

Neural Systems Shared by Visual Imagery and Visual Perception: A Positron Emission Tomography Study

Stephen M. Kosslyn,^{*,†,1} William L. Thompson,^{*} and Nathaniel M. Alpert[‡]

^{*}Department of Psychology, Harvard University, Cambridge, Massachusetts 02138; and [†]Department of Neurology and [‡]Department of Radiology, Massachusetts General Hospital, Boston, Massachusetts 02114

Received April 9, 1997

Subjects participated in perceptual and imagery tasks while their brains were scanned using positron emission tomography. In the perceptual conditions, subjects judged whether names were appropriate for pictures. In one condition, the objects were pictured from canonical perspectives and could be recognized at first glance; in the other, the objects were pictured from noncanonical perspectives and were not immediately recognizable. In this second condition, we assume that top-down processing is used to evaluate the names. In the imagery conditions, subjects saw a grid with a single X mark; a lowercase letter was presented before the grid. In the baseline condition, they simply responded when they saw the stimulus, whereas in the imagery condition they visualized the corresponding block letter in the grid and decided whether it would have covered the X if it were physically present. Fourteen areas were activated in common by both tasks, only 1 of which may not be involved in visual processing (the precentral gyrus); in addition, 2 were activated in perception but not imagery, and 5 were activated in imagery but not perception. Thus, two-thirds of the activated areas were activated in common. © 1997

Academic Press

It has long been believed that mental imagery shares mechanisms with like-modality perception, and much evidence for this inference has been marshaled over many years (e.g., for reviews, see Farah, 1988; Finke and Shepard, 1986; Kosslyn, 1994). One class of evidence has focused on functional commonalities between imagery and perception. For example, Craver-Lemley and Reeves (1987) showed that visual imagery reduces visual perceptual acuity, and this effect depends on whether the image is positioned over the to-be-discriminated stimulus. In addition, Finke *et al.* (1988),

Intraub and Hoffman (1992), and Johnson and Raye (1981) demonstrated that subjects may confuse whether they have actually seen a stimulus or merely imagined seeing it. Another class of evidence has focused on commonalities in the brain events that underlie imagery and perception. For example, Goldenberg *et al.* (1989) used single photon emission computed tomography to compare brain activity when subjects evaluated statements with and without imagery and found increased blood flow in posterior visual areas of the brain during imagery. Convergent results were reported by Farah *et al.* (1988), who used evoked potentials to document activation in the posterior scalp during imagery. Roland *et al.* (1987) also found activation in some visual areas during visual imagery [but not the posterior areas documented by Goldenberg *et al.* (1989), Kosslyn *et al.* (1993), and others; for reviews, see Roland & Gulyas (1994) and accompanying commentaries, as well as Kosslyn *et al.* (1995a)]. Moreover, numerous researchers have reported similar deficits in imagery and perception following brain damage. For example, Bisiach and Luzzatti (1978) found that patients who neglect (ignore) half of space during perception may also ignore the corresponding half of “representational space” during imagery, and Shuttleworth *et al.* (1982) report evidence that brain-damaged patients who cannot recognize faces also cannot visualize faces.

However, results have recently been reported that complicate the picture. In particular, both Behrmann *et al.* (1992) and Jankowiak *et al.* (1992) describe brain-damaged patients who have preserved imagery in the face of impaired perception. Such findings clearly demonstrate that only some of the processes used in visual perception are also used in visual imagery. This is not surprising, given that imagery relies on previously organized and stored information, whereas perception requires one to perform all aspects of figure-ground segregation, recognition, and identification. We would not expect imagery to share “low-level” processes that are involved in organizing sensory input. In contrast, we would expect imagery and perception to share most “high-level” processes, which involve the use of stored information.

¹ To whom reprint requests should be addressed at 832 William James Hall, 33 Kirkland Street, Cambridge, MA 02138.

In this article we use positron emission tomography (PET) to study which brain areas are drawn upon by visual mental imagery and high-level visual perception and which areas are drawn upon by one function but not the other. We are particularly interested in whether imagery and perception draw on more common areas for some visual functions than for others; we use the theory developed by Kosslyn (1994) to help us organize the data in accordance with a set of distinct functions, as is discussed shortly.

We compare imagery and perception tasks that were designed to tap specific components of high-level visual processing. To demonstrate the force of our analysis, we compare an imagery task to a perceptual task that superficially appears very different from the imagery one. Nevertheless, our theory leads us to expect the two tasks to rely on very similar processing. Specifically, the perceptual task requires subjects to decide whether words are appropriate names for pictures. In the baseline condition, the pictures are shown from canonical (i.e., typical) points of view, and hence should be recognized and identified easily; in the comparison condition, the pictures are shown from noncanonical (i.e., unusual) points of view, and hence we expected them not to be recognized immediately, but rather to require additional processing. In this case, we expected the initial input to be used as the basis of a hypothesis, which is tested using top-down processing. Such processing requires accessing stored information about the

distinctive characteristics of the candidate object, shifting attention to the location where such a characteristic should be found, and encoding additional information; if the sought characteristic is encoded, this is additional evidence that the candidate object is in fact present.

Reasoning that such additional processing takes place with noncanonical views, compared to canonical views, Kosslyn *et al.* (1994) predicted—and found—additional activation in a system of brain areas when subjects identified objects portrayed from noncanonical perspectives (these predictions were based on results from prior research, primarily with nonhuman primates; for reviews, see Felleman & Van Essen, 1991; Kosslyn, 1994; Maunsell & Newsome, 1987; Ungerleider & Mishkin, 1982; Van Essen & Maunsell, 1983). According to our theory, the following processes are engaged to a greater extent when one must identify objects seen from noncanonical points of view than when one must identify objects seen from a canonical perspective.

(1) First, if additional information is encoded when one views objects seen from noncanonical perspectives, we expected activation in a subsystem we call the “visual buffer,” which is implemented in the medial occipital lobe (and includes areas 17 and 18); this structure is involved in figure–ground segregation and should be used when additional parts or characteristics are encoded. This structure, in relation to those noted below, is illustrated in Fig. 1.

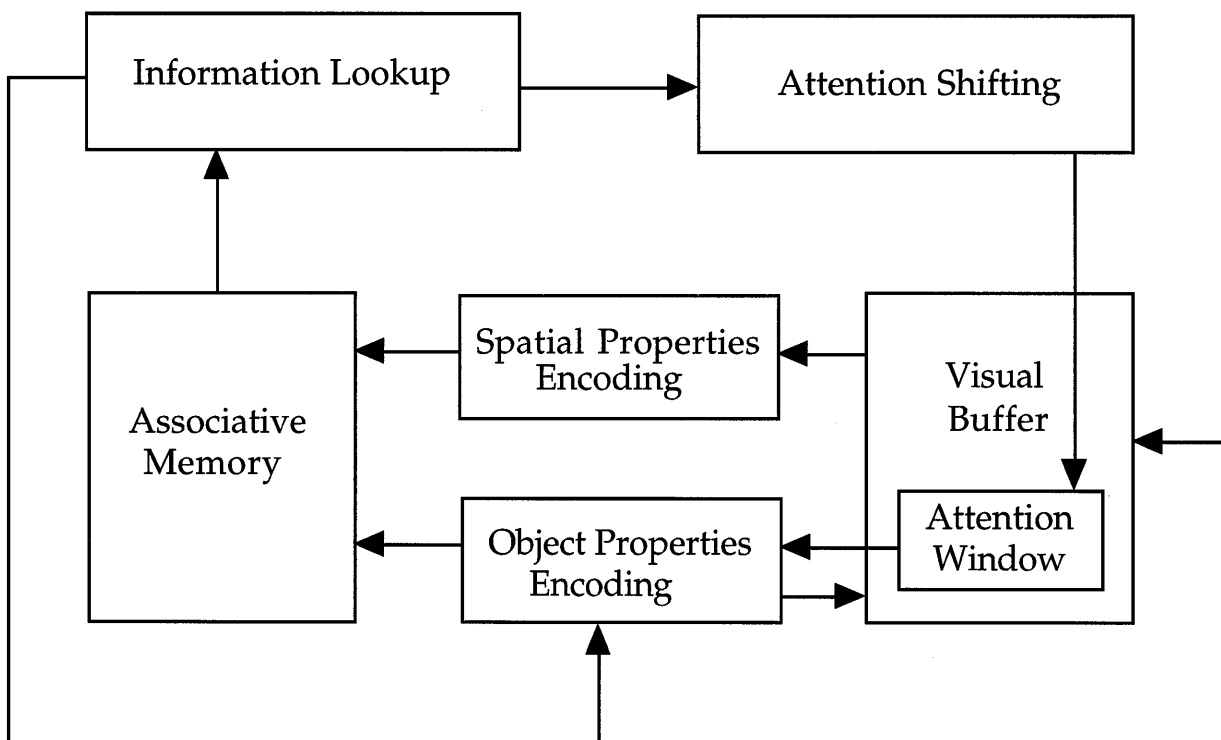


FIG. 1. Six subsystems that are hypothesized to be used in visual object identification and visual mental imagery. See text for explanation.

(2) We also expected additional activation in a subsystem that encodes object properties (shape, color, etc.) and matches them to stored visual memories; this subsystem is implemented in the middle and inferior temporal gyri of humans (e.g., see Bly & Kosslyn, 1997; Gross *et al.*, 1984; Kosslyn *et al.*, 1994; Levine, 1982; Tanaka *et al.*, 1991; Sergent *et al.*, 1992a,b; Haxby *et al.*, 1991; Ungerleider & Haxby, 1994; Ungerleider & Mishkin, 1982).

(3) At the same time that object properties are encoded, spatial properties (such as size and location) are registered; thus, we also expected activation in the inferior posterior parietal lobes (area 7), which encode spatial properties (e.g., Andersen, 1987; Andersen *et al.*, 1985; Corbetta *et al.*, 1993; Ungerleider & Haxby, 1994; Ungerleider & Mishkin, 1982).

(4) The object- and spatial-properties-encoding processing streams must come together at some point; object properties and spatial properties both provide cues as to the identity of an object (moreover, people can report from memory where specific objects are located, which itself is good evidence that the two streams converge). Indeed, Felleman & Van Essen (1991) review evidence that there are complex interconnections between the streams. We posit an “associative memory” subsystem that integrates information about the two kinds of properties and matches such information to stored multimodal and amodal representations; we hypothesize that this subsystem relies on cortex in area 19/angular gyrus (for converging evidence, see Kosslyn *et al.*, 1995b).

(5) If the input does not match a stored representation well, the best-matching representation may be treated as a hypothesis. In such situations, additional information is sought to evaluate the hypothesis. We posit an “information lookup” subsystem that actively accesses additional information in memory, which allows one to search for specific distinctive characteristics (such as a salient part or specific surface marking); we hypothesize that this subsystem is implemented in the dorsolateral prefrontal cortex (for evidence, see Kosslyn *et al.*, 1995b; cf. Goldman-Rakic, 1987).

(6) Finally, an “attention shifting” subsystem actually shifts attention to the location of a possibly distinctive characteristic, which allows one to encode it; this subsystem apparently is implemented in the frontal eye fields, superior colliculus, and superior posterior parietal lobes, as well as other areas (see Corbetta *et al.*, 1993; LaBerge & Buchsbaum, 1990; Mesulam, 1981; Posner & Petersen, 1990). According to the theory, at the same time that one is shifting attention to the location of an expected characteristic, one is also priming the representation of that part or property in the object-properties encoding subsystem; such priming allows one to encode the sought part or property more easily (for evidence of such priming in perception,

see Kosslyn, 1994, pp. 287–289; McDermott & Roediger, 1994). Feedback connections from areas in the inferior temporal lobe to areas 17 and 18 have been identified (e.g., Douglas & Rockland, 1992; Rockland *et al.*, 1992).

We predict that the brain areas that implement each of the foregoing functions also will be activated when visual mental images are generated and used. Specifically, we examined activation in an imagery task used by Kosslyn *et al.* (1993, Experiment 2). On the surface, this task appears nothing like the perceptual object identification task; it does not involve objects or evaluating names. Rather, in the baseline condition subjects see a lowercase letter followed by a 4×5 grid, which contains an X mark in a single cell; they simply respond when they see the grid. In the imagery condition, the subjects receive the identical stimuli used in the baseline condition, but now use the lowercase letter as a cue to visualize the corresponding uppercase letter in the grid, and then decide whether the X mark would have fallen on the uppercase letter if it were actually in the grid. Kosslyn *et al.* (1988, 1993) used this task to study the process of generating images. We expect the six functions described above, which underlie top-down processing during object identification, also to be used to generate a visual mental image. In this case, the information lookup subsystem accesses information about the shape of a letter (including a description of its parts and their locations) stored in associative memory; this information is used by the attention-shifting subsystem to shift attention to the locations where individual segments should be—just as one would look up the location of a distinctive part and shift attention to its position during perceptual top-down hypothesis testing. Only now, one activates the image at the location; we hypothesize that during imagery, the process that primes the representation of an expected object or part during perception primes that representation so strongly that the shape itself is reconstructed in the visual buffer (for a detailed discussion of the possible underlying anatomy, physiology, and computational mechanism, see Kosslyn, 1994). The long-term visual memories apparently are not stored in a spatial structure (e.g., see Fujita *et al.*, 1992), and imagery presumably is used to make explicit the implicit topographic properties of an object. Once the image is present in the visual buffer, we predict that it is thereafter processed the same way that perceptual input is processed: the object and spatial properties are reencoded, and one can identify patterns and properties that were only implicit in a stored visual memory.

In this article, then, we seek to discover whether imagery and perception do in fact draw on many of the same visual areas, and whether there are greater disparities for some of the functions outlined in Fig. 1 than for others. On visual inspection, it appears that

Kosslyn *et al.* (1993, Experiment 2) did find a pattern of activation in their imagery task similar to that observed by Kosslyn *et al.* (1994) in their object identification task; however, it is not clear how to determine exactly how similar the patterns actually were. It is possible that any differences reflect normal between-subject variability, small sample sizes, or statistical error. In the present study we compare the two kinds of processing within the same subjects and compare the patterns of activation directly.

METHOD

Subjects

Six males volunteered to participate as paid subjects. Their mean age was 22 years 2 months, ranging from 19 years 9 months to 27 years 7 months. All but one were attending college or medical school in the Boston area at the time of testing. All subjects were right-handed and all reported having good eyesight. All reported that they were not taking medication or experiencing serious health problems. The subjects were unaware of the specific purposes and predictions of the study at time of testing.

Materials

Object Identification Conditions

Four pictures of each of 27 common objects were drawn. In two of the four versions, the object was portrayed from a canonical point of view, whereas in the other two the object appeared from a noncanonical viewpoint. Examples of the stimuli are illustrated at the top of Fig. 2. The line drawings were scanned in black and white using a Microtek Scanmaker 600ZS scanner to produce files stored in MacPaint format for the Macintosh. After being scanned, the pictures were scaled to 6.35 cm on their longest axis. They thus subtended approximately 7° of visual angle from the subject's vantage point of about 50 cm.

Each drawing was paired with an "entry-level" name (Jolicoeur *et al.*, 1984; see also Kosslyn & Chabris, 1990), which was recorded using Farallon Computing MacRecorder sound digitizer (sampling at 11 kHz, controlled by the SoundEdit program). Two versions of each picture were paired with a word that correctly named it, and two were paired with words that named different objects that had shapes similar to the picture of the target object (which were to be used as distractors). We determined each picture's entry-level name by testing an additional 20 Harvard University undergraduates; these subjects were shown the canonical versions of the pictures and asked to assign a name to each. We accepted the most frequent name assigned to each object as that object's entry-level name provided

that at least 80% of the subjects produced that name for the object.

Additional stimuli were created for the practice section, which included eight trials. The practice trials featured four objects, each presented twice, once paired with a word that correctly named the object and once paired with an inappropriate name. Separate sets of practice objects were created for the canonical and the noncanonical conditions. The 27-object set was divided into subsets of nine objects each. In the original experiment of Kosslyn *et al.* (1994), each of the three subsets was paired with one of the three conditions (baseline, canonical, or noncanonical) for a given counterbalancing group; although we tested the present subjects in the baseline condition (which consisted of viewing nonsense patterns while hearing a name), we do not report those results here. No subject ever saw the same objects (or object derivatives in the case of the baseline task) in more than one condition. Between-subject counterbalancing ensured that each object and word appeared equally often in each of the conditions. In the canonical and noncanonical picture evaluation conditions, the subjects saw a given object four times. Each of the two versions of an object appeared twice in the full block of 36 trials, once paired with the word that correctly named the object and once paired with a distractor name.

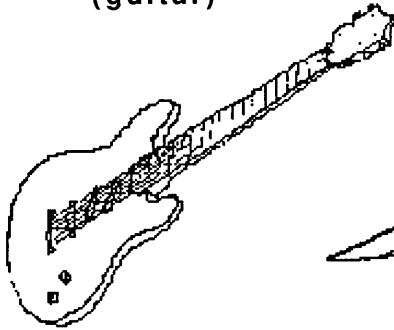
Imagery Conditions

The materials used by Kosslyn *et al.* (1993, Experiment 2) were also used here. Briefly, 16 uppercase letters (A, B, C, E, F, G, H, J, L, O, P, R, S, T, U, Y) and four numbers (1, 3, 4, 7) were used as stimuli. As illustrated at the bottom of Fig. 2, these stimuli appeared in 4 × 5 grids subtending 2.6 (horizontal) by 3.4 (vertical) degrees of visual angle from the subject's viewpoint (at a distance of about 50 cm). On each trial, the stimuli consisted of a small asterisk positioned at the center of the screen, a script lowercase letter also centered on the screen, and a 4 × 5 grid that contained a single X mark in one of the cells. This X probe consisted of two dashed lines 1 pixel wide that connected the opposite corners of the cell; the space between the pixels of the dashed lines was equivalent to 2 pixels in length. Each of the two diagonal segments was approximately 3.5 mm. The stimuli were constructed so that on half the trials the correct response was "yes" and on half it was "no." Furthermore, on half the trials of each type the X fell on (for a yes trial) or next to (for a no trial) a segment that was typically drawn early in the sequence of strokes ("early" trials), and on half the trials the X fell on or near a segment that was typically drawn late in the sequence of strokes ("late" trials; see Kosslyn *et al.*, 1988, for a description of how segment order was determined). The identical stimuli were used in the baseline task and in the

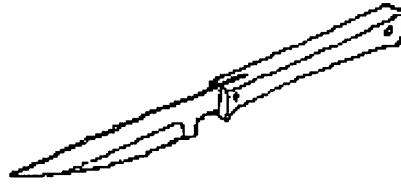
Perception

Canonical Views

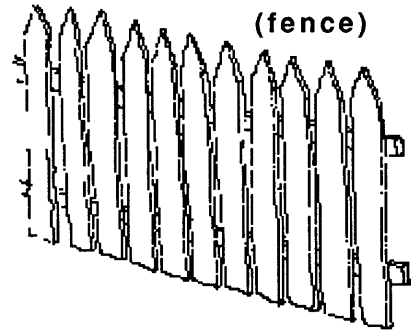
(guitar)



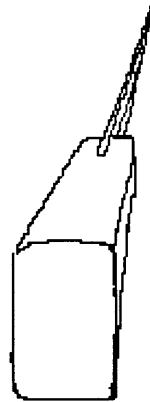
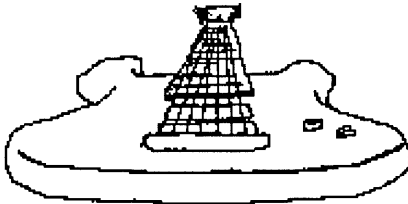
(knife)



(fence)

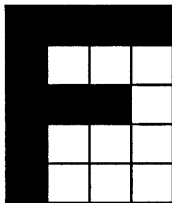


Noncanonical Views



Imagery

Study stimulus



(no study in baseline condition)

Cue



Test stimulus

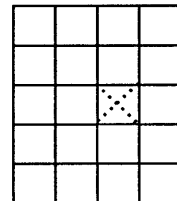


FIG. 2. Examples of stimuli used in the two conditions: (Top) Canonical and noncanonical views of objects, and (bottom) imagery (the same stimuli were used in the test and baseline conditions).

imagery task; only the instructions given to the subject differed.

Task Procedure

The tasks were administered by a Macintosh Plus computer with a Polaroid CP-50 screen filter (to reduce glare). The MacLab 1.91 program (Costin, 1988) displayed the stimuli and recorded subjects' responses and response times. The computer rested on a gantry approximately 50 cm from the subject's eyes; the screen was tilted down so that it faced the subject. A special response apparatus was constructed so that subjects could respond by pressing either of their feet (the feet are controlled by cortex at the very top of the motor strip, which is far removed from areas we predicted to be activated by the experimental tasks).

After being positioned in the scanner, subjects read the instructions for the baseline imagery task. In all conditions, after reading the instructions the subjects were asked to paraphrase them from memory; before continuing, the investigator discussed any misconceptions until the subject clearly understood the task. In all conditions, the subjects were asked to respond as quickly as possible while remaining as accurate as possible. The baseline imagery task was always performed before the subjects studied the uppercase block letters within the grids; this was to ensure that subjects would not visualize the letters during the baseline task. Each trial in the baseline task followed the same sequence. First, a small centered asterisk appeared on the screen for 500 ms, followed by a blank screen for 100 ms, then a lowercase script letter was displayed for 300 ms, followed by another blank screen for 100 ms. Finally, the 4×5 grid containing the X probe appeared for 200 ms. Subjects were asked to respond as quickly as possible after the appearance of the X-marked grid, alternating between left and right foot pedals from trial to trial. A new stimulus was presented 100 ms after each response. The subjects were further instructed to try to empty their minds of all other thoughts and were asked to view the stimuli "without trying to make sense of them or make any connections between them." The subjects were interviewed after the experiment, and all reported not having thought about or made connections between the stimuli, and none reported having visualized any pattern in the grid. Subjects were given 16 practice trials prior to the actual test trials; PET scanning was performed only during the test trials, as described below.

After completing the baseline trials in the imagery condition, the subjects were shown a sheet on which were printed four block digits (1, 3, 4, 7) within 4×5 grids (which were the same size as those that were presented on the computer). The sheet was held up to the screen so that the subjects could study the digits at the same distance that they would later be asked to

visualize them. Under each number was a script version of the digit, and the subjects were asked to become familiar with this digit as well.

Practice trials for the imagery task began as soon as subjects reported knowing which grid cells were filled by each digit and were familiar with each of the corresponding script cues. PET scanning was not performed during these trials. The subjects were told that the script cue would indicate which digit they were to visualize in the grid, and that once they had done so they were to decide whether the X would fall on the digit if it were also in the grid. They were instructed to press the right foot pedal if it would fall on the digit and the left pedal if it would not. They were asked to respond as quickly as possible (not as soon as they saw the X, but as soon as they had reached a decision). A new stimulus was presented 100 ms after each response. Sixteen practice trials were administered; each of the four digits appeared four times, twice in a yes trial and twice in a no trial. The subjects pressed the right pedal for yes and the left for no.

After completing the practice section, the subjects were shown a sheet that illustrated 16 uppercase block letters within grids and were asked to learn the appearance of the letters as well as the corresponding lowercase script cues. Once subjects reported having done so, the main points of the instructions were briefly reviewed and the subjects were told to wait for the investigator's signal to begin. The actual test trials were identical to the practice trials in format, differing only in the type of stimuli visualized. The test trials began 30 s before scanning and continued for a total 2 min. The stimuli were ordered randomly, except that the same type of trial or response (yes/no) could not occur more than three times in succession and the same letter could not occur twice within a sequence of 3 trials. The identical stimulus sequence was used in the imagery condition and the baseline condition. Both the imagery and the baseline tasks included a total of 128 trials, of which subjects completed on average 65 in the baseline condition, with a range of 47–74, and 59 in the imagery condition, with a range of 49–64; there was no significant difference between the number of trials completed by subjects in the baseline versus the imagery condition, $t = 1.04$, $P > 0.3$.

After completing the imagery conditions, the subjects participated in a baseline task for the object-identification conditions (in which they heard familiar words followed immediately by a random meaningless shape and simply responded as quickly as possible as soon as they saw a shape on the screen, alternating right and left foot pedals from trial to trial); these data are not relevant to the present issue, and so are not discussed further here (no stimuli were used in the baseline trials that later appeared in the object-identification trials for a given subject). After performing the baseline task,

each subject read the instructions for the object identification task. When the subject reported that he understood the instructions, eight practice trials were administered. The practice trials contained pictures of four objects, each object appearing twice, once in a yes trial and once in a no trial.

Following this, the actual test trials began 15 s before PET scanning began and continued for 2 min. Because the stimuli were complex, the computer paused to load new trials into memory about 75 s from the beginning of the experiment; thus, beginning the test trials 30 s before scanning, as we did in the imagery tasks, rather than 15 s, would have placed this delay within the actual scanning period. In all trials, a blank screen marked the beginning of each trial and was displayed for 200 ms; a word that named a common object was then read aloud by the computer. At the end of the word, a picture of a common object appeared on the screen. The subject was instructed to press the right pedal if the word was a possible name for the picture and the left pedal if it was not. A new trial was presented 200 ms after the subject responded. The canonical and noncanonical pictures conditions followed the same procedure, except that the pictures depicted objects from a different viewpoint (either canonical or noncanonical, depending on the subject's counterbalancing group). Half the subjects received the canonical pictures first, and half received the noncanonical pictures first; over subjects, each picture-word pair appeared equally often in each condition in the object-identification trials.

All subjects completed all 36 trials of each condition. In the event that scan time still remained, the stimuli from that condition were presented again from the beginning, in the same order as in the first cycle. All subjects reached this second cycle in both the canonical and the noncanonical conditions. The number of trials completed in the canonical condition ranged from 57 to 75 and from 55 to 72 in the noncanonical condition. On average, 69 trials were completed in the canonical condition and 63 trials were completed in the noncanonical condition ($t < 1$).

PET Methods and Procedure

The PET scan procedure has been described in previous publications (e.g., Kosslyn *et al.*, 1994). A brief summary of that material is presented here. A custom-molded thermoplastic headholder stabilized the subjects' heads, and PET slices were parallel to the canthomeatal line. The PET scanner was a GE PC4096+ 15-slice whole body tomograph, with an axial field of 97.5 mm and intrinsic resolution of 6 mm (FWHM) in both the transverse and the axial directions. Transmission scans were performed on each subject to measure the effects of photon attenuation. Subjects were scanned for 1 min while they continuously inhaled tracer quan-

tities of [^{15}O]carbon dioxide gas mixed with room air and performed a cognitive task. Images were reconstructed using a standard convolution back-projection algorithm. After reconstruction, the scans from each subject were treated as follows: A correction was computed to account for head movement (rigid body translation and rotation) using a least-squares fitting technique (Alpert *et al.*, 1996). The mean over all conditions was calculated and used to determine the transformation to the standard coordinate system of Talairach and Tournoux (1988). This transformation was computed by deforming the 10-mm parasagittal brain-surface contour to match the contour of a reference brain (Alpert *et al.*, 1993). Following spatial normalization, scans were filtered with a two-dimensional Gaussian filter, full width at half maximum set to 20 mm.

Statistical Analysis

Statistical analysis followed the theory of statistical parametric mapping (Friston *et al.*, 1991, 1995; Worsley *et al.*, 1992). Data were analyzed with the SPM95 software package (from the Wellcome Department of Cognitive Neurology, London, UK). The PET data at each voxel were normalized by the global mean and fit to a linear statistical model by the method of least squares. The analysis of variance considered cognitive state (i.e., scan condition) as the main effect and subjects as a block effect. Hypothesis testing was performed using the method of planned contrasts at each voxel. This method fits a linear statistical model, voxel-by-voxel, to the data; hypotheses are tested as contrasts in which linear compounds of the model parameters (e.g., differences) are evaluated using t statistics. Data from all conditions were used to compute the contrast error term.

In addition to the four conditions under consideration, each subject also performed two additional baseline tasks, during which they were asked to simply alternate foot pedals when they perceived either a random pattern line drawing or a grid-and-X stimulus that contained a letter; thus, data from six conditions were used to estimate error variance. A z greater than 3.09 was considered statistically significant. This threshold was chosen as a compromise between the higher thresholds provided by the theory of Gaussian fields, which assumes no a priori knowledge regarding the anatomic localization of activations, and simple statistical theories that do not consider the spatial correlations inherent in PET and other neuroimaging techniques.

In order to test for common foci of activation in the imagery condition and the condition in which objects were presented from a noncanonical perspective, we examined the contrast (imagery + noncanonical) - (imagery baseline + canonical). Only activation that is common in imagery and top-down perception at a given point will be significant in this contrast. We also

performed contrasts that allowed us to determine which areas were more activated during imagery than during perception, and vice versa; these contrasts examined the differences between the differences in the respective test and baseline conditions.

RESULTS

Behavioral Results

Cognitive research poses an inherent problem because the phenomena of interest occur in the private recesses of one's mind/brain. Hence, it often is difficult to ensure that subjects were in fact engaged in the kind of processing one wants to study. One solution to this problem is to design tasks that appear to require a particular type of processing and to collect behavioral measures that will reveal whether such processing did in fact occur. We designed our tasks so that there are distinctive behavioral effects, which we would not expect to be present if the subjects did not perform particular types of processing. Specifically, we recorded and analyzed the following types of behaviors, which we took to be "signatures" of the requisite processing; if these measures were significant, we had good reason to infer that the PET results did in fact inform us about the underlying nature of the kind of processing of interest. Separate analyses of variance were performed on data from each task to evaluate error rates and response times, using subject as the random effect.

Imagery Task

Previous research has shown that when subjects generate images of block letters a segment at a time, they require more time or make more errors when evaluating X probe marks that fall on segments farther along the sequence in which the segments typically are drawn; this result did not occur when subjects viewed gray uppercase letters in grids and decided whether X marks fell on them, or when they had generated an image prior to seeing the X mark (e.g., see Kosslyn *et al.*, 1988). Thus, the existence of an effect of early versus late segments on response time or error rates is a good behavioral indicator that subjects were in fact generating images in this task. As expected, when "yes" trials are considered (and we can be certain that subjects had to form the image up to the location of the X probe; see Kosslyn *et al.*, 1988), subjects made fewer errors on early segments than on late ones (14.0% vs 19.9%), $F(1, 5) = 7.42$, $P < 0.05$; there was no speed-accuracy tradeoff, as witnessed by comparable response times in the two types of trials (867 ms vs 829 ms), $F(1, 5) = 1.15$, $P > 0.25$.

Object Identification

If subjects did in fact require an additional processing cycle to encode distinctive characteristics with the

noncanonical stimuli, then they should have required more time to perform these trials (note, however, that this additional time is compensated for by fewer trials, and thus the total "processing time" is comparable in the noncanonical and canonical blocks of trials; see Kosslyn *et al.*, 1994). As expected, subjects made more errors (13.0% vs 4.6%), $F(1, 5) = 5.68$, $P = 0.06$, and required more time (671 ms vs 936 ms), $F(1, 5) = 12.25$, $P < 0.02$, when they evaluated objects seen from a noncanonical viewpoint.

PET Results

The results of performing the contrast (imagery + noncanonical) – (imagery baseline + canonical) are presented in Fig. 3 and Table 1. The jointly activated areas (with $z > 3.09$) can be organized in terms of Fig. 1: First, we found activation in left area 18, which is clearly one of the topographically organized areas that constitute the visual buffer. Moreover, earlier research with the grids task revealed activation in the left hemisphere; however, the coordinates of this region of activation are not close to that found when only the imagery task was examined previously (Kosslyn *et al.*, 1993). We did not find common activation in area 17 in this analysis. We also found evidence of joint activation in regions that are part of the object-properties-encoding system, namely an area in the left occipital-temporal junction, including part of the middle temporal gyrus, and the right lingual gyrus. Next, note that this analysis also revealed joint activation in the inferior parietal lobe bilaterally, which is part of the spatial properties encoding system.

This analysis also revealed a large amount of activation in areas that may implement associative memory. In the left hemisphere, we found a locus in area 19 and another in the angular gyrus, but in the right hemisphere two distinct points in area 19 (one merging into the angular gyrus) were detected. Several of these areas are very close to those found by Kosslyn *et al.* (1995a), who argued that such activation reflected use of associative memory. In addition, replicating earlier results with both tasks, we again found activation in the left dorsolateral prefrontal area. The localization to the left hemisphere is as expected if this area is specifically involved in looking up "categorical" spatial relations information (cf. Kosslyn *et al.*, 1995b). Moreover, we also found activation in areas that are involved in shifting attention, the right superior parietal lobe and the precuneus bilaterally (cf. Corbetta *et al.*, 1993).

The foregoing findings all make sense within the structure of the theory we summarized earlier. However, we also found unexpected activation. Specifically, the left precentral gyrus was jointly activated in both tasks; this area has been implicated in motor-related imagery by Decety *et al.* (1994), and it is possible that

(IMAGERY + NONCANONICAL) - (IMAGEBASE + CANONICAL)

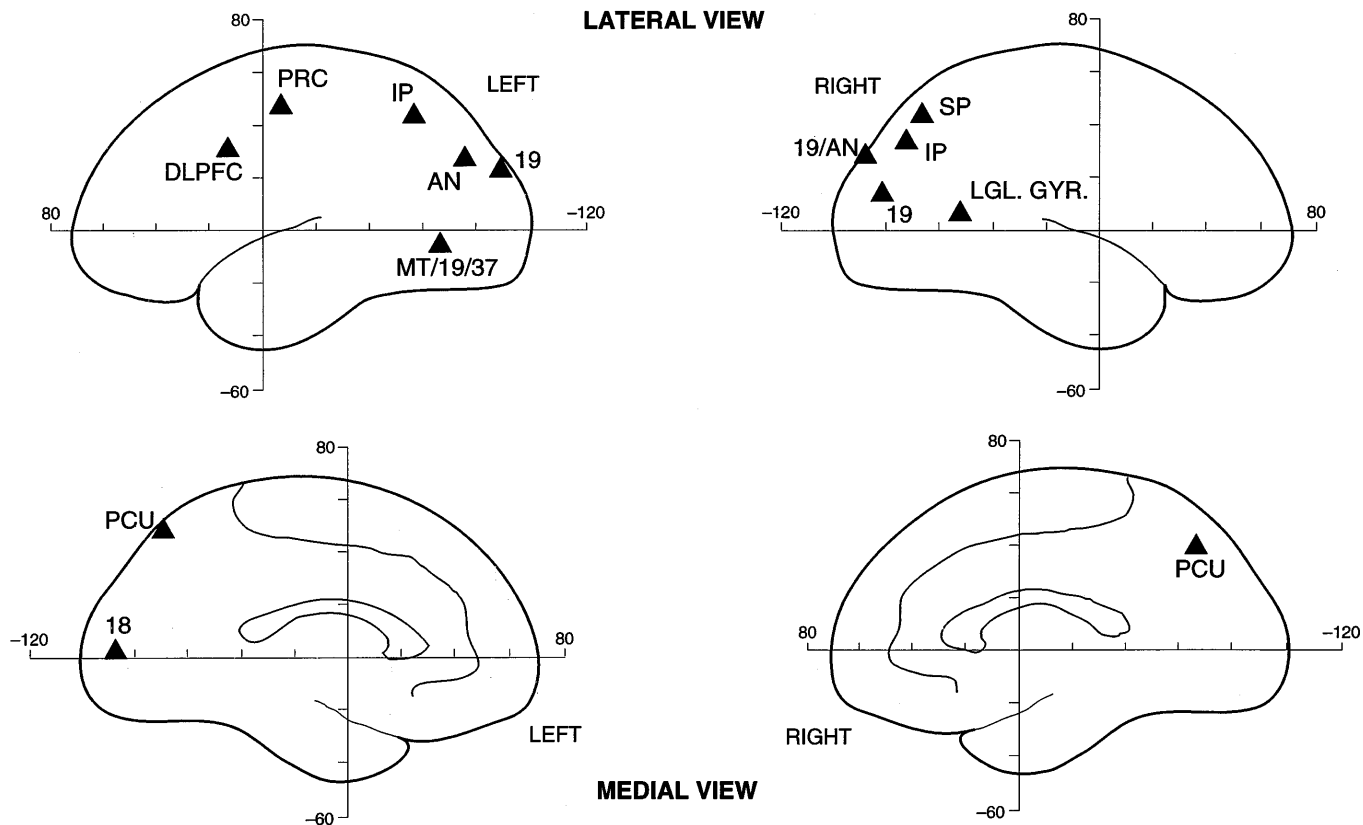


FIG. 3. Results from the contrast analysis in which the noncanonical pictures task and the imagery task were compared to the appropriate baselines. These panels illustrate areas that were activated in common by the two tasks. (Top) Lateral views, (bottom) medial views, (left) left hemisphere, (right) right hemisphere. Each tick mark indicates 20 mm. Axes are centered on the anterior commissure.

there is a motor contribution to both sorts of processing we observed. In their motor imagery condition, Decety *et al.* (1994) asked subjects to imagine themselves grasping objects that appeared on the computer screen with their right hand. In the baseline condition, subjects were simply asked to visually inspect the objects presented to them. For example, subjects in our imagery task may visualize the sequence of strokes by “mentally drawing” them, which is consistent with the fact that segments typically are visualized in the same order that they are drawn (Kosslyn *et al.*, 1988). We must note, however, that it is also possible that this finding reflects activation of the frontal eye fields; there is debate about exactly where this functionally defined region lies in the human brain. Indeed, Paus (1996) reviews the literature from blood flow and lesion studies and concludes that the location of the frontal eye fields is in fact most likely in the vicinity of the precentral sulcus or in the caudalmost part of superior frontal sulcus. This is contrary to the commonly held view that frontal eye fields are located within Brodmann’s area 8.

We next asked the obverse question: which areas were more activated during top-down perception than during imagery and vice versa? Thus, we first subtracted blood flow in the canonical pictures task from that in the noncanonical pictures task, and also subtracted blood flow in the imagery baseline from that in the imagery task. We then compared these two difference maps. Figure 4 and Table 2 present the results of contrasting the map reflecting activation during the imagery task from that reflecting activation during top-down perception. We found more activation in top-down perception than in imagery in two areas in the right hemisphere, the middle temporal gyrus and the orbitofrontal cortex; we also found a trend for activation in the right dorsolateral prefrontal region (with $z = 3.07$, at coordinates 16, 58, 28).

Figure 4 and Table 2 also present the results of contrasting the map reflecting activation during top-down perception with that reflecting activation during imagery. As is evident, we found more activation in two areas in the left inferior parietal lobe and three areas in the right hemisphere: the superior parietal lobe, area

TABLE 1

Coordinates (in mm, Relative to the Anterior Commissure) and *P* Values for Regions in Which There Was More Activation during Both Test Conditions (Top-down Perception and Imagery) Than Both Baseline Conditions

	<i>x</i>	<i>y</i>	<i>z</i>	<i>z</i> score
Left hemisphere regions				
Area 19	-25	-87	24	4.56
Area 18	-33	-86	4	3.24
Precuneus	-8	-79	48	3.20
Angular gyrus	-35	-78	28	3.86
MT/19/37 (occipitotemporal jct.)	-48	-65	-4	4.39
Inferior parietal	-31	-58	44	3.64
Precentral gyrus (area 4)	-36	-5	48	3.34
Dorsolateral prefrontal (area 9)	-45	16	32	3.33
Right hemisphere regions				
Area 19/angular gyrus	31	-86	28	4.62
Area 19	43	-81	16	3.62
Inferior parietal	38	-73	36	3.96
Precuneus	12	-68	40	3.24
Superior parietal (area 7)	22	-65	44	3.28
Lingual gyrus	18	-54	8	3.69

Note. Regions are presented from posterior to anterior. Seen from the rear of the head, the *x* coordinate is horizontal (with positive values to the right), the *y* coordinate is in depth (with positive values anterior to the anterior commissure), and the *z* coordinate is vertical (with positive values superior to the anterior commissure). Only *z* scores greater than 3.09 are presented.

TABLE 2

Coordinates (in mm, Relative to the Anterior Commissure) and *P* Values for Regions in Which There Was More Activation during the One Test Condition, Relative to the Appropriate Baseline, Than the Other

	<i>x</i>	<i>y</i>	<i>z</i>	<i>z</i> score
(Noncanonical – canonical) – (Imagery – imagery baseline)				
Right hemisphere regions				
Middle temporal	53	-44	8	3.84
Orbitofrontal cortex	1	42	-16	3.38
(Imagery – imagery baseline) – (Noncanonical – canonical)				
Left hemisphere regions				
Inferior parietal	-39	-50	44	3.78
Inferior parietal	-50	-45	36	3.65
Right hemisphere regions				
Angular gyrus/area 19	30	-86	28	3.92
Area 19	20	-73	40	3.56
Superior parietal	12	-67	52	3.46

Note. Regions are presented from posterior to anterior. Seen from the rear of the head, the *x* coordinate is horizontal (with positive values to the right), the *y* coordinate is in depth (with positive values anterior to the anterior commissure), and the *z* coordinate is vertical (with positive values superior to the anterior commissure). Only *z* scores greater than 3.09 are presented.

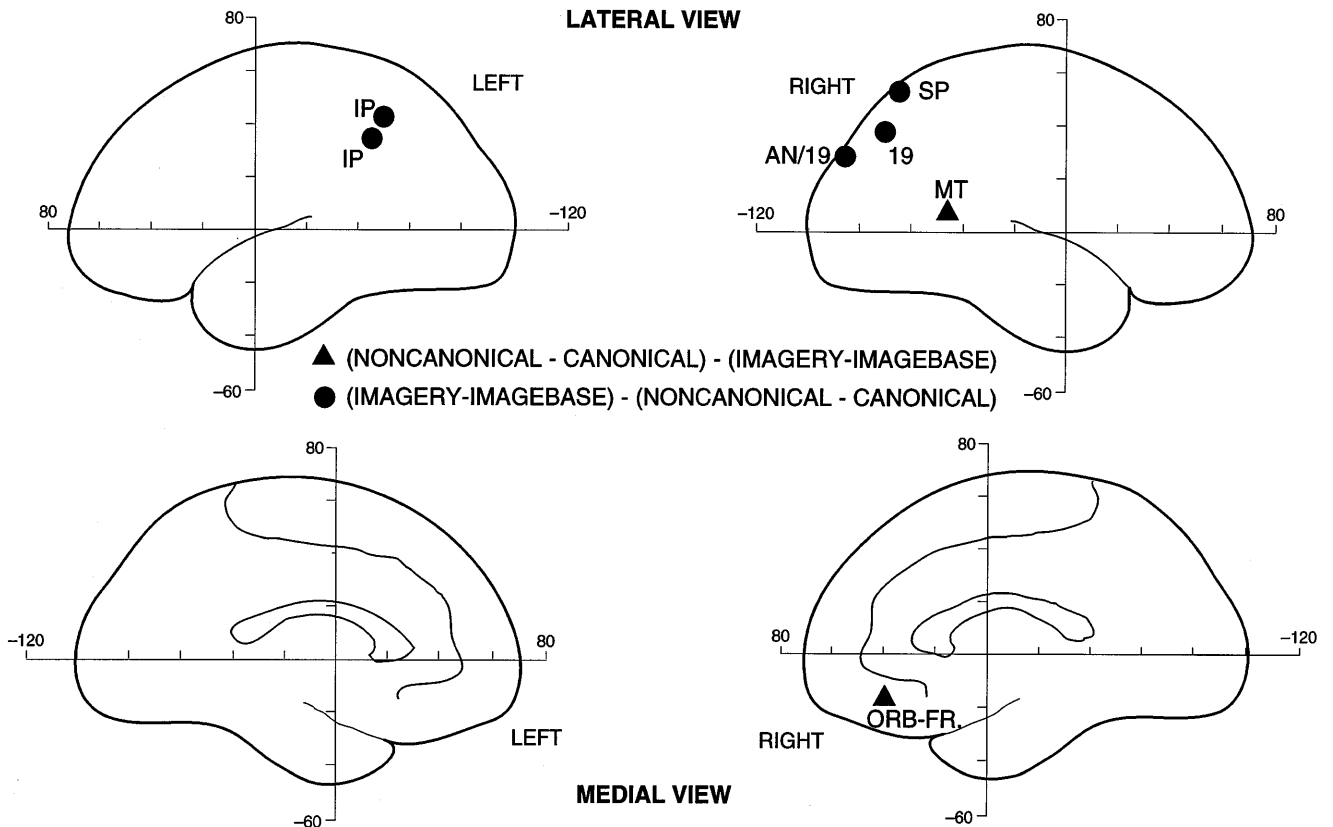


FIG. 4. Results from comparing the noncanonical–canonical pictures difference map with the imagery–baseline difference map. (Top) Lateral views, (bottom) medial views, (left) left hemisphere, (right) right hemisphere. Each tick mark indicates 20 mm. Axes are centered on the anterior commissure.

19, and an area that included parts of area 19 and the angular gyrus.

In short, in this analysis 14 areas were activated jointly by the two tasks, compared to only 2 that were activated in perception and not in imagery and 5 that were activated in imagery but not perception. By this estimate then, 21 areas were activated in total, with two-thirds of them being activated in common.

GENERAL DISCUSSION

In this article we asked which brain areas are drawn upon by visual mental imagery and high-level visual perception and which areas are drawn upon by one function but not the other. We were also interested in whether imagery and perception draw on more common areas for some visual functions than for others, and used the theory developed by Kosslyn (1994) to organize the data in accordance with a set of distinct functions.

As is evident from Figs. 3 and 4, it is clear that the two functions draw on much common neural machinery. Moreover, all but one of the jointly activated areas clearly have visual functions; there can be no question that visual mental imagery involves visual mechanisms. These imagery results are not a consequence of the visual stimuli: the same stimuli appeared in the baseline condition, and hence we were able to remove the contribution of perceiving *per se* from the imagery data. However, the two functions do not draw upon the identical machinery; a reasonable estimate is that about two-thirds of the brain areas used by either function are used in common.

We can use the theory of Fig. 1 to organize the results obtained when we examined which areas were activated more in the perceptual task than the imagery one and vice versa. From this perspective, we found no disparities in activation in areas that implement the visual buffer. However, we found two areas jointly activated that implement the object properties encoding subsystem and one that was activated only during perception (33% disparity), two areas jointly activated that implement the spatial-properties-encoding subsystem and two that were activated only in imagery (50% disparity), four areas jointly activated that implement associative memory and two that were activated only during imagery (33% disparity), one area jointly activated that implements the information lookup subsystem, no area that was activated only during imagery or perception (but a trend for one area to be activated only during perception), and, finally, three areas jointly activated that implement attention shifting and one that was activated only during imagery (25% disparity). These findings suggest that most of the information processing functions are accomplished in slightly different ways in imagery and perception. Moreover,

the spatial-properties-encoding and attention-shifting operations appear particularly likely to operate differently in imagery and perception.

The fact that both imagery and perception drew on processing that was not shared by the other function allows us to understand the double dissociation following brain damage (see Behrmann *et al.*, 1992; Jankowiak *et al.*, 1992): Depending on which of these non-shared areas are damaged, the patient can have difficulties with imagery but not perception or vice versa. However, our finding that about two thirds of the areas are shared by the two functions leads us to expect that imagery and top-down perception should often be disrupted together following brain damage.

One could try to argue that the present results are trivial because the noncanonical picture identification task is actually an imagery task, requiring subjects to mentally rotate the noncanonical object into the canonical perspective. However, there is no evidence that mental rotation is used in this task, and there is evidence that it is not used in this task. First, Jolicoeur (1990) reviewed the cognitive psychology literature on time to name misoriented objects and he explicitly compared these increases in time to those found in classic mental rotation experiments. Jolicoeur notes that the slope of the increase in time to name misoriented pictures is very different from that found in mental rotation experiments, and—unlike standard mental rotation slopes—is often nonmonotonic. Jolicoeur also notes that studies with brain-damaged patients have revealed dissociations between rotation and object identification. The response time differences between the noncanonical and the canonical stimuli in our experiment are in the range of those reported in similar experiments in the past. Thus, the off-line literature suggests that people do not use mental rotation in this task. Second, further buttressing this conclusion, the noncanonical task did not lead to activation in areas that have been demonstrated to be activated during mental rotation (e.g., Cohen *et al.*, 1996). In the Cohen *et al.* study, activation was found in areas 3, 1, 2, and 6, none of which were activated in the perceptual comparison of the present study. Moreover, in the present study, we found activation in the insula and fusiform gyrus, which were not activated in the study of Cohen *et al.* In sum, the rotation tasks involved the motor system to a much greater degree (and perhaps parietal areas as well—there are many more and somewhat stronger foci of activation in rotation), whereas the noncanonical picture identification task drew much more on the ventral system, encoding properties of objects.

However, we must note that one of the areas activated in common, area 8, suggests that eye movements occurred in both tasks. We intentionally presented the pictures in free view because this is usually how

pictures are encountered, and thus we wanted to observe the full range of appropriate processing. In contrast, we presented the imagery stimuli for less time than is required for an eye movement in order to avoid artificially increasing the amount of visual activation caused by the visual cues. Nevertheless, people apparently moved their eyes during imagery, which may function as a "prompt" to help one recall where parts of images belong (see Kosslyn, 1994); in fact, Brandt *et al.* (1989) report that subjects produce eye movements when they visualize objects that are similar to those produced when they perceive them.

To our knowledge, this is the first study to compare visual mental imagery and visual top-down perception, but Mellet *et al.* (1995) have compared spatial imagery and perception and apparently found very different results from ours. Their task involved visualizing and scanning a map or actually seeing and scanning the map. Mellet *et al.* found common activation in the two conditions only in the right superior occipital gyrus. In their perception condition (compared to a resting baseline) they found bilateral activation in primary visual cortex, the superior occipital gyrus, the inferior occipital gyrus, the cuneus, the fusiform/lingual gyrus, superior parietal cortex, the precuneus, the angular gyrus, and the superior temporal gyrus; they also found left-hemisphere activation of the dorsolateral prefrontal cortex (DLPFC) and the inferior frontal gyrus, right-hemisphere activation of the lateral premotor area, and activation of the anterior cingulate and medial cingulate. In contrast, in their imagery condition (compared to a resting baseline), they found right-hemisphere activation of the superior occipital gyrus, left-hemisphere activation of the parahippocampal gyrus, and activation of the supplementary motor area and vermis. Comparing imagery and perception, they found more activation in perception bilaterally in primary visual cortex, the superior occipital gyrus, the inferior occipital gyrus, the cuneus, the fusiform/lingual gyri, and the superior parietal cortex. In contrast, they found more activation in imagery than perception bilaterally in the superior temporal and precentral gyri, as well as in the left DLPFC, in the inferior frontal cortex, and in the supplementary motor area, anterior cingulate, median cingulate, and cerebellar vermis. Note that Mellet *et al.* found no hint of activation in area 17 or 18 in imagery. However, they used a resting baseline, which might explain why they failed to find such activation: Kosslyn *et al.* (1995a) found that such a baseline could remove evidence of medial occipital activation during imagery. However, in a more recent paper Mellet *et al.* (1996) used a different sort of baseline and still failed to find medial occipital activation during imagery. This task involved forming an image of a multiarmed object based on a set of spatial directions. Similarly, Roland *et al.* (1987; see

also Roland & Friberg, 1985) failed to find medial occipital activation when they asked people to imagine turning left or right when walking along a path.

In order to reconcile these disparate findings, it may be useful to distinguish between three types of imagery: First, the kind of imagery involved in the Mellet *et al.* (1996), Roland *et al.* (1987), and Roland and Friberg (1985) studies requires one to preserve spatial relations, but not shapes, colors, or textures. In the framework of Fig. 1, this sort of imagery does not involve the ventral system or the visual buffer, but does make use of processes implemented in the posterior parietal lobes, the DLPFC, the angular gyrus/19, and various areas involved in attention.

Second, in contrast to this "spatial imagery," which has also been suggested by Mellet *et al.* (1996), another sort does not rely on the parietal lobes but rather requires processes implemented in the inferior temporal lobes. Such "figural imagery" arises when stored representations of shapes and their properties (such as color and texture) are activated but produce only a low-resolution topographic representation (the image itself). These representations occur in posterior inferior temporal cortex, which might include coarsely organized topographically mapped areas (there is some evidence that this is true in the monkey; for an overview, see Felleman & Van Essen, 1991; for more details, see DeYoe *et al.*, 1994). Such imagery does not require the visual buffer, but presumably relies on the DLPFC, the angular gyrus/19, and areas that subserve attention. These notions are consistent with findings reported by Fletcher *et al.* (1995). In their PET study, subjects were asked to encode and recall concrete and abstract words. They did not find any activation of medial occipital cortex when concrete words were recalled, which presumably involved imagery (e.g., see Paivio, 1971). However, they did find activation of the right superior temporal and fusiform gyrus (as well as the left anterior cingulate and the precuneus).

The distinction between spatial and figural imagery is consistent with a double dissociation reported by Levine *et al.* (1985); they describe a patient with parietal damage who could visualize objects but not spatial relations and a patient with temporal lobe damage who could visualize spatial relations but not objects.

Finally, "depictive imagery" relies on high-resolution representations in the medial occipital cortex. Such imagery would be useful if one needs to reorganize shapes, compare shapes in specific positions, or reinterpret shapes. Long-term visual memories apparently are stored in the inferior temporal cortex using an abstract code, not topographically (e.g., see Fujita *et al.*, 1992). Thus, high-resolution topographic images would need to be actively constructed if one must reinterpret implicit shapes, colors, or textures. This is the sort of

imagery we focused on in the present study, which relies on the areas that subserve the functions illustrated in Fig. 1.

We suspect that "pure" forms of the three types of imagery are rare. For example, we note that the temporal/fusiform/lingual gyrus area seems to be activated in at least three of the studies that used spatial tasks (Charlot *et al.*, 1992; Mellet *et al.*, 1995; Roland & Friberg, 1985). Roland *et al.* (1987) also report moderate increases in activation in the superior occipital and the posterior inferior temporal region during a spatial imagery (route-finding) task. However, the strongest activations were found in the prefrontal cortex, frontal eye fields, and posterior parietal regions, as well as portions of the right superior occipital cortex that appear to be within area 19. Charlot *et al.* (1992) also found activation of the "left association plurimodal" area (i.e., left inferior parietal areas 39 and 40) during both imagery and verbal tasks, in the high-imagery group. Roland and Friberg (1985) also find selective activation of the parietal lobes, relative to their "backwards counting by 3" and "jingle" tasks. Mellet *et al.* (1995) apparently examined only the superior parietal lobe and found no reliable activation there, although a closer examination of the results reveals that data from different subjects cancel each other out: some subjects had strong activation in superior parietal cortex, whereas others had small increases in blood flow or actual decreases in flow. (Indeed, some of their subjects had up to a 12% increase in the superior parietal lobe during imagery.)

Kosslyn *et al.* (1993) provided evidence that the version of the grids imagery task used in the present study induces depictive imagery. (Indeed, we looked specifically at area 17 in the present imagery results and found $z = 3.0$ in this region of interest, replicating our earlier finding.) Hence, we had reason to expect a close correspondence with the top-down perception task, and our finding that so many areas were activated in common by this type of imagery and top-down perception is probably not a coincidence. Indeed, if one chose at random another task in the PET literature, one would not find anything like this kind of correspondence. For example, we examined results from a task in which subjects were asked whether X marks were on or off visible block letters in grids (see Experiment 2 of Kosslyn *et al.*, 1993), and found that only 4 of the 15 (27%) activated areas in that experiment were within 14 mm (the width of the smoothing filter used in those analyses) of those listed in Table 1. This is remarkable because on the surface this task would appear very similar to the present imagery task and probably does draw on some of the same high-level processes. However, this task involves a large number of bottom-up processes (such as those involved in separating the

letter from the grid lines) that are not required in imagery or top-down perception.

In short, we found similar activation in two superficially dissimilar tasks. Even though the object identification task involves hearing and language comprehension, and the imagery task involves visualization and making a visual judgment, we found that both tasks clearly rely on a core set of common processes. In both cases, a system of areas was activated, and this system presumably carried out computations used in both functions.

ACKNOWLEDGMENTS

This research was supported by Grant N00014-94-1-0180 from the U.S. Office of Naval Research. We thank Avis Loring, Steve Weise, Rob McPeck, and Adam Anderson for technical assistance. We also wish to thank Christopher Chabris for helpful comments.

REFERENCES

- Alpert, N. M., Berdichevsky, D., Levin, Z., Morris, E. D., and Fischman, A. J. 1996. Improved methods for image registration. *NeuroImage* **3**:10–18.
- Alpert, N. M., Berdichevsky, D., Weise, S., Tang, J., and Rauch, S. L. 1993. Stereotactic transformation of PET scans by nonlinear least squares. In *Quantification of Brain Function: Tracer Kinetics and Image Analysis in Brain PET* (K. Uemura, N. A. Lassen, T. Jones, and I. Kanno, Eds.), pp. 459–463. Elsevier, Amsterdam.
- Andersen, R. A. 1987. Inferior parietal lobe function in spatial perception and visuomotor integration. In *Handbook of Physiology, Section 1, The Nervous System. Vol. 5, Higher Functions of the Brain* (F. Plum and V. Mountcastle, Eds.), pp. 483–518. Am. Physiol. Soc., Bethesda, MD.
- Andersen, R. A., Essick, G. K., and Siegel, R. M. 1985. Encoding of spatial location by posterior parietal neurons. *Science* **230**:456–458.
- Behrmann, M., Winocur, G., and Moscovitch, M. 1992. Dissociation between mental imagery and object recognition in a brain-damaged patient. *Nature* **359**:636–637.
- Bisiach, E., and Luzzatti, C. 1978. Unilateral neglect of representational space. *Cortex* **14**:129–133.
- Bly, B. M., and Kosslyn, S. M. 1997. Functional anatomy of object recognition in humans: evidence from positron emission tomography and functional magnetic resonance imaging. *Curr. Opin. Neurol.* **10**:5–9.
- Brandt, S. A., Stark, L. W., Haciosalihzade, S., Allen, J., and Tharp, G. 1989. Experimental evidence for scanpath eye movements during visual imagery. *Proceedings of the 11th IEEE Conference on Engineering, Medicine, and Biology*. Seattle.
- Charlot, V., Tzourio, N., Zilbovicius, M., Mazoyer, B., and Denis, M. 1992. Different mental imagery abilities result in different regional cerebral blood flow activation patterns during cognitive tests. *Neuropsychologia* **30**:565–580.
- Corbetta, M., Miezen, F. M., Schulman, G. L., and Petersen, S. E. 1993. A PET study of visuospatial attention. *J. Neurosci.* **13**:1202–1226.
- Costin, D. 1988. MacLab: A MacIntosh system for psychology lab. *Behav. Res. Methods Instrument. Comput.* **20**:197–200.
- Craver-Lemley, C., and Reeves, A. 1987. Visual imagery selectively reduces vernier acuity. *Perception* **16**:533–614.
- Damasio, H., Grabowski, T. J., Damasio, A., Tranel, D., Boles-Ponto,

- L., Watkins, G. L., and Hichwa, R. D. 1993. Visual recall with eyes closed and covered activates early visual cortices. *Soc. Neurosci. Abstr.* **19**(2):1603.
- DeYoe, E. A., Felleman, D. J., Van Essen, D. C., and McClendon, E. 1994. Multiple processing streams in occipitotemporal visual cortex. *Nature* **371**:151–154.
- Douglas, K. L., and Rockland, K. S. 1992. Extensive visual feedback connections from ventral inferotemporal cortex. *Soc. Neurosci. Abstr.* **18**(1):390.
- Farah, M. J. 1988. Is visual imagery really visual? Overlooked evidence from neuropsychology. *Psychol. Rev.* **95**:307–317.
- Farah, M. J., Peronnet, F., Gonon, M. A., and Girard, M. H. 1988. Electrophysiological evidence for a shared representational medium for visual images and visual percepts. *J. Exp. Psychol. Gen.* **117**:248–257.
- Felleman, D. J., and Van Essen, D. C. 1991. Distributed hierarchical processing in primate cerebral cortex. *Cereb. Cortex* **1**:1–47.
- Finke, R. A., and Shepard, R. N. 1986. Visual functions of mental imagery. In *Handbook of Perception and Human Performance* (K. R. Boff, L. Kaufman, and J. P. Thomas, Eds.), pp. 37-1–37-55. Wiley-Interscience, New York.
- Finke, R. A., Johnson, M. K., and Shyi, G. C.-W. 1988. Memory confusions for real and imagined completions of symmetrical visual patterns. *Memory Cognit.* **16**:133–137.
- Fletcher, P. C., Frith, C. D., Baker, S. C., Shallice, T., Frackowiak, R. S. J., and Dolan, R. J. 1995. The mind's eye—Precuneus activation in memory-related imagery. *NeuroImage* **2**:195–200.
- Friston, K. J., Frith, C. D., Liddle, P. F., and Frackowiak, R. S. J. 1991. Comparing functional (PET) images: The assessment of significant changes. *J. Cereb. Blood Flow Metab.* **11**:690–699.
- Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., and Frackowiak, R. S. J. 1995. Statistical parametric maps in functional imaging: A general linear approach. *Hum. Brain Map.* **2**:189–210.
- Friston, K. J., Worsley, K. J., Frackowiak, R. S. J., Mazziotta, J. C., and Evans, A. C. 1994. Assessing the significance of focal activations using their spatial extent. *Hum. Brain Map.* **1**:214–220.
- Fujita, I., Tanaka, K., Ito, M., and Cheng, K. 1992. Columns for visual features of objects in monkey inferotemporal cortex. *Nature* **360**:343–346.
- Goldenberg, G., Podreka, I., Steiner, M., Willmes, K., Suess, E., and Deecke, L. 1989. Regional cerebral blood flow patterns in visual imagery. *Neuropsychologia* **27**:641–664.
- Goldman-Rakic, P. S. 1987. Circuitry of primate prefrontal cortex and regulation of behavior by representational knowledge. In *Handbook of Physiology, Section 1, The Nervous System, Vol. 5, Higher Functions of the Brain* (F. Plum and V. Mountcastle, Eds.), pp. 373–417. Am. Physiol. Soc., Bethesda, MD.
- Gross, C. G., Desimone, R., Albright, T. D., and Schwartz, E. L. 1984. Inferior temporal cortex as a visual integration area. In *Cortical Integration* (F. Reinoso-Suarez and C. Ajmone-Marsan, Eds.). Raven Press, New York.
- Haxby, J. V., Grady, C. L., Horowitz, B., Ungerleider, L. G., Mishkin, M., Carson, R. E., Herscovitch, P., Schapiro, M. B., and Rapoport, S. I. 1991. Dissociation of object and spatial visual processing pathways in human extrastriate cortex. *Proc. Natl. Acad. Sci. USA* **88**:1621–1625.
- Intraub, H., and Hoffman, J. E. 1992. Reading and visual memory: Remembering scenes that were never seen. *Am. J. Psychol.* **105**:101–114.
- Jankowiak, J., Kinsbourne, M., Shalev, R. S., and Bachman, D. L. 1992. Preserved visual imagery and categorization in a case of associative visual agnosia. *J. Cognit. Neurosci.* **4**:119–131.
- Johnson, M. K., and Raye, C. L. 1981. Reality monitoring. *Psychol. Rev.* **88**:67–85.
- Jolicoeur, P. 1990. Identification of disoriented objects: A dual-systems theory. *Mind Lang.* **5**:387–410.
- Jolicoeur, P., Gluck, M. A., and Kosslyn, S. M. 1984. Pictures and names: Making the connection. *Cognit. Psychol.* **16**:243–275.
- Kosslyn, S. M. 1994. *Image and Brain: The Resolution of the Imagery Debate*. MIT Press, Cambridge, MA.
- Kosslyn, S. M., and Chabris, C. F. 1990. Naming pictures. *J. Visual Lang. Comput.* **1**:77–95.
- Kosslyn, S. M., Alpert, N. M., Thompson, W. L., Chabris, C. F., Rauch, S. L., and Anderson, A. K. 1994. 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., and Buonanno, F. S. 1993. Visual mental imagery activates topographically organized visual cortex: PET investigations. *J. Cognit. Neurosci.* **5**:263–287.
- Kosslyn, S. M., Cave, C. B., Provost, D., and Von Gierke, S. 1988. Sequential processes in image generation. *Cognit. Psychol.* **20**:319–343.
- Kosslyn, S. M., Thompson, W. L., Kim, I. J., and Alpert, N. M. 1995a. Topographical representations of mental images in primary visual cortex. *Nature* **378**:496–498.
- Kosslyn, S. M., Thompson, W. L., and Alpert, N. M. 1995b. Identifying objects at different levels of hierarchy: A positron emission tomography study. *Hum. Brain Map.* **3**:107–132.
- LaBerge, D., and Buchsbaum, M. S. 1990. Positron emission tomography measurements of pulvinar activity during an attention task. *J. Neurosci.* **10**:613–619.
- Levine, D. N. 1982. Visual agnosia in monkey and man. In *Analysis of Visual Behavior* (D. J. Ingle, M. A. Goodale, and R. J. W. Mansfield, Eds.), pp. 629–670. MIT Press, Cambridge, MA.
- Maunsell, J. H. R., and Newsome, W. T. 1987. Visual processing in monkey extrastriate cortex. *Annu. Rev. Neurosci.* **10**:363–401.
- McDermott, K. B., and Roediger, H. L. 1994. Effects of imagery on perceptual implicit memory tests. *J. Exp. Psychol. Learn. Memory Cognit.* **20**:1379–1390.
- Mellet, E., Tzourio, N., Crivello, F., Joliot, M., Denis, M., and Mazoyer, B. 1996. Functional anatomy of spatial mental imagery generated from verbal instructions. *J. Neurosci.* **16**(20):6504–6512.
- Mellet, E., Tzourio, N., Denis, M., and Mazoyer, B. 1995. A positron emission tomography study of visual and mental spatial exploration. *J. Cognit. Neurosci.* **7**:433–445.
- Mesulam, M.-M. 1981. A cortical network for directed attention and unilateral neglect. *Ann. Neurol.* **10**:309–325.
- Paivio, A. 1971. *Imagery and Verbal Processes*. Holt, Rinehart & Winston, New York.
- Paus, T. 1996. Location and function of the human frontal eye field: A selective review. *Neuropsychologia* **34**:475–483.
- Podgorny, P., and Shepard, R. N. 1978. Functional representations common to visual perception and imagination. *J. Exp. Psychol. Hum. Percept. Perform.* **4**:21–35.
- Posner, M. I., and Petersen, S. E. 1990. The attention system of the human brain. *Annu. Rev. Neurosci.* **13**:25–42.
- Rockland, K. S., Saleem, K. S., and Tanaka, K. 1992. Widespread feedback connections from areas V4 and TEO. *Soc. Neurosci. Abstr.* **18**(1):390.
- Roland, P. E., and Friberg, L. 1985. Localization of cortical areas activated by thinking. *J. Neurophysiol.* **53**:1219–1243.
- Roland, P. E., and Gulyas, B. 1994. Visual imagery and visual representation. *Trends Neurosci.* **17**:281–296. [With commentaries]

- Roland, P. E., Erikson, L., Stone-Elander, S., and Widen, L. 1987. Does mental activity change the oxidative metabolism of the brain? *J. Neurosci.* **7**:2373–2389.
- Sergent, J., Ohta, S., and MacDonald, B. 1992a. Functional neuroanatomy of face and object processing: A positron emission tomography study. *Brain* **115**:15–36.
- Sergent, J., Zuck, E., Levesque, M., and MacDonald, B. 1992b. Positron emission tomography study of letter and object processing: Empirical findings and methodological considerations. *Cereb. Cortex* **2**:68–80.
- Shuttleworth, E. C., Syring, V., and Allen, N. 1982. Further observations on the nature of prosopagnosia. *Brain Cognit.* **1**:302–332.
- Talairach, J., and Tournoux, P. 1988. *Co-planar Stereotaxic Atlas of the Human Brain* (translated by M. Rayport). Thieme, New York.
- Tanaka, K., Saito, H., Fukada, Y., and Moriya, M. 1991. Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *J. Neurophysiol.* **66**:170–189.
- Ungerleider, L. G., and Haxby, J. V. 1994. ‘What’ and ‘where’ in the human brain. *Curr. Opinion Neurol.* **4**:157–165.
- Ungerleider, L. G., and Mishkin, M. 1982. Two cortical visual systems. In *Analysis of Visual Behavior* (D. J. Ingle, M. A. Goodale, and R. J. W. Mansfield, Eds.), pp. 549–586. MIT Press, Cambridge, MA.
- Van Essen, D. C. 1985. Functional organization of primate visual cortex. In *Cerebral Cortex* (A. Peters and E. G. Jones, Eds.). Plenum, New York.
- Van Essen, D. C., and Maunsell, J. H. 1983. Hierarchical organization and functional streams in the visual cortex. *Trends Neurosci.* **6**:370–375.
- Worsley, K. J., Evans, A. C., Marrett, S., and Neelin, P. 1992. A three-dimensional statistical analysis for rCBF activation studies in human brain. *J. Cereb. Blood Flow Metab.* **12**:900–918.