Active, passive and snapshot exploration in a virtual environment: influence on scene memory, reorientation and path memory

Florence Gaunet, Manuel Vidal, Andras Kemeny, Alain Berthoz

Research report

Active, passive and snapshot exploration in a virtual environment: influence on scene memory, reorientation and path memory

Florence Gaunet[a,*], Manuel Vidal[a], Andras Kemeny[b], Alain Berthoz[a]

[a]*Laboratoire de Physiologie de la Perception et de l'Action, Collège de France-CNRS, Paris, France*
[b]*Driving Simulation & Virtual Reality Group, Research Department, Renault, Paris, France*

We investigated the importance of active, passive and snapshot exploration on spatial memory in a virtual city. The exploration consisted in traveling along a series of streets. 'Active exploration' was performed by giving directions to the subject who controlled his displacement with a joystick. During 'passive' exploration, the travel was imposed by the computer. Finally, during 'snapshot exploration', simple views of the scene were presented sequentially every 4 m. Travel velocity was the same in all cases. The three visual exploration modes were compared with three spatial memory measures: (1) scene recognition, (2) at the end of the path, reorientation toward the departure point and (3) drawings of the path shape. Scene recognition and estimation of the direction of the starting point of the path were not affected by the mode of exploration. In contrast, reproduction of the shape of the path was affected: the errors of reproduction were greater for the snapshot exploration than for the other two conditions; there was no difference between the other two conditions. These results suggest that (1) 2D image features from a visual scene are memorized. Moreover, (2) pointing towards the origin of the path relies on motion duration integration or a frame of reference integrated during displacement. Finally, (3) drawing the path shape involves a deliberate reconstruction process.

Theme: Neural basis of behavior

Topic: Cognition

Keywords: Spatial environment; Visual exploration; Navigation; Scene recognition; Path integration

1. Introduction

Few human activities are as important for survival as navigation, but currently our understanding of how humans navigate effectively is relatively crude. Navigation is the process of directing one's locomotion toward a goal; it is based on sensory or verbal guidance or on an internalized representation of an environment. One framework that has been useful in conceptualizing navigation assigns distinct roles to the process of navigating with reference to landmarks (*piloting*) and the process of maintaining an updated estimate of one's current position, orientation and velocity between landmarks, i.e. integrating displacements (*path integration*) [6,14,24].

The latter process, *path integration*, excludes methods of determining one's position and orientation that rely upon direct sensing of landmarks; instead, path integration relies upon sensory information that specifies one's self motion [31]. The process of path integration can be further subdivided into updating during linear displacements and updating during rotational displacements. Changes in orientations can be sensed directly by vision and kinesthesis. Vestibular cues from turns and translations of the head and body contribute to path integration [20,21]. Moreover, wayfinders evidently use optic flow as a direct cue to velocity [11,40], and hence distance, for their ability to maintain orientation is reduced if their peripheral vision is occluded [37]. The function of path integration is to encode egocentrically self-location while traveling. However, cumulative errors emerge from path integration.

*Corresponding author. Present address: Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur-CNRS, Orsay, France. Tel.: 133-1-6985-8014; fax: 133-1-6985-8088.
E-mail address: gaunet@limsi.fr (F. Gaunet).

The other process, *piloting*, can be used to minimize these cumulative errors. As it is typically defined, piloting is an egocentric process that relies on direct sensing of the distance and bearing of landmarks in the environment. It is based on aiming toward landmarks ('guidance' in the terms of O'Keefe and Nadel [32]) [14,24]. When using piloting to navigate, at certain locations along the route, we localize ourselves with respect to particular landmarks for the purpose of making decisions about what trajectory to follow next. Integration and recognition of spatial configurations is necessary for decision points to be identified and for the appropriate action to be determined at each decision point [5,11,28]. Scene recognition is thus of principal interest for piloting. When spatial configurations are integrated, different views are perceived and these views are more or less linked by active displacement around the configuration.

The visual information that can be used for memorizing displacement and for scene recognition can be acquired in a variety of ways; it can be obtained both in the course of active exploration of an environment and during passive displacements. Similarly, it can be obtained in a continuous fashion or intermittently. With regard to the roles of active versus passive control, Appleyard [2] submitted hundreds of inhabitants of a city to different types of exploration of urban areas – involving different types of visual information – and then interviewed them. He asked them to draw sketch maps of the entire city and of their local neighborhood. He found that people who mainly drove automobiles to get around the city tended to draw better maps than those who traveled exclusively on buses. This suggests that active exploration plays an important role in determining how vision is used to elaborate spatial knowledge. Although other experiments have contrasted active and passive exploration, the results have been mixed; the same inconsistent results are found when comparing exploration using slides and a movie of the path (for a review, see Ref. [33]). For instance, in a study by Péruch, Vercher, and Gauthier [34], subjects actively (self-initiated movement) or passively (pre-recorded tour) explored a virtual environment composed of target cubes and obstructing inner walls. After exploration, the observer had to reach, using the shortest path, a specified target not visible from the starting point. The performance (mean distance traveled and completion time) was better after active than after passive exploration. Moreover, in a recent study by Attree et al. [3], an active group (yoked with a passive group) visited interconnected rooms in a house, each containing several objects. Memory of spatial layout (relative positioning of objects) was better for active explorers, while passive exploration enhanced object memory. In a similar vein, Wilson et al. [41] compared experimental groups which were active or passive in directing exploration. No difference on pointing error was found between all experimental groups. No studies tackled the role of interactivity and of the control of exploration during visual exploration (active vs. passive exploration) neither the role of continuity of the visual exploration (passive and snapshot exploration) within the same experimental simulation device for studying the two basic tasks of navigation: scene recognition, pointing toward the origin and path shape reproduction. Contrasting these conditions may help to dissociate the effect of the exploratory modes and thus understand the controversial data on the basis of the type of task performed.

The present experiment is aimed at disentangling the relative contribution of visual information, acquired actively, passively and intermittently from the exploration of a large scale environment. It examines how the accuracy of displacement memory and of scene recognition varies as a function of three types of visual exploration. In the first, the subject actively explores the environment. The visual stimulation generated by the exploration is produced by the motor activity of the subject. During the second type of visual exploration, the subject passively explores the environment; the visual stimulation generated by the exploration is not produced by the activity of the subject. Rather, the subject is passively led around an environment, and is submitted only to the visual stimulation relative to the exploration. Finally, the third type involves passive exploration in which only a few static images of the exploratory path are presented. These three modes of exploration, performed in a simulated large-scale urban style environment, can be used to draw inferences about how navigation is controlled in the real world. Active exploration with a joystick share some important aspects with walking in a real environment. There is a tight linkage between visual self-motion information and motor activity, just as there is in the real world. Furthermore, the observer has active control of the exploratory path, and thereby is able to self-direct, in a very natural way, the process of gathering visual information to determine the extent of self-motion memorization and to recognize scenes. Control of action, visuo–motor interaction and decision of action are involved. In contrast, these three components are not involved in passive exploration. Comparing active and passive exploration allows us to gain insight into the role of these components. Finally, during snapshot exploration, visual information for scene recognition and perception of the path trajectory is considerably reduced – intermittent – compared to the condition that involves dynamic passive exploration. The comparison of the latter two conditions allows us to infer the importance of continuous visual stimulation, i.e. to what extent the brain is able to infer spatial knowledge from intermittent exploration. The comparison of the three conditions of exploration and their effect on spatial behavior will allow us to investigate the relative importance of different modes of acquiring visual information for scene recognition, path shape memory and determination of the origin of the path. The results will also have relevance for determining the underlying cognitive processes used to perform the three tasks.

In the present study, the subjects attempted to learn a route through a virtual world in a driving simulator, using one of the three exploration modes mentioned above. Our main interest concerns the role of sensori-motor information in scene recognition and memory of displacements. For that purpose, the decisional component of active exploration was removed: verbal route directions replaced directing activity. The active vs. passive exploration and the passive vs. snapshot exploration comparisons inform, respectively, about the role of motor activity and the visual continuity of exploration. Immediately after being exposed to the route during the exploration phase, the subjects' spatial knowledge of the virtual environment was tested in several ways. In the first test, subjects were presented with successive scenes of the virtual world; the scenes were either taken from the path previously traveled or from other paths of the virtual world, not traveled by the subjects. Subjects used the keyboard to indicate as quickly as possible whether they had seen the scene during the exploration phase. In another test, subjects were exposed to another path using a different exploration mode, and then were asked to orient their point of view towards the starting point of the path, using a joystick. Subjects were next asked to draw the profile of the path traveled on a sheet of paper.

2. Material and methods

2.1. *Subjects*

Subjects were recruited from the University of Paris and were paid for participating in the experiment. They all were fairly experienced at joystick manipulation, and most of them had experience using joysticks to play video games. Their vision was normal, corrected if necessary. They were right-handed.

Forty eight males and females (24 males, 24 females) participated in the experiment, ranging in age from 21 to 38. (age, Mean: 25.9; S.D.: 4.53). Before the experiment, subjects were informed of the overall goals of the experiment but were naive concerning the specific hypotheses. Experiments were undertaken with the understanding and written consent of each subject.

2.2. *Equipment and experimental set up*

2.2.1. *Apparatus*

One PC (Pentium II, 300 MHz, 192 Mb RAM and a 3D accelerator card: 3DFX Voodoo II) and three computer screens were used, one by the subject (Fig. 1) and two by the experimenter. The first screen was used to display the virtual environment to the subject. The frame rate of the continuous stimuli was on average 18 Hz. The second screen displayed the same stimuli and was used by the experimenter to control what the subject experienced visually. Finally, the third screen displayed the software controls. The experimenter setup was located on a table 2 m away from the subject setup. The screen used by the subject (19 in.) was surrounded by a curtain covering the subject's head. No light from the room was visible to the subject. The subject's head was immobilized with a chin-rest, with the eyes 50 cm from the screen. This provided a field of view of 388 wide and 308 high. A joystick was fixed on the table, in front of the subject and 30 cm from the screen. A sheet of paper and a pen were near the joystick.

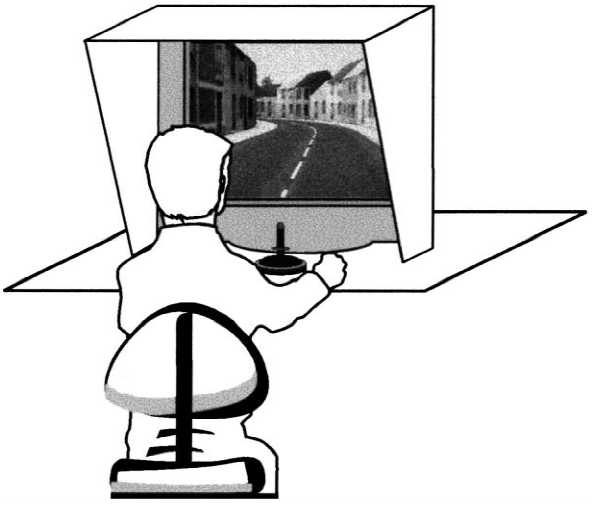

Fig. 1. Experimental set up for the subject: computer screen displaying the stimuli, joystick, chin-rest (not visible) and curtain, here removed from the head of the subject.

2.2.2. *Software*

We used the CaTS driving simulator software developed by the Driving Simulation & Virtual Reality Group, Research Department, Renault. The software simulates navigation along roads in villages, towns, countryside, and highways. Navigation is only possible on roads and footpaths. Software modules manage traffic lights as well as the predefined displacements of vehicles and pedestrians in the environment. Other modules allow the user to define automatic behaviors such as passive displacements along a pedestrian route.

Two different simulated environments (see Fig. 2) were used and one path in each was chosen for the experiment (P1 and P2). The simulated path lengths of P1 and P2 were 437 and 458 m, respectively. Fig. 4 shows few views of the appearance of the environment. This environment was entirely fictional, no tall landmarks could serve as azimuthal reference, and buildings were tall enough to prevent viewing over longer distances.

In the exploration phases, navigation along roads during active exploration involved using a Microsoft joystick – this is the first navigation module. The subject controlled a 'virtual eye' that had a fixed height from the ground (1.80

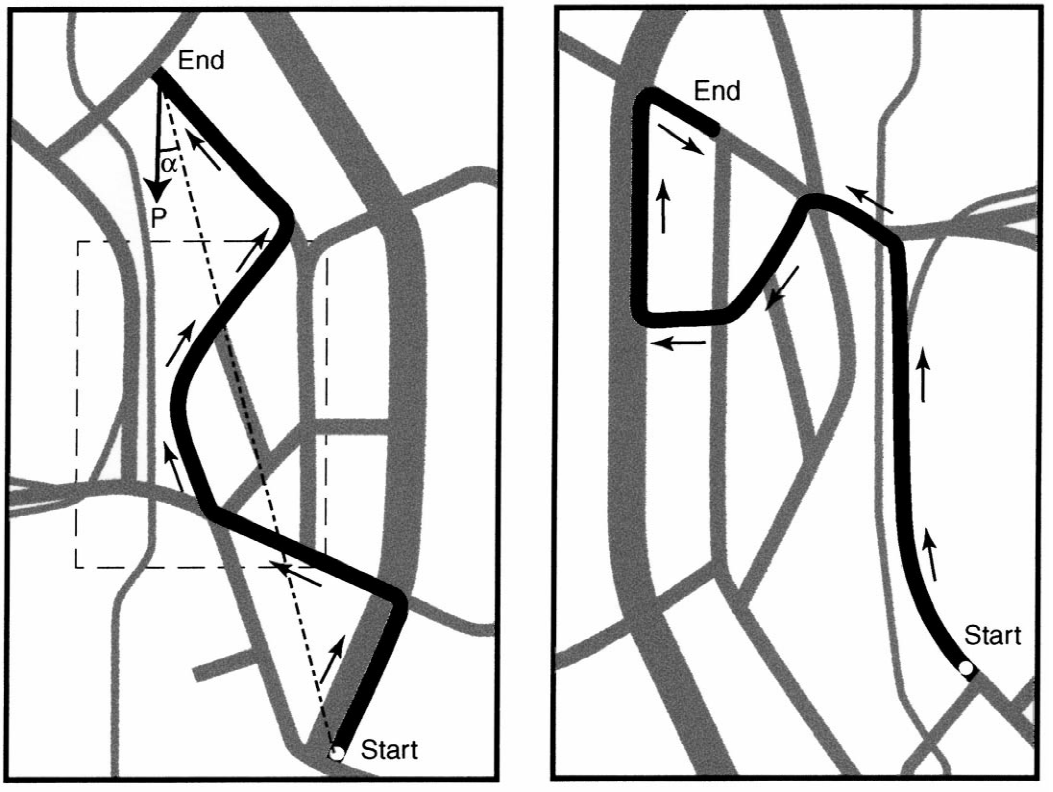

Fig. 2. Top view of the two paths (in black) used for scene recognition (P1, left) and path integration (P2, right). a is the angular error of pointing between the actual direction of the starting point and the direction pointed by the subject during the task. The square represents the part of the space detailed in Fig. 4.

m). As the subject pushed the joystick forward, the simulated speed increased proportionally to a maximum of 15 km/h when the joystick was pushed all the way forward. Although speeds slower than 15 km/h were possible, subjects were required to use the maximum speed. The speed was proportional to the angle of the joystick. The maximum velocity of the turns was 408 by seconds. Navigation during passive exploration involved simulated travel along a path at a constant speed of 15 km/h – this is the second navigation module. A third navigation module was designed for the snapshot exploration condition, which presented static snapshots along a path. A snapshot of the path was presented at regular spatial intervals (960.5 m along the road); each picture was on screen for 2.25 s and a new picture appeared with almost no delay (0.01 s), simulating a displacement speed of 15 km/h. A total of 47 snapshots were used.

In the test phases, a first software module displayed a selected number of static views from the database; some were selected randomly along the path and were called 'known snapshots'. Others were taken from viewpoints that the subject had not experienced in the same database and were called 'unknown snapshots'. Subjects indicated whether or not they had seen each view before by clicking the left or the right button of a mouse. Both accuracy and reaction time were scored. Clicking the mouse to respond triggered the presentation of the following test snapshot. A second software module allowed the subject to orient his/her viewpoint in the virtual world toward a place named by the experimenter by pushing the joystick to the left or to the right; translations were disabled. The final viewpoint orientation of the subject in the virtual world was recorded by the computer.

2.3. *Experimental procedure*

2.3.1. *Procedure*

The subject sat in a chair, with his/her chin on the chin-rest. The curtain was put over the head and body of the subject. The room was dark, so that no light entered the experimental set up of the subject. Fig. 3 describes the time course of the experimental session for each subject. The experimental session lasted between 1 h and 1 h 15 min.

The path P1 was used for the *Scene recognition* task, P2 for *Pointing toward the origin and shape path reproduction* tasks. Each subject performed all three tests, and experienced a different exploration condition for each of the two paths. This yielded six possible combinations of

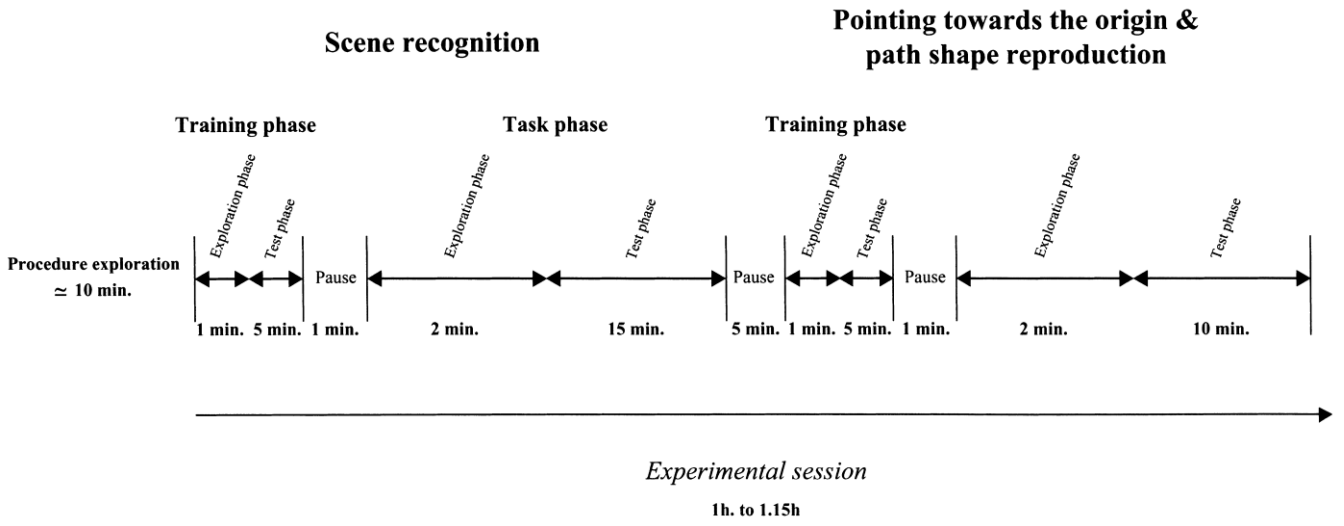

Fig. 3. Schematic representation of the time course of the whole experimental session one subject had to perform. *Scene recognition training* during which the subject experienced a shortened corresponding task; this task consists in experiencing a familiarization phase he/she was trained on earlier (active, passive or snapshot exploration) and a scene recognition phase (one test, 60 snapshots). The subject rested for few minutes. *Pointing towards the origin and path shape reproduction training* during which the subject experienced a shortened corresponding task; these tasks consist in experiencing a familiarization phase he/she was trained on earlier (active, passive or snapshot exploration) and a path integration phase (two tests).

conditions across the two paths: Active–Passive, Passive–Active, Passive–Snapshot, Snapshot–Passive, Snapshot–Active, Active–Snapshot; equal numbers of subjects participated in each combination.

A short training phase, including an exploration and test portion, preceded each exploration phase, but these data were not analyzed. The training phases were performed just before the corresponding stage with the same exploration condition. A different environment database in which two other paths were defined was used for the training. The exploration and test paths in the training phase were considerably shorter than those analyzed in the main experiment.

Exploration phase. Three different exploration phases were used. During *Active exploration*, subjects traveled along a path using the joystick as described above. The experimenter guided the subjects through the city using verbal commands such as: 'turn left', 'take the second street on your right', 'go straight ahead' or 'stop at the next crossroad'. Fifteen meters before each decision point in the route, the experimenter indicated verbally which direction to take, so that the subjects did not slow down their speed. Subjects were told to use the joystick at the maximum speed, which was equivalent to 15 km/h. The experimenter measured the path duration and verified that it corresponded to path duration while moving at 15 km/h, plus or minus 10 s. Path duration was 1 min 45 s for P1 and 1 min 50 s for P2.

During *Passive exploration*, subjects observed a prerecorded passive displacement along a path, at a speed of 15 km/h. Finally, *snapshot exploration* was based on the same principle as *passive exploration* except that the visual display was not continuous but sampled. An example of several views is presented on Fig. 4. The subjects were told about the spatial interval regularly between the views.

Test phase. Three test phases were designed. The first was a *scene recognition test*. It used the first module described in the Software section. The principle of this task is to display snapshots taken along the path and out of the path, but in the same virtual world to keep known (30) and unknown (30) snapshots homogeneous in terms of visual cues. Half of the known snapshots were sampled from the snapshot pool presented during the snapshot exploration. The other half were sampled between two snapshots presented during the snapshot exploration mode. The subjects were told that the scenes to be recognized were randomly chosen along the path. The order of these 60 pictures was counterbalanced between subjects. The subject was asked to press a button of the mouse as quickly as possible without sacrificing accuracy of performance. If the picture was taken from the path previously traveled, the right button had to be clicked, if the picture had not been encountered along the path, the subject had to click left button. Clicking one of these two buttons triggered the next trial. The second test was a *pointing toward the origin*. It used the second module described in the Software section. The subject – located at the end of the virtual path traveled – was asked to orient his viewpoint by moving the joystick toward the direction of the origin of the path. When the subject was satisfied with his/her response, the screen went dark and the joystick was removed. During the third test, the subject had to draw – with a pen – on a sheet of paper the shape of the path, starting by the beginning of the route.

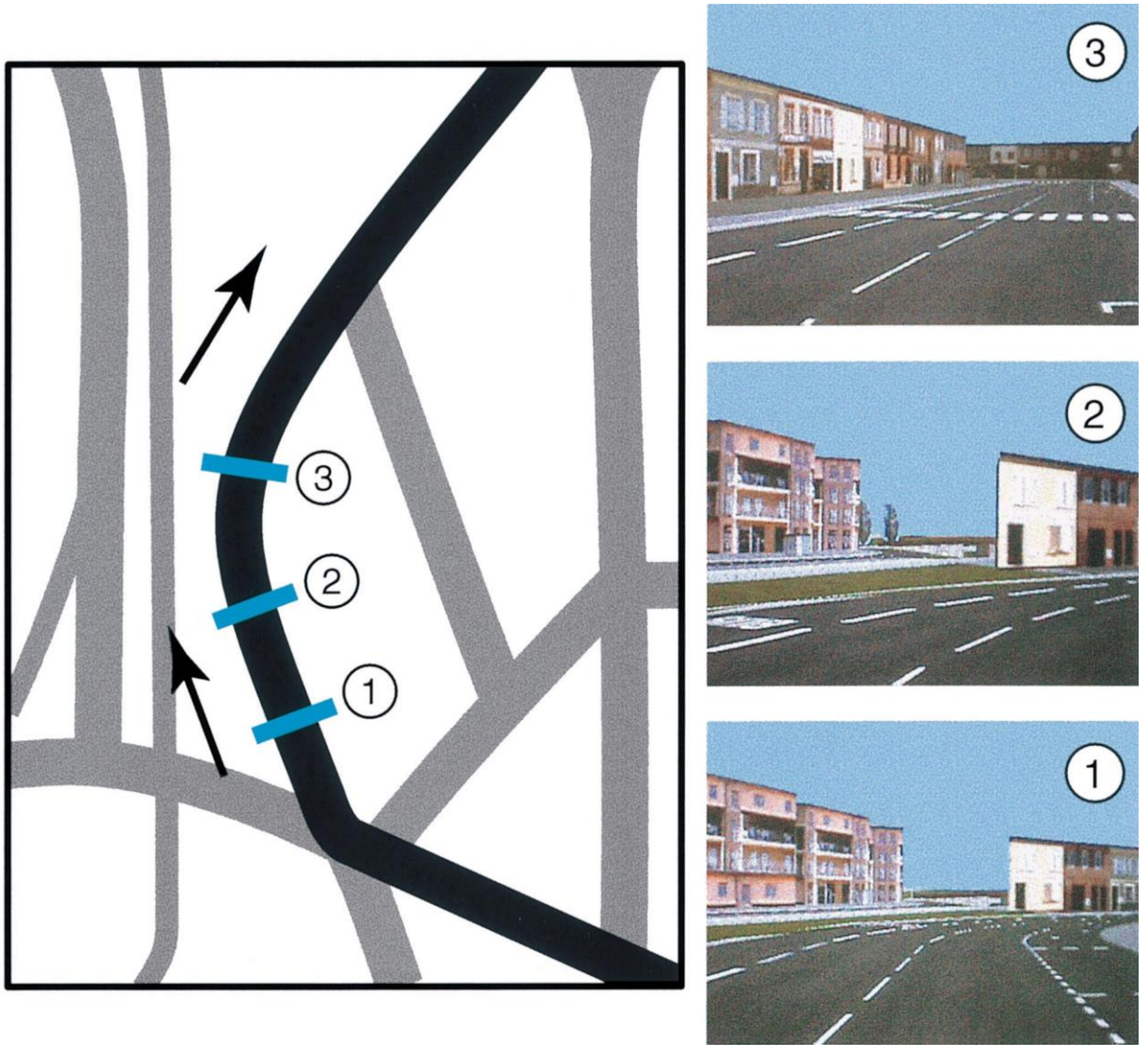

Fig. 4. Snapshot condition of exploration. The left side is a top view of the path and the right side represents three snapshots presented during the snapshot exploration, taken at the places marked on the map view figure.

This work was reviewed and approved by the French national ethics committee (CCPPRB [114-99).

2.4. *Data analysis*

In the scene recognition task, we analyzed the percentage of correct responses obtained for known scenes and the corresponding response latencies. When the subject did not recognize a known image, no corresponding latency was recorded. Percentages of correct responses and corresponding mean latencies were computed for the three experimental conditions.

When subjects oriented their viewpoint – with the joystick – towards the memorized starting point, the corresponding direction in the virtual world was recorded. From the angles a given by the direction of these pointings, we extracted the angular errors (corresponding to the absolute deviation between the real angular value and the one measured) (see Fig. 2). Mean error of direction was computed for the three exploration conditions.

Finally, drawings of path shapes were analyzed. We computed two error indices: one relative to distance reproduction named $\text{score}_{\text{Length}}$, and another one relative to angular reproduction named $\text{score}_{\text{Angle}}$. The way these indices were defined and computed as well as the choices made for their automatic extraction from scanned drawings is given below.

The drawings were first scanned. We have next estimated and marked the positions of the corners between each segment – using the intersection of the most significant tangent of each segment – with a standard graphic program. The figure below (Fig. 5) shows an example of drawing with its marked corners and the corresponding segments in gray.

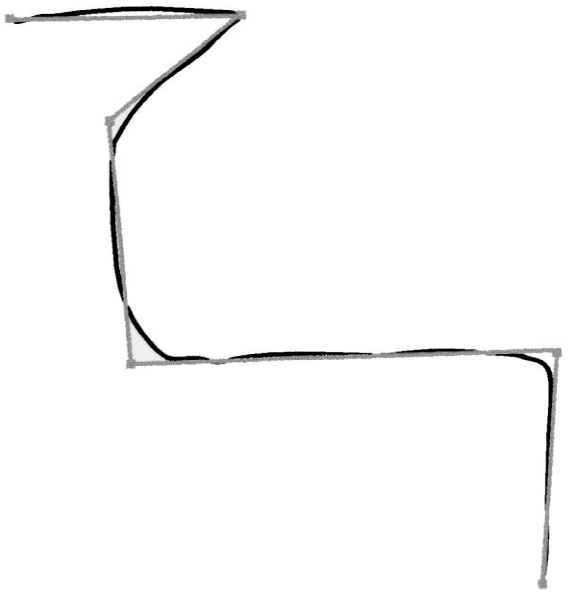

Fig. 5. Example of drawing with its marked corners (dots) and its corresponding segments (in gray).

Those marked drawings were then used as inputs for an automatic processing program that we have developed. It extracted for each drawing the relative length L_i (i5 1, . . . ,5) corresponding to the segment's measured length divided by the total drawing length, and the angular values a_i (i51, . . . ,4) between each segment pair.

Finally, two scores were computed from those previous values:

- $\text{score}_{\text{Length}} = \sum_{i=1}^{5} |L_i - L_i^{o}|$ where L_i^{o} is the real relative length of segment i,
- $\text{score}_{\text{Angle}} = \sum_{i=1}^{4} |a_i - a_i^{o}|$ where a_i^{o} is the real value of the angle i, with $|a_i - a_i^{o}| \in [0°, 180°]$.

Note that the smaller the score, the more the traveled path resembled the drawing.

For the drawings where a corner was missing, the program affected the length of the two adjacent segments half of the value of the single segment represented, and the angular value was set to 08. For the ones where a supplementary segment was drawn, the length and the angular value given by the last segment was compared to a null segment (null length and angle values). The scores were then estimated using these values. Therefore, drawing more or fewer segments than expected automatically increased the scores.

Mean errors of distance and angular reproduction scores were computed for the three exploration conditions.

Analyses of variance (ANOVA) – exploration condition (active exploration, passive exploration snapshot exploration)3sex (male, female) – were conducted on percentages of correct responses, latencies, pointing errors, and distance and direction error scores.

3. Results

3.1. *Scene recognition*

The percentages of correct responses and the mean latencies recorded for the three exploration conditions are shown in Fig. 6. The rates of correct responses (in %) are 74 (S.D.: 9), 69 (S.D.: 12) and 74 (S.D.: 9) and latencies (in seconds) are 3.00 (S.D.: 1.39), 2.82 (S.D.: 0.76) and 2.59 (S.D.: 0.67), respectively, for the active, passive and snapshot exposure. The range of data in the three conditions overlaps for both percentages of correct responses and latencies. These data suggest that the three conditions resulted in similar performance, for both rates of correct responses and latencies. The large standard error and deviation observed in the active condition is due to one

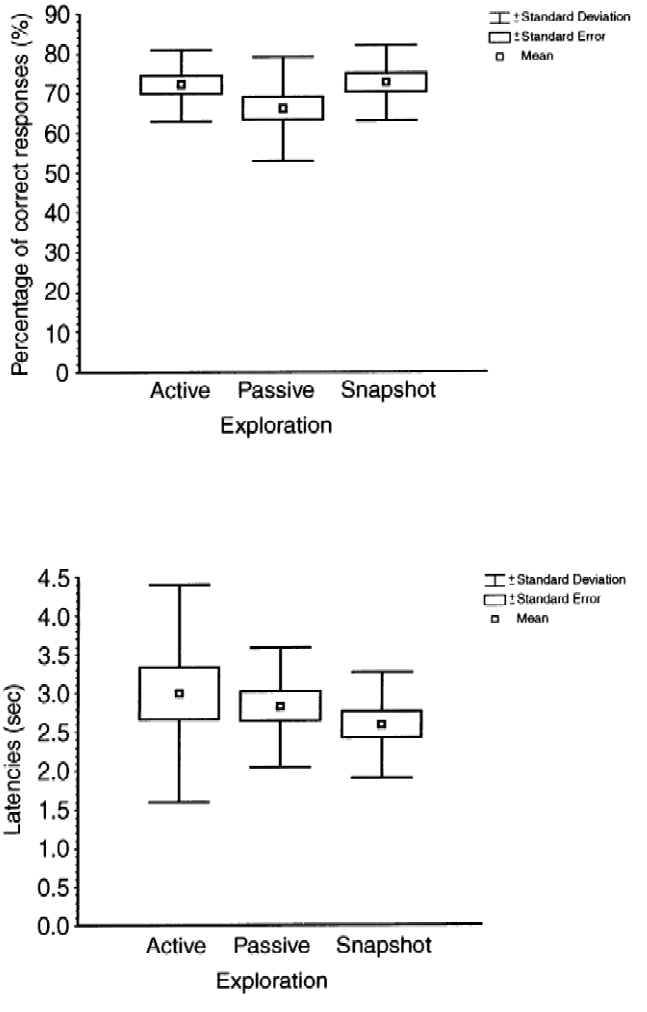

Fig. 6. Presents the percentages of correct responses (above) and the latencies for the correct responses (below), in the three exploration conditions.

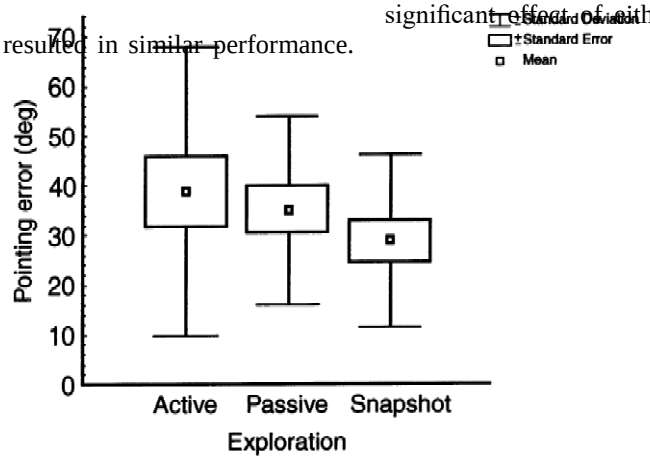

Fig. 7. Mean pointing errors to the starting point, in the three exploration conditions.

subject who took an inordinately long time to respond for one picture. He was faster during the recognition of the other pictures – averaging the mean of the group. We kept his data, and changed it to the mean of his performance plus two times the standard error of the latencies recorded during recognition of the other pictures.

The analysis of variance was performed on the data to test observation of similar performances following the three exploration conditions, for both the percentages of correct responses and latencies. Neither analysis showed a significant effect of either of the two factors, confirming that the three conditions resulted in similar performance.

3.2. *Path integration*

Mean pointing errors for the three exploration conditions are shown in Fig. 7. The means of a (in degrees) are 38.91 (S.D.: 29.09), 35.02 (S.D.: 18.84) and 28.87 (S.D.: 17.39), respectively, for active, passive and snapshot exploration. The data range overlaps for the three conditions of exploration.

The analysis of variance was performed on the pointing errors. Neither of the two factors significantly affected these data.

Fig. 8 shows three examples of the drawings of the path, for each of the three exploration conditions. Note that the drawings following the active and passive exploration resemble more the actual shape of path traveled than the drawing following the snapshot exploration does.

The analysis of mean angular and distance error scores of reproduction are quantifiable indices of these observations. Mean angular and distance errors of reproduction for the three exploration conditions are shown in Fig. 9. Mean angular and distance error scores are 84.02 (S.D.: 42.61) and 91.59 (S.D.: 28.28), 132.82 (S.D.: 45.83) and 27.07 (S.D.: 11.70), 27.43 (S.D.: 10.92) and 38.77 (S.D.: 20.77), respectively, for the active, passive and snapshot exploration. These data suggest that the active and passive explorations led to similar performances for both angular

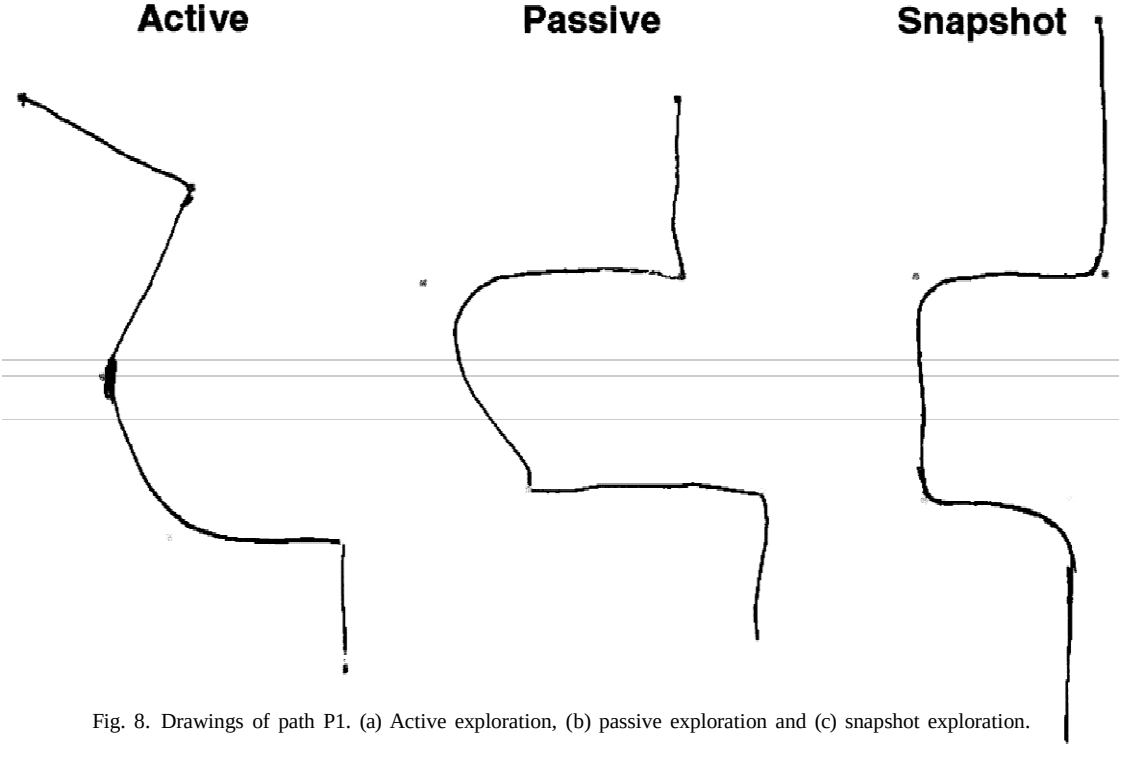

Fig. 8. Drawings of path P1. (a) Active exploration, (b) passive exploration and (c) snapshot exploration.

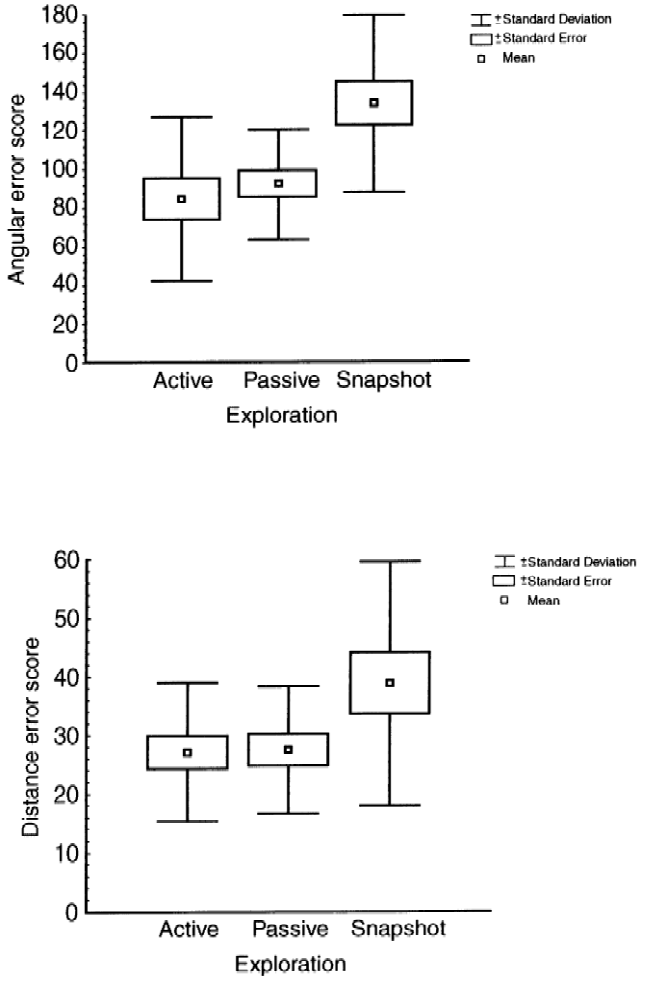

Fig. 9. Mean angular (above) and distance (below) error for drawings, in the three exploration conditions.

and distance error scores: distance and angular error scores are greater for snapshot exploration than in the other two conditions.

The analysis of variance performed on these data confirmed that indeed there was a significant effect of the exploration condition factor for distance reproduction, $F(1,2)55.88$, $P50.005$, and for angle reproduction, $F(1,2)54.53$, $P50.01$. There was no effect of the sex factor. Post-hoc analysis (Tukey test) performed on these data showed that the snapshot exploration was significantly different from the active and passive exploration for distance reproduction, $P50.004$ and 0.03, and for angle reproduction, $P50.026$ and 0.029. Active and passive exploration did not differ significantly for either response measure.

4. Discussion

In this study we showed that accuracy in scene recognition and estimation of the direction of the origin of the path performed in a virtual urban environment were not influenced by the exploration conditions (active, passive and snapshot exploration), i.e. neither by the motor activity that goes with visual exploration nor by the intermittent visual exploration. By contrast the reproduction of the path shape was affected. Total distance and angle reproduction error scores were greater following snapshot exploration than after active and passive explorations, i.e. there was an effect of intermittent visual exploration affect; however the latter two conditions did not differ.

There have been studies concerning the processes underlying *scene recognition* and *displacement memorization*. There are competing theories as to which visual spatial information is encoded to form spatial representations for *scene recognition*. One model considers that the subject moving through the environment extracts geometrical relations between objects and these are used to construct a single explicit structural description [29,30]. Alternatively, mental encoding of a scene would resemble the visual stimulus itself. That is, the representation of the visual scene is, initially at least, pictorial in nature and is determined principally by 2D image features [39]. In this sense the representation encodes spatial relations only implicitly. There exists supporting evidence for both theories and they may correspond to two short-lived mental processes.

The *memory of the displacement* performed in an environment has received considerable research attention. It has been assessed by two principal means [7,18,23,37,38]: orientation towards the origin of locomotion and reproduction of the path shape. In order to point to the origin of the path, maintaining an internal estimate of one's current position and orientation with respect to the origin of locomotion is needed. This can be done by continuously updating the relationships between observer's current position and origin of locomotion, forgetting all previously encoded egocentric relationships. This requires a low-cost memory load. Drawing the shape of a path is another way to externalize integration of the displacement. The length of each segment traveled, as well as changes of direction and the relative spatial organization of the path, must all be memorized and recalled. Based on these differences in the requirements of the two tasks, different cognitive processes could conceivably be used to perform each task. The work of Loomis et al. [26] supports this idea. In a triangle completion task, they found that the latency to initiate the return toward the origin increased with the complexity of the outbound path; this result is incompatible with updating the representation of the origin of the path. Moreover, the fact that, in this study subjects were able to retrace, means that they were able to maintain a history of their route. This study suggests that subjects form some representation of the outbound path in memory while navigating along a path.

What does the effect of the three exploratory modes on scene recognition and integration of the displacement

reveal about *the sensori-motor information processed* and what can we infer from these results about the *cognitive processes* involved in both behaviors? We shall first briefly justify the validity of the use of virtual environments and next compare the range of the data collected in our experiment with similar studies performed in real and virtual environments. The validation of simulation displays for studying spatial cognition has been discussed in many papers (see for instance Refs. [25,33]). These articles review and analyze a great deal of spatial cognition studies using virtual reality technology as a methodological tool. These two articles argue strongly that the most important properties of the spatial representations that underlie spatial behaviors are found in both real and virtual environments. For instance control of action which consists in aiming toward a goal (by navigation, Loomis and Knapp [27]) follow similar rules, in both real and virtual environments. That is also the case for counteralignment effect which consists in making larger errors when using a map misaligned with the environment than when using a map aligned with the environment [34] and for spatial updating which consists in updating relationships between self and objects while moving in the environment [23]. In addition, spatial knowledge acquired in virtual worlds transfers to actual worlds (for a review see Ref. [33]). Consequently, we consider that relevant sensory information is preserved in the present paradigm.

Although there are some procedural differences, the percentages of recognized scenes and the response latencies that we recorded (on average 70% and 2.7 s) are in the same range – though a bit higher – than those recorded by Gaunet and Berthoz [16] (on average 98% and 2.3 s) in a real environment and by Christou and Bültoff [9,10] (on average 85% and 1.6 s). Moreover, the pointing errors we recorded (38.91, 35.02 and 28.878) correspond to the range of performances collected earlier in real world [21,23,36] and in virtual worlds [7,17,19]. It is considerably more difficult to compare path shape reproduction between studies because the indices of errors differ from study to study. The conclusion of these observations is that the performances obtained for scene recognition and path integration are quite representative of performances obtained in other studies, confirming the validity of the tasks performed in this virtual environment.

We shall now discuss the statistical results. We found that the three modes of exploration had no influence on the *recognition of memorized scenes*. This result is in agreement with three previous studies. Gale et al. [13] performed a scene recognition experiment, in which the subjects actually experienced a field or viewed a videotape of the path. The scene recognition rate was similar in both conditions. Moreover, Christou and Bültoff found that active observers who were allowed to perform their movement through a virtual environment recognized about the same percentage of views as observers who viewed a playback of the exploratory path [8] and as observers who viewed snapshots of the environment [9]. In the present experiment, the lack of difference between active and passive explorations suggests that using the hand to control displacement and visuo–motor interaction have little effect on scene recognition. In addition, we found no difference between passive and snapshot explorations and that subjects performed remarkably well at recognizing scenes they had only seen once before. This held true even when the scene to be recognized was seen from an entirely novel viewpoint – in our case, the scenes presented during the testing phase were taken from points of view that were intermediate between the points of view seen during the exploration phase. It is thus possible to extrapolate a scaled spatial scene to another scaled scene, to some extent. Together, these data provide some support to the idea that a pattern from a visual scene is memorized, i.e. the representation of the visual scene is pictorial in nature and is determined principally by 2D image features (cf. the second theory of scene recognition mentioned above). This is the first conclusion of this study.

We were more surprised that no differences emerged from the comparisons of the two first (active and passive) and of the two last (passive and snapshot) conditions of exploration for *orienting towards the origin* task. Indeed, there are many path integration studies involving pointing toward the origin conducted in natural environments with occluded vision showing the importance of activity (for instance [15,22]). So it may be that visual flow only, in such an environment, is enough to update self-position with respect to a place. In contrast, visuo–motor interaction and motor control is not that much important. Moreover, orientation performances following snapshot exploration suggest that even a truncated exploration in such environment allows subjects to keep track of their position with respect to the origin of locomotion. Apparently, structured environments, such as urban environments with many salient landmarks, provide sufficient visual information to support good performance on pointing toward the origin tasks. Such visual information would serve as a mental frame of reference in which the subject can keep track of his displacement [35]. However, another interpretation could be that motion duration extracted during exploration [4] would support the process of keeping track of the displacement.

In contrast, *reproduction of the distances and angles* of the path shown on drawings were influenced: the snapshot exploration led to the worst performance on both response measures; the other two conditions did not differ. This result suggests that a continuous stimulation of visual exploration enhances the ability to reproduce the shape of a path previously traveled. It may be that the visual information was too restricted to support accurate integration of the shape of the path. However, pointing to the origin was not adversely affected in the snapshot exploration condition, suggesting that even during the snapshot condition, all information necessary for the task – updat-

ing, in this case intermittent – was available. The poorer performance for path shape reproduction could reflect the operation of another step of spatial information processing – like a more deliberate cognitive processing – which requires information filling the gaps between snapshots [1,12], by continuously updating the position of landmarks based on interpolation of one's position related to the former snapshot. Finally, whether during exploration the subject used a mental frame of reference or motion duration for displacement updating, these processes were not adequate for path shape reproduction.

Three conclusions can be drawn from the pointing task and path shape reproduction: (a) pointing towards the origin of the path does not rely on the memory of the path shape. (b) Drawing a shape of a path relies on putting together piecemeal information that has to be actively recalled. The path shape is not elaborated as the subject moves within the environment. (c) A reconstruction process occurs and this process requires continuous visual simulation of the path. These conclusions are in agreement with the study by Amorim et al. [1]. In this study, the authors compared the effect of two processing modes on the updating of the location and orientation of a previously viewed object in space during a guided walk without vision. Subjects attempted either to keep track of the object continuously or to estimate the object's perspective only at the terminal vantage point, given the trajectory they walked. During both conditions, similar pointings toward the target were obtained, suggesting that a common representation of the space updated by walking was shared by the two conditions. However, the first condition induced a slow-down of self-paced locomotion velocity for linear and angular movements, and latencies for retrieving the object mental image were shorter in this condition. These results suggest that during blind locomotion the subject can use two cognitive strategies for this purpose: continuously update the location of a previously seen landmark or concentrate on egocentric path integration, and perform the updating when the path is completed. By analogy, because the three exploratory conditions did not affect pointing to the origin of the path but drawing the path shape was, we suggest that, during the first test, observers continuously kept track of the origin position during the walk whereas during the second test, observers deduced the shape of the path at the final viewing position.

In conclusion, this study provides an interpretation concerning the role of the three modes of exploration and, furthermore, on control of action and visuo–motor coupling and on continuity of visual stimulation during visual exploration of a virtual urban environment. They have a different effect according to the ultimate task of knowledge retrieval:

- The representation of the visual scene does not require visuo–motor coupling and continuity of visual stimulation; it is pictorial in nature and is determined principally by 2D image features.
- Pointing towards the origin of a path in a natural environment does not require neither control of action and visuo–motor coupling nor continuity of visual stimulation; it relies on motion duration integration or a frame of reference integrated during displacement.
- Path shape reproduction requires a continuous visual exploration; a deliberate additional cognitive process is involved for this ability.

Any benefit of active exploration found in earlier studies would depend on directing action, suggesting the importance of planning action in some spatial abilities.

Acknowledgements

We thank John Philbeck, Sid Wiener, Michel-Ange Amorim and Mark Wexler for useful comments on earlier versions of this article, Ki-Nam Tran and Alexandre Heidet for computer support and France Maloumian for preparing illustrations. This work was supported by a contract between LPPA and Vehicle Simulation and Perception group (Technocentre-Renault).

References

[1] M-A. Amorim, S. Glasauer, K. Corpinot, A. Berthoz, Updating an object's orientation and location during non-visual navigation: a comparison between two processing modes, Percept. Psychophys. 59 (1997) 404–418.

[2] D. Appleyard, Styles and methods of structuring a city, Env. Behavior 2 (1970) 100–118.

[3] E.A. Attree, B.M. Brooks, F.D. Rose, T.K. Andrews, A.G. Leadbetter, B.R. Clifford, Memory processes and virtual environments: I can't remember what was there, but I can remember how I got there. Implications for people with disabilities, in: P.M. Sharkey (Ed.), Proceedings of the 1st European Conference on Disability, Virtual Reality and Associated Technologies, Reading, UK: ECDVRAT & University of Reading, 1996, pp. 123–132.

[4] A. Berthoz, I. Israël, P. Georges François, R. Grasso, T. Tzuzuku, Spatial memory of body linear displacement: what is being stored?, Science 269 (1995) 95–98.

[5] M. Blades, L.D. Medlicott, Developmental differences in the ability to give route directions from a map, J. Env. Psych. 12 (1992) 175–185.

[6] B.A. Cartwright, T. S Collet, Landmark learning in bees: experiments and models, J. Comp. Physiol. 151 (1983) 521–543.

[7] 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: Teleoperators and Virtual Environments 7 (1998) 168–178.

[8] C.G. Christou, H.H. Bültoff, Scene recognition after active and passive learning, Perception 26 (Suppl.) (1997) 33.

[9] C.G. Christou, H.H. Bültoff, View dependency of scene recognition in a virtual house, Invest. Ophtal. Vis. Sci. 39 (1998) S853.

[10] E.H. Cornell, C.D. Heth, L.S. Broda, Children's wayfinding: response to instructions to use environmental landmarks, Dev. Psych. 25 (1989) 755–764.

[11] J.A. Crowell, M.S. Banks, Perceiving heading with different retinal

regions and types of optic flow, Percept. Psychophys. 53 (1993) 325–337.
[12] M.J. Farell, J.A. Thomson, Automatic updating during locomotion without vision, Quart. J. Exp. Psyc. 51A (1998) 3 637–654.
[13] N.D. Gale, R.G. Golledge, J.W. Pellegrino, S. Doherty, The acquisition and integration of neighborhood route knowledge in an unfamiliar neighborhood, J. Env. Psych. 10 (1990) 3–25.
[14] C.R. Gallistel, The Organization of Learning, MIT Press, Cambridge, MA, 1990.
[15] T. Gärling, T. Mäntylä, J. Säisä, The importance of vision during locomotion for the acquisition of an internal representation of the spatial layout of large scale environment: blindfolded and sighted car passengers with and without a distracting tack learning to localize invisible targets during a town route, Ümea Psychol. Rep. 147 (1978) 1–25.
[16] F. Gaunet, A. Berthoz, Mental rotation of spatial environments, Cog. Brain Res. 9 (2000) 91–102.
[17] F. Gaunet, P. Péruch, M.D. Giraudo, C. Thinus-Blanc, From virtual to actual space: transfer of spatial information, Proceedings of the 4th International Days of Virtual & Actual Worlds Interfaces, 1995, pp. 297–304.
[18] L. Haber, R.N. Haber, S. Penningroth, H. Novak, H. Radgowsky, Comparison of nine methods of indicating the direction to objects: data from blind adults, Perception 22 (1993) 35–47.
[19] D. Henry, T.A. Furness, Spatial perception in virtual environments: evaluating an architectural application, Proceedings of the IEEE Virtual Reality Annual International Symposium (VRAIS), Seattle, Washington, 1993, pp. 33–43.
[20] I. Israël, M. Fetter, E. Koenig, Vestibular perception of passive whole-body rotation about horizontal and vertical axes in humans: goal-directed vestibulo-ocular reflex and vestibular memory-contigent saccades, Exp. Brain Res. 96 (1993) 335–346.
[21] 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.
[22] J. Juurma, K. Suonio, The role of audition and motion in the spatial orientation of the blind and the sighted, Scand. J. Psych. 16 (1975) 209–216.
[23] R. 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, Psych. Sci. 9 (1998) 293–298.
[24] J.M. Loomis, A.C. Beall, Visually-controlled locomotion: its dependence on optic flow, 3-D space perception, and cognition, Ecolo. Psych. 10 (1998) 271–285.
[25] J. M Loomis, J.J. Blascovich, A.C. Beall, Immersive virtual environment technology as a basic research tool in psychology, Behav. Res. Meth. Instrum. Comp. 31 (1999) 557–564.
[26] J.M. Loomis, R.L. Klatzky, R.G. Golledge, J.G. Cicinelli, J.W. Pellegrino, P.A. Fry, Nonvisual navigation by blind and sighted: Assessment of path integration ability, J. Exp. Psych. Gen. 122 (1993) 73–91.
[27] J.M. Loomis, J.M. Knapp, Visual perception of egocentric distance in real and virtual environments, in: L.J. Hettinger, M.W. Haas (Eds.), Virtual and Adaptive Environments, Hillsdale NJ: Erlbaum, in press.
[28] K. Lynch, Image of the City, MIT Press, Cambridge, MA, 1960.
[29] D. Marr, Vision, Freeman, San Francisco, 1982.
[30] D. Marr, H.K. Nishira, Representation and recognition of the spatial organization of three dimensional shapes, Proc. R. Soc. Lond. 200 (1978) 269–294.
[31] M.L. Mittelstaedt, S. Glasauer, Idiothetic navigation in gerbils and humans, Zool. Jb Physiol. 95 (1991) 427–435.
[32] J. O'Keefe, L. Nadel, The Hippocampus as a Cognitive Map, Oxford University Press, London, 1978.
[33] P. Péruch, F. Gaunet, Virtual environments as a promising tool for investigating human spatial cognition, Curr. Psych. Cog. 17 (1998) 881–899.
[34] P. Péruch, J.L. Vercher, G.M. Gauthier, Acquisition of spatial knowledge through visual exploration of simulated environments, Ecolo. Psych. 7 (1995) 1–20.
[35] J.J. Rieser, D.A. Guth, E.W. Hill, Sensitivity to perspective structure while walking without vision, Perception 15 (1986) 173–188.
[36] J.J. Rieser, E.W. Hill, C.R. Talor, A. Bradfield, S. Rosen, Visual experience, visual field size and the development of nonvisual sensitivity to the spatial structure of outdoor neighborhoods explored by walking, J. Exp. Psych. Gen. 121 (1992) 210–221.
[37] M.J. Sholl, From visual information to cognitive maps, in: J. Portugali (Ed.), The Construction of Cognitive Maps, Kluwer, The Netherlands, 1996, pp. 157–186.
[38] A.W. Siegel, The externalization of cognitive maps by children and adults: in search of better ways to ask better questions, in: L.S. Liben, A. Patterson, N. Newcombe (Eds.), Spatial Representation and Behavior Across Life Span: Theory and Application, Academic Press, New York, 1981, pp. 167–194.
[39] M.J. Tarr, Rotating object to recognize them: a case study on the role of viewpoint dependency in the recognition of three-dimensional objects, Psychon. Bull. Rev. 2 (1995) 55–82.
[40] K.A. Turano, X. Wang, Visual discrimination between a curved and straight path of self motion, Exp. Brain Res. 104 (1994) 502–510.
[41] P.N. Wilson, N. Foreman, R. Gillett, D. Stanton, Active versus passive processing of spatial information in a computer simulated environment, Ecolo. Psych. 9 (1997) 207–222.