

Eye movements during visual mental imagery

Fred W. Mast and Stephen M. Kosslyn

It has long been debated whether eye movements play a functional role in visual mental imagery. A recent paper by Laeng and Teodorescu presents new evidence that eye movements are stored as a spatial index that is used to arrange the component parts correctly when mental images are generated.

Eye movements are essential to visual perception. Although we only see with high resolution when an image falls on the central foveal region of the retina (about 2 deg of visual angle), eye movements induce the illusion that we see entire scenes sharply: wherever we shift our gaze, we see with high resolution. Many factors govern our eye movements, but we can divide them into two general classes. Bottom-up processes respond to a sudden change in the visual field (such as a bright flash or movement), which impels us to shift our gaze to that location. Complementing these, top-down processes rely on knowledge to drive sequences of eye movements [1,2]. We are not simply slaves to the vagaries of the environment, but also voluntarily move our eyes to locations of interest. Both bottom-up and top-down processes clearly play a role in eye movements, but the mechanisms underlying top-down processing are only now coming to be understood.

If eye movements exist to allow us to take in the entire visual field with high resolution (albeit a part at a time), why would anyone study eye movements during visual mental imagery? During imagery, there is no sensory visual input to drive eye movements bottom-up, and no scene to be inspected top-down. One possibility is that eye movements are stored in memory along with the images that were encoded at each fixation [3]. If so, then eye movements might play an important role in allowing one to visualize a montage, a composite created on the basis of memories of multiple fixations. That is, eye movements could trigger sequences of memories and could also help us to position correctly each image of a part of the field relative

to other parts. Moreover, the same kinds of processing might operate when one is inspecting an actual scene, seeking key characteristics instead of visualizing them. If so, then studying eye movements during imagery, and comparing them with those that occur during perception, could provide important insights into the nature of top-down processing generally.

In their recent paper, Laeng and Teodorescu demonstrate that there are in fact eye movements during mental imagery and, more importantly, that these eye movements are *functionally* involved in mental imagery processes [4]. Similarities between the eye's scan path in imagery and perception have been described before [5], but Laeng and Teodorescu took an innovative step in their study. They systematically varied the conditions under which the participants encoded the visual stimuli. In a first experiment, the participants viewed a 6 × 6 grid pattern, which contained five black cells (the grid subtended 9.5 deg. of visual angle, as calculated according to the information given in the paper). When initially committing the display to memory, one group was instructed to keep their gaze on a central point on the grid; another group was allowed to examine the pattern freely for the same amount of time (20 s). Eye movements were recorded while participants studied the displays and then later, when they visualized the memorized displays. Laeng and Teodorescu found that the eye movements during imagery reflected the conditions in which the participants had studied the stimuli: the participants who maintained fixation while studying the display made almost no eye movements when later visualizing it; by contrast, the participants who visually explored the pattern while studying it also made eye movements during imagery. Furthermore, the sequence of fixations during imagery was very similar to the sequence that had occurred when the participants studied the pattern initially.

Moreover, the more similar the imagery and perception scanpaths, the better the participants performed in a subsequent spatial memory task. In this task, the participants studied the grid they had memorized and visualized before, but now each of its cells was white and contained a number from 1 to 36. The participants called out five numbers, which they thought corresponded best to the positions previously filled in with black. However, the finding that participants who had similar eye movement patterns during the perception and imagery of the initial display were also more accurate in this test must be evaluated with caution. The participants in the two groups – fixation versus free exploration – performed comparably in this spatial memory task. Thus, eye movements *per se* could not have been responsible for spatial memory. Laeng and Teodorescu suggest that participants in the fixation condition could have scanned the image covertly. This argument is interesting, and there is good evidence that attention can be shifted without eye movements (e.g. Ref. [6]). However, Laeng and Teodorescu did not test this hypothesis in their study.

In a second experiment, Laeng and Teodorescu adjusted the display so that a fixation cross appeared at the center of the screen, and the stimuli appeared in one of the four corners. This procedure was designed to rule out the possibility that participants could adjust their field of attention so that it encompassed the entire stimulus display. In addition to the use of pictures of natural objects instead of grids, this second experiment included a crucial condition: free viewing in perception but fixation during imagery. Based on the previous findings, the investigators hypothesized that this condition should stop participants using memories of eye movements to help reconstruct the image. And in fact, the participants were less accurate when asked about visual details while visualizing the previously studied picture. For example,

one question required the participants to determine how many stripes were on a fish in the picture they had studied. For the retrieval task in the free-viewing perception but fixed imagery condition, the oculomotor links established during perception could not be used in the service of building up a mental image – and this limitation apparently impaired recall.

Laeng and Teodorescu argue persuasively against the view that eye movements during imagery are a meaningless by-product of shifting one's attention across an internally generated image. If this were the case, participants should not be impaired when they visualize whilst maintaining fixation. One might try to claim, still, that such impaired performance is in fact due to the cost of maintaining fixation while visualizing. This alternative account seems unlikely, given other findings in the literature, but further research will be needed to rule out this potential confound conclusively.

Evidence for shared mechanisms in perception and imagery inspires new questions for future research. For example, it would be interesting to know

whether the relationship between imagery and perception persists over time. In Laeng and Teodorescu's study, the participants performed the imagery task 40 s after they studied the stimuli. It is conceivable that the sensorimotor trace is stored only in short-term memory. Ishai and Sagi have shown, for example, that imagery-induced facilitation in a target-detection task decays and is only effective in the first 5 min after the participants saw the stimuli [7]. Another question is whether the eye movement scanpath would be preserved if the perceived stimulus were imagined in an orientation or size other than the one in which it was initially studied. In this case, it would not be possible simply to re-enact the identical motor commands in order to arrange the parts of the image. Yet another question focuses on learning: is it possible to improve individuals' imagery abilities by improving the perceptual encoding process?

To conclude, Laeng and Teodorescu's study adds important information to the growing evidence that many mechanisms involved in perception are also used in mental imagery. These investigators make

a compelling case for the view that eye movements during imagery play a functional role when we generate visual mental images.

References

- 1 Noton, D. and Stark, L.W. (1971) Scanpaths in eye movements during pattern perception. *Science* 171, 308–311
- 2 Noton, D. and Stark, L.W. (1971) Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vis. Res.* 11, 929–942
- 3 Kosslyn, S.M. (1994) *Image and Brain*, MIT Press
- 4 Laeng, B. and Teodorescu, D-S. (2002) Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cogn. Sci.* 26, 207–231
- 5 Brandt, S. and Stark, L.W. (1997) Spontaneous eye movements during visual imagery reflect the content of the visual scene. *J. Cogn. Neurosci.* 9, 27–38
- 6 Posner, M.I., Snyder, C.R.R. and Davidson, B.J. (1980) Attention and the detection of signals. *J. Exp. Psychol. Gen.* 109, 160–174
- 7 Ishai, A. and Sagi, D. (1995) Common mechanisms of visual imagery and perception. *Science* 268, 1772–1774

Fred W. Mast

Stephen M. Kosslyn

Dept of Psychology, Harvard University,
Cambridge, MA 02138, USA.

e-mail: fmast@wjh.harvard.edu

A neural mechanism for non-verbal discourse comprehension

John Kounios

Sometimes, a thousand words isn't worth a picture. A recent study by West and Holcomb used event-related potentials to examine non-verbal, discourse-level processing of stories presented as picture sequences. These complex pictures were presented at a rapid rate, thereby minimizing linguistic mediation. The ERP showed N300 and N400 deflections, hypothesized to reflect image activation and amodal semantic integration, respectively.

Do we think in pictures? Words? Something else? Most people have an opinion about the representational medium of their own cognitive processing. These opinions are typically based on introspection, and could be a reasonably accurate reflection of people's (possibly epiphenomenal) conscious processes. It is not clear, however, whether these

conscious macroprocesses accurately mirror the succession of brief, unconscious microprocesses on which conscious processing might be based. In order to investigate the representational substrate of these cognitive microprocesses, it is therefore necessary to adopt experimental methods capable of partitioning cognition into its components. A seminal new study by West and Holcomb [1] used the event-related brain potential (ERP) technique to make important inroads into this longstanding problem.

Neurocognitive semantics and levels of representation

The 'multiple semantics' view of knowledge representation posits multiple information formats or codes, each format associated with a particular sensorimotor modality. The most famous of the

multiple-semantics theories is the dual-coding theory of Paivio [2], which holds that there are separate representational (processing) systems for pictorial and verbal information. By contrast, the unitary semantics view posits a single, amodal information format underlying cognition [3]. Recent investigations have provided evidence that is arguably consistent with the existence of both amodal and modality-specific representations, although no consensus has yet emerged [4].

One of the complexities of this issue is that there are different (and an unknown number of) *levels* of knowledge representation, each of which may utilize representations with different formats. For instance, semantic memory is at the most fundamental level, as this is the repository of stored word meanings and