

Visual imagery: an interaction between memory retrieval and focal attention

Visual experience can be reproduced mentally by the function of visual imagery. We proposed that visual imagery is implemented by the same neural mechanism that subserves memory retrieval in vision¹. This proposal agrees with the result reported by Roland and Gulyás², that occipito-temporal visual association areas are involved in visual imagery. The main controversial point is their claim that visual imagery, with the exception of very primitive componential imagery, does not require topographical representation. However, we will show that topographical representation might subservise visual imagery when the subject is scrutinizing attentively local features of objects that are stored in memory. One of the critical points in our commentary is the attentional influence on early visual areas during visual imagery. Furthermore, we propose a model that represents visual imagery as an interaction between memory retrieval and focal attention mechanisms.

The facts

Recent developments in human brain-mapping methods such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) have enabled neuroscientists to search for cortical areas that are involved in various cognitive functions^{3,4}. On the involvement of the primary visual area (V1) in visual imagery, Roland and his collaborators^{5,6} have reported no activation in area V1, whereas some PET studies^{7,8} and fMRI studies^{9,10} have shown activation in area V1. We think that these conflicting facts originate from two main sources: the difference in task paradigms (this will be discussed in *Concluding remarks*), and the relatively weak activation of area V1 that might be near the threshold of detection sensitivity, provided that V1 can be activated under a certain imagery condition. It is necessary, therefore, to examine the relative intensities of activation in visual areas of the whole brain, and their changes in task conditions. However, V1 is not the area most strongly activated in imagery tasks, and other functionally important areas are also activated⁷, in agreement with Roland and Gulyás' report. Thus, the narrow window of measurements restricted to the vicinity of V1 and V2

regions^{8,9} needs to be widened. In PET measurements of weak signals that are near the threshold of detection sensitivity, intersubject averaging, used to enhance signal-to-noise ratios, might mask variances among subjects, and thus result in positive or negative results. This averaging procedure assumes a precise transformation algorithm for individually different sulci and gyri. Another important, and often neglected, source of subject variances is the difference in individual mental states; this must be controlled reliably in every mental task. These problems could be overcome partially by the use of fMRI, because its non-invasiveness enables repetition of measurements within the same subject. Recent fMRI data have shown activation of area V1 during visual-imagery conditions in some, but not all, subjects^{9,10}. One reason for this variance is suspected to be uncontrolled mental states of subjects in the imagery tasks: as Le Bihan and colleagues stated in their paper⁹, 'we had no way to screen for attention and quality of the visual-recall performance'. Neglecting the monitoring of task performances cannot be justified if progress in the field of cognitive neuroscience is to continue.

Background

We have reported evidence of the shared mechanisms of memory retrieval and visual imagery in the primate inferotemporal cortex (IT)^{1,11}. We suggest that visual representation is acquired through associative mechanisms¹¹ as well as through tuning mechanisms that specify the fine-form selectivity of IT neurons^{12,13}. Therefore, perceptual aspects of the temporal association area are related closely to its visual representation that can be generated from stored memory. It has been assumed that cognitive effects, such as attention, influence the responses of IT and V4 cells, and that they do not affect V1 cells^{14,15}. However, a recent single-unit study demonstrated clearly the neuronal correlates of focal attention in V1 and V2 as well as V4 cells¹⁶. This significant finding is consistent with a report of response modulations in V1 and V4 cells that were sensitive to behaviorally relevant stimulus features¹⁷. It is notable that these feature-selective effects on V1 cells are smaller or less evident than those on V4

**Commentary by
Kuniyoshi Sakai
and Yasushi
Miyashita**
*Dept of Physiology,
University of Tokyo,
Hongo, Tokyo 113,
Japan.*

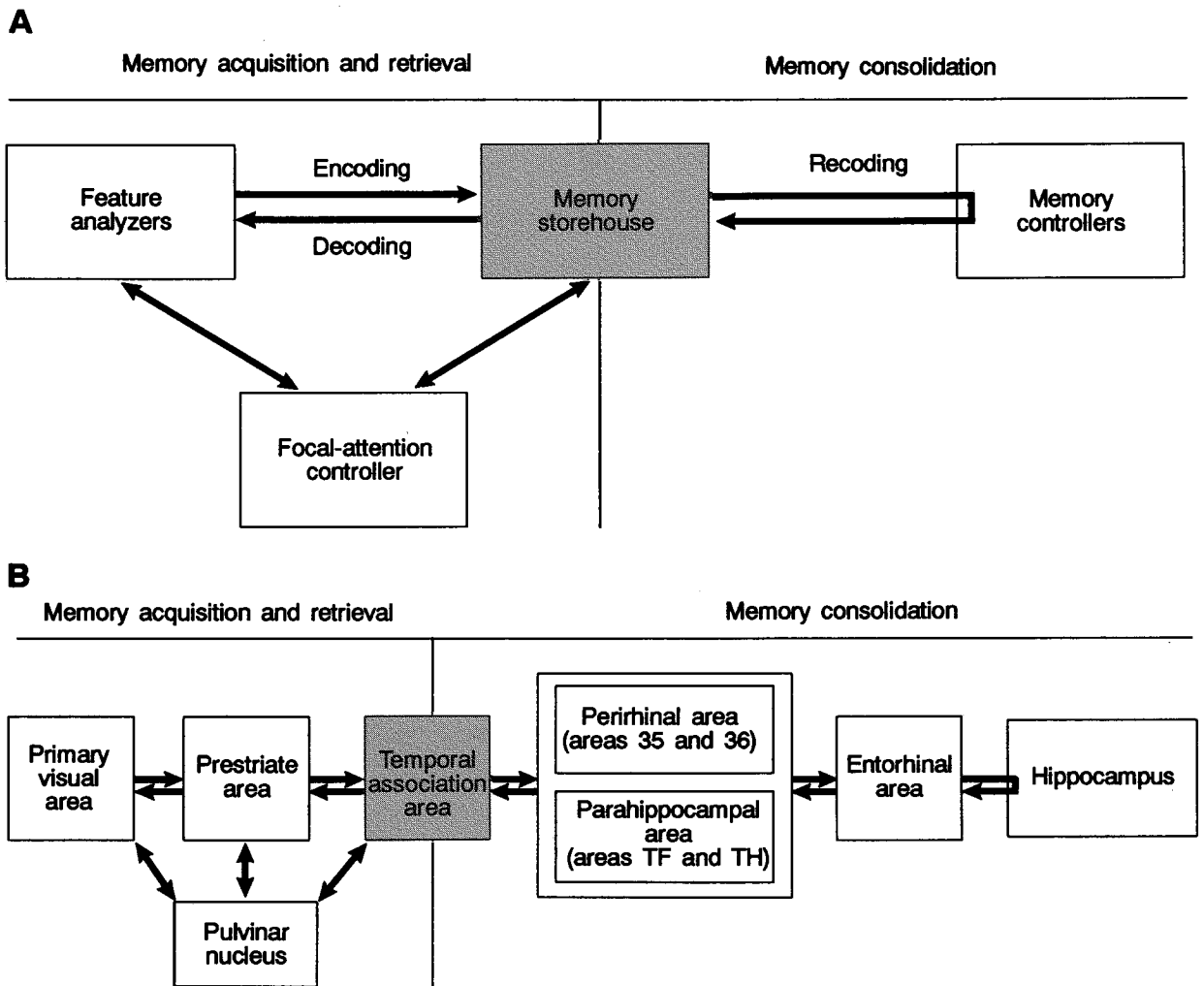


Fig. 1. (A) A model of the cognitive memory system that unifies perception and imagery. The encoding process for memory acquisition proceeds as information flows from feature analyzers to a memory storehouse. In visual perception, a focal-attention controller selects and locates prominent features in a visual field. To consolidate a representation in the memory storehouse, the memory code is recorded by the interaction between the memory storehouse and memory controllers. In visual imagery, the focal-attention controller selects and locates perceptual representations through the decoding process for memory retrieval. **(B)** Structures of the visual memory system for object recognition. Information flow between these structures correspond to information flow indicated in A. Brodmann's areas 35 and 36, and von Economo's areas TF and TH are indicated. Note that most of the pathways between cortical visual areas are anatomically reciprocal, supporting their functional roles in encoding and decoding processes.

cells, which might account for the intensities of activation in visual areas found using PET.

Attentional influence on early visual areas during perceptual processes has been elucidated by a recent computational model proposed by Olshausen and colleagues¹⁸. The main point of their model is that 'control neurons' selectively route topographical information from a window of attention in V1 to higher visual areas. They hypothesized that the pulvinar nucleus might function as these control neurons. The pulvinar nucleus is organized retinotopically and connected reciprocally to all visual areas in the occipito-temporal pathway including areas V1, V2, V4 and IT (Ref. 19). The role of this nucleus in visual attention has been confirmed by PET data²⁰ and lesion studies²¹.

Our proposal is that early visual areas can be influenced by focal attention during visual imagery. Data obtained using PET have consistently indicated

the activation of the pulvinar nucleus in visual-imagery tasks^{5,7}. The possibility of activation in early visual areas due to feature-selective attention has also been supported by PET data, where control conditions were passive response and divided attention²². Therefore, it is likely that attentional components in visual-imagery tasks enhance activities in the pulvinar nucleus or early visual areas, or both. Another piece of evidence supporting our proposal is the mental-imagery deficit caused by lesions in the occipital cortex²³.

Model

A model of the cognitive memory system that we have proposed^{1,24} is shown in Fig. 1A. This scheme is based on structures and functions of the visual-memory system for object recognition²⁵⁻²⁷ (Fig. 1B). In the memory acquisition process, sensory information is transformed into a memory code of

neuronal responses (encoding) with the bottom-up information flow from feature analyzers to a memory storehouse. The primary visual area and the prestriate area serve as feature analyzers in vision. A possible candidate for the memory storehouse is the temporal association area. In visual perception, prominent features in a visual field are selected and located by a focal-attention controller. As noted above, one candidate for the focal-attention controller is the pulvinar nucleus. We hypothesize that perception is implemented by the interaction between memory acquisition (encoding) and focal-attention mechanisms.

To establish a long-lasting representation in the memory storehouse, memory controllers are responsible for consolidation. During this memory consolidation process, the memory code is reorganized dynamically (recoding) by the interaction between the memory storehouse and memory controllers. Medial temporal structures, including the hippocampus, are regarded as memory controllers¹.

In the memory retrieval process, memory codes stored in the memory storehouse are broken down into more elementary attributes (decoding) by feature analyzers through a back-projection pathway. We hypothesize that imagery is implemented by the interaction between memory retrieval (decoding) and focal-attention mechanisms. According to our scheme, visual imagery is generated by top-down activation of perceptual representations that are selected and located by the focal-attention controller. It is also possible that a signal from the focal-attention controller gates the top-down information flow toward early visual areas that function in the decoding process. This model predicts that the extent of visual areas devoted to decoding mental images is controlled dynamically by focal attention.

Concluding remarks

Suppose you are asked to visualize a letter A with your eyes closed. The attentional window would be similar to that indicated by 1 in Fig. 2, but its size and location would vary significantly among subjects. Next, focus your attention mentally on the horizontal bar of a visualized letter A. This time the attentional window would be that indicated by 2, and this window would be centered in front of you. If window 2 highlights 'imaginary' receptive fields near the fovea, the horizontal bar could be an optimum stimulus for some V1 neurons, and these cells would respond to it. In the case of window 2, the topographical representation of a mental image (the geometrical structure of a letter A) must be used and this mental image must be scrutinized. Therefore, topographical representation might subserve visual imagery when the subject is scrutinizing attentively local features (a particular part of an object and its precise location), thereby involving the attentional influence on early visual areas⁷. Early visual areas would not be active when the subject is attending to global features only of a primitive scene or even of complex patterns⁶. Thus, an instruction to subjects, and the strategy of each subject used to

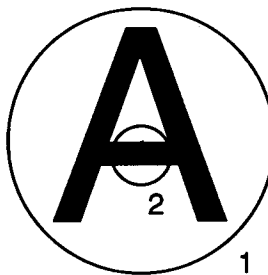


Fig. 2. Examples of focusing of attention. Two attentional windows with different sizes (1 and 2) are indicated. If asked to visualize a letter A with your eyes closed, the attentional window would be similar to that indicated by 1. Next, focus your attention mentally on the horizontal bar of a visualized A. This attentional window (2) would be centered in front of you. In the case of window 2, the topographical representation of a mental image (the geometrical structure of a letter A) must be used and this mental image must be scrutinized.

solve imagery tasks, might well influence the brain activation. We should not forget that our 'usual' visual targets need incessant focusing and shifting of attention, and that this applies to visual imagery as well. The profound mechanisms of human mental processes remain to be challenged with well-designed and well-controlled experiments in cognitive neuroscience.

Selected references

- 1 Sakai, K. and Miyashita, Y. (1993) *Curr. Opin. Neurobiol.* 3, 166–170
- 2 Roland, P. E. and Gulyás, B. (1994) *Trends Neurosci.* 17, 281–287
- 3 Roland, P. E. (1993) *Brain Activation*, Wiley-Liss
- 4 Posner, M. I. and Raichle, M. E. (1994) *Images of Mind*, W. H. Freeman
- 5 Roland, P. E., Eriksson, L., Stone-Elander, S. and Widen, L. (1987) *J. Neurosci.* 7, 2373–2389
- 6 Roland, P. E., Gulyás, B., Seitz, R. J., Bohm, C. and Stone-Elander, S. (1990) *NeuroReport* 1, 53–56
- 7 Kosslyn, S. M. et al. (1993) *J. Cognitive Neurosci.* 5, 263–287
- 8 Damasio, H. et al. (1993) *Soc. Neurosci. Abstr.* 19, 1603
- 9 Le Bihan, D. et al. (1993) *Proc. Natl Acad. Sci. USA* 90, 11802–11805
- 10 Ogawa, S. et al. (1993) *Soc. Neurosci. Abstr.* 19, 976
- 11 Sakai, K. and Miyashita, Y. (1991) *Nature* 354, 152–155
- 12 Sakai, K. and Miyashita, Y. (1993) *Soc. Neurosci. Abstr.* 19, 28
- 13 Sakai, K. and Miyashita, Y. (1994) *NeuroReport* 5, 829–832
- 14 Moran, J. and Desimone, R. (1985) *Science* 229, 782–784
- 15 Posner, M. I. and Dehaene, S. (1994) *Trends Neurosci.* 17, 75–79
- 16 Motter, B. C. (1993) *J. Neurophysiol.* 70, 909–919
- 17 Haenny, P. E. and Schiller, P. H. (1988) *Exp. Brain Res.* 69, 225–244
- 18 Olshausen, B. A., Anderson, C. H. and Van Essen, D. C. (1993) *J. Neurosci.* 13, 4700–4719
- 19 Robinson, D. L. and Petersen, S. E. (1992) *Trends Neurosci.* 15, 127–132
- 20 LaBerge, D. and Buchsbaum, M. S. (1990) *J. Neurosci.* 10, 613–619
- 21 Rafal, R. D. and Posner, M. I. (1987) *Proc. Natl Acad. Sci. USA* 84, 7349–7353
- 22 Corbetta, M., Miezin, F. M., Dobmeyer, S., Shulman, G. L. and Petersen, S. E. (1991) *J. Neurosci.* 11, 2383–2402
- 23 Farah, M. J., Soso, M. J. and Dasheiff, R. M. (1992) *J. Exp. Psychol. Hum. Percept. Perform.* 18, 241–246
- 24 Sakai, K., Naya, Y. and Miyashita, Y. *Learning and Memory* (in press)
- 25 Gross, C. G. (1973) in *Handbook of Sensory Physiology (Central Processing of Visual Information, Vol. VII/3) (Part B: Visual Centers in the Brain)* (Jung, R., ed.), pp. 451–482, Springer-Verlag
- 26 Mishkin, M. (1982) *Philos. Trans. R. Soc. London, Ser. B.* 298, 85–95
- 27 Squire, L. R. and Zola-Morgan, S. (1991) *Science* 253, 1380–1386