

Structural properties of visual images constructed from poorly or well-structured verbal descriptions

MICHEL DENIS and MARGUERITE COCUBE
Université de Paris-Sud, Orsay, France

Previous research has found a linear relation between distance and scanning times for spatial configurations when the spatial configurations were learned perceptually and when they were constructed from well-structured verbal descriptions. The current research replicated the time-distance relation when the images were constructed from repetitions of well-structured descriptions but not when the images were generated from three repetitions of a description that presented information in a random order. Six exposures to the randomly ordered information yielded the expected time-distance relation in image scanning. We posited that additional exposure to the poorly structured information allowed the image to develop the structural coherence and resolution needed to support consistent scanning. Thus, the structure of descriptions can affect the intrinsic structure of images of described objects and hence the mental operations performed subsequently on these images. Another experiment indicated that image coherence and resolution improves even after the verbal description is accurately recalled.

People are able to form visual images of unseen objects that are described verbally. This ability should not be surprising because in everyday exchanges between speakers, substantial knowledge is built up from verbal inputs that people translate or recode in a visual form, hence producing cognitive substitutes for the physical entities described. The implementation of these processes has been investigated when subjects learn spatial descriptions and then perform cognitive tasks that rely on recall of the learned configuration. Typically, the subjects' task is to infer object relations not explicitly stated in the description (e.g., Bryant, Tversky, & Franklin, 1992; Denis & Denhière, 1990; De Vega, 1991; Ehrlich & Johnson-Laird, 1982; Foos, 1980; Franklin, 1992; Franklin & Tversky, 1990; Mani & Johnson-Laird, 1982; Perrig & Kintsch, 1985; Taylor & Tversky, 1992).

A number of empirical arguments converge on the hypothesis that mental imagery plays a prime role in the elaboration of mental representations from spatial descriptions. Images are a key component of mental models that describe spatially extended entities (see Johnson-Laird, 1983; see also Denis & Denhière, 1990; De Vega, 1991). In line with this view, Denis and Cocude (1989; see also Denis, 1991) found that the times to scan between loca-

tions on a spatial representation constructed from verbal descriptions increased with the distance between the two locations (landmarks) in a manner similar to that reported by others, such as Kosslyn, Ball, and Reiser (1978). It has been suggested that these findings might be due to another process—namely, that subjects' knowledge might interfere with the implementation of the scanning process (cf. Goldston, Hinrichs, & Richman, 1985; Intons-Peterson & McDaniel, 1991). This seems unlikely, because mental scanning has been shown to resist empirical efforts to reduce it to a pure consequence of tacit knowledge effects or task demands (see Denis & Carfantan, 1985; Finke & Pinker, 1982, 1983; Jolicoeur & Kosslyn, 1985; Pinker, Choate, & Finke, 1984; Reed, Hock, & Lockhead, 1983).

Denis and Cocude's (1989) results suggest that similar linear time-distance relations occur in mental scanning of representations formed from verbal descriptions and from those formed perceptually. Before drawing firm conclusions, it is important to note that scanning performance showed some differences between verbally based representations (in a text condition) and perceptually based ones (in a map condition). In the text condition, subjects listened to a description of the locations of landmarks specified in the conventional hour-dial terms of flight navigation. For one group of subjects, the text was presented three times; for another group of subjects, it was presented six times. In the map condition, subjects memorized the map of an island that was circular in shape and contained several geographical details located on its periphery.

Subjects who heard the description three times exhibited a lower time-distance correlation coefficient than did those who processed the text six times. In addition, after three learning trials, the absolute scanning times tended to be longer than those in the map condition, but this differ-

This research was supported by DRET Grant 89-242. Katty Corpinot and Philippe Breuil efficiently contributed to data collection, and the assistance of Connie Greenbaum was greatly appreciated in the preparation of the manuscript. The authors wish to thank Manuel De Vega, Johannes Engelkamp, Nancy Franklin, Margaret Jean Intons-Peterson, Mark McDaniel, Barbara Tversky, and Hubert Zimmer for their insightful comments on the research reported in this paper. Requests for reprints should be addressed to Michel Denis, Centre d'Etudes de Psychologie Cognitive, Université de Paris-Sud, Centre d'Orsay, Bâtiment 335, 91405 Orsay Cedex, France.

ence disappeared after six learning trials. Although the verbal information was recalled accurately after three trials, the spatial qualities of the visual representation itself at this time might have been imperfectly consolidated, resulting in a lower time-distance correlation.

These results suggest that images progress toward a stable, high-resolution state, the equivalent of the state attained more rapidly by images derived from perception. Our conjecture was that whatever the input (either perceptual or verbal) from which an image is built up, the final state of this image is similar in nature, and that in both cases the resulting mental representation should be considered to be genuinely visual. The additional assumption was that translating a verbal input into a visual representation requires time and that the representation goes through intermediate states of lower resolution.

Our claims from the previous research were limited because different subjects were involved in the two text conditions. Thus, individual differences, rather than differences in the stage of image coherence, might have produced the results. Accordingly, the purpose of the present experiments was to explore the effects of amount of text learning in a condition in which the same subjects performed the mental scanning task twice—once after three learning trials, then once again after three additional learning trials.

The experiments also explored the effects of discourse structure on scanning performance. In our view, spatial descriptions have their own structural characteristics that affect the processes involved in the construction of mental representations. The order in which information is entered in descriptions affects the on-line construction of representations and hence their availability for retrieval processes. In addition, indeterminateness and referential discontinuities have been shown to hinder the elaboration of visual mental models of spatial configurations (see Denis & Denhière, 1990; Ehrlich & Johnson-Laird, 1982; Mani & Johnson-Laird, 1982).

In the past, we used a well-structured text. Geographical details were presented in a highly predictable, consistent sequence—namely, in clockwise order. Given the above observations, however, it is of theoretical interest to determine whether or not subjects would be able to construct effective and scannable representations from poorly structured text. Accordingly, we constructed a version of the description that was designed to make it difficult for subjects to incorporate details in the outline structure of the island. In this version, sentences were presented in a random sequence. Such a sequence should create more demanding conditions during the formation of an image (although not basically impeding its elaboration) and presumably would take longer to reach the end state of a stable, well-defined image than would well-structured descriptions. If this is so, this would provide a forceful argument that the spatial qualities of images depend on their elaboration conditions.

EXPERIMENT 1

Subjects listened to a short description and were told to generate a visual image of a spatial configuration containing six landmarks. After listening to the description three times, the subjects mentally scanned the distances separating all the pairs of landmarks. Instructions adhered closely to those typical of this paradigm (see Denis & Cocude, 1989; Kosslyn, Ball, & Reiser, 1978). Then, the subjects listened again to the same description of the configuration three more times and subsequently performed the mental scanning task a second time.

Each subject processed one of two versions of the description, with landmarks presented in either a clockwise or a random sequence. The two resulting conditions were thus variations of a between-subject factor.

Method

Subjects. Sixteen undergraduates from the Orsay campus were recruited as subjects. Each subject was randomly assigned to one of the two conditions.

Materials. A text was written that described a fictitious, circular island. It had six features situated around the periphery (harbor, lighthouse, creek, hut, beach, cave), at locations that could be unambiguously defined in the conventional hour-dial terms of flight navigation. The six features were located in such a way that all the distances between pairs of adjacent features were different from each other. The French words for these features were all pronounced as one-syllable words.

In one version of the text, features were introduced in the description in clockwise order, starting with the harbor. The description read as follows (original in French):

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.

The other version of the text resulted from randomly ordering the sentences describing locations of the features. The description also started with the harbor and read as follows:

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.

A tape recording was constructed to present the scanning test. It contained 60 pairs of words. Each feature was named 10 times and was followed 4 sec later by a second word. On 5 of these trials, the second word did not name a feature on the island. The "false" objects were features that could plausibly have been found on the island (meadow, bridge, well, mine, moor; again, the French words for these features were pronounced as one-syllable words). On the other 5 trials, the first word was followed by the name of one of the other five features. Thus, every pair of features occurred twice, alternating the feature that appeared first. Order of pairs was randomized, with the constraints that the same feature could not occur twice in 2 successive pairs, that a "true" feature occurring as the second member of a pair could not occur in the next 2 pairs, and that no more than 3 "true" or 3 "false" trials could occur in a row. Presentation of the second word started a clock. A new trial began 8 sec after the probe word was presented. The test trials were preceded by 8 practice trials (4 "true" and 4 "false"). The practice trials used names of French cities as "true" items. The

whole procedure was driven by a computer program adapted to this experiment.

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 told that they would have to create as vivid and accurate a visual image of the map as possible. The text was presented auditorily three times. Following the second and third text presentations, the subjects were required to form a visual image of the map and to check the exact location of each feature.

At the beginning of the first test phase, the subjects were told that each trial would first consist of hearing the name of a feature on the island. They were to picture the entire map of the island mentally and then were to focus on the feature named. The subjects were told that a few seconds after focusing on the named feature, they would hear another word. If this word named a feature depicted on the map, the subjects were to scan to it and depress 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 feature 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 feature on the map, the subjects were to depress the second button with their non-dominant hand. Response times were recorded. The experimenter interviewed the subjects during the practice trials to make sure that they had followed the instructions about imagery use.

Following the first test, the subjects were invited to resume the learning task to enhance their image of the map of the island. The text was presented three more times, with instructions identical to those used in the first learning phase. Next, the second scanning test was presented. It proceeded in the same way as the first test.

All of the subjects were tested individually. At the end of the experiment, the subjects were interviewed. Two subjects who reported having followed the imagery instructions less than 75% of the time during the test phases were excluded and replaced. In addition, the subjects were asked whether before mentally scanning to the second named feature they had either relied on the location of the feature depicted in their visual image or first revised the hour-coded location of the feature. None of the subjects stated having used this latter procedure.

In each condition (clockwise and random) during the test phase, 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

Only times for correct "true" decisions were analyzed. Times exceeding twice the other time for the same distance were discarded. The error rate was very low (0.8% and 0.6% of the trials in the clockwise and the random conditions, respectively), and errors did not vary systematically with distance scanned.

The analysis of variance (ANOVA) revealed an overall significant effect of distance on scanning times [$F(14,196) = 6.54, p < .001$], with times increasing linearly with increasing distance [$F(1,14) = 26.56, p < .001$]. Scanning times were longer overall on the first than on the second test [$F(1,14) = 13.56, p < .005$]. The overall difference between clockwise and random conditions was not statistically significant [$F(1,14) < 1$]. The marginally significant interaction between scanning tests and text conditions [$F(1,14) = 4.15, p = .065$] occurred because the time decrease from the first to the second test was more marked in the random than in the clockwise

condition. Additional analyses revealed that the intercepts significantly decreased from the first to the second test [$F(1,14) = 20.66, p < .001$] and that the regression line slopes tended to become steeper for the second test [$F(1,14) = 4.39, p < .06$].

Separate analyses were performed for each condition. In the clockwise condition, the analysis showed a consistent increase of scanning times as a function of distance [$F(14,98) = 4.53, p < .001$], with a strong linear component [$F(1,7) = 22.41, p < .005$]. Times were averaged over subjects for each scanning test, and the correlation between times and distances was calculated. The coefficients obtained were $r(13) = .73, p < .01$, for the first scanning test, and $r(13) = .87, p < .01$, for the second scanning test (see Figure 1). There was no significant decrease in times from the first to the second scanning test.

A different pattern of results emerged from the analysis of data in the random condition, although response times increased reliably with distance [$F(14,98) = 2.94, p < .001$], and the linear component was significant [$F(1,7) = 6.64, p < .05$]. The correlation between times (averaged over subjects) and distances was not significant for the first scanning test [$r(13) = .33$] but was significant for the second test [$r(13) = .76, p < .01$] (see Figure 2). Furthermore, there was an overall significant decrease of times from the first to the second scanning test [$F(1,7) = 10.39, p < .025$].

Discussion

Experiment 1 was planned to show a time-distance relationship for mental scanning of an image constructed from a verbal description. It was also designed to compare two learning conditions, which differed in the difficulty of constructing a coherent visual image of the configuration.

The subjects who processed the description designed to place minimal requirements on their processing capacities produced a typical mental scanning time-distance relationship on the first scanning test. These data replicate previously reported findings (Denis & Cocude, 1989) and support the claim that the mechanisms governing mental scanning are valid for representations derived from both perceptual experience and text processing. The findings indicate that even with minimal learning, the scanning process develops in a systematic fashion on the representation built up by the subjects. The results also support the hypothesis that images, even when derived from a verbal description, incorporate metric information.

The same pattern appeared after additional learning, with a slight increase in the time-distance correlation. This finding suggests that although the representation processed during the first scanning test sustained consistent scanning performance, some (moderate) improvement of image coherence occurred with additional text processing. With added study, the scanning process was more consistent, indicating that metric information became more accurately represented in the image. Furthermore, the slight decrease

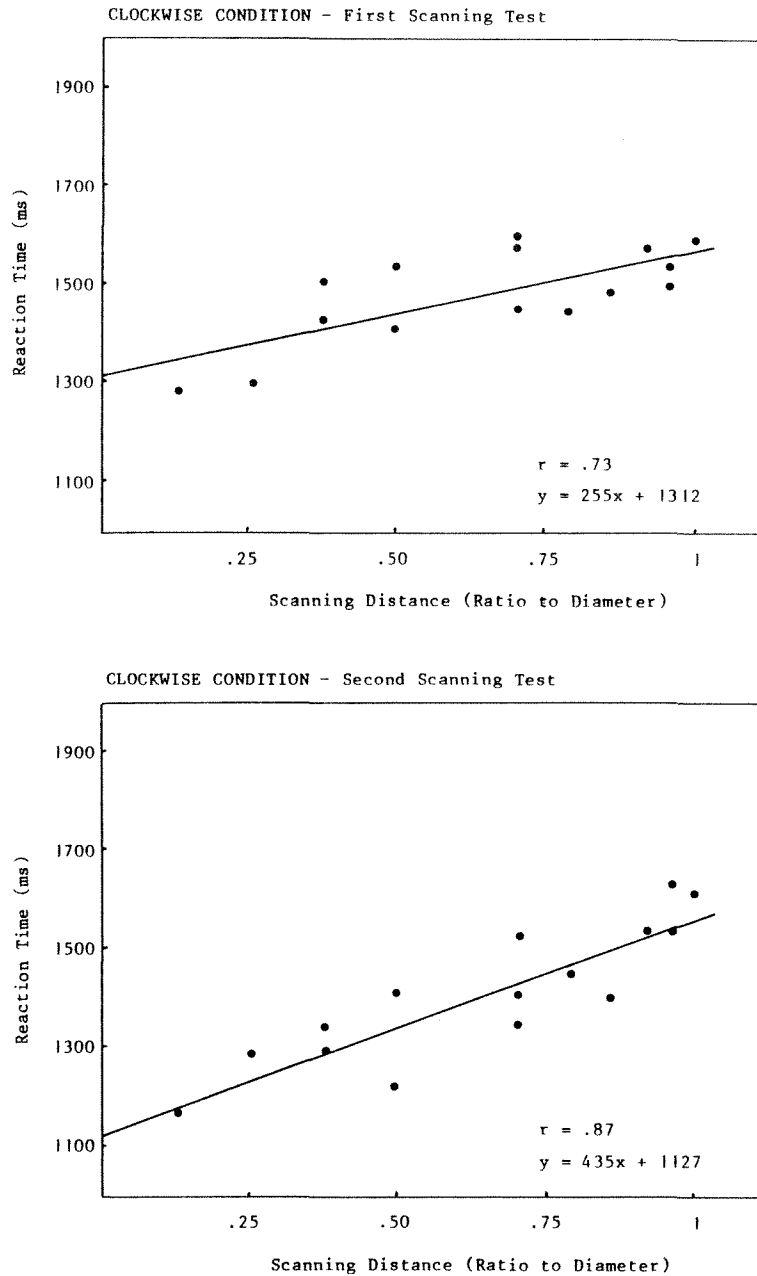


Figure 1. Experiment 1: Reaction time as a function of scanning distance (clockwise condition). Distances are expressed as their ratios to the diameter of the circular island.

of scanning times was consistent with the hypothesis that the representation achieved greater resolution after additional text processing.

Nevertheless, added processing did not increase learning of the verbal description as such. Previous studies have shown that no more than three trials were necessary for subjects to recall accurately the locations of the six features on a blank map (see Denis & Cocude, 1989). Thus,

the improvement that takes place during the second learning phase should be interpreted not with respect to an increase in learning of the wording of the description (which is maximal as early as the first test) but rather as an enhancement of the visual representation built up from the description. The opportunity for subjects to further process their internal representation of the configuration provides the image with higher resolution and more

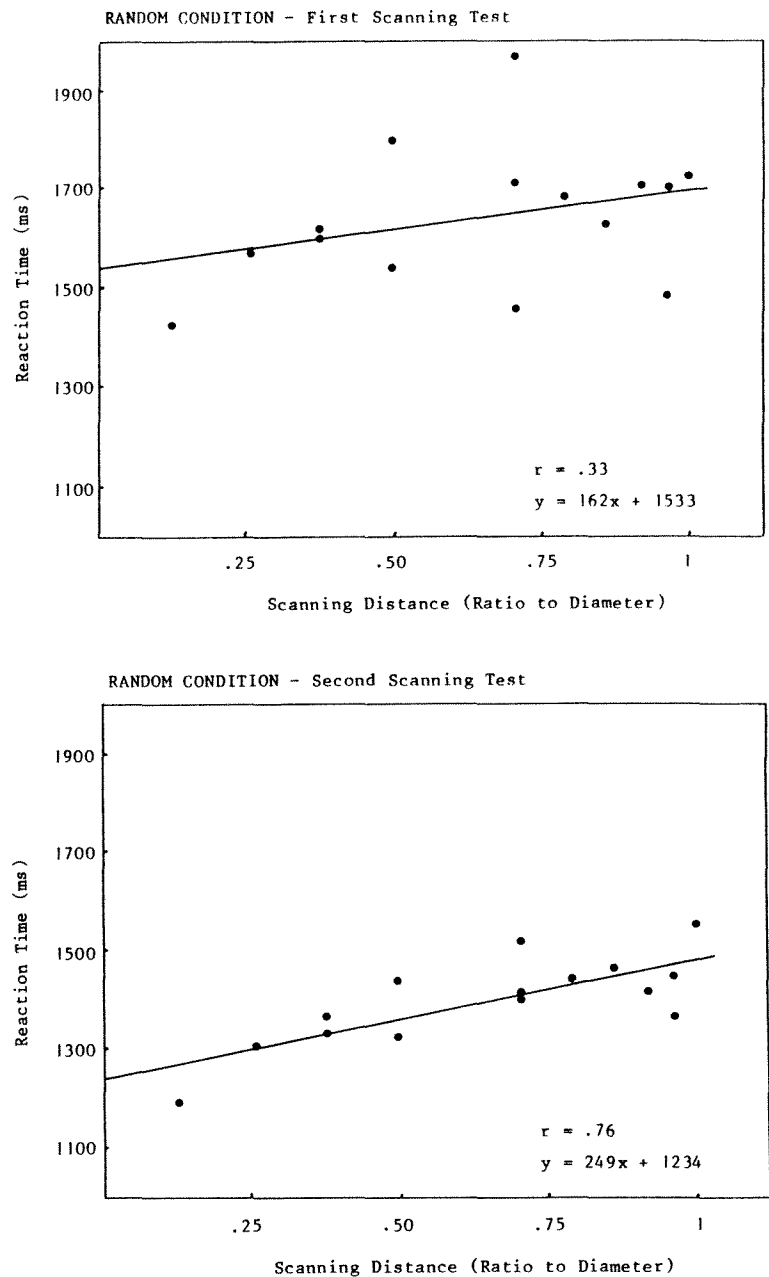


Figure 2. Experiment 1: Reaction time as a function of scanning distance (random condition). Distances are expressed as their ratios to the diameter of the circular island.

fine-grained localization of its constituent parts. As a consequence, mental scanning effects are more clearly evidenced.

The situation differs strikingly for the subjects who processed the random text. We expected this version to retard the construction of a coherent image, especially in the very early steps of learning. In other words, we expected that an image constructed from random text would have poor structural coherence and that this poor structural coherence would be reflected in a low time-distance

correlation. In fact, scanning times in this condition were only weakly related to distance on the first test. The positive correlation became significant after additional processing of the text. Lastly, a marked decrease of the dispersion of points was noticeable between the first and the second scanning test.

Thus, our data support the claim that images generated from descriptions can exhibit genuine metric properties, which are evidenced through the chronometric pattern of mental scanning processes. In addition, the structure of

descriptions affects the elaboration of images. Subjects are capable of compensating for the effects of poor text structure, but this obviously requires additional processing, in contrast to texts that place low requirements on subjects' processing capacities.

EXPERIMENT 2

Experiment 1 showed that at least when the subjects had to process a poorly structured text, the chronometric measures reflecting the structural coherence of the resulting image were significantly affected by additional learning. One question, however, remained unanswered at this point. Was the improvement on the second scanning test due to the additional learning trials following the first test? Or was this improvement an artifact of the subjects' exposure to the first scanning task? Perhaps practicing mental scanning during the first test suffices to facilitate performance on the second occurrence of the scanning task.

To clarify this point, in Experiment 2, two new groups of subjects heard either clockwise or random versions of the description six times and then performed the scanning task. If improvement after the second scanning test in Experiment 1 was due to the presence of the first scanning test, then the time-distance correlation should be lower in Experiment 2 than it was in Experiment 1. In contrast, if improvement during the second scanning test in Experiment 2 was due to added learning, then the time-distance correlation should be about the same after six trials for both experiments.

Method

Subjects. Sixteen undergraduates from the Orsay campus who had not participated in the previous experiment were recruited as

subjects in Experiment 2. Each subject was randomly assigned to one of the two conditions.

Materials. The materials used were identical to those used in Experiment 1.

Procedure. The procedure adhered closely to the procedure used in Experiment 1, except that the subjects listened to the description of the island six successive times before performing the mental scanning task. Three subjects who reported having followed the imagery instructions less than 75% of the time during the test phase were excluded and replaced. In addition, the subjects were asked whether before mentally scanning to the second named feature they had either relied on the location of the feature depicted in their visual image or first revised the hour-coded location of the feature. One subject who stated having used this latter procedure was excluded and replaced.

Results

The overall error rate was very low (3.3% and 1.7% of the trials in the clockwise and the random conditions, respectively), and errors did not vary systematically with distance scanned.

Data from Experiment 2 were submitted to an ANOVA along with the data from the second scanning test in Experiment 1. Factors for this analysis were experiments (1 vs. 2, i.e., with or without intervening scanning test) and texts (clockwise vs. random).

The analysis revealed a significant effect of distance on scanning times [$F(14,392) = 9.70, p < .001$], and times increased linearly with increasing distance [$F(1,28) = 41.94, p < .001$]. No significant difference was found between Experiments 1 and 2, and the overall difference between clockwise and random conditions did not reach significance. Additional analyses did not reveal any significant differences in the intercepts or in the slopes between Experiments 1 and 2.

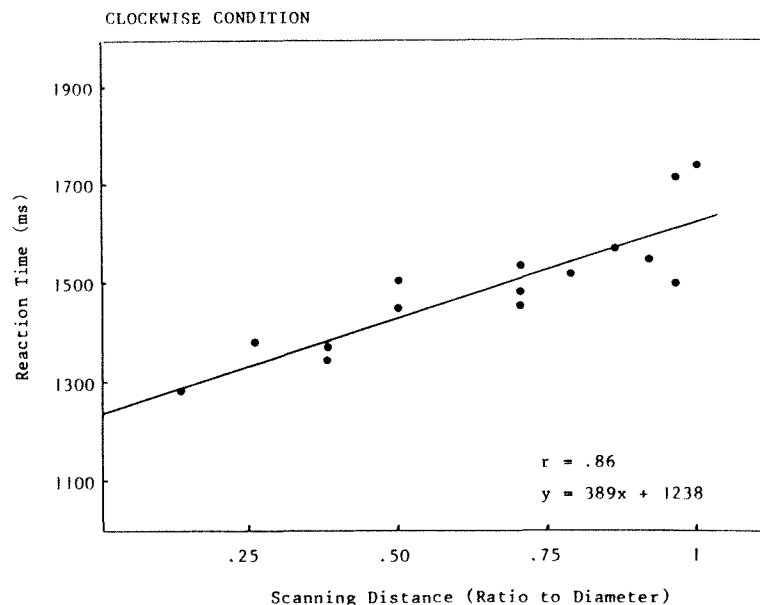


Figure 3. Experiment 2: Reaction time as a function of scanning distance (clockwise condition). Distances are expressed as their ratios to the diameter of the circular island.

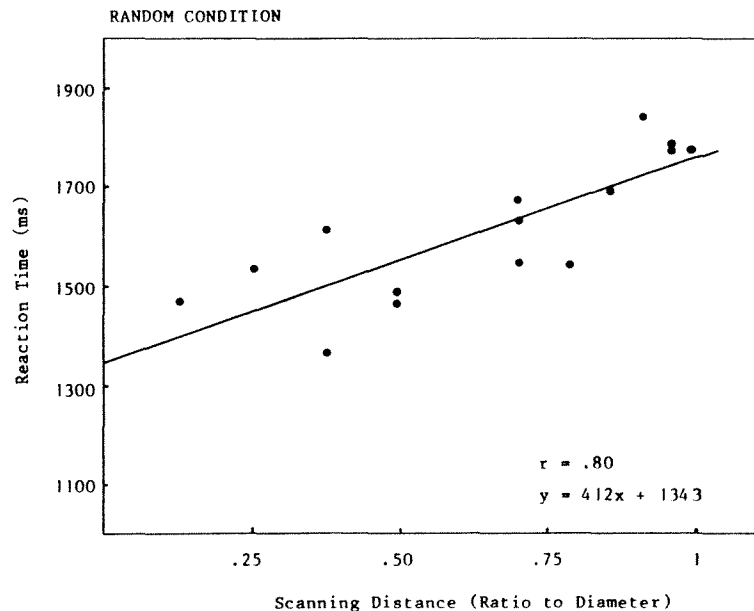


Figure 4. Experiment 2: Reaction time as a function of scanning distance (random condition). Distances are expressed as their ratios to the diameter of the circular island.

Scanning times were averaged over subjects in each condition of Experiment 2. In the clockwise condition, the correlation between times and distances [$r(13) = .86$, $p < .01$; see Figure 3] was similar to the corresponding value in Experiment 1 [$r(13) = .87$]. In the random condition, the correlation [$r(13) = .80$, $p < .01$; see Figure 4] did not differ significantly from the corresponding value in Experiment 1 [$r(13) = .76$].

Discussion

In the absence of any intervening scanning test during learning, scanning performance following six presentations of the description did not differ (and, more specifically, did not decrease) from a condition with an intervening test. This outcome runs counter to the interpretation that the final performance in Experiment 1 after six trials was due to the imposition of a scanning test after the third study trial.

In comparison with the two groups of subjects in Experiment 1, the subjects who did not practice mental scanning in the course of learning had slightly (nonsignificantly) longer scanning times, both in the clockwise and in the random conditions. Their times, in fact, were more similar to the times on the first scanning test of Experiment 1. This measure may reflect either a slight effect of the absence of scanning practice in Experiment 2 or a tendency of the whole group of subjects to be slower than their counterparts in Experiment 1 (although subjects from both experiments were drawn from the same population). More important for the purpose of our experiments, the scanning performance of subjects in Experiments 1 and 2 was quite similar.

EXPERIMENT 3

In Experiment 1, the improvement in time-distance relation from the first to the second scanning test was interpreted as an increase in image resolution and coherence as learning proceeds. Improvement was noticeable in the condition in which the description was presented in the most adequate sequencing (clockwise condition). This pattern was more clear-cut (and statistically significant) in the condition in which text processing was more demanding (random condition). These conclusions were drawn from the results of Experiment 1, with Experiment 2 showing in addition that improvement was not due to an intervening scanning test. We assume that reprocessing of the visual image as learning proceeds is mainly responsible for the increase of image resolution.

A potential objection to this claim is that subjects in the random condition did not learn as much as did subjects in the clockwise condition. In other words, the lower time-distance correlation for the random group might reflect lesser knowledge of the map because of lesser learning of the text describing it.

As mentioned above, previous studies using the same material have shown that text learning is complete within a maximum of three trials (see Denis & Cocude, 1989). However, this argument is based on data collected out of the context of the present series of experiments. Therefore, to shed light on image evolution as text information accumulates, we need information on retention of the text and the spatial properties assumed to be encoded into the image, and these data should be obtained from subjects tested in the same temporal conditions as those used in

Experiments 1 and 2. For this purpose, Experiment 3 required subjects to reproduce the text after three and after six learning trials. In addition, information about image resolution was obtained by asking the subjects to "externalize" their visual image by marking the locations of the six landmarks on an outline map of the island.

This two fold procedure was expected, first, to indicate whether, as observed in previous experiments, text learning can be considered to be complete after three exposures or whether additional learning of text is still required, at least in the group of subjects who learn the random version of the text. Second, if image resolution improves with practice, the accuracy of landmark locations should increase with additional learning, most noticeably for the subjects in the random condition.

Method

Subjects. Twenty-four different undergraduates from the Orsay campus were recruited as subjects.

Materials. The same texts were used. In addition, an outline shape of the island (i.e., a circular shape with a diameter of 130 mm) was printed on sheets of A4 paper.

Procedure. The procedure for the first learning phase was identical to the one in Experiment 1. After three auditory presentations of the description, the subjects first wrote their recall of the description. Then, they were asked to locate each landmark with a pencil mark on the outline map of the island, basing their responses on the visual inspection of their image of the island.

Following this test phase, the subjects received three more trials, with the same instructions as before. Then, the subjects had the second test phase (written recall, followed by location of landmarks on the outline map).

The subjects were tested individually. Each subject was randomly assigned either to the clockwise or to the random condition.

Results

Text recall. At the first test phase, the subjects in the clockwise condition had virtually perfect recall. For the whole group of subjects, only 1 subject produced one inversion of two landmarks, 1 subject omitted one landmark, and 1 subject mislocated one landmark. Thus, the overall rate of correct recall was 96% for this group of subjects. In the random condition, recall was almost equally high. One subject produced one inversion, 2 subjects produced one mislocation, and 1 subject produced two mislocations. The corresponding rate of correct recall was 93%. These figures were considered to be high enough in both groups to preclude any formal statistical comparison.

After three additional learning trials, recall was perfect for all subjects (100%) in both conditions.

Location of landmarks. The distances separating subjects' location and the actual location of each landmark on the map were measured (in mm). These measures of absolute error were entered in an ANOVA, with conditions as a between-subject factor and test phases and individual landmarks as within-subject factors.

Overall error rates tended to be smaller in the clockwise than in the random condition [3.07 vs. 4.38 mm, respectively; $F(1,22) = 3.22, p < .09$] and on the second than on the first test [3.30 vs. 4.15 mm, respectively;

$F(1,22) = 6.77, p < .025$]. Furthermore, a significant interaction emerged between conditions and test phases [$F(1,22) = 5.51, p < .05$], resulting from the fact that the error rate was approximately the same at the two test phases in the clockwise condition [3.11 vs. 3.03 mm; $F(1,11) < 1$], whereas it decreased considerably from the first to the second test in the random condition [5.19 vs. 3.57 mm; $F(1,11) = 8.24, p < .025$].

At the first test phase, there was much less intersubject variability in landmark location in the clockwise than in the random condition. Standard deviations of error measures for individual locations ranged from 1.68 to 2.70 in the clockwise condition and from 3.78 to 7.27 in the random condition. The differences persisted, although they were less marked, at the second test phase, in which standard deviations ranged from 1.65 to 3.13 in the clockwise condition and from 2.27 to 3.68 in the random condition.

Discussion

Data from Experiment 3 reflected a discrepancy between the two measures of amount of learning in the two conditions. Text recall was almost perfect as early as the first test phase, that is, after three learning trials, and no evidence for any strong impairment was observed in the random condition. The near-ceiling recall of the text confirms both the previous data and our speculation that the low time-distance relation evidenced in Experiment 1 at the first test phase in the random condition could not be explained in terms of insufficient memory of the wording of the text.

The pattern of results was quite different when location of landmarks was used as an index of learning. This index was designed to reflect the structural coherence of the image constructed from the text (distinct from purely verbal memory). This index revealed that locations were less accurately marked for the random than for the clockwise conditions at the first test phase (and only at this test phase).

These findings suggest that the initial coherence and resolution of images were lower when subjects processed the random rather than the clockwise description. This effect disappeared with additional learning. The improvement was particularly striking for the random condition, in which subjects probably initially worked from a locationally less accurate image. This pattern of results strengthens the argument that additional learning improves image coherence and resolution.

Finally, Experiment 3 supports the assumption that image-resolution enhancement can continue even when memory for the verbal description of the object has been, or has almost been, completed.

GENERAL DISCUSSION

The research reported in this paper replicated the main outcome of recent extensions of the mental scanning paradigm to images generated from verbal descriptions: The longer the distance between two points of an imaged spa-

tial configuration, the longer the scanning time for this distance. Furthermore, when subjects have an opportunity for additional learning, the time-distance correlation increases, reflecting enhanced structural coherence of the mental representation being processed.

The most novel aspect of the research is the effect of the structure of the description on the metric qualities of the resulting image. When the description was asystematic, the subjects were able to recall the descriptions, and they showed a standard time-distance type of scanning, but the scanning relation was significantly delayed. In contrast, systematic or well-structured descriptions rapidly yielded accurate portrayals of the descriptions and locations, as well as more pronounced time-distance scanning relations. We interpreted the results as indicating that poorly structured descriptions require additional exposure to achieve coherence similar to that produced by well-structured descriptions. Additional learning is necessary for subjects to achieve a consistent representation that incorporates valid metric information. The referential validity of images, that is, their capacity to reflect accurately the objects they refer to, is obviously not an all-or-nothing property, but rather results from stepwise elaboration. Experiment 3 shows that this process continues even after memorization of the verbal description of objects has been completed.

We conceptualize the processes used to translate verbal information into a visual representation incorporating accurate spatial information in the following way. Learning starts with subjects creating an outline shape. Some points are specified and associated with semantic items (landmarks) in the shape. Associating items with metric expressions of their locations is probably not the most demanding component of the process (at least in the case of short descriptions such as those investigated here). More substantial resources are probably needed to create and maintain in working memory a fine-grained representation of the spatial extent of the object and of individual item locations in such a way that every interitem distance is accurate relative to all the other distances.

At the outset, each item (location) probably is not represented in the form of a point, but rather is associated with a region around this point. Consequently, scanning tends to be performed from some (nonfocal) point of a region to some (nonfocal) point of another region. Hence, it is unlikely that any distance has a consistent, unique value; rather, it is more likely that it corresponds to a range of distances. This creates fuzziness in the representation, manifested in scanning by a noisy and reduced time-distance correlation. Learning, then, essentially consists in performing deeper processing of the image, that is, 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 localization 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 almost totally independent of each other, separate processing of each region delays the process of the focusing of individual locations. Therefore, fuzziness of locations tends to persist longer, and more practice is needed on items to attain precise locations (and hence, maximal referential validity).

In conclusion, well-structured descriptive materials clearly facilitate the emergence of patterns of chronometric measures reflecting the structural coherence of visual images. The capacity of images to adequately reflect the objects they represent is thus dependent on the processing load created by the intrinsic structure of the description. It is not enough for a description to be exhaustive or fully determinate to create conditions for the generation of a valid image. The structure of the description itself affects the intrinsic structure of the image of the described object and hence the mental operations performed on this image.

REFERENCES

- BRYANT, D. J., TVERSKY, B., & FRANKLIN, N. (1992). Internal and external spatial frameworks for representing described scenes. *Journal of Memory & Language*, *31*, 74-98.
- DENIS, M. (1991). *Image and cognition*. New York: Harvester Wheatsheaf.
- DENIS, M., & CARFANTAN, M. (1985). People's knowledge about images. *Cognition*, *20*, 49-60.
- DENIS, M., & COUDE, M. (1989). Scanning visual images generated from verbal descriptions. *European Journal of Cognitive Psychology*, *1*, 293-307.
- DENIS, M., & DENHIÈRE, G. (1990). Comprehension and recall of spatial descriptions. *European Bulletin of Cognitive Psychology*, *10*, 115-143.
- DE VEGA, M. (1991, June). *Change of character and change of perspective in narratives describing spatial environments*. Paper presented at the European Workshop on Mental Models, Paris.
- EHRlich, K., & JOHNSON-LAIRD, P. N. (1982). Spatial descriptions and referential continuity. *Journal of Verbal Learning & Verbal Behavior*, *21*, 296-306.
- FINKE, R. A., & PINKER, S. (1982). Spontaneous imagery scanning in mental extrapolation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *8*, 142-147.
- FINKE, R. A., & PINKER, S. (1983). Directional scanning of remembered visual patterns. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *9*, 398-410.
- FOOS, P. W. (1980). Constructing cognitive maps from sentences. *Journal of Experimental Psychology: Human Learning & Memory*, *6*, 25-38.
- FRANKLIN, N. (1992). *Representation of spatial information in described routes: Distance, turns, and objects*. Unpublished manuscript, State University of New York, Stony Brook.
- FRANKLIN, N., & TVERSKY, B. (1990). Searching imagined environments. *Journal of Experimental Psychology: General*, *119*, 63-76.
- GOLDSTON, D. B., HINRICH, J. V., & RICHMAN, C. L. (1985). Subjects' expectations, individual variability, and the scanning of mental images. *Memory & Cognition*, *13*, 365-370.
- INTONS-PETERSON, M. J., & MCDANIEL, M. A. (1991). Symmetries and asymmetries between imagery and perception. In C. Cornoldi &

- M. A. McDaniel (Eds.), *Imagery and cognition* (pp. 47-76). New York: Springer-Verlag.
- JOHNSON-LAIRD, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge: Cambridge University Press.
- JOLICOEUR, P., & KOSSLYN, S. M. (1985). Is time to scan visual images due to demand characteristics? *Memory & Cognition*, **13**, 320-332.
- KOSSLYN, S. M., BALL, T. M., & REISER, B. J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception & Performance*, **4**, 47-60.
- MANI, K., & JOHNSON-LAIRD, P. N. (1982). The mental representation of spatial descriptions. *Memory & Cognition*, **10**, 181-187.
- PERRIG, W., & KINTSCH, W. (1985). Propositional and situational representations of text. *Journal of Memory & Language*, **24**, 503-518.
- PINKER, S., CHOATE, P. A., & FINKE, R. A. (1984). Mental extrapolation in patterns constructed from memory. *Memory & Cognition*, **12**, 207-218.
- REED, S. K., HOCK, H. S., & LOCKHEAD, G. R. (1983). Tacit knowledge and the effect of pattern configuration on mental scanning. *Memory & Cognition*, **11**, 137-143.
- TAYLOR, H. A., & TVERSKY, B. (1992). Spatial mental models derived from survey and route descriptions. *Journal of Memory & Language*, **31**, 261-292.

(Manuscript received August 5, 1991;
revision accepted for publication June 4, 1992.)