Functional Anatomy of High-Resolution Visual Mental Imagery

E. Mellet, N. Tzourio-Mazoyer, S. Bricogne, and B. Mazoyer
UPRES 2127 Université de Caen & CEA-LRC 13V, Caen, France

S. M. Kosslyn
Harvard University

M. Denis
LIMSI CNRS, Orsay, France

Abstract

■ This study had two purposes. First, in order to address the controversy regarding activation of the primary visual area (PVA) during visual mental imagery, regional cerebral blood flow (rCBF) was recorded while subjects performed a task that required high-resolution visual mental imagery. Second, in order to discover whether verbal descriptions can engage visual mechanisms during imagery in the same way as visual stimuli, subjects memorized 3D scenes that were visually presented or were based on a verbal description. Comparison of the results from the imagery conditions to a non-imagery baseline condition revealed no activation in PVA for imagery based on a verbal description and a significant decrease of rCBF in this region for imagery based on visual learning. The pattern of activation in other regions was very similar in the two conditions, including parietal, midbrain, cerebellar, prefrontal, left insular, and right inferior temporal regions. These results provide strong evidence that imagery based on verbal descriptions can recruit regions known to be engaged in high-order visual processing. ■

INTRODUCTION

The functional anatomy of visual mental imagery has been a topic of intense debate, particularly as regards the role of early visual areas in the generation and use of mental images (for example, Roland & Gulyas, 1994). At the center of the debate is the hypothesis that during imagery, primary visual cortex and nearby cortical areas are used to reconstruct the local geometry of the surface of the visualized object or scene (Kosslyn, 1994a). This hypothesis is based on the fact that, thanks to the small receptive fields of its neurons, this area specifies shape with high-resolution (Fox et al., 1986). Indeed a number of researchers have reported that this region is activated during tasks requiring visual mental imagery of objects (Bookheimer et al., 1998; Chen et al., 1998; Damasio et al., 1993; Kosslyn et al., 1993, 1999; Kosslyn, Thompson, Kim & Alpert, 1995; Kosslyn, Thompson, Kim, Rauch & Alpert, 1996; Le Bihan et al., 1993; Menon et al., 1993; Sabbah et al., 1995). Moreover, it has been reported that, as occurs during actual perception, mental images rely on a topographical representation in primary visual cortex: mental imagery of small objects activates the most posterior part of calcarine sulcus, whereas mental imagery of increasingly larger objects activates increasingly more anterior parts of this sulcus (Kosslyn et al., 1993, 1995). However, other laboratories have not found evidence that early visual cortex is activated during visual mental imagery. During various visual mental imagery tasks, such as mental imagery of 2D colored patterns (Roland & Gulyas, 1995), mental exploration of a 2D scene (Mellet, Tzourio, Denis, & Mazoyer, 1995), mental construction of 3D patterns (Mellet et al., 1996), mental navigation (Ghaëm et al., 1997), or mental imagery of objects based on their definitions (Mellet, Tzourio, Denis & Mazoyer, 1998a), no activation of the primary or adjacent early visual cortex was observed. In fact, in some cases the regional cerebral blood flow (rCBF) values were actually lower than those found in the baseline conditions.

The disparity in results could arise from a number of sources. For example, none of the studies that found activation in early visual cortex during visual mental imagery used spatial tasks, whereas most of the studies that failed to find such activation relied upon such tasks. In addition, most of the studies that failed to find activation in these areas relied on stimuli stored in long-term memory, whereas most of those that did

find such activation relied on stimuli learned in the minutes prior the experiment. Moreover, many of the studies that failed to find such activation tested subjects who scored high on spatial abilities tests, whereas none of those that found such activation selected subjects in this way. And the studies differed in many other ways, such as the baseline conditions used. Unfortunately, the corpus of available results does not reveal any clearcut pattern that allows one to infer which factors are crucial and which incidental for producing activation in early visual areas during visual mental imagery.

One recent attempt to reconcile these divergent results posits that the recruitment of early visual areas during visual mental imagery depends on the spatial resolution of the mental images (Kosslyn, Thompson & Alpert, 1997; Thompson & Kosslyn, 1998). As mental images become increasingly vivid and accurate, they come to resemble more closely percepts—and hence it is plausible that such images might recruit early visual areas that may underlie such properties in perception. The present study was designed to determine whether tasks that require high-resolution visual mental images will always engender activation of early visual cortex. In the present task, the subjects must make relatively difficult discriminations, which can be performed well only if the image specifies shape with high-resolution. The discriminations we use are of comparable difficulty to those used in the Kosslyn et al. (1995) study, which did find activation of the primary visual area during imagery. Thus, if we failed to obtain such activation in the present study, this failure cannot be ascribed to this factor.

Another aspect of the task could have an effect on activity in the primary visual cortex during mental imagery: the learning modality. Visual mental images can be created either on the basis of previously stored visual memories of objects or scenes, or on the basis of verbal descriptions. However, previous behavioral studies have shown that visual mental images constructed from verbal descriptions are functionally very much like visual mental images constructed from visual memories (Denis & Cocude, 1992). In a previous neuroimaging study, Mellet et al. found that the construction of mental images from verbal instructions relied on associative visual areas, but did not give rise to activation in PVA. It is possible that early visual areas tend to be recruited only when previously stored visual memories are recalled. In order to clarify this point, our experimental design included the two types of learning.

The goal of the present study was to elucidate the effect of both, requiring the subjects to use high-resolution images and the learning modality on the functional anatomy of visual mental imagery. In order to address these issues, we studied performance and functional anatomy during the same mental imagery task when the

images arise from memories that initially were encoded visually or verbally. Using 15O-labeled water, nine sequential PET measurements of regional rCBF were obtained for each of seven subjects. We conducted three replications of each of three experimental conditions: a high-resolution mental imagery task based on visually learned material, the same task based on verbally learned material, and a baseline task. During the learning phase in the visual learning condition, the subjects were asked to study and memorize scenes prior to scanning. During the learning phase in the verbal learning condition, the subjects were instructed to listen to verbal descriptions of how shapes were to be arranged, and to form and memorize a visual image for each of the described scenes. This was also done prior to scanning. In both imagery conditions, each scene was composed of four simple geometric shapes, arranged on a base; the scenes differed only in the ordering of the elements on the base (see Figure 1A). During the imagery conditions, the subjects visualized one of the scenes, then they heard the labels of two positions on the base and a statement of relative height; they were to consult their image and decide whether the statement correctly characterized the relative heights of the scene above the two named positions on the base. In all cases, the judgments required high-resolution in the images. Finally, in addition to the two imagery conditions, subjects simply heard the stimuli and responded with-
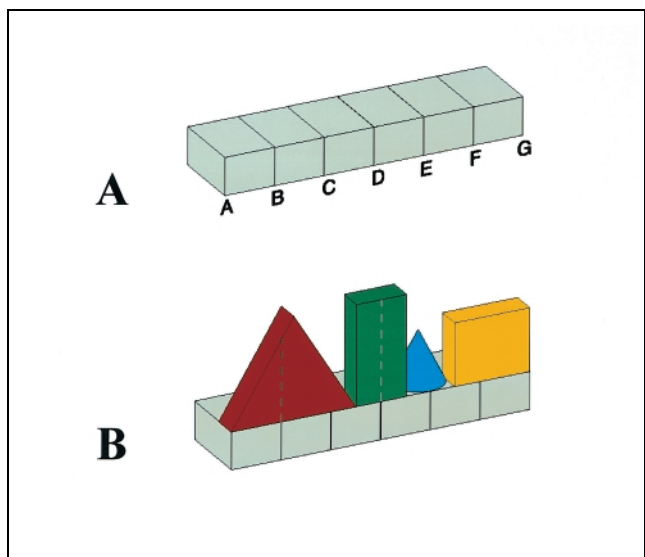

Figure 1. (A) The base on which four geometric shapes were placed. Before the PET experiment, each subject learned the locations specified by the letters. (B) An example of a typical 3D scene that the subject had to learn from a visual or a verbal presentation. During the PET measurements, the subjects, eyes closed, had to form a visual mental image of the scene and evaluate statements about height comparisons, such as ''D lower than B'', heard through earphones. In the present example, the answer would be ''wrong''.

out visualizing or making judgments; this was the baseline condition.

RESULTS

Behavioral Results

During the PET measurements, we recorded the number of correct responses and the response times. The percentage of correct responses in the visual learning condition was 81.5±3.2 %, and in the verbal learning condition it was 80.6±4.8 % (mean±standard error of the mean). In both conditions, the subjects clearly performed better than chance ($p=0.005$ and $p=0.01$, respectively, t-test), and there was no difference in accuracy between the two conditions ($p=0.8$ paired t-test). A slight improvement in the number of correct responses in the third replication of the task was evident in the visual learning condition ($p=0.05$, ANOVA for repeated measures), but no such practice effect was evident in the verbal learning condition ($p=0.7$). The mean response time over subjects was 8.852±0.295 sec (mean±standard error of the mean). The mean response time in the visual learning condition was 8.691±0.412 sec, which did not differ from that in the verbal construction condition (9.014±0.424 sec, $p=0.7$ paired t-test). No replication effect was observed for the mean response time in either condition ($p>0.3$ in both cases).

PET Results

We report below the results of a series of contrasts, each of which reveals different aspects of the neuroimaging results.

Imagery Based on Visual Learning Compared to Baseline

As presented in Table 1 and Figure 2 (left) during the mental imagery task based on visually learned material, the activation exhibiting the largest Z score extended from the precuneus symmetrically in the two hemispheres toward the intraparietal sulcus and in the superior occipital gyrus. This activation was not symmetrical: It had a larger anterior extent in the left intraparietal sulcus, up to the postcentral sulcus.

A large cluster of activation, extending from $z=10$ mm above the AC–PC line to $z=-36$ mm below this

Table 1. High-Resolution Imagery Based on Visual Learning Minus Baseline ($p<.001$)

| Region Size | Anatomical Location of Max. Voxel | Coordinates | | | Z Score |
		x	y	z	
2124	L. intraparietal sulcus*	−26	−58	44	5.8
	L. superior occipital gyrus*	−26	−74	40	5.5
	R. superior occipital gyrus*	34	−78	42	4.5
	R. intraparietal sulcus	16	−74	52	3.7
4332	Ponto-mesencephalic tegmentum*	−2	−30	−8	5.3
	Ponto-mesencephalic tegmentum*	−2	−16	−12	5.0
	L. medial thalamus	−6	−26	10	3.7
	Cerebellar vermis*	−6	−54	−20	5.0
	L. cerebellar hemisphere	−14	−48	−34	4.2
	R. cerebellar hemisphere	46	−64	−34	4.2
58	R. cerebellar hemisphere	32	−66	−50	3.6
37	L. medial thalamus	−8	−4	14	3.8
368	L. anterior insula/inf. frontal gyrus*	−30	22	8	5.1
104	R. middle frontal sulcus*	54	32	24	4.9
	R. inferior frontal gyrus	56	32	16	3.8
44	R. inferior frontal gyrus	60	14	−4	4.0
353	R. superior frontal gyrus	38	54	18	3.9
55	R. inferior temporal gyrus	62	−58	−12	3.6
58	R. anterior cingulate cortex	8	20	34	3.5

Foci of significant normalized regional cerebral blood flow (NrCBF) increases when imagery was based on visual learning compared with the baseline condition. The data, based on seven subjects, are local maxima detected with SPM software. Activated region volumes are given in voxels. Within these regions, the anatomical localization of the maximum Z scores of the voxel is given on the basis of the Talairach and Tournoux atlas (1988), using their stereotactic coordinates in mm. (R., right; L., left).

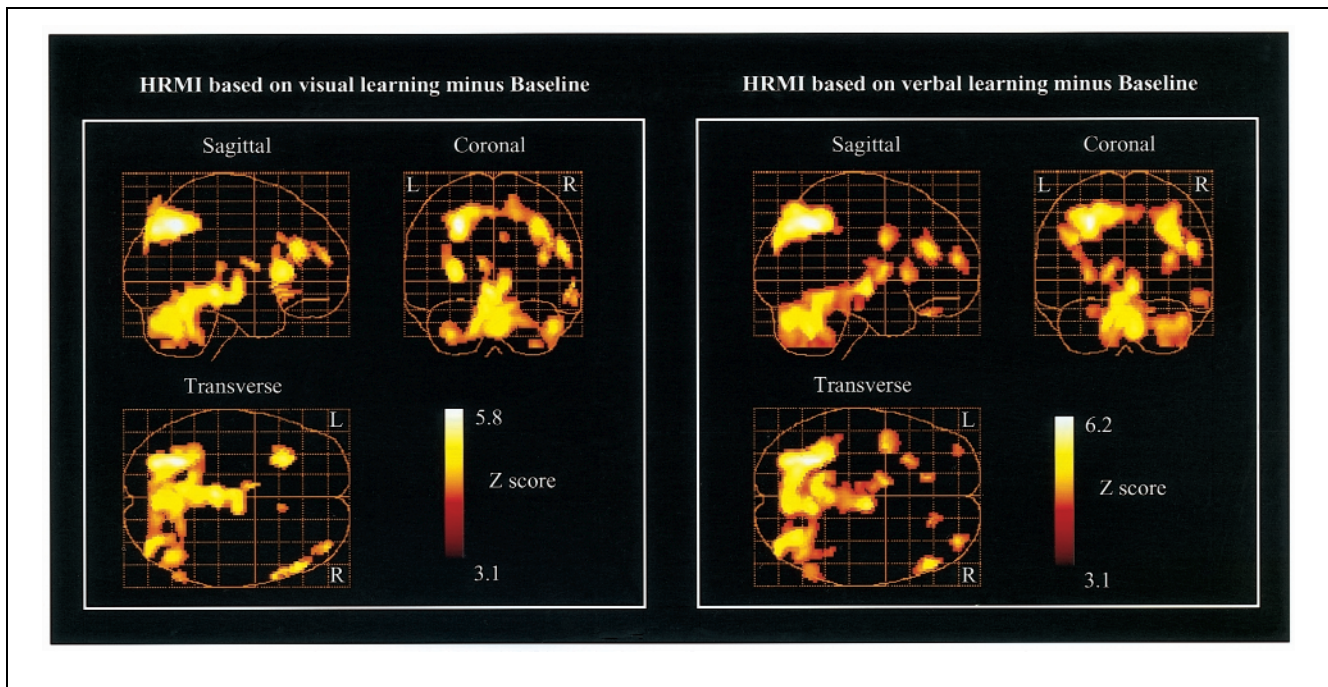*Significant at $p<.05$ corrected for multiple comparisons.

Figure 2. Statistical parametric map revealing the significant activations when subjects formed images based on visual learning compared to the baseline condition (Left) and when they formed images based on verbal learning compared to the baseline condition (Right). The volumes are thresholded to project in three orthogonal directions, sagittal, coronal, and transverse and reach threshold at $Z=3.1$ ($p<.001$; not corrected for multiple comparison). Stereotactic coordinates of local maxima are given in Table 1 and Table 2. HRMI=High Resolution Mental Imagery.

line, encompassed both right and left thalami in their medial part (although the spatial resolution did not allow to distinguish between the medial dorsal nucleus and the medial part of the pulvinar) and spread downward to the ponto-mesencephalic tegmentum, including the midbrain reticular formation, and to the cerebellar vermis and the dentate nuclei of the cerebellum.

A significant increase in rCBF in the left anterior insula was also detected, which extended anteriorly in the depth of the left inferior frontal sulcus. In the opposite hemisphere, a cluster of activation was detected in the superior frontal gyrus together with a distinct cluster that encompassed the middle and the inferior frontal gyrus. In addition, in the ventral system, an activation was detected in the right inferior temporal gyrus.

Finally, a small focus of activation was observed in the right anterior cingulate cortex. Notably, there was no medial occipital activation, in the primary visual cortex (PVA) or other and nearby early visual cortex.

Imagery Based on Verbal Learning Compared to Baseline

As presented in Table 2 and Figure 2 (right), the pattern of activation evident in this comparison was very similar to that found when imagery was based on visually

learned stimuli. Significant activations were located in the precuneus, bilateral intraparietal sulcus and bilateral superior occipital gyrus, as well as in the same subcortical structures as found in the visual learning condition (i.e., the medial part of the thalami, midbrain and mesencephalic tegmentum), extending to the vermis and the cerebellar hemispheres. The right superior middle and inferior frontal gyri were also activated. An additional focus of activation was detected at the intersection of the left middle frontal gyrus and the precentral gyrus. Note that this last activation was also present in the imagery based on visual learning vs. baseline comparison but failed to reach the $p<0.001$ significance level.

We again found activation in the ventral system, in the right inferior temporal gyrus. An activation in the anterior insula was also detected, although it was less extensive than when subjects formed images based on visually learned stimuli. In addition, small clusters of activation were detected bilaterally in the gyrus rectus. Finally, no medial occipital activation, in PVA or nearby cortex, was detected.

Conjunction of the Two Imagery Conditions

In order to quantify the similarities of the patterns of activation in the two imagery conditions, we conducted a conjunction analysis (Price, Moore & Friston,

Table 2. High Resolution Imagery Based on Verbal Learning Minus Baseline ($p<.001$)

| Region Size | Anatomical Location of Max. Voxel | Coordinates | | | Z Score |
		x	y	z	
2385	L. intraparietal sulcus*	−26	−58	44	6.2
	L. superior occipital gyrus*	−26	−72	40	5.7
	L. intraparietal sulcus*	−18	−56	52	5.5
1079	R. intraparietal sulcus*	32	−68	40	5.3
	R. superior occipital gyrus*	38	−82	32	5.1
4598	R ponto-mesencephalic tegmentum*	6	−14	−6	5.5
	Cerebellar vermis*	8	−58	−36	5.3
	Cerebellar vermis*	8	−74	−32	5.1
	L. cerebellar hemisphere*	−12	−50	−30	4.7
	R. cerebellar hemisphere	44	−64	−32	4.1
	R. ponto-mesencephalic tegmentum	6	−28	−10	3.9
	Ponto-mesencephalic tegmentum	−2	−32	−8	3.6
223	L. medial thalamus	−10	0	10	4.2
	L. medial thalamus	−8	−8	12	3.6
176	R. middle frontal gyrus*	52	32	24	4.8
	R. middle frontal sulcus	58	26	20	3.4
	R. inferior frontal gyrus	56	34	12	3.3
285	L. precentral/L. middle frontal sulcus	−38	0	30	4.3
131	R. superior frontal sulcus	34	54	16	4.2
142	L. anterior insula	−28	16	4	4.1
21	L. gyrus rectus	−14	−32	−26	4.0
18	R. gyrus rectus	20	34	−24	3.8
147	R. inferior temporal gyrus	56	−58	−14	3.8

Foci of significant NrCBF increases when imagery was based on verbal learning compared with the baseline condition (for details, see Table 1 legend).

1997).The results are presented in Table 3. This analysis revealed an overall pattern of activation that is very similar to the one described in the two previous sections. We again found bilateral occipito-parietal activation, as well as the medial activation encompassing the medial part of the thalami, midbrain, and mesencephalic tegmentum extending to the vermis and the cerebellar hemispheres. In addition, we found activation in the right superior, middle frontal, and inferior frontal clusters, and in the left intersection of the precentral and middle frontal sulci, as well as the right inferior temporal gyrus.

Comparison of the Two Imagery Conditions

As presented in Table 4, we compared the two imagery conditions to discover which areas were more active when images were based on visual learning than when they were based on verbal learning. This comparison revealed that the two conditions differed only in the locus of a single cluster of activation, in the left anterior insula/left inferior frontal gyrus. Although this activation was present in both imagery conditions when they were compared to the baseline condition, the present comparison revealed that the inferior frontal expansion of this cluster was present only when imagery was based on visual learning.

Table 4 also presents the results of the reverse comparison, observing which areas were more active when imagery was based on verbal learning than when it was based on visual learning. This comparison revealed several clusters of significant rCBF increases, notably in the posterior part of the right calcarine sulcus, which corresponds to the right primary visual area (Figure 3). However, examination of the rCBF values across the three conditions in these clusters (see values in Table 4 and histograms in Figure 3) revealed that blood flow was significantly lower when imagery was based on visual learning than in the other two conditions, with no difference between the verbal learning and baseline conditions. Finally, activations were also detected in the left occipito-parietal sulcus, the left Heschl gyrus and the left caudate nucleus, all of which had the same pattern of rCBF across the three

Table 3. Conjunction Visual/Verbal ($p<.001$)

| Region Size | Anatomical Location of Max. Voxel | Coordinates | | | Z Score |
		x	y	z	
3546	L. intraparietal sulcus*	−26	−58	44	7.3
	L. superior occipital gyrus*	−26	−74	40	6.8
	R. superior occipital gyrus*	34	−78	40	5.4
6300	Ponto-mesencephalic tegmentum*	6	−14	−6	6.1
	Cerebellar vermis*	−6	−54	−22	6.0
	Cerebellar vermis*	−6	−58	−36	5.7
	L. medial thalamus*	−10	0	10	4.5
467	R. middle frontal gyrus*	52	32	26	5.7
	R. superior frontal sulcus*	36	54	18	4.7
	R. inferior frontal gyrus	56	32	16	4.4
306	L. anterior insula*	−28	20	6	5.2
188	L. precentral/middle frontal sulcus	−36	0	32	4.3
175	R. inferior temporal gyrus	60	−58	−12	4.2
41	L. superior frontal sulcus	−38	54	22	3.9
16	R. rectus gyrus	20	34	−24	3.9
56	R. lenticular nucleus	24	16	8	3.6
72	L. cerebellar hemisphere	−32	−58	−40	3.6
19	R. insula	38	16	4	3.5

Conjunction analysis revealing foci of significant NrCBF increases common to imagery task based on visual and verbal learning, as compared to baseline (see Table 1 for details).

conditions (that is, a decrease in rCBF during the visual learning imagery condition).

DISCUSSION

We now return to the two major issues that led us to conduct this study, the question of whether the requirement to use high-resolution imagery will invoke activation in early visual areas, and the relation between imagery based on visually learned stimuli vs. imagery based on verbally learned stimuli. We interpret the results as they bear on these issues.

### Medial Occipital Cortex and High-Resolution Mental Images

One goal of this study was to test the hypothesis that the medial occipital cortex, including the primary visual

Table 4.

| Region Size | Anatomical Location of Max. Voxel | Coordinates | | | Z Score | Adjusted rCBF | | |
		x	y	z		RE	VI	VE
Visual learning minus Verbal learning (p<.001)								
18	L. anterior insula / inf. frontal gyrus	−32	26	10	3.5	68.9	72.4	69.7
Verbal learning minus Visual learning (p<.001)								
99	R. primary visual area	22	−94	−2	4.2	71.7	69.6	72.5
104	L. Heschl gyrus	−44	−26	10	4.0	90.8	88.3	91.4
114	L. occipito-parietal sulcus	−20	−68	12	3.8	70.4	69.0	71.4
20	L. caudate nucleus	12	18	6	3.6	66.1	64.4	66.4

Foci of significant NrCBF increases when imagery was based on visual learning compared with when imagery was based on verbal learning and vice versa (for details, see Table 1 legend).

Figure 3. (a) Statistical parametric map revealing the significant activation clusters when subjects formed imagery based on verbal learning minus imagery based on visual learning (in yellow) and vice versa (in red). The volumes are thresholded to project in three orthogonal directions, sagittal, coronal, and transverse, and reach threshold at $Z=3.1$ ($p<.001$; not corrected for multiple comparison). Stereotactic coordinates of local maxima are given in Table 4. (b) Adjusted blood flow (i.e., based on normalized counts) in the right primary visual area ($x=22$ mm., $y=-94$ mm. $z=-2$ mm.) during the three conditions, which illustrates the drop of rCBF in this area during the imagery task after visual learning. (c) Cluster of rCBF differences in the right posterior primary visual area evidenced by the subtraction VE minus VI. The cluster is superimposed on three contiguous sagittal slices from the template that go through $x=20$ mm to $x=24$ mm. PVA=Primary Visual Area.
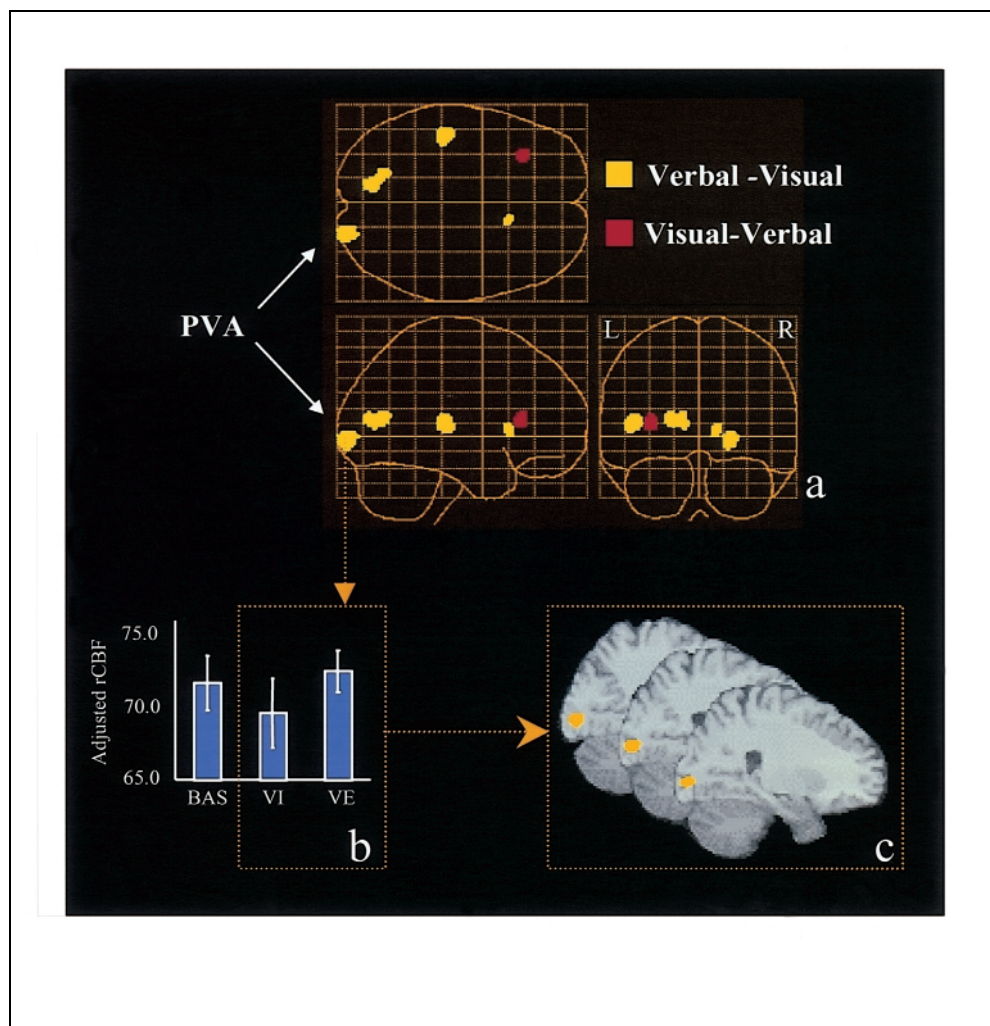
cortex, may be involved in the generation and the inspection of visual mental images, when the pattern must preserve fine-grained shape variation (and thus require high-resolution; see Thompson & Kosslyn, 1998). The present results challenged the hypothesis that these structures are involved when high-resolution imagery is formed and used, given that no significant increase of rCBF was present when high-resolution mental imagery was required to perform the task. However, the fact that we did not find activation of early visual areas does not mean that nothing happened in this region during the task; indeed, PVA was one of the few regions where the learning modality produced a difference in rCBF. It is not clear why a significant deactivation occurred during our imagery task based on visually-learned material, but this result illustrates the complex behavior of such areas during visual mental imagery. One cannot exclude that a top-down modulation may explain this observed decrease in rCBF, when imagery was based on visual presentation (both when compared to baseline and imagery based on verbal presentation, as evident in Table 4). Further investigations will be necessary to understand the different possible roles

of top-down modulation in the generation, and inspection, of visual mental image.

However, probably the most important point here is that high-resolution mental imagery per se does not activate PVA; thus, this variable cannot explain the divergent results between studies noted in the Introduction.

Activation Patterns Compared to Baseline

For the most part, the modality in which the scenes were initially learned did not affect the processes that later allow one to visualize the scene. Apart from the medial occipital cortex, as discussed above, the strong similarity between the results in the two imagery conditions suggests that the subjects used the same strategy to perform the task when they learned the scenes verbally or visually. This result dovetails with those from previous behavioral studies that have shown that visual mental images built from a verbal description exhibit the same metric properties as images derived from perception (Denis & Cocude, 1992; Denis, Goncalves & Memmi, 1995).

Because the present study directly compared the neural correlates of mental imagery learned in each

way in the same subjects, our results provide an anato-mo-functional substrate to these behavioral reports. The similarity in the results from the two imagery conditions indicates that the cerebral structures that underlie the generation and the maintenance of the mental image, as well as the operations performed on the image, did not depend on the modality in which the information was originally acquired. This finding is all the more impressive because the imagery task, performed following visual or verbal learning, included several different steps. First, when they heard the number identifying the scene, subjects had to visually recall the scene and then retain a high-resolution visual image until the comparison statement was delivered. Then, they had to identify the locations corresponding to the named letters, and assess the height difference at the two points. Following this, they had to select and produce the appropriate response. In the following sections we will discuss how memory retrieval, visual working memory, attentional shifting, and high-resolution shape processing could account for the activation that we observed.

Spatial Components and Occipito-Parietal Activation

The bilateral activation of superior occipito-parietal areas could reflect the spatial processing required by the task: The subject had to form a mental image in which each geometric shape was accurately placed on the base and positioned over the corresponding letters. The superior occipital gyrus and the posterior part of the intraparietal sulcus are portions of the dorsal system, which is specialized for the processing of spatial information (Haxby et al., 1994). This pathway is also involved in the spatial processing of visual mental images, as indicated in numerous neuroimaging studies of mental imagery that included a spatial component (Cohen et al., 1996; Kawashima, Roland & O'Sullivan, 1995; Kosslyn et al., 1993, 1998; Mellet et al., 1995; 1996; Tagaris et al., 1997). In this context our results are consistent with the role played by the dorsal system in the spatial processing of visual mental images and show, in addition, that these highly specialized visual areas are recruited independently of the modality in which the information was acquired, visually or verbally.

Object Visual Imagery and the Ventral System

In order to perform our task, the subjects had to form images of objects—and, thus, we would expect activation in the ventral system. The involvement of the ventral system during object imagery has been consistently reported in previous studies (D'Esposito et al., 1997; Kosslyn et al., 1993; Mellet et al., 1996; 1998a; Roland & Gulyas, 1995). The right inferior temporal activation detected in both imagery tasks is thus consistent with the role of this cortical area in the processing of shapes in both visual perception and visual imagery. However, the right lateralization of this activation challenges the leftward functional specialization of the inferior temporal gyrus for mental image generation proposed by others (D'Esposito et al., 1997). It has previously been suggested that the right inferior temporal gyrus may be activated by the generation of complex mental images (Mellet et al., 1998a; Mellet, Petit, Mazoyer, Denis & Tzourio, 1998b), and the images required in the present study appear to qualify as "complex." Indeed, most of the studies based on complex and/or detailed mental imagery reported an activation of the right inferior temporal gyrus (Kosslyn et al., 1993; Mellet et al., 1996; 1998a), whereas those that used simple and/or poorly detailed imagery failed to activate this cortical area (Bookheimer et al., 1998; D'Esposito et al., 1997). Our results are thus consistent with a specific role of the right inferior temporal gyrus in this type of imagery.

Memory Components and Frontal Activations

During the imagery tasks, the frontal lobe was activated in three distinct locations: (1) a right dorsolateral prefrontal area that included an activation in the superior frontal gyrus, middle frontal, and inferior frontal gyri; (2) the intersection of the left middle frontal sulcus and the precentral gyrus; and, (3) the border of the left inferior frontal gyrus and of the anterior insula. These regions may constitute a network that underlies the memory activity involved in visual mental imagery. The right dorsolateral prefrontal activations that we found could be related to retrieval from episodic memory, a crucial prerequisite of our mental imagery task (Buckner, Raichelle, Miezin, & Petersen, 1996; Moscovitch, Kapur, Köhler, & Houle, 1995; Roland & Gulyas, 1995; Tulving, Kapur, Craik, & Moscovitch, 1994). We also found activation in the left middle frontal sulcus. Note that although this activation just failed to reach the $p < 0.001$ significance level in the visual learning condition (the CBF increase was 3.9% as compared to baseline, Z score $= 3.0$), it cannot be considered as specific to the verbal learning condition; indeed, there was no observed difference in activation in this area between the two conditions when we directly compared the results. There are several reasons to assume that this activation may be related to object imagery. An activation in this region has been reported in an imagery task where subjects had to visualize objects on the basis of their dictionary definitions (Mellet et al., 1998a). Moreover, this region has been found activated in various working memory tasks dealing with objects and faces (Cohen et al., 1997; Courtney, Ungerleider, Keil & Haxbyy, 1997; Smith et al., 1995). It is also of interest that a previous study reported a CBF increase in this area, while subjects had to identify objects seen from non-canonical perspectives, which may involve imagery (Kosslyn, 1994a; Kosslyn et al., 1994b).

The activation of the left anterior insula remains difficult to interpret. Most of our subjects reported that, during the mental inspection of the image, they silently rehearsed the comparison term they had just heard. The activation of the left anterior insula could be related to such verbal rehearsal, which is consistent with the finding that this cortical region was activated during verbal working memory studies (Paulesu, Frith, & Frackowiak, 1993; Smith, Jorrides, & Koeppe, 1996).

Mesencephalo-Thalamic and Cerebellar Activations

The imagery task required the subjects to attend to the visualized shapes, and such attention may in turn have drawn on systems that keep the brain aroused. The mesencephalon encompasses a part of the ascending activating reticular system in humans, and it has been demonstrated that it plays a key role, together with the medial thalamus, in maintaining a state of high vigilance and attention (both in vision, for example, Kinomura, Larsson, Gulyás, & Roland, 1996, and audition, for example, Paus et al., 1997). However, in our study, most of the attentional resources were focused on a purely mental scene. Our results, thus, emphasize that the midbrain activating system is not only engaged by attention to sensory input, but also can be involved in absence of perceptual input.

We also observed an important activation of the vermis that extended bilaterally to an output nucleus, the dentate nucleus. A cerebellar contribution to cognitive operations is 1 of the major findings provided by recent imaging studies (see Schmahmann, 1996). Medial activation of the cerebellum has been repeatedly reported during visual episodic memory tasks (Fink et al., 1996; Moscovitch et al., 1995) and during short-term verbal maintenance (Fiez et al., 1996). As noted above, these cognitive processes were required to perform our task. We found cerebellar activations that encompassed bilaterally the dentate nucleus. Among the cerebellar deep nuclei, the dentate seems to be particularly involved in higher cognitive functions (Middleton & Strick, 1994). Although it is not yet possible to assign a specific role to this output nucleus, the present results, together with those from previous studies, support the inference that this region deals with internal representation; indeed, bilateral activations of this nucleus have been described during a pegboard puzzle solving task (Kim, Ugurbil, & Strick, 1994), a sensory discrimination task (Gao et al., 1998), and a conceptual reasoning task derived from the Wisconsin Sorting Card Test (Rao et al., 1997). It is of interest that in nonhuman primates this region is connected via the medial thalamus to the prefrontal cortex (Middleton & Strick, 1994). Consistent with the results of the present study, Andreasen et al. propose that such a network plays a role in humans in the coordination and sequencing of thoughts (Andreasen, Paradiso, & O'Leary, 1998).

EXPERIMENTAL PROCEDURES

Subjects

Seven right-handed healthy male French students (age: 20–25 years) volunteered to participate in this study. All were free from nervous disease or injury and had no abnormality on their T1-weighted magnetic resonance images (MRI). Written, informed consent was obtained from each subject after the procedures had been fully explained. This study was approved by the local Institutional Review Board. In order to ensure optimal homogeneity of the sample of the subjects with respect to their imagery abilities, subjects were selected as ''high imagers'' (or, more specifically, as having high spatial ability) on the basis of their scores on the Minnesota Paper Form Board (MPFB) and on the Mental Rotations Test (MRT); all subjects scored beyond the 75th percentile of a population of 208 male subjects. The mean MPFB score for the subjects was 23.9±1.3 (mean±SD, whole population 19.7±4.4) and their mean MRT score was 15.6±1.8 (whole population 11.8±4.4).

Materials

The subjects were first trained to form a vivid mental image of a base, as illustrated in Figure 1A. This base included letters that demarcated specific positions along it. Scenes were constructed by placing four shapes on the base, a red triangle, a green rectangle, a yellow rectangle, and a blue cone. Figure 1B provides an illustration of one of the scenes. Twenty-four different scenes were thus available, but 16 only were used in the experimental paradigm: four in the training phase prior the PET session (two for training in the visual learning imagery condition and two for training in the verbal learning imagery condition), two in each PET condition repeated three times for each type of learning (12 scenes). In addition to the 16 scenes, 16 sets of nine auditorily delivered comparison statements were prepared. These statements were delivered by computer to headphones. The notion of high-resolution mental images has been operationalized elsewhere as follows: A task requires high-resolution if one needs to attend to distinct portions of shapes that have subtle variations or if the to-be-visualized patterns have parts that subtend less than 0.5° of visual angle and these parts need to be kept distinct from other part (Thompson & Kosslyn, 1998). The high-resolution mental imagery task we designed in our study fulfilled both criteria.

Procedure

Behavioral Procedure

We began with a training phase, prior to scanning. In the initial training, the subjects learned two scenes. As occurred in the subsequent experimental conditions,

the two scenes were identified as "1" and "2." Prior to learning the scenes proper, subjects first memorized the base, including the letters and their precise positions. They then were shown the geometric forms, one at a time, and asked to memorize their appearance.

The subjects were taught two methods of learning the scenes: from a visual presentation and from a verbal description. In the visual learning condition, they saw a complete scene, formed by placing four shapes on the base. The subjects were requested to memorize the scene as accurately as possible. After studying a scene, the subjects closed their eyes and visualized it, and then opened their eyes and compared the image against the actual stimulus, correcting their image as necessary. This procedure was repeated until the subjects felt that their mental image corresponded to the display. The stimuli were left in free view for as long as the subjects required to memorize the display. The same procedure was used with scene "2".

In the verbal learning condition, the subjects never actually saw the scenes. With eyes closed, the subjects were first asked to recall the mental image of the base, including the letters and their precise positions along the base. Then they heard a description of the order in which geometric forms should be arranged on the base for stimulus number 1; for example, the construction of Figure 1B would be described as "red triangle, ... green rectangle, ... blue cone, ... yellow rectangle". The subjects were asked to build and memorize the visual mental image of the corresponding scene. They then heard the description of scene "2", and again built and memorized the corresponding scene. This procedure was repeated until the subjects reported that they had memorized the two scenes (3 times on average).

After the subjects had learned the initial two scenes, whichever the type of learning, they were trained to perform the mental imagery task proper: with eyes closed, they first heard the identifying number of one of the two scenes through earphones, which prompted them to generate a mental image of the whole scene, as accurately as possible, with sharp corners and edges. Four seconds later, they heard a question corresponding to a height comparison term, such as "D lower than B," and had to mentally compare the heights of the two shapes being visualized immediately above the named points (see Figure 1B, dotted lines). If the statement was correct, the subjects were to respond by saying "right"; if it was incorrect, they were to say "wrong." The subjects were asked to respond as quickly as possible, while still remaining as accurate as possible. Once the subjects had responded, the identification number of the other scene was delivered after a lag of 750 msec, the subjects were then to generate the corresponding mental image, and four seconds later they heard a new comparison term.

The subjects were interviewed to ensure that they understood the nature of the learning procedures and task, and any confusions were corrected and questions answered.

PET Procedure

PET conditions. There were three PET conditions: Imagery after visual learning, imagery after verbal learning, and a baseline. For the two imagery conditions, the subjects memorized two scenes during the 15 min prior to scanning. They used the same method to memorize the scenes as was taught during the training phase. During the PET measurements, the subjects performed the imagery task, with the cues and comparison statements being delivered through earphones. The subjects had to respond by saying "right" or "wrong" into a microphone connected to a computer. After each response, the computer recorded the response time, then 750 msec later delivered the identification number of the other scene, and 4 sec later a new comparison statement. The investigator recorded whether each response was correct or incorrect. Each condition consisted of nine comparison statements, alternating from one scene to the other. A different set of nine comparison terms was used in each of the conditions in each replication, so that the subjects could not know in advance which judgments they would need to perform. Thus, the subjects could not simply perform the judgments and memorize them during the learning phase. In addition to the two imagery conditions, there was a baseline task. During the baseline task, the subjects closed their eyes, listened to randomly chosen comparison statements delivered every seven seconds, and alternatively said "right" and "wrong" after each term. The subjects were instructed not to produce mental images. Although the subjects had, at this point, already gone through the training phase, and, thus, knew the nature of the task, they all reported that they could refrain from forming images or performing the task during the baseline measurements. All conditions were conducted with eyes closed in total darkness.

Image data acquisition. For each rCBF measurement, 63 2.425 mm thick contiguous brain slices were acquired simultaneously on an ECAT HR+PET camera. A black tent-like chamber was set up all around the PET tomograph, to ensure that the subjects were in total darkness. Emission data were acquired in 3D mode. Tasks were started 20 sec before the intravenous bolus injection of 8 mCi of 15O-labeled water. A single 90-sec scan was acquired and reconstructed (including a correction for head attenuation using a measured transmission scan) with a Hanning filter of 0.5 mm.$^{-1}$ cut off frequency and a pixel size of 2×2 mm^2. The between-scan time interval was 15 min, and the condition order was the same in all subjects, specifically:

baseline, imagery after visual learning, and imagery after verbal learning, and this series was replicated three times.

Data analysis. After automatic realignment (AIR, Woods, Grafton, Holmes, Cherry, & Mazziotta, 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 (scaling). The comparisons across conditions were made by way of *t*-statistics. Statistical parametric maps corresponding to comparisons between the three conditions were generated with the 1996 version of SPM (Friston et al., 1995). As indicated earlier, the experimental protocol was designed to use the baseline condition as a control condition for both imagery tasks. The two imagery conditions were also compared directly. For each comparison, the voxel amplitude t-map was transformed in a Z volume that was thresholded at $Z_0 = 3.09$, which corresponds to a 0.001 confidence level (without correction for multiple comparisons).

The conjunction analysis was performed using the same SPM package. It revealed the common sites of activation in the two imagery conditions compared to the baseline condition (Price et al., 1997). This type of analysis required at least two pairs of tasks. In our study, the first pair corresponded to the two imagery tasks, whereas the second pair consisted of two different replications of the baseline condition. As in the subtraction analysis, the resulting Z volume was thresholded at $Z_0 = 3.09$.

Acknowledgments

REFERENCES

Andreasen, N. C., Paradiso, S., & O'Leary, D. S. (1998). "Cognitive Dysmetria" as an integrative theory of schizophrenia: A dysfunction in cortical–subcortical–cerebellar circuitry? *Schizophrenia Bulletin, 24,* 203–218.

Bookheimer, S. Y., Zeffiro, T. A., Blaxton, T. A., Gaillard, W. D., Malow, B., & Theodore, W. H. (1998). Regional cerebral blood flow during auditory responsive naming: evidence for cross-modality neural activation. *NeuroReport, 9,* 2409–2413.

Buckner, R. L., Raichle, M. E., Miezin, F. M., & Petersen, S. E. (1996). Functional anatomic studies of memory retrieval for auditory words and visual pictures. *The Journal of Neuroscience, 16,* 6219–6235.

Chen, W., Kato, T., Zhu, X. -H., Ogawa, S., Tank, D. W. & Ugurbil, K. (1998). Human primary visual cortex and lateral geniculate nucleus activation during visual imagery. *NeuroReport, 9,* 3669–3674.

Cohen, J. D., Peristein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., & Smith, E. E. (1997). Temporal dynamics of brain activation during a working-memory task. *Nature, 386,* 604–607.

Cohen, M. S., Kosslyn, S. M., Breiter, H. C., DiGirolamo, G. J., Thompson, W. L., Anderson, A. K., Bookheimer, S. Y., Rosen, B. R., & Belliveau, J. W. (1996). Changes in cortical activity during mental rotation: A mapping study using functional magnetic resonance imaging. *Brain, 119,* 89–100.

Courtney, S. M., Ungerleider, L. G., Keil, K., & Haxby, J. V. (1997). Transient and sustained activity in a distributed neural system for human-working memory. *Nature, 386,* 608–611.

D'Esposito, M., Detre, J. A., Aguirre, G. K., Stallcup, M., Alsop, D. C., Tipett, L. J., & Farah, M. J. (1997). A functional MRI study of mental-image generation. *Neuropsychologia, 35,* 725–730.

Damasio, H., Grabowski, T. J., Damasio, A. R., Tranel, D., Boles-Ponto, L. L., & Watkins, G. L. (1993). Visual recall with eyes closed and covered activates early visual cortices. *Neuroscience Abstracts, 19,* 1-603 (Abstract).

Denis, M. & Cocude, M. (1992). Structural properties of visual images constructed from poorly, or well-structured, verbal descriptions. *Memory and Cognition, 20,* 497–506.

Denis, M., Goncalves, M. R., & Memmi, D. (1995). Mental scanning of visual images generated from verbal descriptions: Toward a model of image accuracy. *Neuropsychologia, 33,* 1511–1530.

Fiez, J. A., Raife, E., Balota, D. A., Schwarz, J. P., Raichle, M. E., & Petersen, S. E. (1996). A positron emission tomography study of the short-term maintenance of verbal information. *The Journal of Neuroscience, 16,* 808–822.

Fink, G. R., Markowitsch, H. J., Reinkemeier, M., Bruckbauer, T., Kessler, J., & Heiss, W. D. (1996). Cerebral representation of one's own past: Neural networks involved in autobiographical memory. *The Journal of Neuroscience, 16,* 4275–4282.

Friston, K. J., Holmes, A., Worsley, K., Poline, J. -B., Frith, C. D., & Frackowiak, R. S. J. (1995). Statistical parametric maps in functional imaging: A general approach. *Human Brain Mapping, 2,* 189–201.

Gao, J. -H., Parson, L. M., Bower, J. M., Xiong, J., Li, J., & Fox, P. T. (1998). Cerebellum implicated in sensory acquisition and discrimination rather than motor control. *Science, 272,* 545–547.

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.

Haxby, J. V., Horwitz, B., Ungerleider, L. G., Maisog, J. M., Pietrini, P., & Grady, C. L. (1994). The functional organization of human extrastriate cortex: A PET–rCBF study of selective attention to faces and locations. *The Journal of Neuroscience, 14,* 6336–6353.

Kawashima, R., Roland, P. E., & O'Sullivan, B. T. (1995). Functional anatomy of reaching and visuo-motor learning: A positron emission tomography. *Cerebral Cortex, 2,* 111–122.

Kim, S. G., Ugurbil, K., & Strick, P. L. (1994). Activation of a cerebellar output nucleus during cognitive processing. *Science, 265,* 949–951.

Kinomura, S., Larsson, J., Gulyás, B., & Roland, P. E. (1996).

Activation by attention of the human-reticular formation and thalamic intralaminar nuclei. *Science*, *271*, 512–515.

Kosslyn, S. M. (1994a). *Image and brain*. Cambridge: MIT Press.

Kosslyn, S. M., Alpert, N. M., Thompson, W. L., Chabris, C. F., Rauch, S. L., & Anderson, A. K. (1994b). Identifying objects seen from different viewpoints: A PET investigation. *Brain*, *117*, 1055–1071.

Kosslyn, S. M., Alpert, N. M., Thompson, W. L., Maljkovic, V., Weise, S. B., Chabris, C. F., Hamilton, S. E., Rauch, S. L., & Buonanno, F. S. (1993). Visual-mental imagery activates topographically-organized visual cortex: PET investigations. *Journal of Cognitive Neuroscience*, *5*, 263–287.

Kosslyn, S. M., DiGirolamo, G. J., Thompson, W. L., & Alpert, N. M. (1998). Mental rotation of objects versus hands: Neural mechanisms revealed by positron emission tomography. *Psychophysiology*, *3*, 151–161.

Kosslyn, S. M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J. P., Thompson, W. L., Ganis, G., Sukel, K. E., & Alpert, N. M. (1999). The role of area 17 in visual imagery: Convergent evidence from PET and rTMS. *Science*, *284*, 167–170.

Kosslyn, S. M., Thompson, W. L., & Alpert, N. M. (1997). Neural systems shared by visual imagery and visual perception: A positron emission tomography study. *Neuroimage*, *6*, 320–334.

Kosslyn, S. M., Thompson, W. L., Kim, I. J., & Alpert, N. M. (1995). Topographical representations of mental images in primary visual cortex. *Nature*, *378*, 496–498.

Kosslyn, S. M., Thompson, W. L., Kim, I. J., Rauch, S. L. & Alpert, N. M. (1996). Individual differences in cerebral blood flow in area 17 predict the time to evaluate visualized letters. *Journal of Cognitive Neuroscience*, *8*, 78–82.

Le Bihan, D., R. Turner, T. A. Zeffiro, C. A. Cunod, P. Jezzard & V. Bonnerot. (1993). Activation of human primary visual cortex during visual recall: A magnetic resonance imaging study. *Proceedings of the National Academy of Sciences of the United States of America*, *5*, 11802–11805.

Mellet, E., Tzourio, N., Denis, M., & Mazoyer, B. (1995). A positron emission tomography study of visual, and mental, spatial exploration. *Journal of Cognitive Neuroscience*, *7*, 433–445.

Mellet, E., Tzourio, N., Crivello, F., Joliot, M., Denis, M., & Mazoyer, B. (1996). Functional anatomy of spatial mental imagery generated from verbal instruction. *The Journal of Neuroscience*, *16*, 6504–6512.

Mellet, E., Tzourio, N., Denis, M., & Mazoyer, B. (1998a). Cortical anatomy of mental imagery of concrete nouns based on their dictionary definition. *NeuroReport*, *9*, 803–809.

Mellet, E., Petit, L., Mazoyer, B., Denis, M., & Tzourio, N. (1998b). Reopening the imagery debate: Lessons from functional anatomy. *Neuroimage*, *8*, 129–139.

Menon, R. S., Ogawa, S., Tank, D. W., Ellermann, J. M., Merkele, H., & Ugurbil, K. (1993). Visual-mental imagery by functional brain MRI. In D. Le Bihan, R. Turner, M. Mosley, & J. Hyde (Eds.), *Functional MRI of the brain: A workshop presented by the society of magnetic resonance in medicine and the society for magnetic resonance imaging*. Arlington, VA: Society of Magnetic Resonance in Medicine.

Middleton, F. A. & Strick, P. L. (1994). Anatomical evidence for cerebellar and basal ganglia involvement in higher cognitive funtion. *Science*, *266*, 458–461.

Moscovitch, M., Kapur, S., Köhler, S., & Houle, S. (1995). Distinct neural correlates of visual long-term memory for spatial location and object identity: A positron emission tomography study in humans. *Proceedings of the National Academy of Sciences of the United States of America*, *92*, 3721–3725.

Paulesu, E., Frith, C. D., & Frackowiak, R. S. J. (1993). The Neural Correlates of the Verbal Component of Working Memory. *Nature*, *362*, 342–345.

Paus, T., Zatorre, R. J., Hofle, N., Carmanos, Z., Gotman, J., Petrides, M., & Evans, A. C. (1997). Time-related changes in neural systems underlying attention and arrousal during the performance of an auditory-vigilance task. *Journal of Cognitive Neuroscience*, *9*, 392–408.

Price, C. J., Moore, C. J., & Friston, K. J. (1997). Subtractions, conjunctions and interactions in experimental design of activations studies. *Human Brain Mapping*, *5*, 264–272.

Rao, S. M., Bobholz, J. A., Hammeke, T. A., Rosen, A. C., Woodley, S. J., Cunningham, J. M., Cox, R. W., Stein, E. A., & Binder, J. R. (1997). Functional MRI for subcortical participation in conceptual reasonning skills. *NeuroReport*, *8*, 1987–1993.

Roland, P. E. & Gulyas, B. (1994). Visual imagery and visual representation. *Trends in Neurosciences*, *17*, 281–286.

Roland, P. E. & Gulyas, B. (1995). Visual memory, visual imagery, and visual recognition of large field patterns by human brain: Functional anatomy by positron emission tomography. *Cerebral Cortex*, *1*, 79–93.

Sabbah, P., Simond, G., Levrier, O., Habib, M., Trabaud, V., Murayama, N., Mazoyer, B. M., Briant, J. F., Raybaud, C., & Salamon, G. (1995). Functional magnetic resonance imaging at 1.5 T during sensorimotor and cognitive task. *European Neurology*, *35*, 131–136.

Schmahmann, J. D. (1996). From movement to thought: Anatomic substrates of the cerebellar contribution to cognitive processing. *Human Brain Mapping*, *4*, 174–198.

Smith, E. E., Jonides, J., & Koeppe, R. A. (1996). Dissociating verbal and spatial working memory using PET. *Cerebral Cortex*, *6*, 11–20.

Smith, E. E., Jonides, J., Koeppe, R. A., Awh, E., Shumacher, E. H., & Minoshima, S. (1995). Spatial vs. object working memory: PET investigations. *Journal of Cognitive Neuroscience*, *7*, 337–356.

Tagaris, G. A., Kim, S. G., Strupp, J. P., Andersen, P., Ugurbil, K., & Georgopoulos, A. P. (1997). Mental rotation studied by functional magnetic resonance imaging at high field (4 Tesla): Performance and cortical activation. *Journal of Cognitive Neuroscience*, *9*, 419–432.

Thompson, W. L. & Kosslyn, S. M. (1998). Neural systems activated during visual mental imagery: A review and meta-analyses. In A. W. Toga & J. C. Mazziotta, (Eds.), *Brain mapping: The applications*. New York: Academic Press.

Tulving, E., Kapur, S., Craik, F. I. M., Moscovitch, M., & Houle, S. (1994). Hemispheric encoding/retrieval asymmetry in episodic memory: Positron emission tomography findings. *Proceedings of the National Academy of Sciences of the United States of America*, *91*, 2016–2020.

Woods, R. P., Grafton, S. T., Holmes, C. J., Cherry, S. R., & Mazziotta, J. C. (1997). Automated image registration: I. General methods and intrasubject validation. *Journal of Computer-Assisted Tomography*, *22*, 139–152.