Aspects of a Cognitive Neuroscience of Mental Imagery

STEPHEN M. KOSSLYN

Although objects in visual mental images may seem to appear all of a piece, when the time to form images is measured this introspection is revealed to be incorrect; objects in images are constructed a part at a time. Studies with split-brain patients and normal subjects reveal that two classes of processes are used to form images—ones that activate stored memories of the appearances of parts and ones that arrange parts into the proper configuration. Some of the processes used to arrange parts are more effective in the left cerebral hemisphere and some are more effective in the right cerebral hemisphere; the notion that mental images are the product of right hemisphere activity is an oversimplification.

PERHAPS THE MOST FUNDAMENTAL INSIGHT OF CONTEMPORARY cognitive science is the discovery that mental faculties can be decomposed into multicomponent information-processing systems. Although mental faculties such as "memory," "thinking," "imagery," and so on intuitively may seem to be single abilities, they are not. How visual mental imagery is being analyzed into distinct processing components and how these functionally characterized components are coming to be identified with brain structures is the subject of this article. Only one facet of imagery is considered here, namely the way visual mental images are generated from stored information.

Mental imagery has played a key role in many theories of mental function, both historically and currently (*1–3*). Imagery consists of brain states like those that arise during perception but occurs in the absence of the appropriate immediate sensory input; such events are usually accompanied by the conscious experience of "seeing with the mind's eye," "hearing with the mind's ear," and so on. Visual imagery is a particularly useful place to begin in that it clearly draws on some of the mechanisms also used in visual perception (*2–5*), and the anatomy and physiology of vision is becoming relatively well understood (*6, 7*). Evidence for the use of common mechanisms in imagery and like-modality perception abounds. For example, visual perception is more difficult than auditory perception when one is simultaneously holding a visual mental image, and vice versa when one is holding an auditory mental image (*8*). In addition, some visual illusions also appear in visual imagery (*4*). Indeed, there is emerging evidence that visual areas of the brain are selectively activated during visual mental imagery (*5*).

Generating hypotheses about the processing that underlies imagery is aided by consideration of three kinds of factors. First, it is necessary to begin by characterizing the behavior of imagery mechanisms. Without such information, there is nothing to explain. Second, because a theory of human information processing is in fact a theory about how the brain functions, it is useful to have some knowledge of the underlying neural substrate. Given that imagery shares some modality-specific perceptual mechanisms, facts about the anatomy and physiology of the visual system can be used in generating hypotheses about the processing underlying imagery. Third, it is useful to perform an analysis of what would be required to build a system that would produce the observed behavior. The use of these three kinds of factors is illustrated in the following section.

Generating Visual Mental Images

Probably the most obvious behavioral property of the imagery system is that images are not present all the time, but only occur in specific circumstances. For example, if one is asked to decide whether the uppercase letters of the alphabet have only straight lines or contain any curved lines, images of the letters are likely to be used. These images come to mind only when one begins to perform the task. The question to be considered here is, what is the nature of the processing that produces mental images?

Behavioral characterization. When asked, most people report that images of simple objects, such as letters or line patterns, seem to pop into mind all at once. However, when the time course of image formation is charted, such introspections are revealed to be incorrect: imaged patterns are built up a part at a time. Consider the following task. First, observe the letter in the grid at the upper left of Fig. 1. If that letter were present in the grid at the upper right, would it cover the X mark? In these experiments, subjects first memorized a set of such block letters, which varied from two (L) to five segments (G). The subjects later were shown a blank grid with a lowercase letter beneath it, and were asked to decide whether the corresponding uppercase version of the letter—if drawn in the grid as previously seen—would fill the cells occupied by two such X marks. On half the trials the letter would have covered both X marks, whereas on the other half it would have covered only one (the other was in a cell that would have been adjacent to the letter). Subjects were told to respond as quickly as possible while being as accurate as possible; response time and accuracy were measured.

The key to this method is that the two probe marks appeared in the grid only 500 milliseconds after the lowercase cue letter was presented. Given that up to 250 ms are necessary to read a letter cue (*9*), and about 250 ms are required to move one's eyes up from the cue, there was not enough time to finish forming the image before the probes appeared. Hence the time to respond should in part reflect the time to form the image (*10*).

The author is professor of psychology at Harvard University, Cambridge, MA 02138.

The first result of interest was that the response times increased with the visual complexity (number of segments) of the queried letter (mean slope, 133 ms per segment; SE, 28 ms; $P < 0.0005$). Although this result suggests that more complex forms require more time to image, it could instead reflect the time to search for the probe marks. Thus, it is important that complexity had greatly reduced effects when these subjects evaluated probe marks with the figure actually present (mean slope, 10 ms; SE, 4 ms; $P < 0.05$); the results in the imagery task do not reflect only search and evaluation time. In addition, in another condition the probe marks were eliminated, and subjects now were asked simply to read the lowercase cue and form an image of the corresponding uppercase version in the grid; as soon as the image was fully formed, the subjects were to press a key. These times also increased for letters with more segments, and did so to a similar degree in this task and the image evaluation task (mean slope, 100 ms per segment; SE, 19 ms; $P < 0.0005$; $P > 0.1$ for the comparison of the two slopes). Thus, there is reason to infer that differences in response times in the experimental task reflect differences in image formation time (*11*).

In addition to varying the complexity of the stimuli, the positions of the probe marks were varied along the individual letters. If the image is being constructed a segment at a time, then some probes ought to require more time to reach than others. A separate group of 25 subjects was asked to copy the block letters into empty grids, and the order in which the segments were drawn was covertly observed; the order was highly consistent, with five of the letters being drawn in the same way by 100% of the subjects, and the remainder being drawn in the same way by at least 75% (when these letters were drawn differently it was always in the order of a single segment). As is illustrated in Fig. 2, more time was required in the image evaluation task when the "farthest" probe mark fell on a segment typically drawn later in the sequence (mean slope, 178 ms per segment; SE = 35 ms; $P < 0.0005$). This effect of probe position did not occur in the perception control task (mean slope, 2 ms per segment; SE, 7 ms; $P > 0.25$). Similar results were found for novel two-dimensional patterns (*11*) and three-dimensional shapes (*12*).

Thus, it appears that patterns in images are built up by activating parts individually and that parts are imaged in roughly the order in which they are typically drawn. These inferences were supported by a host of additional experiments controlling for various alternative accounts. For example, it was possible that the effect of probe position was due to scanning an imaged pattern in search of the probes (which might be different than inspecting a figure that is actually present). If so, then farther probes should require more time to evaluate than nearer ones, even when one has formed the image in advance of the probe; this did not occur. It was also possible that the effects reflect patterns of eye movements; nevertheless, they persisted even when subjects fixated on the center of the screen while performing the task (*11, 12*).

Additional research has been conducted to discover what factors determine the nature of the parts, and has shown that principles of perceptual organization also determine the part structure of images (*13*). That is, it has been known since the early part of this century that we see lines and regions as being organized into "perceptual units." For example, the pattern "------" is seen as a line (grouped by the "law of good continuation"), not six isolated dashes; "XXX XXX" is seen as two units (grouped by the "law of proximity"), not six solitary X's; and XXXooo is seen as two units (grouped by the "law of similarity"), not simply three X's and three o's. Similarly, lines that form a symmetrical pattern or that form enclosed areas tend to be grouped as units (*14*). In the block letter stimuli used in the image generation experiments, adjacent filled cells will form a unit (a bar) as per the law of good continuation. There is good evidence that these sorts of units are not only perceived, but also are stored in memory (*13*).

Thus, given that visual mental images are formed by activating previously stored perceptual information, it is easy to formulate a hypothesis about why images are constructed a part at a time: namely, when originally viewed the parts were stored individually and hence they are later activated into an image individually. But even so, the data suggest that parts are activated sequentially. Why are they not simply activated all at once to reconstruct the entire object in the image?

Neurological constraints. One possible reason why parts are imaged sequentially hinges on the way parts and spatial relations among them might be stored in memory. Ungerleider and Mishkin (*7*) summarize evidence for "two cortical visual systems" in primates (Fig. 3). The ventral system runs from area OC (primary visual cortex) through area TEO down to the inferior temporal lobe. This system has been identified with the analysis of shape ("what"). The dorsal system runs almost directly from circumstriate area OB to OA and then to PG (in the parietal lobe). This system has been identified with the analysis of location ("where").

Three sorts of data have been marshalled to support Ungerleider and Mishkin's claims. First, neuroanatomical investigations have documented the existence of the separate pathways. Indeed, each pathway has now been decomposed into connections among numerous distinct areas (*6, 7*).

Second, neurophysiological investigations of monkey brains have revealed that cells in both systems are sensitive to visual input, but have different functional properties. For example, cells in the inferior temporal lobe are sensitive to shape (often being highly tuned for specific shape properties), color, and have very large receptive fields that almost always include the fovea (*15, 16*). In contrast, cells in the parietal lobe are not particularly sensitive to shape or color, rarely include the fovea in their receptive fields, are sensitive to direction of motion, and some cells in this region respond selectively to an object's location (as gated by eye position) (*17*).

Third, behavioral data provide dramatic evidence of the distinct visual functions of the two systems. When the temporal lobes are ablated but the parietal lobes are spared, animals are severely impaired in learning to discriminate among patterns, but are relatively unimpaired in learning to discriminate among locations. In contrast, when the parietal lobes are ablated and the temporal lobes are spared, the animals are severely impaired in learning to discriminate locations, but are relatively unimpaired in learning to discriminate among patterns; corresponding results have also been reported in humans after stroke (*18*).

Information-processing analyses. The observation that "what" and "where" are processed separately during perception leads to an explanation of why parts are imaged sequentially if the shape of each part is stored separately, and a part's location is specified relative to another part. If so, then one needs to have the reference part already activated before one can know where a subsequent part belongs in an image. When generating an image of a block letter F, for example, one might have encoded that there is a vertical line on the left, a horizontal line connected at its left side to the top of the vertical line, and another horizontal line connected at its left side to the vertical line midway down (94% of the subjects we have observed print the segments in this order). When forming the image, then, the vertical line on the left is a prerequisite for the other two lines; the locations of the horizontal lines are specified relative to the vertical line. Thus, some parts should be imaged before others. Finally, because one needs to attend to a specific place on the reference part in order to place a new part, and focal attention is restricted to only a single region of space at a time (*19*), only one part can be imaged at a time.

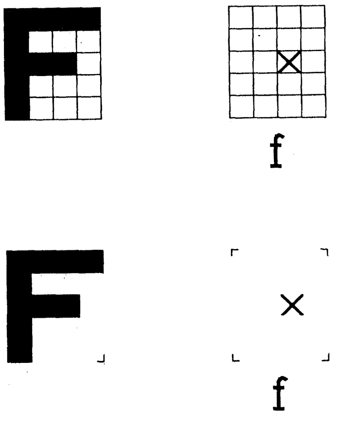

Fig. 1. (**Top left**) A letter formed by selectively filling in cells of a matrix. (**Top right**) Subjects were shown a lowercase cue and asked whether the corresponding uppercase block letter would occupy the cells containing one or two X marks. Because only 500 ms were allowed between presentation of the lowercase cue and the probe marks, which is not enough time to read the cue and finish forming the image, the decision times in part reflect the time to form the image in the matrix. (**Bottom**) An alternative way of presenting the stimuli, which was expected to induce a different method of arranging parts in the image.

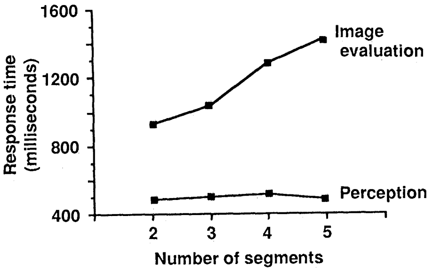

Fig. 2. The mean time to evaluate probe marks in the image evaluation and perception tasks when the farthest probe mark was on the first, second, third, or fourth segment typically drawn. The image evaluation task required deciding whether probes would have fallen on a letter, and the perception task required deciding whether probes actually did fall on a letter.

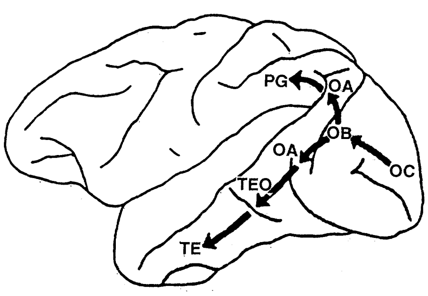

Fig. 3. The dorsal and ventral systems of the primate brain. [Adapted from (*32*)]

Consider first the assumption that individual parts of shapes are stored separately. When is this likely to occur? An important fact about the visual system is that it operates at multiple scales of analysis; it is sensitive to coarse overall features and to fine-grained features, and one can attend to a given scale. However, there appears to be a trade-off between scope and resolution (*20, 21*). For example, one can take in an entire human form, but not see much in the way of details about the face, or one can attend to individual features of the face at high resolution but lose the overall pattern. When one attends to details, the laws of perceptual organization serve to parse objects into parts, as was noted earlier (*13*). Attending to details is necessary to distinguish among similar objects that have different parts, such as letters; thus, parts should be encoded separately for such objects.

Consider next the assumption that part locations are stored relative to other parts. This assumption follows from an analysis of what would be required to build a machine that recognizes shape: for purposes of perceptual recognition, it is of limited use to store a single pattern to be matched against input as a template. Many objects (such as a dog, scissors, or a person) can assume a very large number of distinct shapes (as, for example, when a dog is scratching, sleeping, running, jumping, and so on). Similarly, many objects assume multiple variants on a shape, such as a letter of the alphabet (which comes in many fonts). In such cases, there are so many distinct shapes, and new ones occur all the time, that shapes may often arise that do not correspond to one previously seen and stored in memory. Thus, encoding a shape as a single pattern may not lead to a match with a previously stored shape.

A more effective way to represent such shapes is to extract specific properties that will not change when the object assumes a new configuration or a shape-variant appears. One such invariant is the type of spatial relation between parts. For example, no matter how a dog is contorted, the parts remain connected in the same way. I refer to this type of representation as categorical because equivalence classes of relations are specified. "Connected to" (or "above," "below," "inside," "next to," and so on) does not correspond to a particular topographic configuration, but rather specifies a large category of such configurations (for example, the foreleg remains "connected to" the upper leg no matter how the leg is bent or stretched. With such categorical representations, part locations are specified relative to other parts (*20*). If the appropriate categorical relations are used, a description of the arrangement of an object's parts will be the same across its various contortions and variations and, hence, will be useful for recognition.

Neuropsychological Hypothesis Testing

Images thus appear to be formed a part at a time because (i) shapes of individual parts of objects are stored separately, (ii) spatial relations among parts are stored separately from shapes, (iii) spatial relations specify location relative to other parts, (iv) stored parts and relations are used to form mental images, and (v) only a single reference point on a prior part can be located at a time. Hence, because parts can only be added when the reference part is present and the reference location on it has been found, parts will be imaged sequentially.

The separation of the storage of parts and spatial relations suggests a possible distinction between two classes of processes—ones that activate stored visual shapes and ones that gain access to and use stored spatial relations to arrange those shapes correctly. One way to garner evidence for this initial, rather coarse decomposition of the processing underlying image generation is to show that the two kinds of processing can be dissociated during image formation. Farah *et al.* and Kosslyn *et al.* have done just this in a series of experiments examining image generation in commissurotomy patients (*22, 23*).

One task we used required subjects to make judgments about letters. Letters were the initial stimuli of choice in part because of the evidence that they are imaged a segment at a time. The task was to decide from memory whether specific uppercase letters are composed only of straight lines (for example, A and H) or include at least one curved line (for example, B and D) (*24*); unless one has performed the task many times or intentionally memorized the responses, imagery is used to make this judgment (*23*). There were two critical assumptions in our experiments. First, we assumed that the shapes of letters are stored as segments and categorical spatial relations among them. Letters come in many different fonts, and one wants to recognize new instances; thus it is efficient to store categorical representations of the spatial relations among parts, which will apply to a wide range of different topographic positions. Second, we assumed that categorical relations are language-like (indeed, they almost always can be labeled by a word or two). This assumption was critical for our experiments because one uncontroversial fact about the functional specialization of the cerebral hemispheres is that for right-handed people the left hemisphere is superior to the right at producing and using language (*18*).

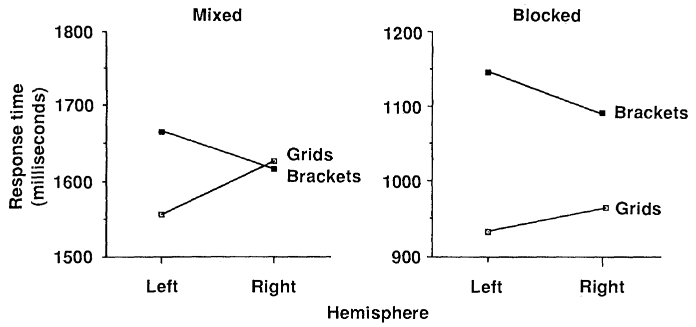

Fig. 4. The mean time required by normal right-handed people to evaluate single probe marks in lateralized grids. The grids were complete (grids) or reduced to four brackets at the corners (brackets) (see Fig. 1). (**Left**) Results when each letter was imaged only once before a new letter was presented (mixed trials). (**Right**) Results when the letter was imaged eight times in succession after the cue was presented (blocked trials). In mixed trials a lowercase cue was presented beneath the grid; in blocked trials the cue was presented at the central fixation point at the beginning of the trials.

Therefore, we expected that the left cerebral hemisphere would be better at generating images of letters, assuming that multiple parts would be composed by use of categorical spatial relations. In contrast, we expected no difference between the hemispheres in the ability to form images of single shapes, when separate representations of spatial relations would not be used; visual memories of shape should be equally available in both hemispheres. If we could document such a dissociation and implicate image generation differences as its cause, we would have evidence for a distinction between a mechanism that activates stored shapes and one that uses stored spatial relations to arrange them.

Documenting an Image-Generation Deficit

We tested these hypotheses by lateralizing lowercase letter cues. The lateralization procedure consisted of asking subjects to focus on a fixation point, and then to view stimuli presented 1.5° to the left or right of fixation. Because the cues were only presented for 150 ms, the subjects could not move their eyes to examine them, ensuring that the letter was projected onto only one half of the retina. Because the left half of each retina sends input only to the left cerebral hemisphere, and the right half of each retina sends input only to the right cerebral hemisphere, lateralizing stimuli in this way allowed us to provide input to the separate hemispheres (*25*).

The most detailed examination, which was conducted with patient J.W., will be summarized here (*23*). This patient had undergone surgery about 3 years before testing (for the treatment of otherwise intractable epilepsy), and magnetic resonance scanning revealed that his corpus callosum was fully sectioned. Because his corpus callosum was transected, information presented to one hemisphere was not available to the other hemisphere.

We began by lateralizing lowercase letters, and asking J.W. to judge whether the corresponding uppercase version had any curved lines. He pressed one key when he decided that the uppercase letter had at least one curved line and another if it had only straight lines. The results were striking: J.W.'s left hemisphere responded perfectly, whereas his right hemisphere was correct only 70% of the time; in a replication experiment, the left hemisphere again was perfect, whereas the right was correct only 65% of the time (*26*).

This finding is not enough to document an image generation deficit per se. A rather lengthy series of control experiments was required to eliminate various alternative accounts. The most simple control consisted of presenting J.W. with the uppercase letters themselves, and asking him to judge them as they appeared on the screen. Both hemispheres responded correctly on at least 97.5% of the trials. This result indicates that the right hemisphere deficit in the imagery task was not due to its being unable to perform the judgment, to encode letters, or produce appropriate key presses. In other control experiments we showed that both hemispheres could understand the correspondence between the lowercase cues and the uppercase letters, could retain images long enough to interpret their shapes, and could perform multipart tasks.

A selective deficit. We expected the right hemisphere to have difficulty generating images of letters if parts are activated sequentially and categorical representations of spatial relations are used to arrange the parts into the image. However, to argue that the difficulty lies in one class of processes and not others, it must be demonstrated that the right hemisphere can perform some imagery tasks, but not those in which the parts must be arranged with the use of categorical relations.

We reasoned that images should be constructed from parts whenever the task requires evaluation or comparison of parts; in these cases, high-resolution images of parts are necessary, and these representations were encoded and stored individually. In contrast, arranging parts should not be necessary to perform a task requiring imaging the overall shape of an object, if such a shape was encoded as a single perceptual unit. That is, even though such a template has little use in perceptual recognition for generalizing over variations in an object's shape, we expected relatively low-resolution patterns of overall shape to be encoded because they provide information about how an object is oriented in space, which is useful for navigation. If so, then we expected both hemispheres to be able to image such a pattern representing a single object. Although such a pattern would have a relatively low resolution, we reasoned that it should be sufficient to perform tasks that do not require comparing or evaluating parts (and hence high-resolution images of individual parts are not necessary).

In one task, J.W. was asked to decide which of two similar-sized animals (for example, a goat or a hog) was the larger, as seen from the side at the same distance; this task previously had been shown to require imagery (*27*). The name of one of the animals was lateralized, and hence only one hemisphere had the opportunity to perform the task. Only one error occurred during the entire testing session. This high level of performance is worrisome, however, in that it may reflect a "ceiling effect." The task may be so easy that it is insensitive to differences in hemispheric processing. Thus, we devised a second imagery task that did not require assembling parts, but which was considerably more difficult. J.W. was asked to decide whether named objects (book, nose, and buckle, for example) are taller than they are wide. This was a difficult task, given the stimuli we used, and resulted in overall worse performance. Nevertheless, both hemispheres could perform the task at better than chance levels of performance (50%), and did so equally well (70.8% versus 66.7% correct for the left and right hemispheres, respectively).

One could argue that all that has been shown is a difference for letters versus words. Thus, we conducted another task with the animal names presented in the size judgment experiment, but this task required comparing locations of parts, which we assumed required relatively high-resolution images of the parts—entailing the use of individually stored parts and spatial relations. We now asked J.W. to decide whether the named animal's ears protrude above the top of its skull (for example, an ape and a sheep versus a cat and a mouse). The left hemisphere performed correctly on 87.5% of the trials, whereas the right hemisphere performed correctly on only 45% of the trials. In short, the problem was not limited to letters, but apparently to tasks that involve juxtaposing parts.

Convergent evidence for the distinction between processes that activate images and that arrange parts is also available in the clinical literature. For example, Deleval *et al.* (*28*) describe a patient who experienced impaired imagery following left hemisphere damage. This patient claimed, "When I try to image a plant, an animal, an object, I can recall but one part, my inner vision is fleeting, fragmented; if I'm asked to imagine the head of a cow, I know that it has ears and horns, but I can't revisualize their respective places. In the same way, I cannot determine how many fingers a frog paw has, even though I have manipulated this animal each day in the laboratory. . . ."

Contrasting Left and Right Hemisphere Abilities

A second split-brain patient also provided evidence for a functional dissociation between processes that activate images and processes that arrange parts in images, with the left hemisphere being superior when the latter processes were required. However, only two subjects were tested, and these patients may have atypical cerebral organization due to years of severe epilepsy and the disconnection of the cerebral hemispheres at the time of testing. Thus it is important to show that the inferences about component processes and their neural realization generalize to normal people.

To obtain such converging evidence, a group of normal right-handed Harvard University students was tested in a variation of the grids imagery task described above (see the top panels of Fig. 1). After subjects memorized the block letters, they saw lateralized grids with a lowercase cue beneath them (64 trials, half presented in each visual field). The letters were presented in mixed order, with a letter not repeated in fewer than four trials. The grids now contained only a single "X" probe, as is illustrated in Fig. 1. Because the corpus callosa of these subjects is intact, information presented to one hemisphere will be transmitted to the other; thus, the primary measure of interest here was the time to respond: response times should be fastest when the hemisphere that receives the initial input is more effective in processing (*25*).

Consistent with the findings from the split-brain patients, these subjects evaluated the probes more quickly when the grid was presented to the right visual field (and hence was projected onto the left side of each retina and was seen first in the left cerebral hemisphere). However, it was possible that this result only reflected the left hemisphere's greater facility at reading the lowercase cues. Hence, as a control, an additional group of students was tested in a modified version of the task. The lowercase cue was now presented in the center (replacing the fixation point), not beneath the grid; after this, an empty grid with one probe mark was lateralized eight times (four in the left field, and four in the right, with no more than three trials appearing in the same field in a row and with the probe mark in a different location on each trial). Each new series of trials began with a different lowercase cue appearing in the center. As is evident in Fig. 4, although these subjects responded faster overall in this blocked design, a left hemisphere advantage was nevertheless obtained. Thus, the left hemisphere was shown to be better than the right hemisphere even in normal subjects at performing this multipart image generation task (*29*).

These experiments with normal subjects pushed the information-processing analyses one step further. The tests done with the split-brain patients hinged on the assumptions that categorical relations are used to arrange segments of letters and relative positions of animal parts in images, and that such relations are processed better in the left hemisphere. However, categorical representations cannot be the only method used by the brain to store spatial relations: what is a virtue for recognizing semirigid objects is a drawback for distinguishing among subtly different multipart shapes or for reaching or navigating. For these tasks, one needs to know the actual metric spatial relations among parts or objects. Knowing only that an object is "next to" the wall will not help one very much to find it and pick it up. For these sorts of tasks, the coordinates of an object must be internally represented. In short, information-processing considerations lead to the hypothesis that the brain can store spatial relations in two ways, either in terms of a category or in terms of more precise coordinates (*30*).

If this is so, then there should be two ways of forming images of a multipart object—by using either categorical or coordinate stored spatial relations to arrange parts. Given the evidence that the right hemisphere is more efficient at representing and processing metric spatial relations (*18*), the right hemisphere therefore should be better than the left when parts must be arranged in precise positions in an image. To test this idea, an additional group of students was tested in a modified version of the grids task; this task was the same as the grids task except that the internal lines were removed and only brackets at the four corners were depicted (as is illustrated in the bottom panels of Fig. 1). After subjects memorized the block letters as they appeared within the brackets, probe marks within the brackets were lateralized (lower right corner of Fig. 1), and the subjects were asked to decide whether the probe would fall on the letter were it within the brackets as previously shown. When grid lines are present, a categorical representation of how segments are connected is adequate; the grid lines are a crutch for placing segments properly in accordance with a description. In contrast, when only four corner brackets are present, more precise representations of segment location are necessary to determine whether an imaged letter would cover the X mark. Thus, it was expected that a process that uses coordinate representations to arrange parts would be recruited in this task, and that this process would be more effective in the right hemisphere.

The results from both the "grids" and "brackets" conditions are illustrated in Fig. 4. As expected, the subjects were faster when the brackets stimuli were presented to the left visual field, and hence were seen first in the right hemisphere. These results were obtained when lowercase cues were presented beneath the brackets ("mixed" presentation) or when an additional group saw them in the center before eight consecutive trials with that letter ("blocked" presentation); in both cases, these results are in sharp contrast to those obtained with grids, when categorical relations were presumably adequate.

In order to consider whether the results with the brackets were due to a right-hemisphere superiority at localizing the probe marks, and not due to image generation per se, a separate group of subjects was given an analog of the task that did not require image generation but did require encoding the probe location and comparing it to an uppercase letter. A probe X was lateralized within brackets (for 150 ms) and then replaced by an uppercase letter (as illustrated at the lower left of Fig. 1, for 100 ms). The task was to judge whether the X mark would have fallen on the letter, had they been superimposed. The letter served to mask the X, requiring subjects to encode its location into memory to be compared to the locations of the letter segments. As expected, the right hemisphere was superior when the metric location of the X had to be stored. However, this right hemisphere advantage was 3.2 times too small to account for the right hemisphere advantage for the brackets imagery task.

In short, both hemispheres can form images of the components, but the hemispheres apparently differ in the preferred way of arranging them. These results from normal subjects not only provide support for the inferences drawn from the split-brain subjects, but also provide evidence for a second means by which parts can be arranged in images (*31*).

Conclusions

In this article, I have illustrated how one can discover structure in mental abilities where none was obvious. After first examining behavior during task performance, facts about the brain and information-processing analyses can lead to relatively subtle hypotheses about processing. These hypotheses are testable in part by examining selective impairments in neuropathological populations. With this approach, it was found that the act of generating a visual mental image involves at least two classes of processes—ones that activate stored shapes and ones that use stored spatial relations to arrange shapes into an image. The discovery that the left hemisphere is better at arranging shapes when categorical information is appropriate, whereas the right hemisphere is better when coordinate information is necessary, suggests that the processes that arrange parts can be further decomposed into two classes that operate on different sorts of information.

The findings that under some circumstances the left cerebral hemisphere is better at mental imagery is counterintuitive to many. The left hemisphere has traditionally been identified with language, and the right with imagery (*22, 23, 28–30*). However, neither hemisphere can be said to be the seat of mental imagery: imagery is carried out by multiple processes, not all of which are implemented equally effectively in the same part of the brain.

REFERENCES AND NOTES

1. A. Paivio, *Imagery and Verbal Processes* (Holt, Rinehart, & Winston, New York, 1971).
2. S. M. Kosslyn, *Image and Mind* (Harvard Univ. Press, Cambridge, 1980).
3. R. N. Shepard and L. A. Cooper, *Mental Images and Their Transformations* (MIT Press, Cambridge, 1982).
4. R. A. Finke and R. N. Shepard, in *Handbook of Perception and Human Performance*, K. R. Boff, L. Kaufman, J. P. Thomas, Eds. (Wiley, New York, 1986), vol. 37, p. 1.
5. M. J. Farah, *Psychol. Rev.*, in press.
6. D. C. Van Essen, in *Cerebral Cortex*, A. Peters and E. G. Jones, Eds. (Plenum, New York, 1985), vol. 3, p. 259.
7. L. G. Ungerleider and M. Mishkin, in *Analysis of Visual Behavior*, D. J. Ingle, M. A. Goodale, R. J. W. Mansfield, Eds. (MIT Press, Cambridge, 1982), p. 549.
8. S. J. Segal and V. Fusella, *J. Exp. Psychol.* **83**, 458 (1970).
9. C. W. Eriksen and B. A. Eriksen, *ibid.* **89**, 306 (1971).
10. This method was adapted from one developed to compare imagery and perception by P. Podgorny and R. N. Shepard, *J. Exp. Psychol.: Hum. Percept. Performance* **4**, 21 (1978).
11. S. M. Kosslyn, C. B. Cave, D. A. Provost, S. M. Von Gierke, *Cog. Psychol.*, in press.
12. J. D. Roth and S. M. Kosslyn, *Cog. Psychol.*, in press.
13. S. M. Kosslyn, B. J. Reiser, M. J. Farah, S. J. Fliegel, *J. Exp. Psychol. Gen.* **112**, 278 (1983); S. K. Reed and J. A. Johnsen, *Mem. Cog.* **3**, 569 (1975).
14. J. E. Hochberg, *Perception* (Prentice-Hall, New York, 1964); L. D. Kaufman, *Sight and Mind* (Oxford Univ. Press, New York, 1974); I. Biederman, *Psychol. Rev.* **94**, 115 (1987).
15. C. G. Gross, C. J. Bruce, R. Desimone, J. Fleming, R. Gattass, in *Cortical Sensory Organization II: Multiple Visual Areas*, C. N. Woolsey, Ed. (Humana, Clinton, NJ, 1984), p. 187.
16. R. Desimone, T. D. Albright, C. G. Gross, C. J. Bruce, *J. Neurosci.* **4**, 2051 (1984).
17. R. A. Andersen, G. K. Essick, R. M. Siegel, *Science* **230**, 456 (1985).
18. E. DeRenzi, *Disorders of Space Exploration and Cognition* (Wiley, New York, 1982); H. Hecaen and M. L. Albert, *Human Neuropsychology* (Wiley, New York, 1978).
19. C. J. Downing and S. Pinker, in *Attention and Performance XI*, M. I. Posner and O. S. I. Marin, Eds. (Erlbaum, Hillsdale, NJ, 1985), p. 177; D. LaBerge, *J. Exp. Psychol.: Hum. Percept. Performance* **9**, 371 (1983); M. I. Posner, C. R. R. Snyder, B. J. Davidson, *J. Exp. Psychol.: Gen.* **109**, 160 (1980); A. M. Treisman and G. Gelade, *Cog. Psychol.* **12**, 97 (1980).
20. D. Marr, *Vision* (Freeman, San Francisco, 1982).
21. H. Egeth in *The Psychology of Learning and Motivation*, G. H. Bower, Ed. (Academic Press, New York, 1977), p. 277; C. W. Eriksen and J. D. St. James, *Percept. Psychophys.* **40**, 225 (1986); J. Jonides, *Bull. Psychonomics Soc.* **21**, 247 (1983); G. L. Shulman and J. Wilson, *Perception* **16**, 89 (1987).
22. M. J. Farah, M. S. Gazzaniga, J. D. Holtzman, S. M. Kosslyn, *Neuropsychologia* **23**, 115 (1985).
23. S. M. Kosslyn, J. D. Holtzman, M. J. Farah, M. S. Gazzaniga, *J. Exp. Psychol. Gen.* **114**, 311 (1985).
24. This task is a variant of one developed by R. J. Weber and J. Castleman, *Percept. Psychophys.* **8**, 165 (1970).
25. J. G. Beaumont, Ed., *Divided Visual Field Studies of Cerebral Organization* (Academic Press, New York, 1982).
26. Both response time and accuracy were always measured in all of these experiments to ensure that errors were not made because of hurried decisions. Such possible speed-accuracy tradeoffs do not undermine the inferences drawn from any of the response time or accuracy results described here.
27. S. M. Kosslyn, G. L. Murphy, M. E. Bemesderfer, K. J. Feinstein, *J. Exp. Psychol.: Gen.* **106**, 341 (1977).
28. J. Deleval, J. De Mol, J. Noterman, *Acta Neurol. Belg.* **83**, 61 (1983) (M. J. Farah and O. Koenig, translators). See also M. J. Farah, D. N. Levine, R. Calvanio, *Brain Cog.*, in press; M. J. Farah, *Cognition* **18**, 245 (1984).
29. For additional converging results, see M. J. Farah, *Neuropsychologia* **24**, 541 (1986).
30. See S. M. Kosslyn [*Psychol. Rev.* **94**, 148 (1987)] for additional explication of, and evidence for, this distinction.
31. Only the results from the first set of trials are summarized here; additional trials were administered to assess possible changes in strategy with practice, as is describe by S. M. Kosslyn, J. R. Feldman, V. Maljkovic, S. Hamilton, unpublished manuscript.
32. M. Mishkin, L. G. Ungerleider, K. A. Macko, *Trends Neurosci.* **6**, 414 (1983).
33. Supported by NIMH grant MH 39478, ONR contract N00014-85-K-0291, and AFOSR contract 88-0012. Requests for reprints should be sent to S. M. Kosslyn, 1236 William James Hall, 33 Kirkland Street, Cambridge, MA 02138. I thank J. Gabrieli, O. Koenig and V. Maljkovic for critical readings, S. Hamilton and C. Moheban for technical assistance, and my collaborators and students cited herein for stimulating discussions and critical reviews of earlier works.