IMAGINATION, COGNITION AND PERSONALITY, Vol. 23(2 & 3) 163-171, 0203-2004

MENTAL SCANNING OF IMAGES CONSTRUCTED FROM VISUAL EXPERIENCE OR VERBAL DESCRIPTIONS: THE IMPACT OF SURVEY VERSUS ROUTE PERSPECTIVE*

VANESSA CHABANNE, PH.D.

PATRICK PÉRUCH, PH.D.

EMI-U INSERM 9926, Marseille, France

MICHEL DENIS, PH.D.
Groupe Cognition Humaine, LIMSI-CNRS, Orsay, France

CATHERINE THINUS-BLANC, PH.D.
EMI-U INSERM 9926, Marseille, France

ABSTRACT

Mental scanning was used to assess the metric properties of mental spatial representations derived from visual experience or the processing of a verbal description, and either from survey or route perspective. Participants were asked to mentally scan their images of a spatial environment they had learned in one of the following four conditions: Visual-Survey, Visual-Route, Verbal-Survey, and Verbal-Route. No difference was found between the scanning times of the visual and verbal conditions, but scanning times were shorter after survey than route acquisition, and they consistently increased as a function of the Euclidean distances between locations in the environment, with steeper slopes in the route conditions. These results demonstrate that mental spatial representations derived from different sources and perspectives are endowed with similar properties and preserve the Euclidean characteristics of the original environment, but that they are easier to access when they have been constructed from a survey perspective.

*This research was supported by the Programme Cognitique of the French Ministère délégué à la Recherche et aux nouvelles Technologies (Grant COG 141b).

INTRODUCTION

In the domain of spatial cognition, many studies have been devoted to the conditions of acquisition, storage and use of mental spatial representations (e.g., [1]). The relevance of these representations depends on their isomorphism with the corresponding spatial environments. For instance, the metric features of spatial representations may have important consequences on orienting behaviours that involve estimating and comparing distances. Whereas topological relationships are relatively preserved [2], the metric features have been found to be uncertain in these representations [3-5], regardless of the methods that were used to evaluate the accuracy of mental representations.

However, another body of data issued from the field of mental imagery strongly suggests that spatial representations hold precise metric properties [6, 7]. Mental scanning is the paradigm that has contributed the most to support claims about the structural properties of visual images. It is well established that when a person scans across the visual image of a previously learned configuration, scanning times are linearly related to the distances scanned. This finding demonstrates that the visual image is organized in such a way as to reflect the spatial structure of the configuration, in particular with respect to the relative distances between features. The mental scanning paradigm was also used as a means of assessing the properties of images constructed from verbal descriptions. With an adaptation of materials from Kosslyn et al. [7], Denis and Cocude [8] found the same scanning regularities following the verbal description of a configuration than after its visual inspection, suggesting that in both situations, information was structured in a similar way.

Besides the modality of acquisition (visual or verbal), the perspective in which a configuration is memorized may also induce differences in the nature of the resulting representations. For instance, the visual processing of a configuration seen from a survey perspective is quite different from the sequential integration of information collected from a route perspective. Whether spatial knowledge resulting from either route or survey perspectives has distinct properties has been questioned in several studies which have provided discrepant results [9-11]. The present study considered visual and verbal acquisitions under either route or survey perspective and addressed two main issues. First, are the spatial properties of representations constructed from route information similar to those derived from survey experience? Poorer performance is generally expected after route acquisition since, in this case, constructing a mental image requires on-line transformation of three-dimensional (egocentric) spatial information into a two-dimensional (allocentric) mental representation. In the present experiment, the metric properties of the environment should be more clearly reflected in representations constructed from survey than from route perspective. Second, are the properties of mental images resulting from visual experience similar to those derived from listening to corresponding verbal descriptions? Given the primary role of vision in the processing of spatial information [12], visual experience should help construct representations that are easier to process during a subsequent task than a verbal description. However, the capacity of people to use verbal descriptions to build representations containing veridical metric information is well documented and opens the possibility for visual and verbal information to be expressed in functionally similar representations, at least for configurations of low complexity [8, 13].

METHOD

Participants

Forty graduate students (20 women and 20 men), aged between 19 and 31 years, with normal or corrected-to-normal vision, were randomly assigned to one of the four conditions described below, with the same number of men and women within each group.

Materials

The experiment comprised a learning phase and a testing phase. Four learning conditions were constructed, resulting from the combination of the modality of acquisition (Visual or Verbal) and the perspective of presentation (Survey or Route). Specific materials were created for each condition.

Learning

For the Visual-Survey condition, a map of a circular garden was constructed (15 cm in diameter), with six geographical details represented as pictures (flowers, tree, bench, hedge, sandpit, well) located at the periphery (see Denis and Cocude [8]). The six details were located in such a way that all the distances between pairs of adjacent details differed from each other. For the Visual-Route condition, a video presentation of the trip along the path surrounding the garden was displayed on a computer screen at a normal walking pace. The complete trip around the garden started from (and terminated at) the flowers (in clockwise direction) and lasted 45 seconds. For the Verbal-Survey condition, a text was written describing the shape of the garden and the location of the six details based on hour codes used in aerial navigation in the same order followed by the tour in the Visual-Route condition (e.g., "The flowers are at 11 o'clock. The tree is at 1 o'clock, etc."; as in Denis and Cocude [8]). For the Verbal-Route condition, a verbal description of the circular travel was constructed in a similar fashion as in the video presentation. The details were named when the observer arrived at each one of them, so that the time intervals between the naming of successive details was proportional to the distance between the details in the video of the Visual-Route condition, and the duration of travel was the same.

Testing

For the tests, a tape-recording session was used, containing 60 pairs of words. In each pair, a detail was named and was followed four seconds later by a second word. On half of these trials ("false"), the second word referred to the same detail ("bench-bench"), whereas on the other half ("true"), the second word was the name of one of the other details. Every pair of details (15 different distances, ranging from 1.4 to 10 units, where the diameter of the circular garden was assigned the value of 10) occurred twice, alternating the details that were named first ("bench-flower"/"flower-bench"). The order of the pairs was randomized, with the constraints that the same detail could not occur twice in two successive pairs and that no more than three "true" or three "false" trials could occur successively.

Procedure

The participants received instructions and practice trials before the main experimental manipulation.

Learning

In the Visual-Survey condition, the participants were given orally the name of the six geographical details before studying the map for one minute. In the Visual-Route condition, the participants were shown the video of the travel around the garden; the video was presented three times. In the Verbal-Survey condition, the participants were given a description of the map of the garden; the text was read aloud by the experimenter three times. In the Verbal-Route condition, the participants listened to an oral description of the travel around the garden; the description was presented three times via listening to a tape. Preliminary tests had revealed that, except in the Visual-Survey condition, three learning trials were necessary. In each learning condition, the participants were invited to create a visual image as vivid and accurate as possible.

Testing

Following the learning phase, all participants completed the same mental scanning task. They were told that each trial would first consist of hearing from a tape the name of one detail. They would have to form a visual image of the garden map and then focus on the detail named. Second, they would hear the name of another detail. If this word named a different detail, they would have to mentally scan across their visual image to this second location. The participants were required to indicate when scanning had been completed by pressing a button which stopped a timer triggered by the onset of the second detail named. If the second detail named was the same, the participants had to press the button immediately.

RESULTS

Scanning times were first treated in a mixed analysis of variance with Gender (Men versus Women), Modality (Visual versus Verbal) and Perspective (Survey versus Route) as between-participant factors, and Distance ($n = 15$) as within-participant factor. The analysis revealed a significant effect of Perspective, $F(1, 116) = 64.48$, $p < 10^{-4}$, with shorter scanning times for Survey (mean: 7013 ms, *SD:* 366 ms) than for Route condition (mean: 8201 ms, *SD:* 928 ms), and for Distance, $F(11, 72) = 8.74$, $p < 10^{-4}$, with times increasing with increasing distances. No effect was found for Gender, nor for Modality (Visual: mean = 7511 ms, $SD = 759$ ms; Verbal: mean = 7703 ms, $SD = 535$ ms), and there was no significant interaction.

Response times were averaged over participants, and correlations (Pearson r, $df = 13$) between times and distances were calculated for each learning condition (see Figure 1). For all groups, a positive correlation was observed between scanning times and distances (Visual-Survey: .79, $p < .01$; Visual-Route: .71, $p < .01$; Verbal-Survey: .61, $p < .05$; Verbal-Route: .72, $p < .01$). The distributions of correlations were normalized using Fisher's r-to-z transformations and their confidence intervals were compared. These correlations did not differ from each other, although it should be noted that the highest value was found in the condition where the image had been built from the visual acquisition of a survey map of the configuration, which was expected to be the situation where the construction of the visual image is the easiest.

The comparison of the regression equations revealed that slopes were clearly different between the conditions [14]. The overall pattern was that slopes were steeper in the scanning responses following route acquisition than following survey acquisition, reflecting more processing time for every additional step in mental scanning. The differences were significant for Visual-Survey versus Visual-Route, $t(26) = 16.58$, $p < .001$) and Verbal-Survey versus Verbal-Route, $t(26) = 5.39$, $p < .001$). The comparison of the visual and verbal conditions did not reveal any significant difference when acquisition had been made under the survey perspective, $t(26) = 1.19$. There was a significant difference between Visual-Route and Verbal-Route, $t(26) = 2.61$, $p < .02$.

DISCUSSION

The mental scanning paradigm has been successfully used in the domain of mental imagery, as a tool for assessing the analogous features of visual images of previously seen configurations or their verbal descriptions [15]. In the present study, the paradigm was applied to cases more directly relevant to spatial cognition, with the aim of determining whether mental representations constructed from visual experience or from the description of a travel through a spatial environment had the same analogous metric properties. The first finding was that scanning times were positively correlated with the length of the mentally

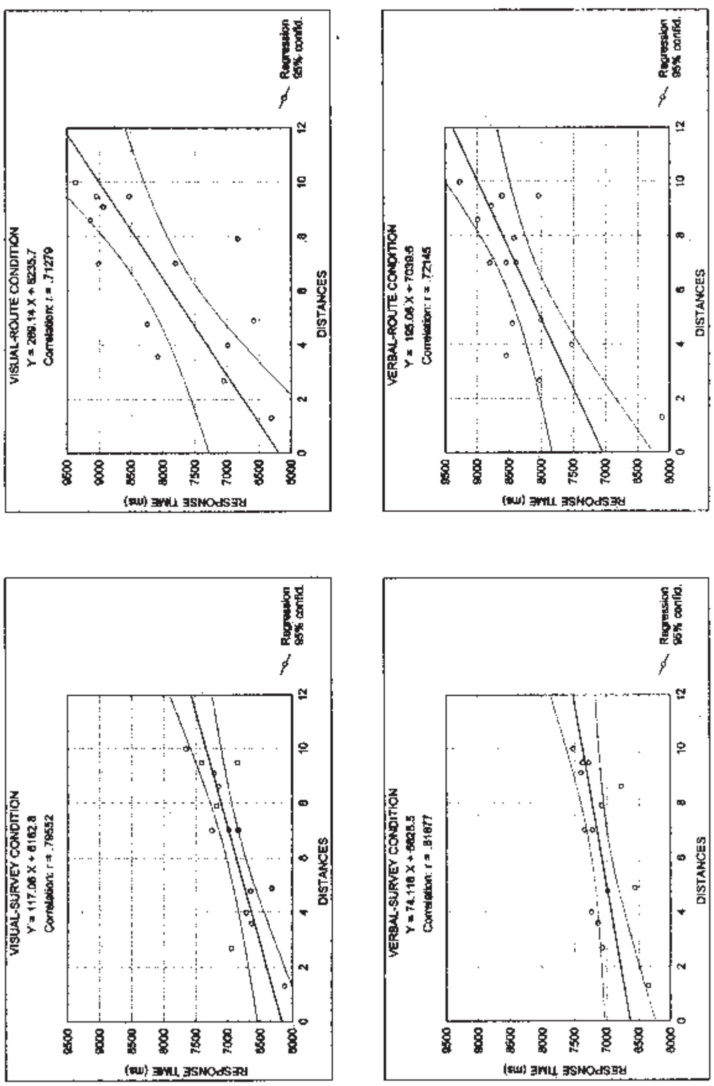

Figure 1. Scanning times as a function of scanning distances for the Visual-Survey, Visual-Route, Verbal-Survey, and Verbal-Route conditions.

travelled distances. The results in the survey conditions are in agreement with those of previous studies [8]. However, in the present experiment, scanning times were shorter when the configuration was presented from a survey perspective than when the participants had visually travelled around it or heard the description of the travel.

The comparison of regression equations showed differences of slopes among conditions. Slopes were steeper in the route than in the survey conditions, indicating that scanning times increased with distances significantly more in the former than in the latter conditions. While slopes did not differ from each other in the two survey conditions, the slope for the Visual-Route group was found to be steeper than the slope of the Verbal-Route group.

This pattern of data partly confirms our expectation that learning an environment from a survey perspective would result in more accurate mental images than learning it from a route perspective. In regard to the learning modality, our data support the finding that images constructed from verbal descriptions may exhibit similar metric properties than those reconstructing a previous visual experience. Indeed, the metric accuracy of representations is reflected by the positive correlations between scanning times and the length of scanned distances. These correlations are significant in all four conditions, and they do not differ from each other. This finding strongly suggests that the representations constructed by the participants of all four groups held metric properties that are characteristic of similar representations.

However, a detailed analysis of scanning times and slopes suggests that learning conditions had, to some extent, an effect on the participants' responses. In particular, one may wish to account for the reasons for which longer response times were recorded in the route conditions. A first explanation may be offered in terms of differential cognitive load. The construction of a representation while travelling imposes greater cognitive load than memorizing a survey map, this resulting from the fact that the formats of the original input and of the final representation are structurally different from each other. When learning is made from a route perspective, individual pieces of information are processed sequentially and have to be maintained in short-term memory in order for the learner to construct a global representation. A perspective shift is required when spatial information is organized, and consequently a change of reference frame takes place. During navigation, information is organized with respect to the body (egocentric) referent, and then has to be re-organized along a bird's eye view perspective. The cognitive load is still higher when the travel is described verbally, since the durations of intervals between locations have to be translated into spatial information, in addition to the shift of reference frame and perspective. As a result, not only are response times longer after the acquisition of spatial information from a route perspective, but every additional unit of distance that is scanned mentally calls for longer processing than is the case after acquisition from a survey perspective. This explanation may be correct if one assumes that on each test trial, the participants reconstruct their representations from the information provided during training. But the participants were required to create visual images during learning itself. Thus, the cognitive load hypothesis is certainly relevant to the construction of representations, not to the tests.

An alternate explanation may be formulated in terms of scaling. A survey or a route perspective during acquisition may lead to differences in the internal scaling of the spatial representation of the layout, with route representations being scaled up and thus more extended than survey representations, which are more likely to be confined and well delimited. Such differences in scaling would provide a sound explanation for the differences of scanning times per distance unit, that is, slope differences.

Lastly, it cannot be excluded that the representations constructed from route perspective keep some sequential features of the initial learning experience. Then, instead of scanning systematically along straight lines connecting two places, participants may follow the circular path, which would account for their longer scanning times. This possibility could be tested by recording eye movements during the tests.

ACKNOWLEDGMENTS

The authors thank Marie-Pascale Nesa and Anne Bret-Fontaine who participated in the data collection, Mark May and Francesca Pazzaglia for their helpful comments on a previous version of the manuscript, and Sven Joubert for making editing suggestions.

REFERENCES

1. R. G. Golledge, Environmental Cognition, in *Handbook of Environmental Psychology*, D. Stockols and I. Altman (eds.), Wiley, New York, 1987.
2. K. Lynch, *The Image of the City*, MIT Press, Cambridge, Massachusetts, 1960.
3. M. D. Giraudo and J. Pailhous, Distortions and Fluctuations in Topographic Memory, *Memory & Cognition*, *22*, pp. 14-26, 1994.
4. T. P. McNamara, Mental Representations of Spatial Relations, *Cognitive Psychology, 18*, pp. 87-121, 1986.
5. B. Tversky and D. J. Schiano, Perceptual and Conceptual Factors in Distortions in Memory for Graphs and Maps, *Journal of Experimental Psychology: General, 118*, pp. 387-398, 1989.
6. R. A. Finke and S. Pinker, Spontaneous Imagery Scanning in Mental Extrapolation, *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*, pp. 142-147, 1982.
7. S. M. Kosslyn, T. M., Ball, and B. J. Reiser, Visual Images Preserve Metric Spatial Information: Evidence from Studies of Image Scanning, *Journal of Experimental Psychology: Human Perception and Performance, 4*, pp. 47-60, 1978.
8. M. Denis and M. Cocude, Scanning Visual Images Generated from verbal Descriptions, *European Journal of Cognitive Psychology, 1*, pp. 293-307, 1989.

9. W. Perrig and W. Kintsch, Propositional and Situational Representations of Text, *Journal of Memory and Language, 24*, pp. 503-518, 1985.
10. H. A. Taylor and B. Tversky, Spatial Mental Models Derived from Survey and Route Descriptions, *Journal of Memory and Language, 31*, pp. 261-292, 1992.
11. P. W. Thorndyke and B. Hayes-Roth, Differences in Spatial Knowledge Acquired from Maps and Navigation, *Cognitive Psychology, 14*, pp. 560-589, 1982.
12. C. Thinus-Blanc and F. Gaunet, Representation of Space in Blind Persons: Vision as a Spatial Sense? *Psychological Bulletin, 121*, pp. 20-42, 1997.
13. M. Denis, M. R. Gonçalves, and D. Memmi, Mental Scanning of Visual Images Generated from Verbal Descriptions: Towards a Model of Image Accuracy, *Neuro-psychologia, 33*, pp. 1511-1530, 1995.
14. K. W. Smillie, *An Introduction to Regression and Correlation*, The Ryerson Press, Toronto, 1966.
15. M. Denis and S. M. Kosslyn, Scanning Visual Mental Images: A Window on the Mind, *Current Psychology of Cognition, 18*, pp. 409-465, 1999.

Direct reprint requests to:

Patrick Péruch
Laboratoire de Neurophysiologie et Neuropsychologie (EMI-U INSERM 9926)
Université de la Méditerranée, Faculté de Médecine Timone
27, Boulevard Jean-Moulin, 13385 Marseille Cedex 5
France
e-mail: Patrick.Peruch@medecine.univ-mrs.fr