition (Fig. 2A, Stationary condition), both the subject and the game tray remained in the same static position in the environment, so that both egocentric (i.e., centered on the individual) and allocentric (i.e., centered on the environment) spatial frames of reference were coherent and could be used to encode and remember the image locations. In the second condition (Fig. 2B, Subject and tray rotate), subjects rotated 90° around the table and rotated the game tray containing the images 90° in the same direction along with them between each move. In this condition, an egocentric coding of space could be used to encode and remember the image locations on the game tray, but because the tray rotated, the image locations were not stable within the room and an allocentric spatial representation of the image locations within the environment could not be used to perform the task. In the third condition (Fig. 2C, Subject rotates), subjects rotated 90° around the game tray between each move, and thus egocentric information could not be used to remember the image locations. Instead, in this condition the image locations remained stable with respect to the room and thus subjects could keep track of their own movements and the allocentric position of the viewed images within the surrounding environment in order to identify matching pairs. In the fourth condition (Fig. 2D, Tray rotates), subjects remained stationary, but they had to rotate the game tray by 90° between each move. In this condition, neither an egocentric coding of space (i.e., centered on the individual) nor an allocentric representation of the image locations within the room (i.e., centered on the surrounding environment) could enable task performance. Nonetheless, a relational representation of the image locations on the game tray could be encoded (as was the case in the other three experimental conditions) and then mentally rotated by the subject each time they rotated the game tray. This representation, however, was disconnected from the allocentric spatial frame of reference provided by the surrounding environment (i.e., the room). The cognitive demands imposed by this last condition are thus similar to those imposed by mental rotation and virtual navigation tasks, in which part of the environment is moved, but the experimental subject remains stationary within the experimental room.

2.3. Subjects

In Experiment 1, we tested twenty normal adult subjects (10 males, 10 females; mean age: 23 years; range 20–29 years) in two different locations: one group of subjects (4 males, 4 females) were students from the University of Fribourg (Switzerland); another group of subjects (6 males and 6 females) were students at the University of Lausanne (Switzerland). In Experiment 2, we tested twenty normal adult subjects (10 males, 10 females; mean age: 23 years; range 21–28 years), all students at the University of Fribourg (none of these subjects had participated in Experiment 1). Human subjects research was approved by the Inter cantonal Ethics Committee for Jura, Neuchâtel, Fribourg (Neuchâtel, Switzerland) and the Ethics Commission of Vaud (Lausanne, Switzerland). Research was performed in accordance with the NIH guidelines for the use of human subjects in research.

2.4. Data analysis

Subjects' spatial memory performance was measured by the number of moves (defined as the sequential turning of two images) necessary to find all pairs of images (8 pairs in Experiment 1, 12 pairs in Experiment 2), as well as by the time per move (i.e., the time spent to choose and turn two blocks; determined by dividing the total trial duration by the number of moves). Data are presented as group average ± standard error of the mean in the different testing conditions across all sessions. We also present normalized data based on the average value (number of moves or time per move) per session per individual: i.e., the individual score for condition X in session Y was divided by the average score for the four trial conditions of the same session (Y) for that individual. As the normalized data are unit-less, they enable us to compare the different experimental conditions based on parameters that would otherwise not be comparable (number of moves with 8 or 12 pairs of images; number of moves versus time per move versus memory capacity).

In Section 3.3, we present a mathematical model with which to estimate memory capacity based on the number of moves executed by each individual subject. We performed non-parametric Mann–Whitney and Friedman repeated measures analyses of variance to compare gender and testing conditions. Post hoc pairwise multiple comparisons were performed with the Wilcoxon signed rank test; statistical significance was defined as $p < 0.05$.

3. Results

3.1. Experiment 1: 8 image pairs

The average number of moves necessary to find the 8 pairs of images did not differ between men and women ($z = 0.038$, $p = 0.9698$). Across all subjects, however, the average number of moves necessary to find the 8 pairs of images was clearly influenced by the testing condition (Chi-square(3) = 28.879, $p < 0.001$; Table 1). Access to an egocentric representation of the position of the identified images (Stationary and Subject & tray rotate [S & T] conditions) enabled better short-term spatial memory performance as compared to when an egocentric frame of reference was unreliable

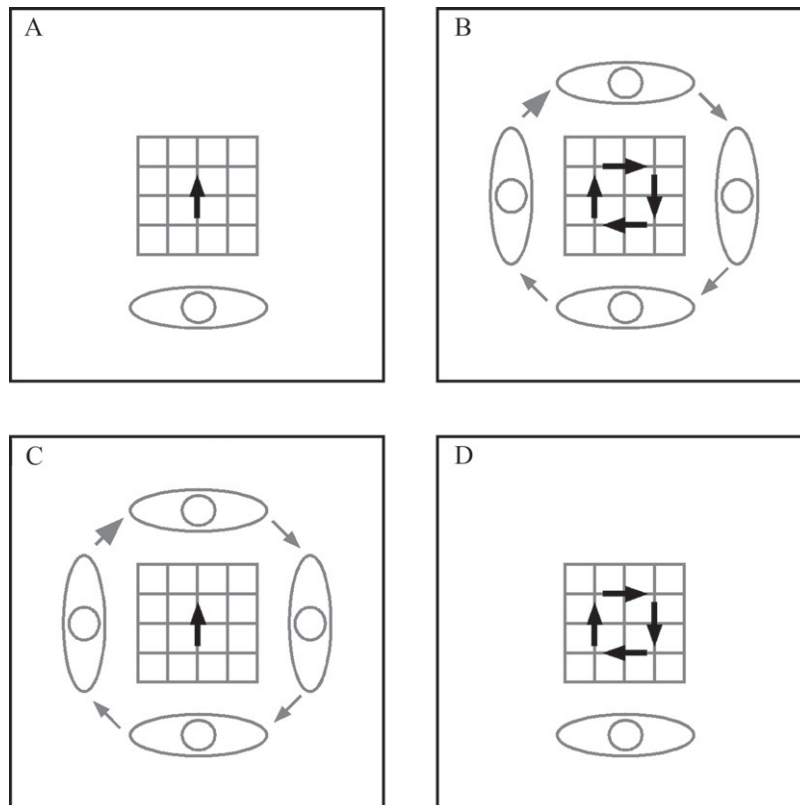


Fig. 2. Experimental testing conditions. A. Stationary: both the game tray and the subject remained in the same position throughout the trial. B. Subject and tray rotate: both the subject and the game tray rotated 90° in the same direction between each move. C. Subject rotates: the game tray remained in the same position, but the subject rotated 90° around the apparatus between every move. D. Tray rotates: the subject remained in the same position, but rotated the game tray 90° between each move.

Table 1
Number of moves necessary to find the 8 pairs of identical images (Experiment 1).

	Stationary	S & T rotate	Subject rotates	Tray rotates
Raw score				
Men	14.55 ± 0.42	15.13 ± 0.45	16.23 ± 0.90	16.60 ± 0.87
Women	14.52 ± 0.51	14.73 ± 0.51	16.33 ± 0.92	16.70 ± 0.85
Average	14.54 ± 0.32	14.93 ± 0.33	16.28 ± 0.62	16.65 ± 0.59
Normalized score				
Men	0.94 ± 0.02	0.97 ± 0.02	1.03 ± 0.02	1.06 ± 0.02
Women	0.94 ± 0.01	0.95 ± 0.01	1.04 ± 0.02	1.06 ± 0.02
Average	0.94 ± 0.01	0.96 ± 0.01	1.04 ± 0.01	1.06 ± 0.01

(Subject rotates and Tray rotates conditions; Stationary = Subject and tray rotate < Subject rotates = Tray rotates; all $p < 0.01$). Statistical analyses conducted on the normalized data yielded similar results (Chi-square(3) = 28.788, $p < 0.001$; Stationary = Subject and tray rotate < Subject rotates = Tray rotates; all $p < 0.01$).

The average time per move (i.e., the time taken to choose and turn two blocks) did not differ between men and women ($z = 0.454$, $p = 0.6501$), but across all subjects was also influenced by the testing condition (Chi-square(3) = 42.480, $p < 0.001$; Table 2).

Table 2
Time per move in the four different testing conditions (Experiment 1: 8 image pairs).

	Stationary	S & T rotate	Subject rotates	Tray rotates
Raw score (s)				
Men	4.51 ± 0.39	6.02 ± 0.57	6.04 ± 0.62	6.25 ± 0.61
Women	4.38 ± 0.26	5.75 ± 0.39	5.66 ± 0.36	5.88 ± 0.40
Average	4.44 ± 0.23	5.89 ± 0.34	5.85 ± 0.35	6.07 ± 0.36
Normalized score				
Men	0.80 ± 0.02	1.06 ± 0.01	1.05 ± 0.01	1.09 ± 0.01
Women	0.82 ± 0.01	1.06 ± 0.01	1.04 ± 0.01	1.08 ± 0.01
Average	0.81 ± 0.01	1.06 ± 0.01	1.05 ± 0.01	1.09 ± 0.01

Unsurprisingly, the time per move was lowest in the stationary condition, as compared to all three other conditions, as neither the subject nor the game tray moved in this condition. More interestingly, the time per move in the Subject rotates condition was not different from that in the Subject and tray rotate condition, but both were lower than in the Tray rotates condition (Stationary < Subject and tray rotate = Subject rotates < Tray rotates; all $p < 0.05$). Statistical analyses conducted on the normalized data yielded similar results (Chi-square(3) = 42.540, $p < 0.001$; Stationary < Subject and tray rotate = Subject rotates < Tray rotates; all $p < 0.05$). Thus, even though the two conditions (Subject rotates and Tray rotates) did not differ in the number of moves necessary to solve the game, in the absence of a reliable egocentric representation of the image positions, an allocentric representation of the position of the identified images within the room (Subject rotates condition) enabled a shorter decision time than a relational representation of the position of the identified images relative to the board and each other, that is disconnected from the rest of the environment (Tray rotates condition).

3.2. Experiment 2: 12 pairs

The average number of moves necessary for men and women to find the twelve pairs of images did not differ ($z = 1.02$, $p = 0.30$). Across all subjects, however, the number of moves necessary to find the 12 pairs was clearly influenced by the testing condition (Chi-square(3) = 47.30, $p < 0.0001$; Table 3). In this experiment, as in Experiment 1 (Section 3.1) access to an egocentric representation of the position of the identified images (Stationary and Subject & tray rotate conditions) enabled a better short-term spatial memory performance as compared to when an egocentric frame of reference was unreliable (Subject rotates and Tray rotates conditions).

Table 3
Number of moves necessary to find the twelve pairs of identical images (Experiment 2).

	Stationary	S & T rotate	Subject rotates	Tray rotates
Raw score				
Men	26.25 ± 0.76	26.55 ± 1.08	30.80 ± 1.77	35.85 ± 2.92
Women	24.69 ± 0.81	25.47 ± 0.98	28.56 ± 1.75	31.94 ± 2.46
Average	25.47 ± 0.57	26.01 ± 0.72	29.68 ± 1.24	33.90 ± 1.91
Normalized score				
Men	0.89 ± 0.02	0.90 ± 0.02	1.03 ± 0.01	1.18 ± 0.04
Women	0.91 ± 0.02	0.93 ± 0.01	1.02 ± 0.03	1.14 ± 0.04
Average	0.90 ± 0.02	0.91 ± 0.01	1.03 ± 0.02	1.16 ± 0.03

In addition, the number of moves differed between the two conditions in which an egocentric representation of the position of the images did not enable the completion of the task. Short-term spatial memory performance was better when the subjects had access to an allocentric representation of the image locations relative to the room in the Subject rotates condition, as compared to when the subjects did not in the Tray rotates condition (Stationary = Subject and tray rotate < Subject rotates < Tray rotates; all $p < 0.01$). Statistical analyses conducted on the normalized data yielded similar results (Chi-square(3) = 47.40, $p < 0.0001$; Stationary = Subject and tray rotate < Subject rotates < Tray rotates; all $p < 0.01$).

The average time per move (i.e., the time taken to choose and turn two blocks) did not differ between men and women ($z = 0.832$, $p = 0.405$), but across all subjects was also clearly influenced by the testing condition (Chi-square(3) = 43.26, $p < 0.0001$; Table 4). As in Experiment 1 (Section 3.1), the time per move was lowest in the stationary condition, as compared to all three other conditions, as neither the subject nor the game tray moved in this condition. More interestingly, the time per move in the Subject rotates condition was not different from that in the Subject and tray rotate condition (Stationary < Subject and tray rotate = Subject rotates < Tray rotates; all $p < 0.05$). Statistical analyses conducted on the normalized data yielded similar results (Chi-square(3) = 43.82, $p < 0.0001$; Stationary < Subject and tray rotate = Subject rotates < Tray rotates; all $p < 0.05$). Thus, as was found in Experiment 1 with 8 image pairs, in the absence of a reliable egocentric representation of the image positions, an allocentric representation of the position of the identified images within the room (Subject rotates condition) enabled a shorter decision time than a relational representation of the position of the identified images relative to the board that is disconnected from the rest of the environment (Tray rotates condition).

3.3. Modeling the memory capacity observed while playing the Memory game

The results presented above reveal interesting differences between testing conditions in the number of moves (number of choices to find the 8 or 12 matching pairs) and the time taken to make a choice (time per move). Here, we present a model to

Table 4
Time per move in the four different testing conditions (Experiment 2: 12 image pairs).

	Stationary	S & T rotate	Subject rotates	Tray rotates
Raw score (s)				
Men	4.52 ± 0.39	5.88 ± 0.41	5.79 ± 0.44	6.04 ± 0.46
Women	4.33 ± 0.38	5.41 ± 0.47	5.26 ± 0.44	5.64 ± 0.45
Average	4.42 ± 0.27	5.65 ± 0.31	5.53 ± 0.31	5.84 ± 0.31
Normalized score				
Men	0.81 ± 0.02	1.06 ± 0.01	1.04 ± 0.01	1.09 ± 0.01
Women	0.84 ± 0.01	1.05 ± 0.01	1.02 ± 0.01	1.10 ± 0.01
Average	0.82 ± 0.01	1.06 ± 0.01	1.03 ± 0.01	1.09 ± 0.01

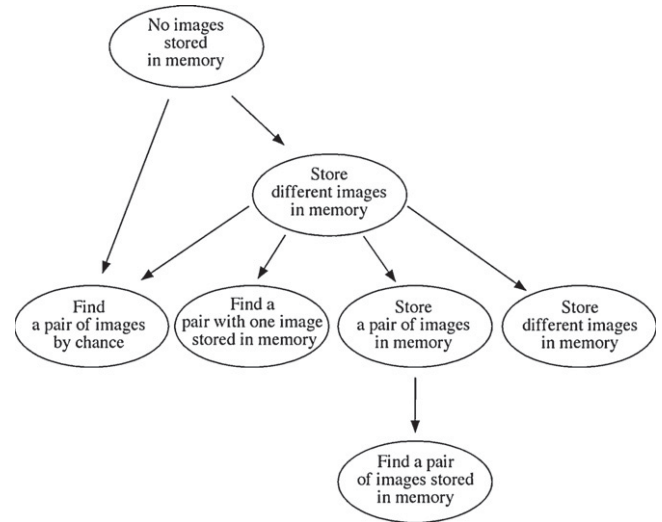


Fig. 3. A schematic outlining how the game can proceed. Downward-pointing arrows represent a move in the game during which two cards are turned sequentially.

estimate the subjects' short-term spatial memory capacity, i.e., the number of card locations that can be held in memory of experimental subjects performing the task in each of the four different testing conditions.

Our mathematical model assumes that a subject can remember exactly L cards amongst all cards that have been seen previously. The parameter L models the memory capacity (or length), and is unknown. The idea consists in testing null hypotheses like $H_0: L = L_0$ against the alternative hypothesis $H_1: L \neq L_0$ for various choices of L_0 . We will assign a memory capacity L to each subject by choosing the parameter L for which the related null hypothesis is not rejected, hence having a large p -value and also yielding the highest posterior probability (see below). The state of the game after k moves is denoted by $X_k = (n, l, *)$, where n is the number of pairs of cards which remain in the game, l is the number of cards memorized by the player, and $*$ is a binary value used to differentiate the two following cases: d , only distinct (unpaired) cards in memory, and p , a pair amongst the cards in memory. The end of the game is given by the random number of moves T , corresponding to the last move (see below), which is the number of moves necessary to find the N matching pairs. The level and the p -value of this test are obtained by computing the probability $PC(k)$ that a subject having a memory capacity of L finds the N matching pairs after k moves.

In order to solve this problem, we have developed a full mathematical treatment using the Markov chain theory, which is based on the following conditions: N pairs of identical cards are shuffled and spread face down on a table. At each step in the game, the player turns two cards sequentially. If the two cards are identical, they are removed from the table. If not, the cards are replaced face down in their original position. Fig. 3 represents the different events that can occur at each step during the game; the game ends when all pairs have been found. The player's score is equal to the number of moves it took to find all matching pairs. The mathematical model is further based on the following assumptions:

3.3.1. Memory processes

(1) the model assumes that the player has at his disposal L memory slots, which he can fill with encountered cards and their positions. (2) The player's memory is perfect; he does not make any errors. (3) When all of the L slots are filled, the player must delete a card from memory in order to memorize a new one. (4) Matching pairs removed from the game are forgotten.

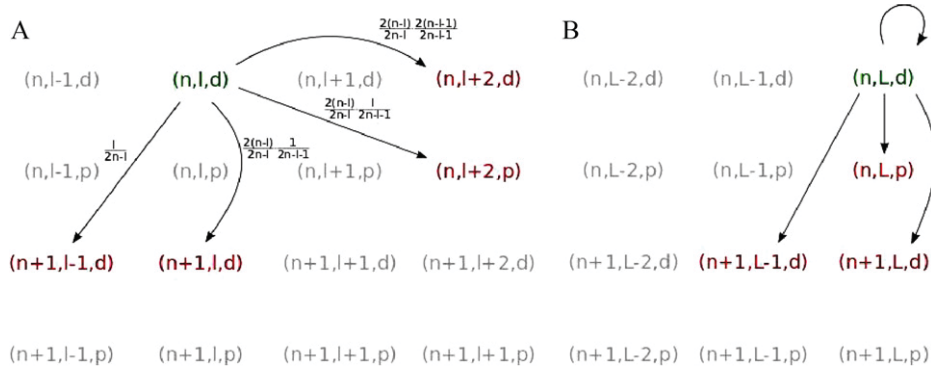


Fig. 4. A. Example of transitions when the memory limit has not been reached. B. Example of transitions when the memory limit has been reached.

3.3.2. Procession of the game

(1) At each step in the game, the player turns two cards, one after the other (Fig. 3). (2) If the player has only distinct, unpaired cards in his memory, he chooses at random a card from amongst those that are not in his memory. (3) If this randomly chosen card corresponds to a card in his memory, he chooses its matching pair. (4) If the randomly chosen card does not correspond to a card in his memory, he turns a second card at random. (5) If the second card corresponds to the first one, he has identified a pair and removes the two cards from the game. (6) If the second card corresponds to a different card already in his memory, he places the two new cards in his memory, amongst which is a pair. If maximal memory capacity is reached, the pair is memorized preferentially. (7) If the second card corresponds neither to the first one, nor to one already in memory, then the player simply places two more cards in his memory (until maximal memory capacity is reached). (8) Finally, if at any point the player has a matching pair in memory, he identifies and removes this pair

The probability with which certain events occur depends on three factors: (1) the number of remaining pairs; (2) the number of memorized cards; and (3) whether or not the player has a pair in his memory. Hence, to summarize, this game can be compared to a Markov chain $X_k = (n, l, *)$ where n is the number of pairs of cards which remain in the game; l is the number of cards in the player's memory; and $*$ = d : only distinct, unpaired cards in the player's memory, or $*$ = p : a matching pair exists amongst the cards in the player's memory; and k is the number of moves.

All the possible transitions and transition probabilities are defined by these hypotheses (Fig. 4). The full mathematical description of the related stochastic process is provided in the Appendix.

3.3.3. Statistics

We further developed a statistical test based on the above Markov chain to determine whether a given player has a particular memory capacity. We considered the following hypotheses: H_0 : memory of length $L = L_0$ versus H_1 : $L \neq L_0$. Each player repeatedly plays M games, thus yielding the sample (T_1, \dots, T_M) , where T denotes the number of moves needed to recover the N cards. We then compute the likelihood of the random sample as $PL(T_1, \dots, T_M) = PL(T_1) \times PL(T_2) \times \dots \times PL(T_M)$, and consider the likelihood ratio statistics $R(L_0; T_1, \dots, T_M) = \max_{L \neq L_0} (PL(T_1, \dots, T_M) / PL_0(T_1, \dots, T_M))$. The null hypothesis H_0 : $L = L_0$ is rejected in favor of the alternative hypothesis H_1 : $L \neq L_0$ when $R(L_0; T_1, \dots, T_M)$ is larger than some threshold value for some significance level. The threshold is fixed for a given level using Monte-Carlo simulations based on the exact distribution $PL_0(k)$. We proceed similarly to get the related p -value (Fig. 5).

We also provide a Bayesian estimation by assuming a uniform distribution for the parameter L , that is we assume that L is

uniformly distributed amongst all possible values. The posterior distribution of L given the observed data $\{T_1, \dots, T_M\}$ is obtained from Bayes' formula as $P(L|T_1, \dots, T_M) = P(L; T_1, \dots, T_M) / P(T_1, \dots, T_M) = PL(T_1, \dots, T_M) / PL_0(T_1, \dots, T_M)$. More details are provided in the Appendix.

We can thus provide a probability plot that yields the probability that L is the memory capacity given the observed sample $\{T_1, \dots, T_M\}$. The results from the likelihood ratio test and the Bayesian analysis generally agree: the posterior distribution attains its maximum value at some L that corresponds in most cases to the value for which the related p -value is large (Fig. 5).

3.3.4. Memory capacity

3.3.4.1. Experiment 1: 8 pairs. We used the predictive value of the model (Bayesian analysis) to estimate each individual subject's memory capacity in the different testing conditions. The estimated memory capacity computed with our model (Table 5) mirrors the actual number of moves necessary to find all image pairs observed in our experiment (Table 1). The memory capacity did not differ between men and women ($z = 0.076$, $p = 0.9397$), but was clearly influenced by the testing condition (Chi-square(3) = 27.610, $p < 0.001$). Subjects' memory capacity was greater in the two conditions in which egocentric information could be used to remember the position of identified images, as compared to the two conditions in which an egocentric representation of the position of the identified images was unreliable (Stationary = Subject and tray rotate > Subject rotates = Tray rotates; all $p < 0.01$). Indeed, individuals retained on average 0.58 more items in memory in the egocentric conditions (average of 3.33 in the Stationary and Subject & Tray rotate conditions), as compared to when an egocentric frame of reference was unreliable (average of 2.75 in the Subject rotates and Tray rotates conditions).

3.3.4.2. Experiment 2: 12 pairs. The estimated memory capacity computed with our model (Table 6; Bayesian analysis) mirrors the actual number of moves necessary to find all 12 pairs of images observed in our experiment (Table 3). The memory capacity did not differ between men and women ($z = 1.250$, $p = 0.2114$), but was clearly influenced by the testing condition (Chi-square(3) = 42.388, $p < 0.0001$). Subjects' memory capacity was greater in the two conditions in which egocentric information could be used to remember the position of identified images, as compared to the two conditions in which an egocentric representation of the position of the identified images was unreliable. Moreover, subjects' memory capacity was greater in the Subject rotates condition, as compared to the Board rotates condition (Stationary = Subject and tray rotate > Subject rotates > Tray rotates; all $p < 0.01$).

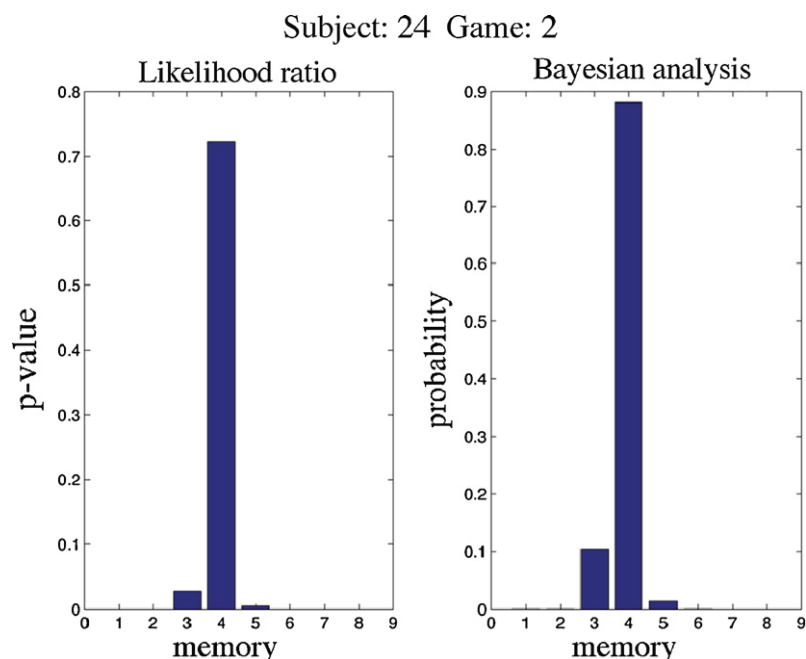


Fig. 5. Memory capacity estimates for subject 24 in the “Subject rotates” condition. A. Likelihood ratio: various estimates of the memory capacity, and their probability value to be true. A p -value smaller than 0.05 is necessary to reject the null hypothesis $H_0: L = L_0$ in favor of the alternative hypothesis $H_1: L \neq L_0$. B. Bayesian analysis: probability that L is the true memory capacity as determined by the total number of moves made by the player. In this example, a memory capacity of 4 is the only likelihood ratio that is not rejected (A), and it has the highest probability according to the Bayesian analysis as well (B).

Table 5
Estimated memory capacity in different testing conditions (Experiment 1: 8 image pairs).

	Stationary	S & T rotate	Subject rotates	Tray rotates
Raw score				
Men	3.43 ± 0.22	3.07 ± 0.18	2.85 ± 0.20	2.71 ± 0.24
Women	3.48 ± 0.22	3.34 ± 0.29	2.70 ± 0.22	2.74 ± 0.21
Average	3.46 ± 0.15	3.21 ± 0.17	2.78 ± 0.15	2.73 ± 0.16

4. Discussion

Our study, using a new experimental design in a real-world environment, revealed no gender differences, but fundamental differences in task performance, processing time and short-term spatial memory capacity of young adult human subjects tested in egocentric and allocentric coordinates. Subjects’ performance and short-term memory capacity were highest and largely similar in the two conditions (Stationary, and Subject and tray rotate) where an egocentric spatial representation enabled subjects to remember the locations of identified images (Experiments 1 [Section 3.1] and 2 [Section 3.2]). When egocentric coding was unreliable, subjects’ performance and short-term memory capacity were lower, and moreover, differed based on whether the image locations remained stable in the experimental room (Subject rotates) or were disconnected from the surrounding environment and remained stable solely in relation to each other on the testing tray (Tray rotates; Experiment 2).

Table 6
Estimated memory capacity in different testing conditions (Experiment 2: 12 image pairs).

	Stationary	S & T rotate	Subject rotates	Tray rotates
Raw score				
Men	3.67 ± 0.20	3.75 ± 0.30	3.05 ± 0.26	2.48 ± 0.29
Women	4.32 ± 0.30	3.92 ± 0.33	3.45 ± 0.32	3.00 ± 0.33
Average	3.99 ± 0.19	3.84 ± 0.22	3.25 ± 0.21	2.74 ± 0.22

In order to assess the effects of memory load on performance, different groups of subjects were tested under different memory load conditions. Indeed, whereas subjects tested with 12 pairs of cards demonstrated a significant difference in the number of moves necessary to complete the game, the time per move, and memory capacity when comparing the Subject rotates and Tray rotates conditions, subjects tested with 8 pairs of cards (Experiment 1) only showed a differential performance in the time per move when comparing these two conditions. Although the lack of a significant difference in the number of moves and short-term memory capacity between these two conditions in Experiment 1 was likely due to a floor effect, the differential performance of subjects in Experiment 2 (with 12 pairs of cards) demonstrates that memory load does indeed impact performance in the Memory game task as implemented here.

In sum, in our task, short-term spatial memory performance was greatest when an egocentric spatial frame of reference (i.e., centered on the individual) enabled subjects to encode and remember the image locations. When egocentric information was not reliable, short-term spatial memory performance was greater when an allocentric representation of image locations based on their relationships with distant objects in the surrounding environment was available, as compared to when the image locations could be encoded using only the relationships between the individual images on the testing tray, and whose positions were incoherent with distant objects in the surrounding environment.

4.1. Comparison with previous studies

Experiments carried out by Simons and Wang [15] evaluated the ability of young adults to detect the displacement of one object from an array of five physical objects located on a tabletop placed in front of the subjects. Similar to what we found in the current study, they reported that viewpoint changes caused by rotating the array of objects in front of a subject were not equivalent to viewpoint changes caused by the subject moving around the array. When distant objects in the surrounding environment were visible (i.e., lit

experimental room; Simons & Wang's Experiment 1), display rotation produced a decrease in the recognition of the displaced object, whereas moving the observer to produce the same amount of viewpoint change did not. When subjects moved around the array in the dark, so distant objects present in the surrounding environment were not visible (Experiment 2), subjects' recognition performance was decreased as compared to the condition in which they moved in a lit environment (Experiment 1). This suggests that the locations of the objects displayed on the tabletop were encoded in an allocentric spatial representation including their interrelationships with distant objects in the surrounding environment (see also [14] for similar evidence). Interestingly, in both Experiments 1 and 2 by Simons and Wang [15], there was a larger decrease in spatial recognition performance when the array was rotated in front of the subject who viewed the array from the same unchanged position, as compared to the condition in which the subject moved around the table to a different viewing position. Their tabletop rotation condition was similar to our current condition in which stationary subjects rotated the tray (Tray rotates), where we found task performance and memory capacity to be the lowest, and processing time to be the longest, as compared to the three other testing conditions. Accordingly, in Simons and Wang's study, performance in the tabletop rotates condition was also lower compared to the condition in which the tray was rotated at the same time and in the same direction as the subject moved (Experiments 1 and 2), thus leading to an unchanged viewpoint of the array of objects from the subject's new position in the room (i.e., enabling an egocentric coding of object locations). In a third experiment, subjects were passively displaced on a fast spinning, wheeled chair after viewing the array, so that they could not use self-generated movement information (i.e., vestibular, proprioceptive and somesthetic) to keep track of their own displacement within the illuminated experimental room. In this condition, subjects' performance was lower than when the subjects actively displaced themselves, but was significantly higher as compared to when a change in viewpoint was produced by the rotation of the array itself (Experiments 1 and 2).

Follow-up experiments carried out by Wang and Simons [16] investigated whether the results of their first study were due to the subjects' lack of active control over the displacement of the viewed array. In their second study subjects could thus either watch, or, similar to our experiment, actively move the array themselves (albeit with a pole that was attached to and extended out from the table). The results from this second series of experiments [16] essentially replicated their initial findings [15]. As compared to the control condition, the ability of subjects to detect changes in the location of a single object from an array of five physical objects placed on a tabletop was lower even when the subject produced the rotation of the array, but did not differ when the subject was displaced around the array that remained stable in relation to the surrounding environment (i.e., the experimental room), or when subjects were moved passively (i.e., wheeled without disorientation) to a new viewing position.

In sum, Simons and Wang's experiments suggested a clear hierarchy in the reliance on different spatial frames of reference to encode object locations. Similar to our own experiments, the rotation of the array of physical objects placed in front of a subject, within a stable surrounding environment, produced the greatest impairment in task performance. In contrast to our current findings, however, the displacement of their subjects while the array of objects remained stable within the testing room produced no impairment in task performance in the presence of distal environmental cues (as compared to when neither the subjects or the array moved). These differential results might simply be due to a lack of sensitivity of their task [15,16], as our current findings clearly showed decreased performance and lower short-term spatial memory capacity in the Subject rotates condition in

which egocentric and allocentric spatial frames of reference were dissociated, and only allocentric, but not egocentric, information was available for subjects to encode and remember the image locations.

Burgess et al. [14] carried out a study using a design similar to that used by Simons and Wang, in order to further evaluate the reliance on an allocentric spatial representation encoding the object locations in relation to distant external cues present in the surrounding environment. Although their experimental design included a number of additional conditions, we limit our discussion to the four conditions (no change, subject, subject and table, table; [14]) that are directly relevant for our own results. Consistent with our current findings, performance was highest in their no change condition (corresponding to our stationary condition); similarly, performance was lowest when the array of objects was rotated out of view of the subject who remained in the same position between presentation and testing (table condition; NB, the subject was informed that the array of objects was rotated). Consistent with the results of Simons and Wang and our current findings, Burgess et al.'s experiments revealed that rotating an array of objects in front of an observer is not equivalent to the observer moving around the array of objects. We will consider the implications of these findings below.

In addition, in the Burgess et al.'s experiment, although performance was lower in the subject rotates conditions as compared to the no change condition, it was, surprisingly, significantly higher than the performance observed in the subject and table rotate condition (again, subjects were informed of the rotation, but could not see it happening). Their results therefore contrast with our own findings that the stationary, and the subject and tray rotate conditions (i.e., two conditions in which the egocentric coding of image locations was reliable) did not differ in terms of task performance or short-term spatial memory capacity, and that these parameters were significantly higher in these conditions than in the other two conditions in which an egocentric coding of the image locations was not reliable. Such differences might be due to the fact that the subjects of our current study manipulated the images to be remembered and were not simply viewing an array of objects from a (relatively short) distance. Efferent copies of motor programs, as well as proprioceptive information regarding arm and hand positions might contribute to the preferential coding of object locations in an egocentric coordinate system, at least over a short period of time (see [8] for similar arguments). This could explain how the relative weight of different spatial frames of reference (i.e., egocentric versus allocentric) can vary based on differences in experimental design or testing conditions.

One additional difference between our experiment and the Simons and Wang and Burgess experiments [14–16] was the fact that during our experiments, although the images were always turned faced down when viewpoint changes were taking place, the experimental tray was always kept within the subject's view. In contrast, for both Simons and Wang's and Burgess' experiments, the experimental array was always out of the subjects' view when viewpoint changes were taking place. This lack of visual contact with the experimental apparatus during manipulation may have ultimately necessitated twice as much updating for subjects once the array was displayed again. Specifically, in the subject and tray (table) rotate condition, the ability of subjects to visually track the movement of the tray (table) while moving might facilitate the updating process, and thus explain why we saw no difference between this condition and the stationary condition in our experiments; it might also explain why Burgess et al. saw poorer performance in the subject and table rotate condition (the subject must update their movement as well as that of the table) relative to the subject rotates condition (the subject must only update his own movement).

In sum, in accordance with the results of previous studies discussed above, our current findings obtained with a new experimental design in a real-world situation demonstrate fundamental differences in processing time and short-term human spatial memory capacity that are dependent on the availability or coherence of spatial information derived from different sensory modalities encoded in egocentric or allocentric coordinates.

4.2. As the world turns: a neurobiological perspective

The major aim of our study was to determine whether human subjects' reliance on different spatial frames of reference (i.e., egocentric versus allocentric) affects their performance, decision time and memory capacity in a short-term spatial memory task performed in the real world. A second aim was to determine and compare the short-term spatial memory processes in two allocentric conditions that have often been considered interchangeable when studying human spatial memory: namely (1) when a subject moves about in a stationary environment, and (2) when part of the environment moves in front of a stationary subject. In this section, we briefly discuss the results of studies on hippocampal place cells in rodents and path integration in humans, in order to consider the implications of our current findings in a neurobiological context.

Electrophysiological studies in rats have shown that pyramidal neurons in the CA1 region of the hippocampus not only fire in relation to a rat's location in an open-field environment ("place cells"; [20]), but also integrate converging inputs from different sensory modalities to form relational representations of experience [20,21]. In the spatial domain, the receptive fields of these neurons, called place fields, represent a portion of the environment in which the rat moves and where an individual neuron fires preferentially. Place cell firing is the result of the integration of spatial information derived from different sensory modalities, which in normal conditions are coherent and contribute to the high selectivity of place cell firing activity [22]. When different sources of spatial information are incoherent (e.g., the movement of distal visual objects in the surrounding environment versus self-generated movement information derived from vestibular, proprioceptive and motor efferent copy information), the control of place cell activity depends on the relative discrepancy between the different sources of information [23]. Interestingly, when some self-generated movement information is eliminated, e.g., by moving a rat passively in the environment, place fields are larger than when a rat is actively moving in the same environment [24]. Similarly, when the environment is rotated around a stationary rat, and thus only the movement of distal visual objects suggests movement, the spatial information content per spike of place cells is reduced by 50% [25]; i.e., an experiment similar to virtual reality, or the experience one can have when sitting in a stationary train, watching an adjacent departing train. Thus, altogether these experiments indicate that the ensemble of spatial information derived from different sensory modalities contributes to the selectivity of place cells firing, and that the highest selectivity is obtained when all sources of spatial information are coherent, as is normally the case in the real world.

Recent experiments on path integration in humans also suggest that multiple sources of redundant spatial information are weighted and integrated based on their respective reliability to support behavior [26,27]. In a virtual reality study by Kearns et al. [27], although subjects wearing virtual reality goggles were able to use optical flow alone to complete a path integration task (i.e., moving through the virtual environment using a joystick and return to their starting point after a right triangle outward journey), the subjects' responses became more consistent if they were actively walking. Indeed, they exhibited a decreased variability in performance when self-generated movement information (generated by the subjects moving simultaneously in the real environment) was consistent

with the visual information displayed via the virtual reality goggles. Similarly, recent experiments by Campos et al. [26] demonstrated that although adult humans can use optical flow alone to estimate the distance travelled in a flat open outdoor environment, self-generated movement information contributes more to the correct estimation of distance than optical flow.

In sum, these studies indicate that experimental investigations of human spatial behavior in real world environments are critical, and provide a necessary perspective to interpret findings obtained in artificial environments such as virtual reality, and comprehend the functions of different regions of the human brain in spatial learning and memory.

5. Conclusion

Our current findings, together with the work of others discussed above, suggest that classical neuropsychological tests and virtual reality investigations of human spatial memory reflect limited approximations of human spatial memory in the real world. We certainly do not argue against the importance of conducting clinical or experimental investigations of human memory using standard neuropsychological tests or virtual reality technology. However, experiments carried out with tabletop designs or virtual reality technology artificially restrict the sensory information normally available to human subjects which is fundamental to building multimodal, relational representations of their surrounding environment. The results of such experiments and their contributions to further our understanding of human spatial memory processes should therefore be considered with these limitations in mind. Nevertheless, it is imperative that experiments in which subjects can move about freely in a real-world environment, and therefore perceive and integrate coherent visual, vestibular, proprioceptive, motor efferent copy, somesthetic and auditory information, be carried out in order to fully understand human spatial memory processes in real life [13].

Acknowledgements

This research was supported by the Swiss National Science Foundation: grants PMPDP3-122844 and PMPDP3-128996 to PBL and grants PP00A-106701 and PP00P3-124536 to PL. We thank all of the participants, as well as Céline Favrod, Lucas Jaquet, Melody Laeser, Cindy Lawrence, Nicolas Horowitz, and Chana Thuraiatnam for technical assistance.

Appendix A. Supplementary data

Supplementary data associated with this article can be found

References

- [1] O'Keefe J, Nadel L. The hippocampus as a cognitive map. Oxford: Clarendon Press; 1978.
- [2] Schenk F, Grobety M-C, Lavenex P, Lipp H-P. Dissociation between basic components of spatial memory in rats. In: Alleva E, Fasolo A, Lipp H-P, Nadel L, Ricceri L, editors. Behavioural brain research in naturalistic and semi-naturalistic settings. NATO ASI series, series D, behavioural and social sciences. The Netherlands: Kluwer; 1995. p. 277–300.
- [3] Lavenex P, Schenk F. Influence of local environmental olfactory cues on place learning in rats. *Physiology & Behavior* 1995;58:1059–66.
- [4] Lavenex P, Schenk F. Integration of olfactory information in a spatial representation enabling accurate arm choice in the radial arm maze. *Learning & Memory* 1996;2:299–319.
- [5] Lavenex P, Schenk F. Olfactory cues potentiate learning of distant visuospatial information. *Neurobiology of Learning and Memory* 1997;68:140–53.
- [6] Lavenex P, Shiflett MW, Lee RK, Jacobs LF. Spatial versus nonspatial relational learning in free-ranging fox squirrels (*Sciurus niger*). *Journal of Comparative Psychology* 1998;112:127–36.

- [7] Banta Lavenex P, Amaral DG, Lavenex P. Hippocampal lesion prevents spatial relational learning in adult macaque monkeys. *Journal of Neuroscience* 2006;26:4546–58.
- [8] Banta Lavenex P, Lavenex P. Spatial memory and the monkey hippocampus: not all space is created equal. *Hippocampus* 2009;19:8–19.
- [9] Morris RGM, Garrud P, Rawlins JNP, O'Keefe J. Place Navigation impaired in rats with hippocampal-lesions. *Nature* 1982;297:681–3.
- [10] Bird CM, Burgess N. The hippocampus and memory: insights from spatial processing. *Nature Reviews Neuroscience* 2008;9:182–94.
- [11] Bohbot VD, Iaria G, Petrides M. Hippocampal function and spatial memory: evidence from functional neuroimaging in healthy participants and performance of patients with medial temporal lobe resections. *Neuropsychology* 2004;18:418–25.
- [12] King JA, Burgess N, Hartley T, Vargha-Khadem F, O'Keefe J. Human hippocampus and viewpoint dependence in spatial memory. *Hippocampus* 2002;12:811–20.
- [13] Banta Lavenex P, Lavenex P. Spatial relational learning and memory abilities do not differ between men and women in a real-world, open-field environment. *Behavioural Brain Research* 2010;207:125–37.
- [14] Burgess N, Spiers HJ, Paleologou E. Orientational manoeuvres in the dark: dissociating allocentric and egocentric influences on spatial memory. *Cognition* 2004;94:149–66.
- [15] Simons DJ, Wang RF. Perceiving real-world viewpoint changes. *Psychological Science* 1998;9:315–20.
- [16] Wang RF, Simons DJ. Active and passive scene recognition across views. *Cognition* 1999;70:191–210.
- [17] Moffat SD. Aging and spatial navigation: what do we know and where do we go? *Neuropsychology Review* 2009;19:478–89.
- [18] McBurney DH, Gaulin SJ, Devineni T, Adams C. Superior spatial memory of women: stronger evidence for the gathering hypothesis. *Evolution and Human Behavior* 1997;18:165–74.
- [19] Sykes Tottenham L, Saucier D, Elias L, Gutwin C. Female advantage for spatial location memory in both static and dynamic environments. *Brain and Cognition* 2003;53:381–3.
- [20] O'Keefe J, Dostrovsky J. The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Research* 1971;34:171–5.
- [21] Eichenbaum H. Hippocampus: cognitive processes and neural representations that underlie declarative memory. *Neuron* 2004;44:109–20.
- [22] Wiener SI, Berthoz A, Zugaro MB. Multisensory processing in the elaboration of place and head direction responses by limbic system neurons. *Brain Research Cognitive Brain Research* 2002;14:75–90.
- [23] Knierim JJ, Kudrimoti HS, McNaughton BL. Interactions between idiothetic cues and external landmarks in the control of place cells and head direction cells. *Journal of Neurophysiology* 1998;80:425–46.
- [24] Gavrilo VV, Wiener SI, Berthoz A. Discharge correlates of hippocampal complex spike neurons in behaving rats passively displaced on a mobile robot. *Hippocampus* 1998;8:475–90.
- [25] Terrazas A, Krause M, Lipa P, Gothard KM, Barnes CA, McNaughton BL. Self-motion and the hippocampal spatial metric. *Journal of Neuroscience* 2005;25:8085–96.
- [26] Campos JL, Byrne P, Sun HJ. The brain weights body-based cues higher than vision when estimating walked distances. *European Journal of Neuroscience* 2010;31:1889–98.
- [27] Kearns MJ, Warren WH, Duchon AP, Tarr MJ. Path integration from optic flow and body senses in a homing task. *Perception* 2002;31:349–74.