Neural Basis of Mental Scanning of a Topographic Representation Built from a Text

E. Mellet, S.Bricogne, F. Crivello, B. Mazoyer, M. Denis[1] and N. Tzourio-Mazoyer

CNRS UMR 6995, Université de Caen and Université Paris V, BP 5227, 14074 Caen Cedex and [1]Groupe Cognition Humaine, LIMSI-CNRS, BP 133, 91403 Orsay Cedex, France

Humans have the ability to build and to inspect an internal visual image of an environment built from a verbal description. We used positron emission tomography (PET) to investigate the brain areas engaged in the mental scanning of a map that subjects built from the reading of a descriptive text. This task engaged a parieto-frontal network known to deal with spatial representations. Additional activations were evidenced in the angular gyrus and in Broca's and Wernicke's areas. In order to examine the neural impact of the learning modality, these PET results were compared to those obtained in another group of six subjects who performed a similar mental scanning task on a topographic representation built from visual inspection of a map. Both scanning tasks engaged the parieto-frontal network. However, the bilateral activation of the angular gyrus as well as the involvement of language areas appeared specific to the mental scanning of the topographic representation built from textual information. On the other hand, the right medial temporal lobe was activated only when a map had been visually learned. These results suggest that although both tasks involved visuo-spatial internal representation, a trace of the learning modality remained present in the brain.

Introduction

One of the main functions of language is to allow people to exchange information about objects or scenes that are remote from their immediate environment. People thus use a topographical representation that can be seen as the knowledge that they acquire about an environment, beyond the range of their immediate perception (Aguire and D'Esposito, 1999). It is a common situation for speakers (or writers) to try to convey their internal visual memories of an object to their addressees in the form of verbal descriptions. Reciprocally, listeners (or readers) are capable of translating that linguistic information into the form of an internal mental image. By re-expressing linear verbal information in the form of internal visuo-spatial representations, they provide themselves with a mental model where information is displayed in a form that preserves the spatial structure of the described configurations (Denis, 1996; Johnson-Laird, 1996). As a result, in addition to constructing a linguistic or propositional representation of the text itself (Van Dijk and Kintsch, 1983), readers tend to build representations endowed with properties that make them structurally analogous to the described configurations.

It has been suggested that the representation built from a verbal description shares some structural properties with the mental image of the environment built from visual perception. It is now widely accepted that the topographic representation built from a text not only includes the spatial relations that have been explicitly described in the text, but, as for topographic representation built from perception, allows people to deduce a number of non-stated relations (Perrig and Kintsch, 1985; Taylor and Tversky, 1992a,b). Empirical support was provided by studies showing that when people perform mental scanning across spatial configurations containing several landmarks, a positive correlation between scanned distances and scanning times is obtained in both cases, i.e. whether images were built from the visual perception of a map or the processing of a verbal description of that map where the relative distances between landmarks could be inferred (Denis and Cocude, 1992; Denis and Kosslyn, 1999).

The existence of features common to representations built from visual and verbal inputs raises the hypothesis that both types of representation share some neural components. However, the topographic knowledge obtained through verbal information differs in several respects from that acquired through visual experience. In particular, metric information, which is naturally included in the topographic representation built visually, may be absent or incomplete in a text description. Thus, the structural isomorphism between a physical environment and its mental counterpart may not be achieved when mental scanning (or some other form of mental processing) is performed on an environment that has been verbally described. Consequently, the neural components involved in the storage and the retrieval of topographic knowledge acquired from these two modalities might not be identical.

In a recent work (Mellet *et al.*, 2000a), we have shown that mental scanning of a visually learned map involves a parieto-frontal network and a region in the right medial temporal lobe which straddles the right hippocampus and the adjacent entorhinal cortex. These results were in agreement with other neuroimaging studies that had explored the neural components involved in the encoding and retrieval of topographic representations constructed from a visual input. These works have demonstrated that the medial temporal lobe is involved in the encoding and the retrieval of topographic knowledge visually acquired (Aguirre *et al.*, 1996; Maguire *et al.*, 1996, 1997, 1998a,b; Aguirre and D'Esposito, 1997; Ghaëm *et al.*, 1997; Gron *et al.*, 2000; Ino *et al.*, 2002; Pine *et al.*, 2002; Shelton and Gabrieli, 2002). In most of these studies, the construction of the topographic representation derived from navigation in virtual or in actual environments. This favored a route representation of the environment. We have recently shown that the medial temporal lobe is active during the retrieval and the mental exploration of an environment acquired by map learning (Mellet *et al.*, 2000a). However, another study reported an activation in this region during encoding in a route perspective but not in a survey perspective, suggesting an effect of the perspective during the encoding phase (Shelton and Gabrieli, 2002).

The areas involved in the storage and retrieval of topographic representations built from verbal input remain largely unknown. In the visual imagery domain, we have shown that generating mental images from verbal descriptions of objects that had never been seen before relies on a neural network close to that mobilized when the objects were visually encoded (Mellet *et al.*,

© Oxford University Press 2002. All rights reserved.

1996, 2000b). In spite of the linguistic nature of the information from which the mental image were generated, no language area was found activated during these mental imagery tasks. However, the objects used were relatively simple and did not involve the high spatial organization required by the representation of an environment or a map. Moreover, the linguistic stimuli used in these studies were restricted to single words or simple verbal instructions.

The research reported here had three objectives. First, we collected chronometric measures reflecting the mental scanning of two distinct visuo-spatial configurations which came to readers' knowledge through the sole vehicle of two texts. Although no explicit metric information was available in the description of the environments, we expected that readers would not leave their representation of the environments devoid of such metric information. However, whether such a representation would exhibit a time/distance linear relationship remained to be established. We thus collected evidence for the correlation between distances subjectively generated by readers and the times devoted to mental scanning of these distances. The second objective was to identify, using positron emission tomography (PET), the neural networks involved during the mental scanning of maps built from the reading of descriptive texts. More specifically, the questions addressed were whether visuo-spatial areas and medial temporal regions would be activated and whether the language areas would demonstrate any activity during mental scanning following verbal learning of the configuration. The last objective was to compare the results of the present study with results collected in a previous study, where people performed mental scanning across a similar map configuration that had been learned perceptually (Mellet *et al.*, 2000a). In particular, we wanted to establish whether common structures would be activated in mental scanning of map configurations irrespective of the type of input (visual or verbal) used to construct the visuo-spatial representation.

Materials and Methods

Subjects

Six healthy right-handed male volunteers (19–25 years old) participated in this study. Handedness was assessed using the Edinburgh questionnaire (83–100%). All subjects were free from nervous disease or injury and had no abnormalities on their T_1-weighted magnetic resonance images (MRI). In order to ensure optimal homogeneity of the sample with respect to their imagery abilities, subjects were selected as high visuo-spatial imagers on the basis of their scores on the Mental Rotations Test (Vandenberg and Kuse, 1978); all subjects scored beyond the 75th percentile of a population of 100 male subjects (corresponding scores from 14 to 20).

Texts

Two texts describing distinct environments (the surroundings of a park and the surroundings of a village) were adapted from an earlier study (Taylor and Tversky, 1992a). Some landmarks were added in order to increase the number of routes to be followed in the test phase (texts are given in the Appendix). The texts described the environment in a survey perspective, i.e. using the canonical terms 'north', 'south', 'east' and 'west'. Each text included eight salient landmarks considered as anchors because they were presented at the beginning of the text and were cited at least twice (Ferguson and Hegarty, 1994). In addition, both texts included further secondary landmarks. All the segment routes were defined from the eight anchors, resulting in 56 different routes available. Note that the texts did not provide any metric distance between the landmarks.

Learning Phase

Three hours prior to the PET experiment, the subjects were requested to read the two texts until they considered that they had an accurate representation of the environment described. There were no limitations on time taken or the number of times they read each text, but these data were collected.

After the reading phase of the description of one environment, the accuracy of the topographic representation built by the subjects was assessed using a questionnaire that included 18 locative statements for each text (verbatim, paraphrased and inferential statements). Each statement was presented on a computer screen and the subjects responded by pressing a key.

Task Design and PET Procedure

This PET study included three conditions: a mental scanning task performed on the mental image of the environment built from the text (referred to as MS-Text for 'mental scanning from text' modality) and two control conditions, namely 'listening to abstract words' (LAW) and a resting state condition (REST). All were performed with eyes closed in total darkness, a black and opaque cloth covering the whole camera.

During MS-Text, the subjects, eyes closed, were to generate a survey image of the environment as accurately as possible. They were then given through earphones the name of two anchor landmarks (for example, 'church', 'school') and had then mentally to scan across the path connecting them. To help participants, instructions invited them to imagine a laser dot following the path connecting the two landmarks. When the first landmark name was given, a time count was started. When the second landmark was reached, the subjects had to press a button with their right index finger, this action interrupting the time count, which gave the time devoted to mental scanning and released the auditory delivery of a second pair of landmark names.

During LAW, the subjects had to listen to a pair of abstract nouns delivered every 5 s and to press a button with their right index finger after each couple of words. This condition shared the listening to pairs of nouns and the motor response with the mental scanning task. It did not include any mental imagery or mental scanning component.

During REST, subjects were in total darkness with eyes closed and were only instructed to keep their eyes closed, to relax, to refrain from moving and to avoid any structured mental activity such as counting, rehearsing, etc. This state has been widely used as a basic control condition in our laboratory (Zago *et al*. 2001). It has recently been proposed as an appropriate baseline physiological state (Gusnard *et al.*, 2001).

Twelve normalized regional cerebral blood flow (NrCBF) measurements (time acquisition, 90 s) were obtained from each subject on an ECAT Exact HR + PET camera, replicating four times the series of three experimental conditions (MS-Text, LAW and REST) in randomized order. Each environment was then scanned twice, different pairs of landmarks being used.

Debriefing of the Subjects

During the PET session, after each mental scanning condition, the subjects were asked to assess the vividness of the mental images they produced from 0 (no mental image), to 5 (clear and vivid mental images). The subjects were also asked whether they retrieved sentences from the texts during mental scanning and whether they relied on a verbal strategy to perform the task.

After the PET session, the subjects were invited to draw maps of the environments that they had learned from the texts. They were instructed that the maps should reflect as closely as possible their mental representations of the described environments. Their drawings were used to determine if landmarks were sited on the maps at places compatible with their respective descriptions. Actually, the maps did not contain any incoherence or discrepancy in terms of landmark positions. The most noticeable feature was, in fact, the great variability among participants regarding the distances between landmarks. Because we needed measures of these distances (in order to compute their possible relationships with scanning times), they were normalized as the ratio of their absolute value to the half-perimeter of the drawn map for each participant (from the old lighthouse to the hunting clubhouse for the Park text and from the Roman bridge to the castle for the Village text).

Data Analysis

After automatic realignment – AIR (Woods *et al.*, 1997) – the original brain images were transformed into the standard stereotactic Talairach space using the MNI template. The images were smoothed using a Gaussian filter of 12 mm FWHM, leading to a final smoothness of 15 mm FWHM. The rCBF was normalized within and between subjects using a proportional model. The comparisons across conditions were made by way of *t*-statistics. Statistical parametric maps corresponding to comparisons between conditions and between studies (see below) were generated with the 1999 version of SPM (Friston *et al.*, 1995). For each comparison, the voxel amplitude *t*-map was transformed in a *Z*-volume that was thresholded at $P < 0.05$ (corrected for multiple comparisons). However, because we had a specific hypothesis regarding the involvement of the language areas and of the medial temporal lobe, we also searched for activation at a lower threshold ($P < 0.001$, uncorrected for multiple comparisons). This was done for Broca's and Wernicke's areas and for the medial temporal lobe. The anatomical labeling of the activated areas was performed using an automated anatomic parcellation of the MNI template (Tzourio-Mazoyer *et al.*, 2002).

Comparison with a Previous Study of the Neural Basis of Mental Scanning from a Visually Learned Map

In order to compare the brain areas involved in the scanning of a mental map built from a descriptive text to those engaged in the task of mental scanning of a map that had been learned visually, we conducted a between-study analysis including the results of the present study and those of a previous work (Mellet *et al.*, 2000a). This work was carried out with another sample of six subjects with high visuo-spatial abilities. The map that the subjects learned represented a park and included seven colored dots that were connected by paths. The mental scanning task proper was quite similar to that described above. With eyes closed, the subjects had to visualize the map as accurately as possible, including the seven dots; they were then given the names of two colored dots (e.g. 'red', 'blue') through earphones and had then to imagine a laser dot following the path segment drawn on the original map between the two dots. Once the second dot was reached, the subjects had to press a button with their right index finger, this action releasing the auditory delivery of a further pair of dot names. This study will be referred to as MS-Map.

Between-study comparisons included a conjunction analysis of the two mental scanning tasks versus REST and the following comparisons: (MS-Text – REST) versus (MS-Map – REST) and the reverse comparison, i.e. (MS-Map – REST) versus (MS-Text – REST). In order to avoid 'false' activation due to deactivation in the second contrast, each interaction was masked by the main effect thresholded at 0.05 uncorrected, for example (MS-Map – REST) versus (MS-Text – REST) was masked by (MS-Map – REST).

Results

Behavioral Results

Training Phase (before the PET Session)

The subjects took, on average, 10.7 min (range, 6–18 min) to read the description of the Park environment and 13.0 min (range, 8–20 min) to read the description of the Village environment. Each text was read 3.0 times on average.

The subjects were able to answer correctly 82% (range 40–100%) of the locative statements they had to assess for the first environment and 90% (range 78–100%) for the second environment.

Debriefing of the Subjects (after the PET Session)

All subjects reported using a mental imagery strategy to perform the mental scanning task from texts. They rated the mental image of the environments as clear and vivid (4.3; range, 3–5). Although one subject reported that he occasionally recalled some sentences of the text during the mental scanning task, none of the participants said they had relied on a verbal strategy to perform the task.

Map Drawings and Chronometric Data Analysis

All subjects were able to draw detailed maps from the mental representations they had built (see Fig. 1 for an example). The subjects were, on average, able to locate 7.6 of the 'anchor' landmarks (range, 6–8). One subject drew an inverted map of the park environment (north–south inversion).

For each subject we computed a regression between the time spent mentally to scan the different route segments and their normalized length (from individual map drawings). The results, shown in Table 1, revealed that four out of six subjects exhibited

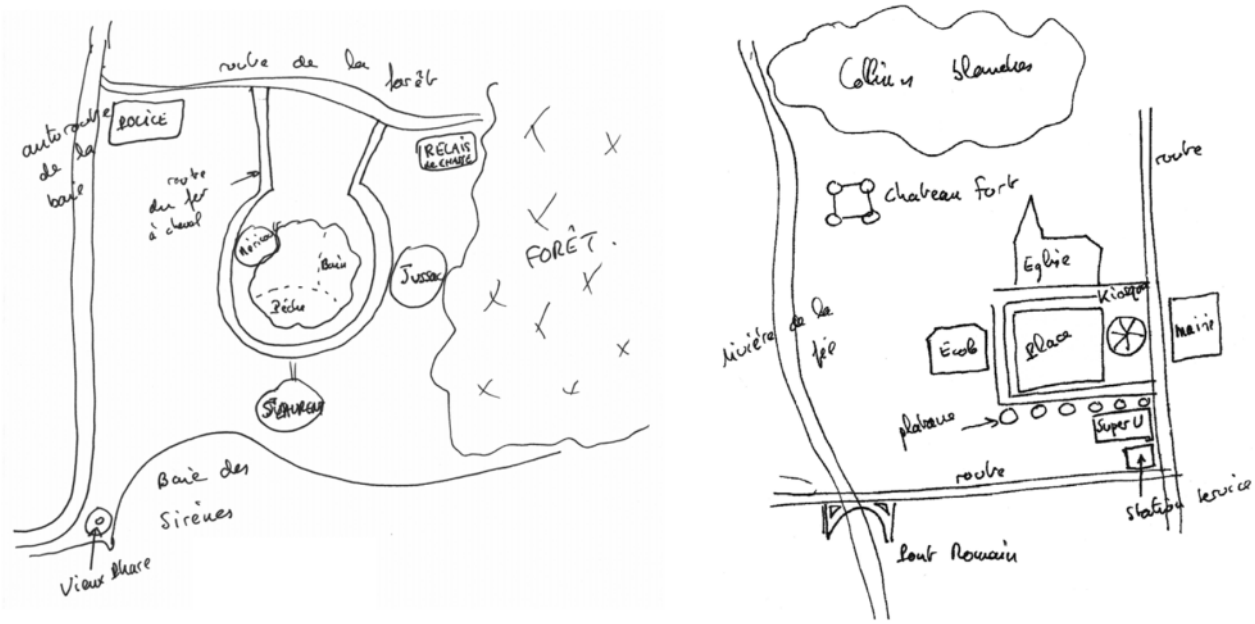

Figure 1. Two maps corresponding to the two environments drawn by a subject (subject 2) after the PET experiment.

a significant positive correlation, while two subjects did not. Figure 2 illustrates the regression for the six subjects.

In the same vein, we computed a regression for each subject for the MS-Map group (Table 2). Four out of six subjects exhibited a positive correlation.

PET Results

MS-Text – REST (Table 3)

The most significant activations were located in the parietal lobe including the precuneus, spreading from the intraparietal sulcus to the angular gyrus. A cluster of activation was detected in between the very internal part of the right anterior insula and the right putamen.

The left middle temporal gyrus also presented an activation. Note that this last cluster could not be considered as strictly unilateral because a contralateral activation in the homologous region was detected at a more permissive statistical threshold ($P < 0.001$, uncorrected). Moreover, at this threshold, the cluster was not confined to the middle temporal gyrus, but spread over the superior temporal sulcus. An activation was also detected in Broca's area in the left inferior frontal gyrus.

In this contrast, activations of the anterior part of the cingulate sulcus and, at a subcortical level, a bilateral activation of the thalami, of the vermis and of the right cerebellar hemisphere were also evident. A bilateral activation of the left superior frontal sulcus at its intersection with the precentral sulcus was also present. Note that this mental imagery task did not elicit any activation in early visual areas. Rather, a strong deactivation was detected in the calcarine sulcus ($x = 0$, $y = -90$, $z = 8$, $P < 0.001$, corrected for multiple comparisons).

MS-Text – LAW (Table 3)

The parietal activations remained present when the mental scanning task was compared to the word listening task, namely the bilateral activations of the precuneus, the internal part of the intraparietal sulcus and the angular gyrus. In this contrast, the posterior activations extended to the superior part of the left middle occipital gyrus and to the occipito-parietal sulcus. This was due to the fact that these regions were deactivated during the word listening task (data not shown).

Temporal activations were no longer present when mental scanning was compared to word listening. This supports the assumption that these regions were mainly implicated in the word processing component common to both tasks. The putamen/insular, the anterior cingulate and the cerebellar activations were also cancelled out in this contrast, revealing

Table 1
Correlation between the time and distance for the MS-Text group

Subject	No. of segments (mean of two environments)	*r*	*P*
1	43	0.36	0.02
2	42	0.70	<0.001
3	27	0.41	0.03
4	45	0.23	0.14
5	49	0.60	<0.001
6	58	0.12	0.35

Individual correlation coefficients between mental maps scanning duration and route segment length for the MS-Text group. The distances between landmarks were measured on the maps drawn by the subjects after the PET measurements.

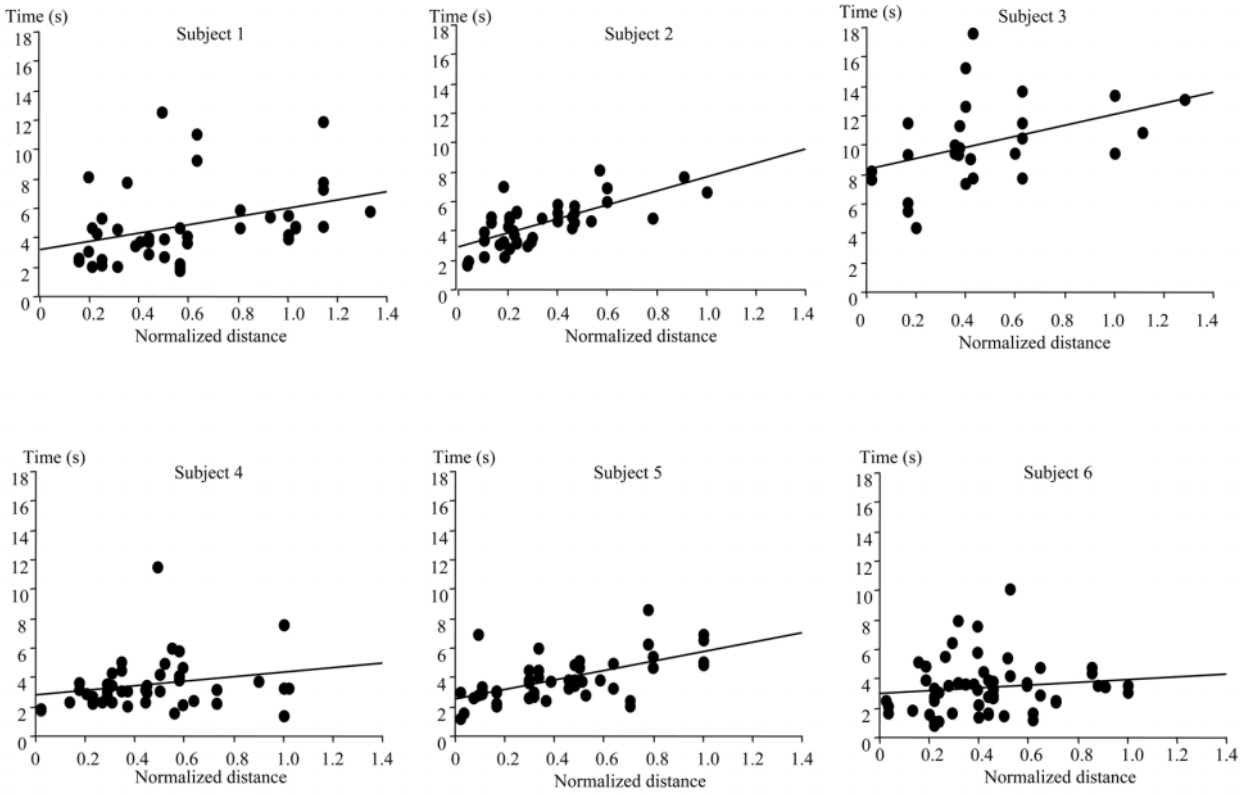

Figure 2. Linear regression analysis between mental scanning duration and route segment length for The six subjects. Subjects 1, 2, 3 and 5 exhibited a significant positive correlation (see Table 1), thus complying to the classical isomorphism between real and imagined map. On the other hand, subjects 4 and6 did not exhibit such a correlation.

Table 2
Correlation between the time and distance for the MS-Map group

Subjects	No. of segments	*r*	*P*
1	16	0.66	0.005
2	31	0.62	<0.001
3	31	0.63	0.001
4	31	0.19	0.31
5	31	0.56	0.001
6	31	0.12	0.51

Individual correlation coefficients between mental map scanning duration and route segment length for the MS-Map group. The distances between landmarks were measured on the map visually learned by the subjects.

Table 3
MS-Text– REST and MS-Text – LAW

Anatomical location of max. voxel	Coordinates			*Z*-score
	x	*y*	*z*	
MS-Text – REST				
R. precuneus/intraparietal sulcus	14	–74	52	7.3
L. precuneus	–6	–70	42	6.7
R. angular gyrus	42	–76	34	5.8
R. angular gyrus	38	–64	46	5.3
L. angular gyrus	–32	–78	40	5.4
R. occipito-parietal sulcus	28	–60	26	4.9
R. anterior insula/putamen	30	22	2	5.0
L. anterior cingulate sulcus	–6	18	42	4.7
L. superior frontal sulcus	–30	6	48	4.6*
R. superior frontal sulcus	36	4	52	3.1*
L. inferior frontal gyrus	–48	20	24	2.9**
L. middle temporal gyrus	–64	–44	–4	4.4*
L. middle temporal gyrus	–72	–34	0	4.8
Vermis	2	–76	–28	5.5
R. cerebellum	42	–56	–34	5.1
L. thalamus	–4	–12	0	5.0
R. thalamus	10	–10	0	5.0
MS-Text – LAW				
R. precuneus	14	–70	52	7.5
L. precuneus	–4	–66	50	6.3
L. intraparietal sulcus	–20	–74	48	6.5
L. angular gyrus	–34	–80	30	6.3
R. angular gyrus	42	–78	32	6.9
L. inferior parietal lobule	–38	–52	46	4.7
R. occipito-parietal sulcus	26	–64	32	5.2
R. occipito-parietal sulcus	26	–58	20	5.1
L. middle occipital gyrus	–28	–78	40	6.4
L. occipito-parietal sulcus	–18	–64	22	5.7
L. superior frontal sulcus	–26	2	50	4.8
R. superior frontal sulcus	28	–4	52	4.0

Foci of significant normalized regional cerebral blood flow (NrCBF) increases when mental scanning from text (MS-Text) were compared to the rest condition (REST, upper part) or to the listening of abstract words (LAW, lower part). The data are local maxima detected with SPM software at $P = 0.05$, corrected for multiple comparisons, except $*P < 0.001$ (uncorrected) and $**P = 0.002$ (uncorrected). Within these regions, the anatomical localization of the maximum *Z* of voxel is given on the basis of the MNI template, using stereotactic coordinates in the Talairach space in millimeters. R., right; L., left.

that they were not specific to the mental scanning condition. The left superior frontal sulcus presented an activation located at its intersection with the precentral sulcus. Note that this activation was also present in the right hemisphere when the mental scanning condition was compared with REST at a less rigorous statistical threshold ($P < 0.001$, uncorrected).

Finally, it should be stressed that, unlike most topographic scanning or navigation neuroimaging studies, no activation was detected within the medial temporal lobe during mental scanning from text when compared either to rest or passive abstract listening, even when lowering the statistical threshold.

Table 4
MS-Text and MS-Map between-study analysis

Anatomical location of max. voxel	Coordinates			*Z*-score
	x	*y*	*z*	
Conjunction MS-Text – REST and MS-Map – REST				
L. precuneus/intraparietal sulcus	–10	–74	50	6.0
R. superior parietal	18	–72	52	4.8
L. pre-SMA/cingulate	–10	14	38	5.6
L. pre-SMA	–10	6	50	4.9
L. superior frontal sulcus	–22	0	60	4.7*
R. superior frontal sulcus	38	4	56	4.5*
R. insula/putamen	32	20	8	5.4
L. insula/putamen	–26	22	6	4.9
R. superior temporal sulcus	64	–20	2	5.1
L. superior temporal sulcus	–62	–22	0	5.0
MS-Text – MS-Map				
L. middle occipital/angular gyrus	–40	–78	34	4.7
R. angular gyrus	38	–62	44	3.6*
L. middle temporal gyrus	–64	–42	0	4.1*
L. inferior frontal gyrus	–48	20	20	3.9*
MS-Map – MS-Text				
R. hippocampus	32	–14	–22	3.0*

Upper part: conjunction analysis revealing foci of significant NrCBF increases common to mental scanning from text learning and from map learning as compared to rest. Middle and lower parts: foci of significant difference between mental scanning from text and map learning (see Table 3 footnote for details).

Between Study Comparison

Conjunction Analysis of MS-Text and MS-Map (Table 4)

This analysis revealed the brain areas active during mental scanning whichever the modality from which the topographic representations were built (i.e. from reading texts or from visual learning of a map). An antero-posterior network was active in both MS-Text and MS-Map conditions. This network included the intraparietal sulcus and the precuneus for its posterior part and the superior frontal sulcus, bilaterally, and the pre-SMA (extending downward to the cingulate cortex) for its anterior part. Additional activated areas were found bilaterally in between the very internal part of the anterior insula and the putamen and bilaterally in the superior temporal sulcus.

MS-Text – MS-Map (Table 4)

The only cluster of activation evidenced at the 0.05 corrected statistical threshold straddled the left middle occipital gyrus and the left angular gyrus. In order to check whether this activation was strictly unilateral or not, we searched for a controlateral activation using a less stringent statistical threshold ($P < 0.001$, uncorrected for multiple comparisons). An activation of the right angular gyrus was evidenced at this threshold. Figure 3 (first and second rows) illustrates the individual CBF variation values in the left and right angular gyrus for both MS-Text and MS-Map groups. It shows that all subjects in the MS-Text group presented an increase in these regions, while most of the MS-Map subject exhibited a decrease in the same region. At $P = 0.001$, uncorrected for multiple comparisons, we found that a region in the left superior temporal sulcus and a left inferior frontal region known to be related to language processing were indeed activated during MS-Text, but not during MS-Map. Examination of the individual CBF variation indicated that these

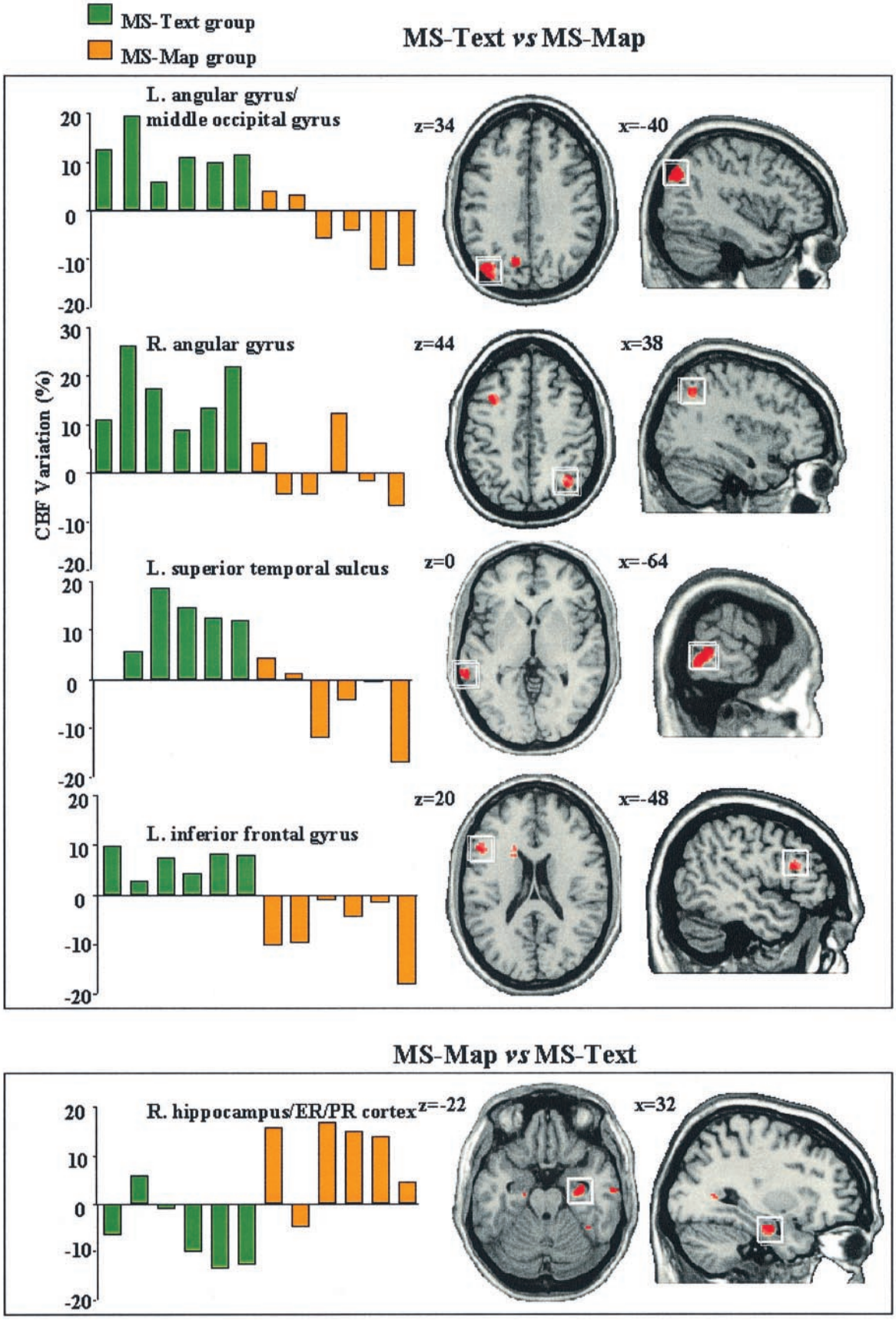

Figure 3. Left: histograms of the individual normalized rCBF values (averaged across trials) for the MS-Text (in green) and the MS-Map (in orange) groups in the angular gyri, the left inferior frontal gyrus (corresponding to Broca's area), the left superior temporal sulcus (corresponding to Wernicke's area) and the right medial temporal lobe. Right: corresponding slices from the MNI template, *z*- and *x*-values are given in millimeters.

average results were found for most or all subjects (Fig. 3, third and forth rows).

MS-Map – MS-Text (Table 4)

No cluster of activation was evidenced at a $P < 0.05$ threshold corrected for multiple comparisons. When looking at a lower statistical threshold ($P < 0.001$, uncorrected for multiple comparisons), an activation was present in the anterior part of the right medial temporal lobe. This activation encompassed the hippocampus and the entorhinal cortex and extended laterally to the banks of the collateral sulcus, corresponding to the perirhinal cortex (Insausti *et al.*, 1998). Figure 3 (bottom row) shows the individual CBF variations in this region. It shows that all MS-Text subjects but one exhibited a CBF decrease, while the MS-Map group showed the reverse pattern, i.e. all subjects but one demonstrated a regional CBF increase.

Discussion

The fact that the subjects from the TEXT group produced accurate drawings of the maps during the post-PET session demonstrated that they had built a veridical representation of the environments described. Moreover, it supports the assumption that the subjects were able to convert the descriptive text into an image-like representation when required. Note, however, that large individual variations were evident. As a matter of fact, four subjects (1, 2, 3 and 5) exhibited a correlation between mental scanning times and inter-landmark distances as assessed on their drawings. Interestingly, a comparable variability was observed in the MAP group. The time/distance correlation obtained here replicates the findings from a number of mental scanning studies. All of these studies share the feature that participants built mental images with explicit constraints regarding the metrics of the configuration (whether distances were simply displayed in a map or implied from a verbal description of the map (Denis and Kosslyn, 1999). In the present study, the TEXT participants were not exposed to such information, so that each of them built a representation incorporating individual metric values. The results show that individual scanning times are well adjusted for the majority of the participants to the metric of the representation constructed by each individual person. Although they produced correct drawings as well, the two remaining subjects (4 and 6) did not exhibit such a correlation. This was also the case for two subjects belonging to the MAP group. This could indicate that these subjects built and scanned a less stabilized representation in which metric information was not encoded in a consistent manner. The fact that in both MAP and TEXT groups, a similar number of subjects did not show a time distance/correlation suggests that the modality of learning and the absence of explicit metric information have little influence on the spatial properties of the representation. Note that all subjects were selected for their good spatial imagery abilities and that the individual behavioral differences observed in both mental scanning tasks are thus unlikely to be ascribed to individual differences in imagery abilities.

Areas Activated in Both Mental Scanning based on Map and Text Inputs Match the Spatial Imagery Brain Areas

A first result is that mental scanning of an environment verbally described gave rise to a parieto-frontal network activation that has already been described in mental scanning of a visually learned environment (Mellet *et al.*, 2000a). This network, including the intraparietal sulcus, the pre-SMA and the superior frontal sulcus, is also activated during various spatial mental imagery tasks (Mellet *et al.*, 1996, 2000b). This parieto-frontal network was also detected in the conjunction analysis which included mental scanning from both text and map learning. We have suggested that this parieto-frontal network constitutes the smallest set of regions necessary to deal with spatial representation, including spatial working memory and spatial mental imagery (Mellet *et al.*, 2000a). The results of the present study support this claim. Moreover, the parieto-frontal activation gives a neural substrate to the hypothesis that when people use verbal information to build an internal topographic representation, they rely on analog processes which encode information in an internal spatial array rather than a purely text-based representation (Mani and Johnson-Laird, 1982). Note, however, that the subjects were selected as 'high imagers' and might be particularly inclined to create an analog visuo-spatial representation derived from the text. The anatomo-functional similarities between the two mental scanning tasks might have been less important in 'low spatial imagers'. One should also stress that early visual areas were not activated in these spatial imagery tasks, with the calcarine sulcus even being strongly deactivated during the MS-Text condition. These results are in line with recent evidence that spatial mental imagery does not involve early visual cortex (Mellet *et al.*, 1998; Thompson and Kosslyn, 2000; Mazard *et al.*, 2002).

Mental Scanning of a Topographical Representation Built from Text Reading Selectively Activates the Angular Gyrus and Broca's and Wernicke's Areas

The comparison between the mental scanning of representations constructed from a map and from a text revealed that, when the topographic representation was built from a text, a parietal activation extended downward to a region that straddles the angular gyrus and the middle occipital gyrus. Although this region has been implicated in various cognitive processes, its role still remains unclear.

First, this region has been shown to be involved in spatial processing. More specifically, it has been suggested that the left angular gyrus is involved in categorical spatial processing (bearing on verbal markers such as 'above', 'behind', etc.), while the right angular gyrus is involved in metric spatial processing, which quantitatively specifies distances between items (Baciu *et al.*, 1999). In our study, both types of spatial processing may have been used. In particular, the subjects who exhibited a significant time/distance correlation are thought to use metric representations, while those who did not show this correlation may have relied on categorical spatial processing in which no metric information is required. However, the individual profiles of rCBF (Fig. 2) did not support this hypothesis since the two types of subjects did not differ in term of lateralization of angular gyrus activation. Moreover, according to this hypothesis, a right lateralized activation of the angular gyrus should have been evident in the subjects who had learned the map, since such a mental scanning task strongly relies on metric spatial processing. On the contrary, this region showed a CBF decrease in most subjects.

A second hypothesis stems from the fact that the angular gyrus belongs to heteromodal cortex and seems to be involved in processes that combine symbolic and analog representations. For example, the left angular gyrus is engaged in reading tasks that combines visuo-spatial and linguistic processing. It is also involved in calculations which merge symbols (figures and number) and analog representations of magnitude (Dehaene, 1992; Chochon *et al.*, 1999; Zago *et al.*, 2001). In the present study, the specific implication of the angular gyrus in the MS-Text condition may reflect the transformation of symbolic

information into analog 'map-like' information used during mental scanning. In the same vein, the activation of Broca's area in the left inferior frontal gyrus and of Wernicke's area in the posterior part of the left superior temporal gyrus could also be related to the verbal nature of the original input. These results suggest that Broca's and Wernicke's areas could be activated even when the task does not explicitly include any language components. These activations demonstrate that areas belonging to the language network and presumably active during the construction and the encoding of the representation (i.e. the mental map), were also involved during the retrieval of this representation. These activations could indicate that there is a representation of the text encoded in the brain, from which the analog 'picture-like' representation used during the mental scanning task is derived. Moreover, these results are in line with neuroimaging findings suggesting that some of the brain regions active during the encoding of specific piece of information are reactivated during retrieval of this information (Nyberg *et al.*, 2000; Wheeler *et al.*, 2000).

Mental Scanning of a Topographical Representation Built from Text Reading does not Involve Medial Temporal Lobe Areas

Another noticeable difference between mental scanning of representations constructed from either map or text is that no activation was detected within the medial temporal lobe in mental scanning following text learning, whereas such activation was present (albeit slight) in mental scanning following map learning. One could argue that the absence of hippocampal activation might be due to the use of a resting baseline as a reference condition. As a matter of fact, it has been shown that this region can be activated during rest (Stark and Squire, 2001). In the present study, this interpretation seems unlikely because neither was the hippocampus activated when MS-Text was compared to the abstract word listening condition (LAW). On the other hand, a hippocampal activation was evident in the MS-Map condition, even when compared to the resting baseline. This difference could then be reasonably ascribed to an effect of the learning modality.

Activation of the hippocampus and of the parahippocampus has been reported in the very large majority of neuroimaging studies that investigated the neural basis of topographic knowledge in humans. Although the exact location (i.e. hippocampus proper or parahippocampus) remains an open question, it is admitted that the medial temporal lobe plays a role in encoding and retrieving topographic information. Some authors have proposed that the hippocampus (more specifically the right hippocampus) provides an allocentric representation of space, i.e. a cognitive map (Maguire *et al.*, 1998a). Others have suggested that in humans, the parahippocampus is the key region for the representation of an environment in associating landmarks and their spatial relationships (Aguirre *et al.*, 1998; Aguirre and D'Esposito, 1999). However, all these neuroimaging studies investigated topographic representations which were built from visual inputs. As a matter of fact, this part of the medial temporal lobe has been shown to be involved in mental imagery following visual perception (Kreiman *et al.*, 2000). The fact that it was not engaged in the MS-Text condition and even deactivated in five subjects indicates that the medial temporal lobe plays a more limited role than generally admitted and that it is crucial only when the information from which the topographic representation is built is visual or when accurate metric information is provided.

Conclusion

The present study provides a neural basis to the claim that multiple representations of the environment might coexist in the brain (Johnson-Laird, 1983; Perrig and Kintsch, 1985; Taylor and Tversky, 1992a). First, a representation corresponding to an analog mental model of the described environment engaged the parieto-frontal network common to the processing of images constructed from both visual and verbal learning. Furthermore, a representation of the language used in the text is likely supported by a left inferior frontal and a left posterior temporal area belonging to a semantic language network. The angular gyrus and the occipito-parietal junction may play a role in the coordination of these two types of representations. In the present study, the absence of medial temporal lobe involvement when a mental map is built from text reading is the most salient difference with situations involving the learning of visually presented topographic information.

Notes

The authors are deeply indebted to their colleagues V. Beaudouin, O. Thirel and G. Perchey for their invaluable help in tracer production and to L. Petit for his thoughtful comments. The authors also thank B. Tversky for her advice in the adaption of the texts. This research was supported in part by a grant from the GIS 'Science de la Cognition'.

Address correspondence to Emmanuel Mellet, Groupe d'Imagerie Neurofonctionnelle, GIP Cyceron, Bd H. Becquerel, BP 5229, 14074 Caen Cedex, France. Email: mellet@cyceron.fr.

References

Aguirre GK, D'Esposito M (1997) Environmental knowledge is subserved by separable dorsal/ventral neural areas. J Neurosci 17:2512–2518.

Aguirre GK, D'Esposito M (1999) Topographical disorientation: a synthesis and taxonomy. Brain 122:1613–1628.

Aguirre GK, Dettre JA, Alsop DC, D'Esposito M (1996) The parahippocampus subserves topographical learning in man. Cereb Cortex 6:823–829.

Aguirre GK, Zarahn E, D'Esposito M (1998) Neural components of topographical representation. Proc Natl Acad Sci USA 95:839–846.

Baciu M, Koenig O, Vernier MP, Bedoin N, Rubin C, Segebarth C (1999) Categorical and coordinate spatial relations: fMRI evidence for hemispheric specialization. Neuroreport 10:1373–1378.

Chochon F, Cohen L, Vandemoortele PF, Dehaene S (1999) Differential contributions of the left and right inferior parietal lobules to number processing. J Cogn Neurosci 11:617–630.

Dehaene S (1992) Varieties of numerical abilities. Cognition 44:1–42.

Denis M (1996) Imagery and the description of spatial configurations. In: Models of visuospatial cognition (De Vega M, Intons-Peterson MJ, Johnson-Laird PN, Denis M, Marscharked M, eds), pp. 128–197. New York: Oxford University Press.

Denis M, Cocude M (1992) Structural properties of visual images constructed from poorly or well-structured verbal descriptions. Mem Cognit 20:497–506.

Denis M, Kosslyn SM (1999) Scanning visual mental images: a window on the mind. Curr Psychol Cognit 18:409–465.

Ferguson EL, Hegarty M (1994) Properties of cognitive maps constructed from texts. Mem Cognit 22:445–473.

Friston KJ, Holmes A, Worsley K, Poline J-B, Frith CD, Frackowiak RSJ (1995) Statistical parametric maps in functional imaging: a general approach. Hum Brain Mapp 2:189–210.

Ghaëm O, Mellet E, Crivello F, Tzourio N, Mazoyer B, Berthoz A, Denis M (1997) Mental navigation along memorized routes activates the hippocampus, precuneus, and insula. Neuroreport 8:739–744.

Gron G, Wunderlich AP, Spitzer M, Tomczak R, Riepe MW (2000) Brain activation during human navigation: gender-different neural networks as substrate of performance. Nat Neurosci 3:404–408.

Gusnard DA, Raichle ME, Raichle ME (2001) Searching for a baseline: functional imaging and the resting human brain. Nat Rev Neurosci 2:685–694.

Ino T, Inoue Y, Kage M, Hirose S, Kimura T, Fukuyama H (2002) Mental

navigation in humans is processed in the anterior bank of the parieto-occipital sulcus. Neurosci Lett 322:182–186.
Insausti R, Juottonen K, Soininen H, Insausti AM, Partanen K, Vainio P, Laakso MP, Pitkanen A (1998) MR volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices. Am J Neuroradiol 19:659–671.
Johnson-Laird PN (1983) Mental models. Cambridge, MA: Harvard University Press.
Johnson-Laird PN (1996) Images, models and propositional representations. In: Models of visuospatial cognition (De Vega M, Intons-Peterson MJ, Johnson-Laird PN Denis, M, Marscharked M, eds), pp. 90–127. New York: Oxford University Press.
Kreiman G, Koch C, Fried I (2000) Imagery neurons in the human brain. Nature 408:357–361.
Maguire EA, Frackowiak RSJ, Frith CD (1996) Learning to find your way: a role for the human hippocampal formation. Proc R Soc Lond B Biol Sci 263:1745–1750.
Maguire EA, Frackowiak RSJ, Frith CD (1997) Recalling route around London: activation of the right hippocampus in taxi drivers. J Neurosci 17:7103–7110.
Maguire EA, Burgess N, Donnett JG, Frackowiak RSJ, Frith CD, O'Keefe J (1998a) Knowing where and getting there: a human navigation network. Science 280:921–924.
Maguire EA, Frith CD, Burgess N, Donnett JG, O'Keefe J (1998b) Knowing where things are: parahippocampal involvement in encoding object locations in virtual large-scale space. J Cogn Neurosci 10:61–76.
Mani K, Johnson-Laird PN (1982) The mental representation of spatial descriptions. Mem Cognit 10:181–187.
Mazard A, Mazoyer B, Etard O, Tzourio-Mazoyer N, Kosslyn SM, Mellet E (2002) Impact of fMRI acoustic noise on the functional anatomy of visual mental imagery. J Cogn Neurosci 14:172–186.
Mellet E, Tzourio N, Crivello F, Joliot M, Denis M, Mazoyer B (1996) Functional anatomy of spatial mental imagery generated from verbal instruction. J Neurosci 16:6504–6512.
Mellet E, Petit L, Mazoyer B, Denis M, Tzourio N (1998) Reopening the imagery debate: lessons from functional anatomy. Neuroimage 8:129–139.
Mellet E, Bricogne S, Tzourio-Mazoyer N, Ghaëm O, Petit L, Zago L, Etard O, Berthoz A, Mazoyer B, Denis M (2000a) Neural correlates of topographic mental exploration: the impact of route versus survey perspective learning. Neuroimage 12:588–600.
Mellet E, Kosslyn SM, Mazoyer N, Bricogne S, Denis M, Mazoyer B (2000b) Functional anatomy of high resolution mental imagery. J Cogn Neurosci 12:98–109.
Nyberg L, Habib R, McIntosh AR, Tulving E (2000) Reactivation of encoding-related brain activity during memory retrieval. Proc Natl Acad Sci USA 97:11120–11124.
Perrig W, Kintsch W (1985) Propositional and situational representations of text. J Mem Lang 24:503–518.
Pine DS, Grun J, Maguire EA, Burgess N, Zarahn E, Koda V, Fyer A, Szeszko PR, Bilder RM (2002) Neurodevelopmental aspects of spatial navigation: a virtual reality fMRI study. Neuroimage 15:396–406.
Shelton AL, Gabrieli JD (2002) Neural correlates of encoding space from route and survey perspectives. J Neurosci 22:2711–2717.
Stark CE, Squire LR (2001) When zero is not zero: the problem of ambiguous baseline conditions in fMRI. Proc Natl Acad Sci USA 98:12760–12766.
Taylor HA, Tversky B (1992a) Spatial mental models derived from survey and route description descriptions. J Mem Lang 31:261–292.
Taylor HA, Tversky B (1992b) Descriptions and depictions of environments. Mem Cognit 20:483–496.
Thompson WL, Kosslyn SM (2000) Neural systems activated during visual mental imagery: a review and meta-analysis. In: Brain mapping: the applications (Toga AW, Mazziotta, JC, eds). New York: Academic Press.
Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15:273–289.
Van Dijk TA, Kintsch W (1983) Strategies of discourse comprehension. New York: Academic Press.
Vandenberg SG, Kuse AR (1978) Mental rotations, a group test of three-dimensional spatial visualization. Percept Mot Skills 47:599–604.
Wheeler ME, Petersen SE, Buckner RL (2000) Memory's echo: vivid remembering reactivates sensory-specific cortex. Proc Natl Acad Sci USA 97:11125–11129.
Woods RP, Grafton ST, Holmes CJ, Cherry SR, Mazziotta JC (1997) Automated image registration: I. General methods and intrasubject validation. J Comput Assist Tomogr 22:139–152.
Zago L, Pesenti M, Mellet E, Crivello F, Mazoyer B, Tzourio-Mazoyer N (2001) Neural correlates of simple and complex mental calculation. Neuroimage 13:314–327.

Appendix: Texts Read by the Subjects before the PET Experiment

Environment 1

One of the most famous grape harvest festival happens every year in a village called Courteval. Four main landmarks limit Courteval and its surroundings: the White Hills, the Fairy's River, the National Highway and the Cicadas Road. The northern limit is formed by the White Hills. The Fairy's River runs from north to south and forms the western limit of the region. The southern limit is formed by the National Highway. The National Highway spans the Fairy's River by means of the old Roman bridge. The eastern limit is formed by the Cicadas Road that links the National Highway to the White Hills. Northwards, at the end of the Cicadas Road, at the bottom of the White Hills, there is a medieval castle. The village of Courteval mainly extends westwards to the Cicadas Road, at the northern side of the crossroads between this road and the National Highway. Courteval is built around four streets that surround a park. On the east side of the park there is a bandstand. The east side of the park is bordered by the Cicadas Road. The three other sides of the park are bordered by three streets. The south side is bordered by Plane Trees Street. There are hundred-year-old trees along this street. The west side of the park is bordered by Jules Ferry Street. The north side of the park is bordered by Saint-Christophe Street. Three important buildings are on the other side of the streets facing the park: the town hall, the church and the school. The town is on the other side of Cicadas Road and faces the east side of the park. The church is on the other side of Saint-Christophe Street and faces the north side of the park. The school is on the other side of Jules Ferry Street and faces the west side of the park. Northwest of the intersection between the National Highway and Cicadas Road there is a gas station. East of the village, on the edge of the National Highway, there is a supermarket.

Environment 2

The theme park of Aigrettes Lake is an ideal place for people attracted by outdoor activities. The area where the park is located is delimited on its four sides by a common forest, Sirens Bay, the Bay Highway and the Forest Road. Eastwards, the region is bordered by the common forest. Southwards, the region is limited by Sirens Bay. Two highways also constitute the limits of the area, the Bay Highway and the Forest Road. The Bay Highway goes from north to south and constitutes the western border of the area. The Bay Highway leads to Sirens Bay. Southwards, at the end of the Bay Highway, on the edge of Sirens Bay, there is an old lighthouse. The Bay Highway is also the main acess to the region. The northern limit of this region is formed by the Forest Road that stretches over 60 kilometers and links the Bay Highway to the common forest. The police station is located at the north-west crossroad of the Forest Road and the Bay Highway. Eastwards, at the end of the Forest Road, where the common forest starts, there is a hunting club house. The Aigrettes Lake is a big boating lake located at the centre of the area. On the eastern shore of the lake there is an area reserved for swimming. At the most southern point of the lake there is an area reserved for fishing. The road that follows the edge of the lake is Horseshoe Road. The Horseshoe Road is connected to the Forest Road at two points, one at 20 kilometers west of the common forest and the other at 20 kilometres east of the Bay Highway. The Horseshoe Road is the only road to come into the theme park of Aigrettes Lake. Three villages are located in the vicinity of Horseshoe Road. Eastward of the lake, nearby the part of Horseshoe Road, which is close to the common forest, is the village called Juzac. Southward of the lake, midway between Horseshoe Road and Sirens bay, is the village called Saint Laurent. On the western shore of the lake, between the lake and Horseshoe Road is a village called Méricourt. Mericourt faces the swimming area on the other side of the lake.