Pergamon

Neuropsychologia, Vol. 33, No. 11, pp. 1511–1530, 1995

0028–3932/95 $9.50 + 0.00

0028–3932(95)00078–X

MENTAL SCANNING OF VISUAL IMAGES GENERATED FROM VERBAL DESCRIPTIONS: TOWARDS A MODEL OF IMAGE ACCURACY

MICHEL DENIS,* MARIE-ROSE GONÇALVES and DANIEL MEMMI

LIMSI-CNRS, Université de Paris-Sud, Orsay, France

(*Received* 7 *June* 1994; *accepted* 22 *December* 1994)

Abstract—Recent mental scanning experiments have shown that subjects are not only able to construct mental images from verbal descriptions, but that these images have structural properties similar to those of images derived from perception. In addition, the specific sequencing of a description can affect the internal structure of images of described objects, in particular their metric properties. Discontinuous descriptions require additional exposure to achieve the structural coherence of images constructed from continuous descriptions. Thus, the capacity of images to reflect accurately the objects they refer to is not an all-or-nothing property, but rather results from stepwise elaboration. This study describes a quantitative model designed to account for the gradual process of image elaboration and the progressive increase in image accuracy. The model posits that the location of a landmark mentioned in a description is not represented as a sharp point in the mental image, but is instead associated with a region around this point. Learning the description essentially consists of progressively narrowing each 'region of uncertainty' associated with a landmark to its exact location. Additional experimental data were collected to provide a more fine-grained understanding of image elaboration. Computations of the regions of uncertainty associated with each landmark were used to develop a computer program simulating the whole mental scanning protocol, which provided support for our account of image accuracy.

Key Words: mental scanning; visual images; image elaboration; image accuracy; verbal descriptions.

INTRODUCTION

A major goal of mental imagery research is to describe the structural and functional similarities between images and the perceptual events from which they were formed. Experiments generally compare the cognitive processes involved in visually perceiving an object or forming a mental image of this object. The many similarities between these two situations lend weight to the hypothesis that imagery is an analog form of representation, endowed with specific properties which differentiate it from other forms of mental representation [6, 14, 21, 22, 26, 36]. Not only do visual imagery and perception share several functional properties, but both systems apparently have common underlying brain structures. Several neuropsychological studies have provided support for this assumption [1, 3, 11]. The involvement of cortical structures common to imagery and perception in the

*Address for correspondence: Groupe Cognition Humaine, LIMSI-CNRS, Université de Paris-Sud, BP 133, F-91403 Orsay Cedex, France.

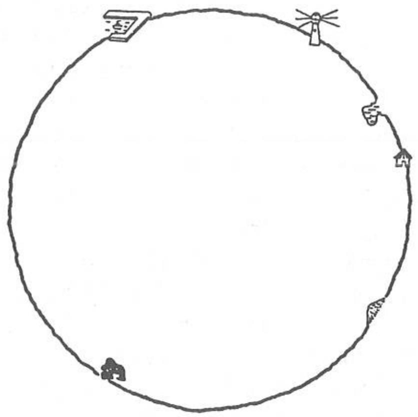
visual modality is also supported by studies on evoked potentials [13] and regional cerebral blood flow [5, 18, 24, 30].

However, several findings have raised doubts about the hypothesis that visual imagery and visual perception use the very same neural substrate. The most significant are those showing that brain-damaged patients with severely impaired object recognition may have fully preserved visual imagery [2]. On the other hand, tests on neuropsychological patients with visual imagery disorders have shown that these subjects may exhibit no perceptual disorder [20]. This double dissociation between imagery and perception in brain-damaged patients may simply reflect the fact that visual imagery depends on brain structures which perception does not require. But it is also possible that selective damage to early visual areas impairs visual perception while leaving visual imagery intact, if the areas subserving visual imagery are considered to be a subset of those active in visual perception [34]. Data indicating that the early visual areas are activated in visual imagery argue for the strong hypothesis that the visual areas subserving visual imagery are identical to those subserving visual perception [24, 27]. However, these areas are not activated during visual imagery in all subjects, and it is possible that they are activated only by tasks that require high-resolution images [35]. Recent evidence argues against the involvement of early visual areas in visual imagery, suggesting that neuronal populations in temporo-occipital and parieto-occipital areas represent objects and scenes during imagery [34], in particular when subjects mentally scan visual images [30].

A recent extension of imagery research has been the investigation of images which are not constructed from a perceptual experience of specific objects, but from a text or a verbal description of these objects. Several experiments have shown that individuals are not only able to construct mental images from verbal descriptions, but that these images have properties which are similar to those of images derived from perception. By 'similar', we not only mean that these representations have the same informational content, but more importantly that they may serve the same cognitive functions. For example, the responses of subjects who must compare distances mentally after memorizing a map have many features in common with those of subjects who read a text describing this map [9]. Similar results have been obtained for mental scanning. The mental inspection of intervals separating landmarks on a geographical configuration usually shows a positive correlation between time and distance. This correlation is interpreted as reflecting the structural isomorphism of the image with the described object, even when this image has been elaborated from text [7, 25, 33].

Little information is available on the brain structures that are involved in imagery based on verbal inputs. It may be assumed that the same neural substrate is activated whatever the origin of an image (either derived from visual perception or constructed from a description of a never-seen object). In the only neuroimaging (SPECT) study available on visual imagery for described objects [5], subjects were invited to construct the visual image of a spatial configuration from a purely verbal description. Subjects were then required to mentally explore their visual image of the configuration. The results showed that relative to resting, there was a significant increase in blood flow in the left temporo-occipital cortex when subjects imagined and mentally scanned the configuration. This result is compatible with the idea that visual representations constructed from visual experience and those constructed from descriptions involve the same cortical areas.

A recent extension of the mental scanning paradigm was designed to study image construction at different points in the learning of a text [8]. The text described the map of a fictitious island which was circular in shape, with six landmarks situated at the periphery (Fig. 1). The locations of the landmarks were unambiguously defined in the conventional clock-dial terms of aerial navigation. Two versions of the description were compared. One group of subjects listened to a text that described the landmarks on the map in clockwise order. The other group of subjects heard a text that introduced landmarks in a random order. Subjects in both groups listened to the description three times before undergoing a mental scanning test. Text learning was then resumed, with subjects listening three more times to the same description, then doing a second mental scanning test. Mental scanning was executed according to the following instructions: Subjects were asked to reconstruct the visual image of the map and mentally focus on a feature that was given to them orally. They then had to mentally scan across the map to another specific feature. Subjects were required to indicate the moment when scanning had been completed by depressing a button which stopped a timer triggered by the onset of the second feature name. (In some



Clockwise Version

The island is circular in shape. Six features are situated at its periphery. At 11 o'clock, there is a harbor. At 1, there is a lighthouse. At 2, there is a creek. Equidistant from 2 and 3, there is a hut. At 4, there is a beach. At 7, there is a cave.

Random Version

The island is circular in shape. Six features are situated at its periphery. At 11 o'clock, there is a harbor. At 4, there is a beach. At 1, there is a lighthouse. At 7, there is a cave. Equidistant from 2 and 3, there is a hut. At 2, there is a creek.

Fig. 1. Map of an island and two versions of a text describing it.

cases, the feature named did not belong to the map, and subjects then had to depress another button.)

Analysis of response times revealed that, in the first scanning test, the subjects who processed the clockwise text produced responses reflecting high structural coherence and resolution of their images, as indicated by significant positive correlation between times and distances. After additional exposure to the same text, subjects' responses in the second scanning test resulted in slight increase in time–distance correlation and a decrease in absolute scanning times (Fig. 2). The pattern of results was strikingly different for the subjects who processed the random text. The results of the first scanning test revealed no sign of internal structure in the subjects' mental images. Response times were very long, and there was no significant correlation between scanning times and distances. The situation changed markedly after three more exposures to the text. Scanning times were much shorter and, more importantly, there was a significant positive correlation between times and distances (Fig. 3).

These data support the claim that images generated from verbal descriptions can have metric properties which are revealed through the typical chronometric pattern of mental scanning, just as was the case in scanning tasks involving images constructed from perceptual inputs [25]. This study also demonstrates that the structure of a description can affect the intrinsic structure of images of described objects, and hence the mental operations performed subsequently on these images. When the description was poorly structured, the subjects eventually showed a standard time–distance correlation, but this effect was delayed. In contrast, a well-structured description (or at least, a description whose structure both conformed to most subjects' expectations and placed minimal requirements on their processing capacities) rapidly yielded accurate portrayals of the described objects, as well as more pronounced time–distance scanning relations. Therefore, poorly structured descriptions require additional exposure to achieve image coherence similar to that produced by well-structured descriptions. We interpret these results as indicating that the referential validity of images (i.e. their capacity to reflect accurately the objects they refer to), is not an all-or-nothing property, but rather results from stepwise elaboration. An additional experiment also showed that when subjects were

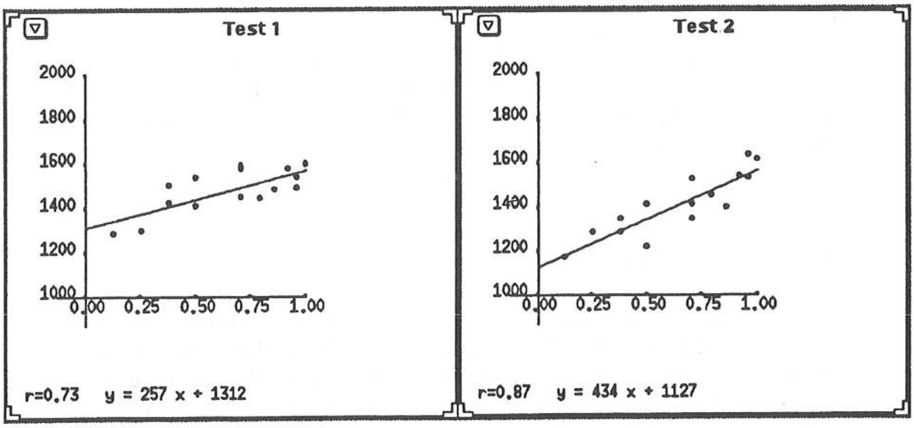


Fig. 2. Response time (msec) as a function of scanning distance (Clockwise condition). In this figure and those which follow, distances are expressed as their ratios to the diameter of the circular island.

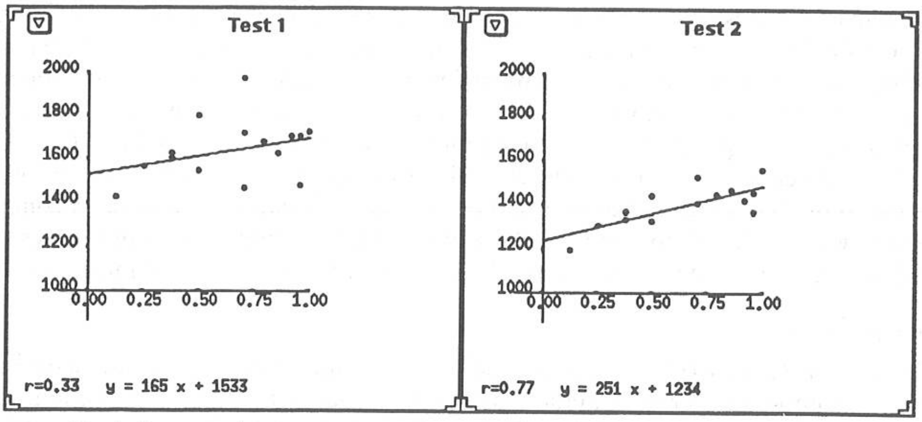


Fig. 3. Response time (msec) as a function of scanning distance (Random condition).

simply asked to indicate the locations of landmarks by placing a point on the outline map of the island, the errors had the same properties as those inferred from the image scanning data, indicating that learning did improve image coherence and resolution [8].

In short, our claim is that there is a need for imagery research to concentrate on image accuracy, a feature that has not yet received much consideration. Mental images, as representations, refer to external entities (either experienced perceptually or learned from descriptions). Therefore, a key property of images is their 'referential validity', in particular, the fact that the intervals between any two points in the representation realistically reflect the intervals between the corresponding points in the source entity. Referential validity, as far as metric information is embodied in the representation, requires that intervals among a set of tokens be represented consistently and accurately. This situation justifies the increased concern of imagery researchers for accuracy as an intrinsic property of visual images, and for the factors (including non-perceptual factors, and individual differences) which hinder or enhance it.

A QUANTITATIVE MODEL OF THE DENIS AND COCUDE (1992) DATA

We have developed a quantitative model and a simulation based on the results reported in Denis and Cocude [8]. The model was designed to account for the gradual process of image elaboration and the progressive increase in image accuracy. There have been repeated claims of the heuristic value of complementing empirical studies of mental imagery with well-grounded computational approaches, but few attempts have been made to substantiate these claims. Kosslyn and Shwartz [28] argued for simulation as a heuristic research tool which constrains theories by forcing scientists to be explicit at every step of the formulation of these theories. They created a computer program to simulate the mental events that occur when a person generates, inspects and transforms mental images of simple objects. Glasgow and Papadias [17] proposed a knowledge representation scheme for computational imagery. The scheme incorporates a long-term memory representation and two working memory representations, corresponding to the distinct visual and spatial components of mental imagery. The three representations, and a set of primitive functions for visual and spatial reasoning, are specified using a formal theory of arrays.

The aim of the present paper is to contribute to this modeling effort. In the first part, the data obtained in the experiments of Denis and Cocude [8] are used as the basis for modeling the processes involved in stepwise image elaboration. In the present context, it was not our purpose to elaborate a formal model with strong explanatory power, but rather to propose a quantitative way to account for and explain a set of data. The second part of the paper describes an experiment which was designed to collect a new set of data expected to provide a more fine-grained understanding of image elaboration. A computer program was developed to simulate the experimental protocol and estimate resulting scanning times. The findings are discussed with respect to the concept of image accuracy.

Image accuracy

This section summarizes our understanding of the processes used to translate verbal information into a visual representation incorporating accurate spatial information. In the experimental situation investigated by Denis and Cocude [8], learning starts with subjects creating an outline shape (a circle). Points (locations) in the image are specified and associated with semantic items (landmarks) in the shape. Each location is probably not initially represented as a sharp point, but rather is associated with a region around this point. Consequently, scanning tends to be performed from some (non-focal) point of a region to some (non-focal) point of another region. Hence, it is unlikely that any distance has a consistent, unique value; more likely, it corresponds to a range of distances. This source of variability is thought to create fuzziness in the representation, manifested in the scanning process by a noisy and reduced time–distance correlation. According to this view, learning essentially consists of performing more extensive processing of the image, narrowing each region associated with landmarks to its exact location. Iterative focusing proceeds up to a terminal state in which landmarks are associated with the precise (focal) points of the outline shape.

During learning, when subjects have just processed a given sentence and then engage in the processing of the next incoming sentence, movement to a proximal part of the configuration (as in the Clockwise condition) places them in a more optimal condition than does movement to a distal part (as in the Random condition). The risks of erroneous or imprecise location are lower, and the landmark being processed may be more rapidly associated with a more restricted region. On the contrary, when subjects have to process landmarks that are distant from each other due to their random order of appearance in the description, separate processing of each region delays the process of focusing individual locations. Therefore, the locations tend to remain imprecise for longer, and more practice is needed to attain precise locations (and hence, maximal referential validity). This account is consistent with repeated demonstrations of less efficient processing of spatial descriptions which do not preserve referential continuity [6, 10].

Our model building started with the assumption that each landmark is not represented in a subject's working memory as a precise point, but rather as a region centered on this point. The region represents the possible range of the landmark's location at a given time of learning. The size of these unfocused (or 'fuzzy') regions is expected to vary inversely with the degree of image elaboration. The closer the image is to its ultimate step of elaboration (with maximal stability and structural coherence), the more restricted these regions are. Considering each distance, we computed the theoretical length of each individual 'mental travel'. This allowed us to draw the region of uncertainty corresponding to each landmark for each subject. The data were averaged across the whole sample of subjects to see if we obtained values which would support the notion of regions of uncertainty whose sizes vary as a function of individual sites, of the amount of learning of the description (that is, first or second scanning test), and of the type of text (clockwise vs random).

Developing a quantitative model

Our first step in developing a quantitative model of the data was to decide on the shape of regions of uncertainty. In accordance with the above-mentioned option of considering these regions centered on a point, we decided to model uncertainty as a circle, which fits with the assumption of uncertainty being evenly distributed in all directions around a focal point. We also assumed that the same circular shape should be used for all six landmarks. The objective was then to generate a graphic representation which would express the size of the region of uncertainty associated with each landmark for each subject in each scanning test. Thus, we looked for a way of measuring the diameter of each circle.

Individual scanning times for every distance separating a pair of landmarks were used to compute these values. The underlying logic was that these times could be used in conjunction with a measure of the subject's individual scanning speed in order to compute the distances theoretically scanned during these times. Each such distance was then compared with the actual distance separating the corresponding pair of landmarks, which made it possible to calculate the amount of error affecting landmark location. It should be noted that the word 'distance', in this context, does not refer to absolute measures of spatially extended entities, but rather to relative measures among a set of intervals. As a reference value, the actual longest distance, 'lighthouse–cave' (in fact, the diameter of the island) was given the value of 1.00. All other distances were expressed as ratios to this reference value (thus, the value is 0.13 for 'creek–hut', 0.26 for 'creek–lighthouse', etc.).

Each individual subject's scanning speed was determined using the subject's performances at the second scanning test as a plausible set of reference measures. These were assumed to be the closest to the terminal state of image construction. Rather than calculating an average value of all 15 times–distance ratios (which would have been far from satisfactory since these ratios themselves vary as a function of distance, i.e. the longer the distance, the faster the scanning speed), we calculated the regression lines between these ratios and actual distances. The speed of an individual subject was then given by the formula:

$$\text{speed} = a \times \text{distance} + b.$$

The average slope of the regression line (a) was 6.35×10^{-3} for the subjects in the Clockwise condition, and 6.75×10^{-3} for the subjects in the Random condition. These values reflect a similar speed increase in both conditions.

Each individual scanning time for a given subject was used to compute a corresponding estimated distance based on these individual speed values. This distance was compared to the actual distance and the corresponding error was calculated. In this procedure, the scanning time for a given distance was an average value for both scanning directions (e.g. 'lighthouse–harbor' and 'harbor–lighthouse'). The amount of error was thus distributed evenly between the two landmarks. Therefore for a given landmarks, five error values were calculated (since there were six landmarks, any given landmark was involved in five distances). The error was expressed in terms of the ratio to the diameter of the island (1.00). Corresponding values were then entered as radii of circles reflecting the region of uncertainty associated with each landmark. The regions of uncertainty for each individual subject were expressed graphically as circles centered on the exact point of each landmark. Averaging values over subjects for each scanning test in each condition resulted in graphic representations for the Clockwise (Fig. 4) and Random conditions (Fig. 5).

Inspection of the data first revealed that the sizes of regions of uncertainty (i.e. the reverse of accuracy) varied widely among landmarks. Overall, landmarks tended to be located most accurately (i.e. by smallest regions of uncertainty) in the part of the map which was richest in landmarks. This result is consistent with the assumption that landmark location is favored when a landmark has close neighbors, with the consequence that several landmarks act as reference points for each other. Isolated landmarks did not benefit from this factor, and they were less accurately located at the outset. A second

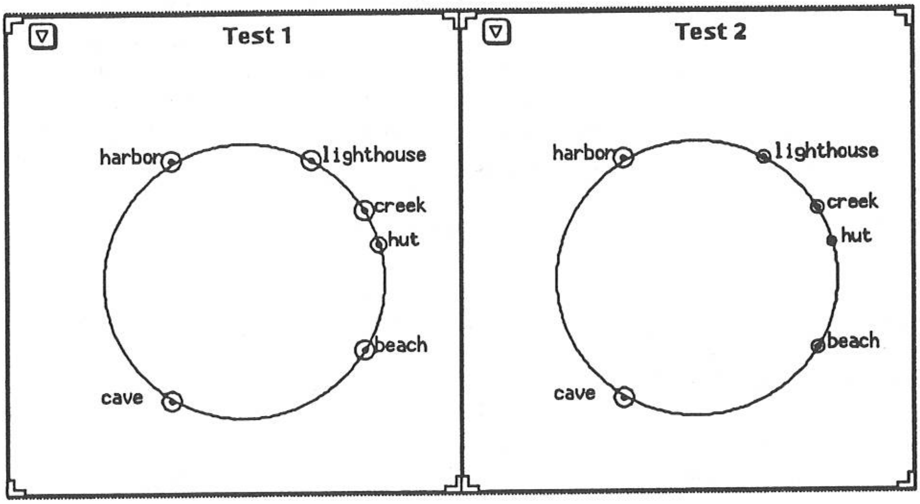


Fig. 4. Graph of regions of uncertainty (Clockwise condition).

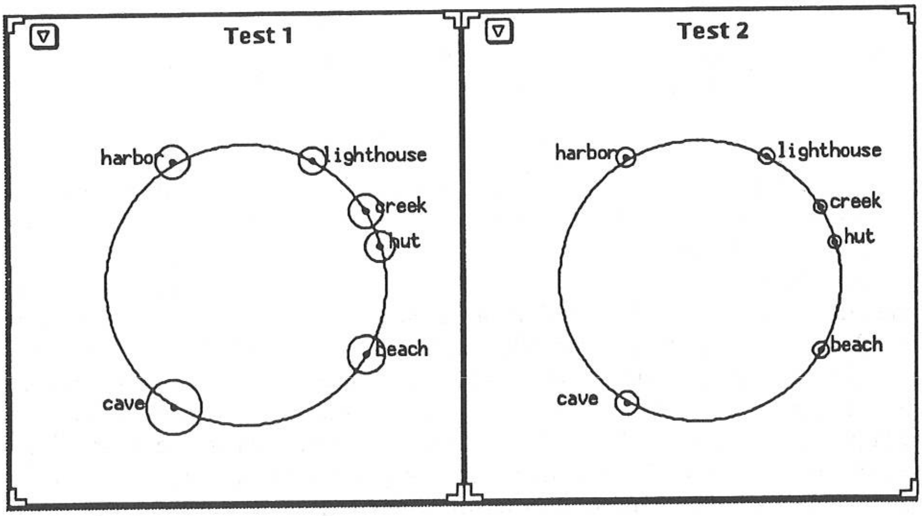


Fig. 5. Graph of regions of uncertainty (Random condition).

feature of interest is that exposure to the description produced a substantial reduction in the size of regions of uncertainty. During the second scanning test, that is, after additional learning, regions of uncertainty were much more restricted than at the first test. Lastly, text structure was also an important factor. At the first scanning test, in the Clockwise condition, the sizes of regions of uncertainty were overall smaller than in the Random condition. This is consistent with our theory that poorly structured text generated a representation wherein regions of uncertainty were larger than in the case of well-structured text. Learning had much more visible effects for the poorly structured text, in that the second test resulted in a dramatic reduction in the regions of uncertainty.

An analysis of variance (ANOVA) was performed on error values (radii of regions of uncertainty). Error values were larger overall on the first test than on the second [F (1, 14) = 8.12, $P < 0.025$]. They significantly differed among landmarks [F (5, 70) = 7.66, $P < 0.001$], and there was a significant interaction between conditions and landmarks [F (5, 70) = 2.75, $P < 0.05$]. Separate analyses were also performed for each condition. In the Clockwise condition, the mean radius of regions of uncertainty was 0.040 at the first test, and 0.032 at the second test. This difference was not significant. In the Random condition, the mean radius decreased from 0.069 to 0.034 between the first and second test. This decrease was significant [F (1, 7) = 5.87, $P < 0.05$], and the difference among landmarks was significant [F (5, 35) = 6.19, $P < 0.001$].

The pattern reflected by these analyses extends the meaning of our previous data. The main feature is the contrast between the moderate (non-significant) improvement in image accuracy for subjects who were exposed to a well-structured text, and the large improvement shown by subjects who processed the random description. All subjects benefited from additional learning following the first scanning test, but those in the Random condition demonstrated the most marked increase in image accuracy after additional text processing.

Testing the model

The implementation reported above essentially consisted in computing subjects' scanning speeds, distances assumed to be scanned, the error affecting landmark location, and radii of regions of uncertainty for each landmark. Note that the map, landmarks, and regions of uncertainty all remained implicit in the implementation. They were represented by numerical values in a sequence of computations, and not by actual data structures in computer memory.

These computations were used to develop a computer program that could simulate the whole experimental protocol and estimate the resulting scanning times. This simulation was executed on every reconstruction of individual mental images, and for each subject in turn. The input data used were the radii of regions of uncertainty for each subject and for the test under consideration. For each scanning interval, we randomly selected a point in the regions of uncertainty of both relevant landmarks. The actual straight line distance between these two points was calculated using the formula:

$$\text{distance}(\text{point}_1, \text{point}_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

where (x_i, y_i) were the Cartesian coordinates of point i. The subject's scanning speed was then used to compute a corresponding scanning time. A specific time (for a given interval and a given subject) was based on the ratio of the corresponding distance (point_1, point_2)

to the subject's scanning speed. All the distances scanned were calculated for each subject and the scanning times for each distance were averaged. Lastly, the correlation between average scanning times and actual distances (on the map) was calculated (as in the Denis and Cocude [8] experiment).

The whole procedure was executed six times. The main variables from run to run were the points selected in the regions of uncertainty. As a result, the scanning times calculated were different from each other, as were the correlation coefficients. Table 1 shows the values of correlation coefficients (r) resulting from the six runs of the program, for each condition and each scanning test. The average values in the Clockwise condition were 0.87 for the first scanning test and 0.91 for the second. Corresponding data in the Denis and Cocude [8] experiment were 0.73 and 0.87, respectively. In the Random condition, the average values resulting from the six runs were 0.29 for the first scanning test and 0.81 for the second. Corresponding experimental values were 0.33 and 0.76, respectively. While the absolute values resulting from the simulations differed slightly from the experimental data, the overall pattern was in fact quite similar: (a) There was a consistent relationship between times and distances; (b) in the first scanning test, the correlation was lower in the Random than in the Clockwise condition; and (c) the correlation values increased from the first to the second scanning test, but more markedly in the Random condition. Overall, these results indicate that our quantitative model correctly reflects the psychological assumptions that motivated its implementation.

A NEW EXPERIMENT AND COMPUTER SIMULATION

While we were primarily interested in how image accuracy increases at different steps of learning, the data we used for implementing the simulation had an obvious limitation. Indeed, only two steps were investigated, one at the mid-point of learning, and the other at a step which corresponded to the terminal learning state. Although these data have the merit of accounting for scanning at some given time before the terminal state, the single view of the scanning process (i.e. first scanning test) can hardly be considered to reflect the

Table 1. Correlation coefficients resulting from six simulations based on the Denis and Cocude (1992) data, for each condition and each scanning test

Simulation No.	First test	Second test
Clockwise condition		
1	0.80	0.92
2	0.90	0.91
3	0.88	0.88
4	0.87	0.88
5	0.86	0.96
6	0.91	0.92
Random condition		
1	0.45	0.81
2	0.47	0.78
3	−0.23	0.83
4	0.18	0.80
5	0.30	0.71
6	0.57	0.94

dynamics of the process in its very early stages. We therefore replicated the whole procedure, paying more attention to the early steps of image elaboration.

A new experiment was designed with a new set of subjects. In the Denis and Cocude [8] experiment, the subjects heard the description six times, and the first scanning test was given after three exposures and the second test after three more exposures. The new experiment used the same number of exposures, but there was a total of three mental scanning tests. The protocol was as follows: The subjects listened to the verbal description twice, then did the first scanning test; they listened to the description twice more, then did the second scanning test; they listened to the description twice more, then did the third (last) scanning test. This procedure provided the possibility of three successive 'views' of the mental representation under construction.

Given the length of the experimental session, we focused on the Random condition, in which the most dramatic effects should in principle occur, especially between the first and second tests. The first test was expected to reveal less image structure than the first test in the original Denis and Cocude [8] study, and we hypothesized that the performance in the second test should reach a value somewhere between the first and second tests of the original study. The most significant increase in image accuracy was expected to occur between the first and second tests, and subsequent change (between the second and third tests) was likely to be smaller. This expectation was based on the fact that (a) the first test would occur earlier than the first test of the previous study (that is, at an early step of image construction where the visual image is still far from attaining structural coherence) and (b) the second test would occur later than the first test of the previous study (that is, closer to the terminal step, when the image should be close to its final, most stable state). In line with our previous findings [7, 8], it was reasonable to predict that greater effects should be obtained between the first and second tests than between the second and third tests, since the period between the first and second tests still consisted of building the spatial structure of the visual image, whereas in the period between the second and third tests, additional processing of the description mainly helped subjects to refine a visual image which was close to its stable state.

Method

Subjects. Fourteen undergraduate volunteers (4 male, 10 female) were recruited.

Materials. The text used was the random version from the Denis and Cocude [8] experiment (see Fig. 1 above). The text described the map of a fictitious, circular island, with six landmarks situated around the periphery ('harbor', 'lighthouse', 'creek', 'hut', 'beach', 'cave'). The French words for these landmarks were all pronounced as one-syllable words.

A tape recording was used to present the scanning test. It contained 60 pairs of words. Each landmark was named 10 times and was followed 4 sec later by a second word. On five of these trials, the second word did not name a landmark on the island. The 'false' items were landmarks that could plausibly have been found on the island ('meadow', 'bridge', 'well', 'mine', 'moor'). In the other five trials, the first word was followed by the name of one of the other five landmarks. Thus, every pair of landmarks occurred twice, alternating the landmark that appeared first. The order of pairs was randomized, with the following three constraints: (a) the same landmark could not occur twice in two successive pairs; (b) a 'true' landmark occurring as the second member of a pair could not occur in the next two pairs; (c) no more than three 'true' or three 'false' trials could occur in a row. The presentation of the second word started a timer. A new trial began 8 sec after the probe word was presented. The test trials were preceded by eight practice trials (four 'true' and four 'false'). The practice trials used names of French cities as 'true' items.

Procedure. At the beginning of the first learning phase, the subjects were told that they would hear a description of the map of an island. They were also told that they would have to create as vivid and accurate a visual image of the map as possible. The text was presented auditorily twice. Following the second presentation, the subjects were required to form a visual image of the map and revise the exact location of each landmark.

At the beginning of the first test phase, the subjects were told that each trial would first consist of hearing the name of a landmark on the island. They were to picture the entire map of the island mentally and then focus on the landmark named. The subjects were told that a few seconds after focusing on the named landmark, they would hear another word. If this word named a landmark present on the map, the subjects were to scan to it and press a button with their dominant hand when they reached it. The scanning was to be accomplished by imagining a black speck zipping along the shortest straight line from the first landmark to the second. The speck was to move as quickly as possible, while still remaining visible. If the second word of a pair did not name a landmark on the map, the subjects were to press the second button with their non-dominant hand. Response times were recorded. The experimenter interviewed the subjects during the practice trials, making sure that they had followed the instructions about imagery use.

Following the first test, the subjects were asked to resume the learning task to enhance their image of the map of the island. The text was presented two more times, with instructions identical to those used in the first learning phase. The second scanning test was then given. It proceeded in the same way as the first test. A final learning phase began, with the text being presented two more times with the same instructions. The experiment finished with the third scanning test, using the same procedure and instructions as for the previous two.

Subjects were tested individually and were interviewed at the end of the experiment. Two subjects who reported having followed the imagery instructions less than 75% of the time during the test phases were excluded and replaced. The subjects were also asked whether they had either relied on the location of the landmark depicted in their visual image or first revised the hour-coded location of the landmark before mentally scanning to the second named landmark. One subject who reported having used this latter procedure was excluded and replaced.

During the test phases, half of the subjects processed the items according to the randomized order defined above, whereas the other half processed the second half of the items and then the first half.

Results and discussion

Only the times for the correct 'true' decisions were analyzed. Times exceeding twice the other time for the same distance in a given scanning test were discarded. The error rate was very low (first test: 3.8%; second test: 0.2%; third test: 0.5%). Errors did not vary systematically with distance scanned.

The subjects' times for the different distances were analyzed first. The ANOVA revealed an overall significant effect of distance on scanning times [$F\ (14, 182) = 2.81$, $P<0.001$]. Scanning times were overall longer on the first than on the second test [$F\ (1, 13) = 12.07$, $P<0.005$]. There was a further decrease of times between the last two tests [$F\ (1, 13)=9.26$, $P<0.01$]. In addition, times were averaged over subjects for each scanning test, and the correlation between times and distances was calculated. The resulting coefficient was not significant for the first scanning test [$r\ (13) = 0.13$], but it reached significance for the second and third tests [$r\ (13)=0.58$, $P<0.025$, and $r\ (13)=0.76$, $P<0.005$, respectively] (Fig. 6).

The results are in agreement with those of the Denis and Cocude [8] study. They also provide a more detailed view of the dynamics of the process. The results at the first scanning test revealed no structure at all in the image under construction. Response times were extremely long. Two exposures were not enough to construct a coherent, accurate image, since mental exploration did not reveal any sign of the metric properties of the imagined map. Two more exposures changed the situation dramatically. Subjects performing mental scanning produced chronometric patterns which at this point of learning revealed that their images had an internal structure where metric information was fairly correctly represented. The times were also significantly shorter, a finding compatible with the assumption that the images had an internal structure which was more readily available. The process was even more marked after two additional exposures. The numerical values at the third test (scanning times as well as correlation coefficients) all indicated further improvement in the internal structure of the image.

Although it is difficult to compare the data with those of the Denis and Cocude [8] study because the numbers of subjects were not the same in both studies, the correlation

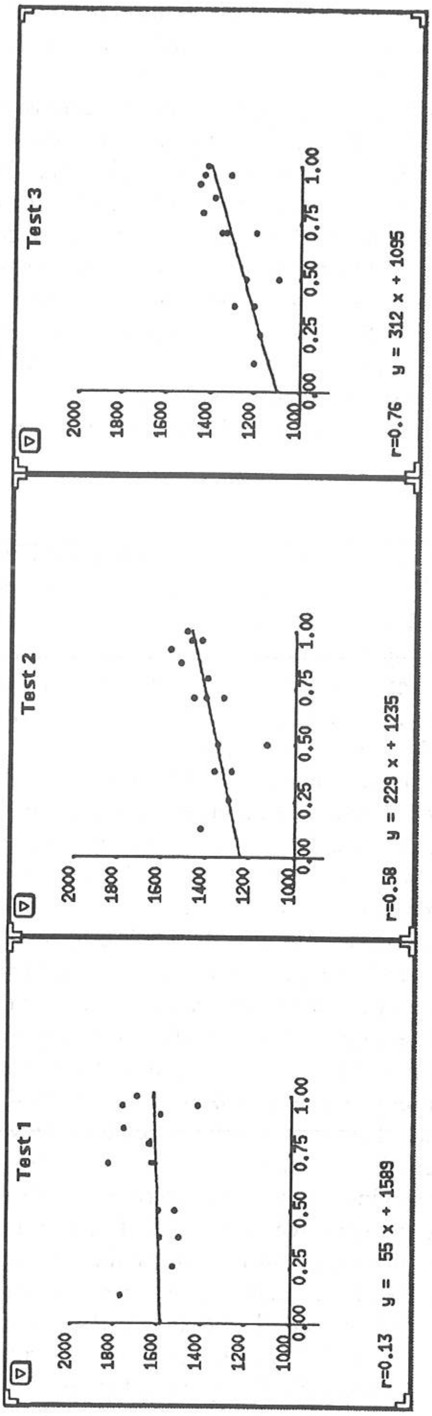


Fig. 6. Response time (msec) as a function of scanning distance (three scanning tests).

obtained for the first scanning test in the former study (0.33) was intermediate between those of the first two scanning tests of the present study (0.13 and 0.58). Furthermore, the correlation coefficients reflected by the last scanning test were of the same order of magnitude in both studies (0.76).

A potential objection may be that the improvement in successive scanning tests was not due to additional learning which occurred between the tests, but to the subjects' experience of the scanning task itself. While the practice of scanning may have had some effect on subsequent performance, this factor is unlikely to account for the observed changes in the pattern of responses. The Denis and Cocude [8] study included a control experiment (Experiment 2) which examined this specific point and did not provide any support for this explanation. The control condition consisted of measuring scanning performance following six presentations of the description without any intervening scanning test during learning. The absence of any difference between this condition and the original one (in terms of scanning times and correlation coefficients) ran counter to the interpretation that the final performance after six trials was due to the imposition of a scanning test after the third study trial.

A further test of the model

The new set of experimental data shows that there was a significant change in the mental representation constructed by subjects between the first and second scanning tests, and that there was a (less marked) additional improvement between the second and third tests. Our quantitative model of image accuracy was applied to this new set of data, to see whether the analysis would reflect a stepwise restriction of regions of uncertainty in the mental image, and whether such restriction would be more evident between the first and second tests than between the second and third.

The procedure described above was run with the new set of data. First, we calculated each subject's scanning speed, taking as reference measures those provided by the last (in this case, third) scanning test. Each individual 'distance' and the corresponding 'error' were computed for each subject, and the resulting values were averaged over subjects to produce a graph of the regions of uncertainty for the three scanning tests (Fig. 7). Inspection of these figures reveals that the sizes of uncertainty regions varied among landmarks. As was pointed out in the previous analysis, locating the landmark which was farther from the others ('cave') produced the greatest error. The other feature of interest is that the size of regions of uncertainty decreased after each new exposure of the subjects to the description. This decrease occurred for all regions, although it was more marked for some of them, and the hierarchy among the six landmarks remained the same throughout the experiment. As expected, there was a greater decrease between the first and second tests than between the second and third. The fact that there was a slight additional restriction of regions of uncertainty between the second and third tests confirms that even at this point of the learning sequence where the verbal description is perfectly memorized [8], there can still be further improvement in image coherence and resolution.

An ANOVA on error values confirmed that the mean radius decreased considerably from test to test (first test: 0.116; second test: 0.065; third test: 0.049). This effect was significant [$F(2, 26)=17.17$, $P<0.001$]. Radii were overall larger on the first than on the second test [$F(1, 13)=16.32$, $P<0.005$]. There was a further decrease in radii between the last two tests [$F(1, 13)=5.09$, $P<0.05$]. The analysis also revealed that the radii differed significantly among landmarks [$F(5, 65)=11.28$, $P<0.001$]. Pairwise comparisons at the

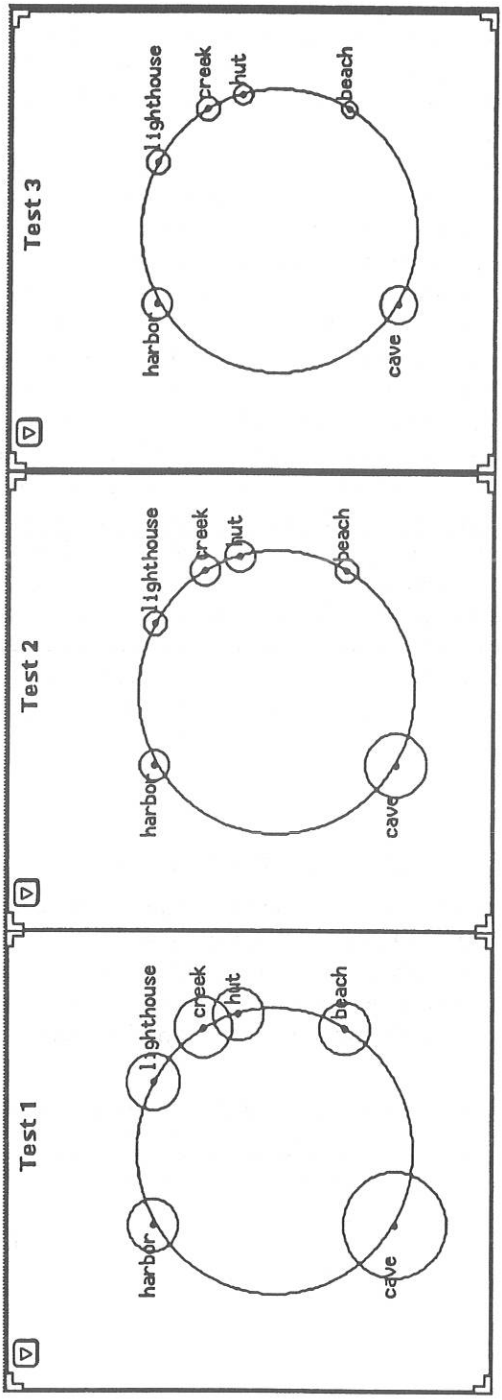


Fig. 7. Graph of regions of uncertainty (three scanning tests).

first and second tests showed that this effect was entirely due to the fact that the radius for the item 'cave' was significantly larger than those of all five other landmarks ($P<0.005$, for the first test; $P<0.05$, for the second test), which did not differ from each other. At the third test, there was no longer any difference between 'cave' and 'harbor' (all remaining differences, $P<0.01$ or less), and 'harbor' and 'lighthouse' differed from 'beach' ($P<0.025$ and $P<0.05$, respectively).

Another feature of interest stems from an examination of individual differences. Figure 8 shows the individual patterns of Subjects 4 and 7. Subject 4 showed a general trend which conformed to the average results, but the resolution of the image was achieved quite early. There was only a moderate improvement between the first and second tests, as was also the case between the second and third tests. In contrast, Subject 7 apparently experienced great difficulty at the outset. There was a considerable error, with large overlapping of regions of uncertainty. The image structure improved significantly between the first and second tests, then performance remained stable at the third test.

The last step of our approach consisted of simulating the whole experiment on the reconstructions of successive mental images of each subject. The procedure was the same as before: random selection of points in the regions of uncertainty for every distance to scan, followed by computation of scanning times based on individual subjects' scanning speed. Ten runs of the computer program were conducted. Table 2 shows the resulting values of r for all three scanning tests. Average values were -0.07, 0.81 and 0.93, for the three successive tests. These figures reflect a more contrasted pattern than those based on behavioral data, but they clearly exhibit the same predicted trend. After an initial stage essentially characterized by the lack of internal structure in the image, substantial improvement resulted from additional learning, with a slight further improvement due to overlearning.

GENERAL DISCUSSION

Image accuracy is one of a set of concepts usually invoked to refer to the visual properties of images. Other such related concepts are those of 'vividness', 'clarity' and 'resolution' [29, 31]. The properties referred to by these labels have been considered mainly in experimental contexts involving self-reports. However, these notions probably have genuine cognitive counterparts that can be tested experimentally.

Table 2. Correlation coefficients resulting from 10 simulations based on the new set of data, for each scanning test

Simulation No.	First test	Second test	Third test
1	−0.09	0.70	0.91
2	−0.03	0.68	0.95
3	0.01	0.75	0.97
4	−0.28	0.81	0.93
5	0.10	0.88	0.96
6	−0.21	0.82	0.89
7	0.11	0.93	0.93
8	0.39	0.92	0.94
9	−0.41	0.79	0.89
10	−0.31	0.87	0.93

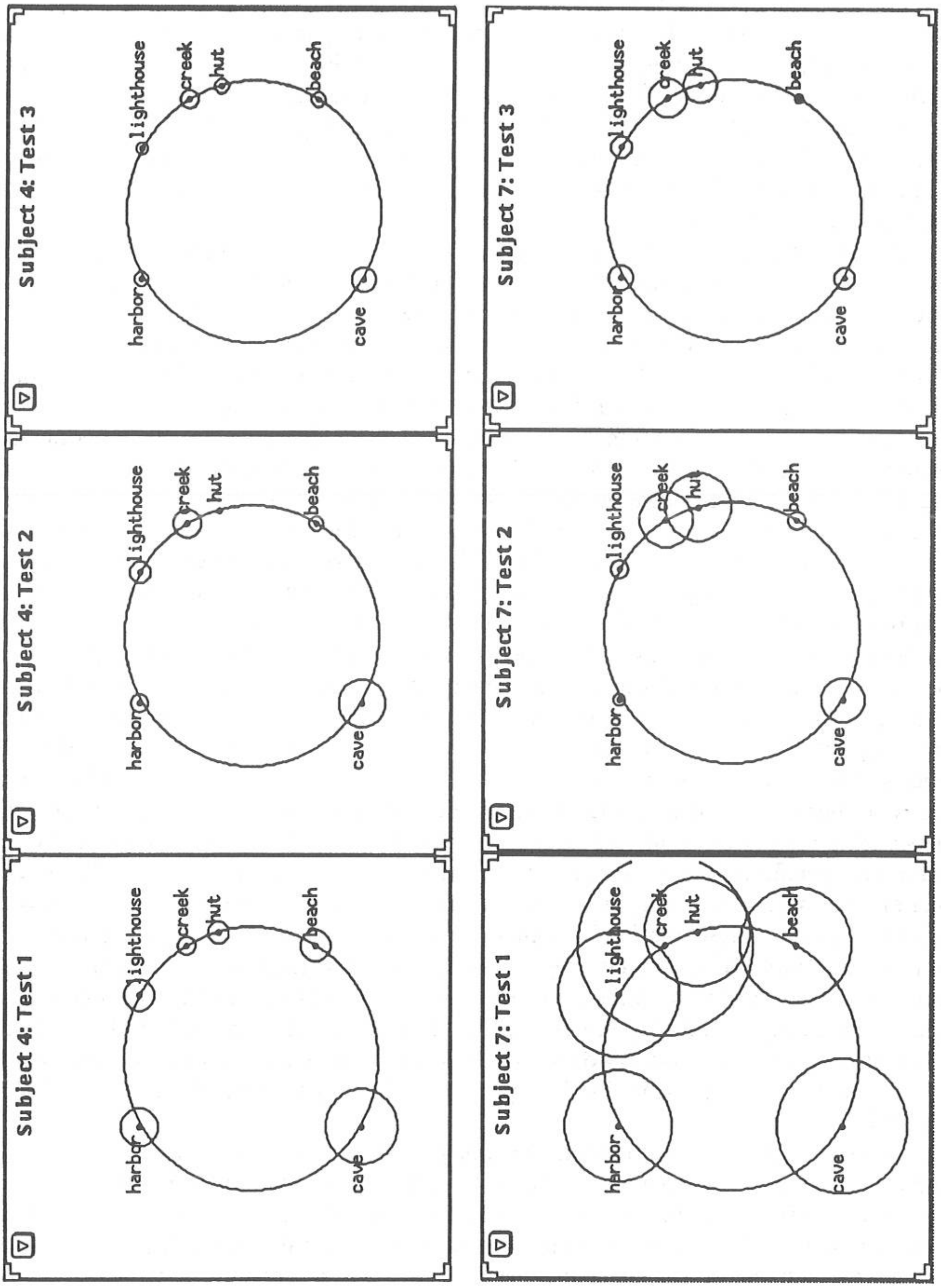


Fig. 8. Graph of regions of uncertainty for two individual subjects.

This family of concepts is in fact present in the best developed contemporary theories of mental imagery [14, 23, 26]. Kosslyn [23] posits that images correspond to patterns of activation in a visual buffer, and that visual resolution for the features of images depends on the field of the visual buffer where activation occurs. In particular, image resolution is lower in the peripheral visual field. Results reported by Finke and Kosslyn [15] and Finke and Kurtzman [16] show that peripheral acuity in visual imagery is limited by the same types of neural constraints that limit peripheral acuity in visual perception. Our interpretation of mental scanning data suggests that accuracy may not have to do only with the periphery of the visual buffer. More generally, it has to do with the lack of constraining information in any region of the visual buffer.

A distinction is usually made between the *visual* characteristics of mental images (such as limits on their resolution) and their *spatial* properties. This distinction is supported by a number of empirical as well as neurological arguments [12, 26]. However, such a distinction should not prevent researchers from considering domains where the two sets of properties might interact. The construction of mental representations of spatial configurations containing distinguished points, either based on perception or on the processing of a verbal description, is an example of these domains. In the experiments discussed above, landmarks contribute to the spatial structure of the constructed mental configuration, but this structure obviously also depends on such visual properties as the resolution and accuracy of landmark locations (the size of corresponding regions). Variations in the 'fuzziness' of landmarks, a visual property, are correlated with spatial structuring of visual images. Distance, and thus location, one of the most important spatial properties [4], could thus be affected by visual characteristics.

Another aspect on which existing theories of imagery obviously need more elaboration is the dynamics of the visual qualities of images. For instance, it would be relevant to consider practice as a factor likely to affect imagery acuity. The research reported above clearly suggests that the internal structure of visual images depends on the amount of practice subjects had during the construction of these images. This approach differs from the early studies on mental scanning, which all focused on the process once the image had been perfectly memorized and was fully available. Although this step was necessary to validate the specific methodology of mental scanning, it is now important to proceed further and investigate how imagery processes develop. Connectionist approaches provide promising avenues of research in the domain of image construction and processing [21]. An important feature of such modeling is that the data structures which were implicit in the implementation described in the present research (i.e. the map, landmarks, and regions of uncertainty) can be explicitly represented in the model. Following the work reported here, we have developed a connectionist model, in which the image is directly represented by a neural net and mental scanning is implemented by propagation of activation within the net [19].

It is probably too extreme to claim that image accuracy is in all cases a property of primary importance. An image needs only to be as precise or accurate as is just necessary for adequate performance in the current task. Different degrees of image accuracy may then correspond to different needs in the context of a given task. These degrees may also be correlated with the degree of involvement of early visual cortical areas. These areas would be more likely activated in tasks requiring attentive inspection of high-resolution visual images, or those placing stringent demands on topographic processes [27, 32]. Another relevant issue for further investigation is whether accuracy of memory for location is correlated with accuracy of memory for shape. Neuropsychology suggests that they may not be. If not, findings on the accuracy of spatial location could not be easily generalized to shape.

Lastly, the concept of accuracy in imagery research allows us to take into account the notion of variations in 'fuzziness' of the different parts of an image. The fact that some portions of an image are more sharply defined than others corresponds to most subjects' intuition and self-report. By modeling such concepts as accuracy or fuzziness in the domain of mental imagery, it becomes easier to account for differential sharpness of the subparts of an image than in classic studies on mental scanning.

Acknowledgements—This research was supported by a grant from the National Center of Scientific Research (Cogniscience Program, National Thematic Project on "Representation of Space"). The authors are grateful to Sophie Cocude for collecting the new experimental data reported in this article and to Jean-François Jodouin for helpful advice and comments during the preparation of the paper. A first draft of the paper was written during the first author's research stay at the Faculty of Psychology of the University of La Laguna, Tenerife, Spain. We thank Manuel de Vega, Jose M. Diaz, Marlene Behrmann, and two anonymous reviewers for their comments and suggestions on an earlier version of the paper.

REFERENCES

1. Basso, A., Bisiach, E. and Luzzatti, C. Loss of mental imagery: A case study. *Neuropsychologia* **18,** 435–442, 1980.
2. Behrmann, M., Winocur, G. and Moscovitch, M. Dissociation between mental imagery and object recognition in a brain-damaged patient. *Nature* **359,** 636–637, 1992.
3. Bisiach, E. and Berti, A. Hemineglect and mental representation. In *Cognitive and Neuropsychological Approaches to Mental Imagery*, M. Denis, J. Engelkamp and J. T. E. Richardson (Editors), pp. 387–391. Martinus Nijhoff, Dordrecht, The Netherlands, 1988.
4. Cave, K. R., Pinker, S., Giorgi, L., Thomas, C. E., Heller, L. M., Wolfe, J. M. and Lin, H. The representation of location in visual images. *Cognit. Psychol.* **26,** 1–32, 1994.
5. Charlot, V., Tzourio, N., Zilbovicius, M., Mazoyer, B. and Denis, M. Different mental imagery abilities result in different regional cerebral blood flow activation patterns during cognitive tasks. *Neuropsychologia* **30,** 565–580, 1992.
6. Denis, M. *Image and Cognition.* Harvester Wheatsheaf, New York, 1991.
7. Denis, M. and Cocude, M. Scanning visual images generated from verbal descriptions. *Eur. J. cognit. Psychol.* **1,** 293–307, 1989.
8. Denis, M. and Cocude, M. Structural properties of visual images constructed from poorly or well-structured verbal descriptions. *Mem. Cognit.* **20,** 497–506, 1992.
9. Denis, M. and Zimmer, H. D. Analog properties of cognitive maps constructed from verbal descriptions. *Psychol. Res.* **54,** 286–298, 1992.
10. Ehrlich, K. and Johnson-Laird, P. N. Spatial descriptions and referential continuity. *J. verb. Learn. verb. Behav.* **21,** 296–306, 1982.
11. Farah, M. J. The neurological basis of mental imagery: A componential analysis. *Cognition* **18,** 245–272, 1984.
12. Farah, M. J. Is visual imagery really visual? Overlooked evidence from neuropsychology. *Psychol. Rev.* **95,** 307–317, 1988.
13. Farah, M. J., Weisberg, L. L., Monheit, M. and Péronnet, F. Brain activity underlying mental imagery: Event-related potentials during mental image generation. *J. cognit. Neurosci.* **1,** 302–316, 1989.
14. Finke, R. A. *Principles of Mental Imagery.* MIT Press, Cambridge, Mass., 1989.
15. Finke, R. A. and Kosslyn, S. M. Mental imagery acuity in the peripheral visual field. *J. exp. Psychol.: Human Percept. Perform.* **6,** 244–264, 1980.
16. Finke, R. A. and Kurtzman, H. S. Mapping the visual field in mental imagery. *J. exp. Psychol.: Gen.* **110,** 501–517, 1981.
17. Glasgow, J. I. and Papadias, D. Computational imagery. *Cognit. Sci.* **16,** 355–394, 1992.
18. Goldenberg, G., Podreka, I., Steiner, M., Willmes, K., Suess, E. and Deecke, L. Regional cerebral blood flow patterns in visual imagery. *Neuropsychologia* **27,** 641–664, 1989.
19. Gonçalves, E., Memmi, D. and Denis, M. A connectionist model of mental imagery. In *Proceedings of the Conference on Neural Networks and their Applications*, pp. 341–350. Marseilles, France, December 1994.

20. Guaraglia, C., Padovani, A., Pantano, P. and Pizzamaglio, L. Unilateral neglect restricted to visual imagery. *Nature* **364,** 235–237, 1993.
21. Intons-Peterson, M. J. Integrating the components of imagery. In *Models of Visuospatial Cognition*, M. de Vega and M. Marschark (Editors). Oxford University Press, New York, in press.
22. Jeannerod, M. The representing brain: Neural correlates of motor intention and imagery. *Behav. Brain Sci.* **17,** 187–202, 1994.
23. Kosslyn, S. M. *Image and Mind.* Harvard University Press, Cambridge, Mass., 1980.
24. 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. Visual mental imagery activates topographically organized visual cortex: PET investigations. *J. cognit. Neurosci.* **5,** 263–287, 1993.
25. Kosslyn, S. M., Ball, T. M. and Reiser, B. J. Visual images preserve metric spatial information: Evidence from studies of image scanning. *J. exp. Psychol.: Hum. Percept. Perform.* **4,** 47–60, 1978.
26. Kosslyn, S. M. and Koenig, O. *Wet Mind: The New Cognitive Neuroscience.* Free Press, New York, 1992.
27. Kosslyn, S. M. and Ochsner, K. N. In search of occipital activation during visual mental imagery. *Trends Neurosci.* **17,** 290–292, 1994.
28. Kosslyn, S. M. and Shwartz, S. P. A simulation of visual imagery. *Cognit. Sci.* **1,** 265–295, 1977.
29. Marks, D. F. Imagery and consciousness: A theoretical review from an individual difference perspective. *J. mental Imagery* **1,** 275–290, 1977.
30. Mellet, E., Tzourio, N., Denis, M. and Mazoyer, B. A PET study of visual and mental spatial exploration. *J. cognit. Neurosci.*, in press.
31. Morris, P. E. and Hampson, P. J. *Imagery and Consciousness.* Academic Press, London, 1983.
32. Moscovitch, M., Behrmann, M. and Winocur, G. Do PETS have long or short ears? Mental imagery and neuroimaging. *Trends Neurosci.* **17,** 292–294, 1994.
33. Pinker, S., Choate, P. A. and Finke, R. A. Mental extrapolation in patterns constructed from memory. *Mem. Cognit.* **12,** 207–218, 1984.
34. Roland, P. E. and Gulyás, B. Visual imagery and visual representation. *Trends Neurosci.* **17,** 281–287, 1994.
35. Sakai, K. and Miyashita, Y. Visual imagery: An interaction between memory retrieval and focal attention. *Trends Neurosci.* **17,** 287–289, 1994.
36. Shepard, R. N. and Cooper, L. A. *Mental Images and Their Transformations.* MIT Press, Cambridge, Mass., 1982.