

Reprint Series
23 June 1995, Volume 268, pp. 1772-1774

SCIENCE

Common Mechanisms of Visual Imagery and Perception

Alumit Ishai* and Dov Sagi

Common Mechanisms of Visual Imagery and Perception

Alumit Ishai* and Dov Sagi

Detection of a visual target can be facilitated by flanking visual masks. A similar enhancement in detection thresholds was obtained when observers imagined the previously perceived masks. Imagery-induced facilitation was detected for as long as 5 minutes after observation of the masks by the targeted eye. These results indicated the existence of a low-level (monocular) memory that stores the sensory trace for several minutes and enables reactivation of early representations by higher processes. This memory, with its iconic nature, may subserve the interface between mental images and percepts.

Visual imagery is the invention or recreation of a perceptual experience in the absence of retinal input. Brain imaging studies implicate activity in cortical visual areas during visual imagery (1, 2), yet the neural mechanisms that subserve "seeing with the mind's eye" are controversial (3, 4). The degree to which the same neural represen-

tations are involved in both visual imagery and visual perception is unclear. Earlier studies have shown that visual imagery interferes with perception (Perky effect) (5). Visual imagery can facilitate letter detection by increasing expectation (6), yet there is no evidence for direct facilitatory interactions between imagery and perception. In order to test whether visual imagery can induce a facilitatory effect on visual perception, we used a lateral masking detection paradigm (7, 8), in which human observers

Department of Neurobiology, Brain Research, Weizmann Institute of Science, Rehovot 76100, Israel.

*To whom correspondence should be addressed.

performed imagery tasks that were preceded by perceptual tasks.

Six observers that were naïve to the task and one of the authors (A.I.) participated in the experiments (9). The task of each observer was the detection of a Gabor target under three conditions: (i) a perception condition, in which the target was flanked by two peripheral Gabor masks (Fig. 1A, upper panel); (ii) a control condition, in which the masks were absent; and (iii) an imagery condition, in which the observer had to detect the target while imagining the absent masks (Fig. 1A, lower panel). Contrast detection thresholds were measured as a function of target-to-mask distance (Fig. 2). In the perception condition, two zones were seen: suppression, that is, an elevation of threshold within eccentricity of twice the Gabor target wavelength (2λ); and enhancement, that is, a threshold reduction, reaching a maximum at three times the target wavelength (3λ). In the control condition, both suppression, which is due

in part to integration within receptive fields, and enhancement, which is due to integration outside receptive fields, were not apparent. In the imagery condition, the suppression disappeared and a facilitatory effect was seen. The imagery-induced facilitation also reached a maximum at 3λ , the optimal target-to-mask distance [a threshold reduction of 0.1 ± 0.01 logarithmic unit, as compared with 0.2 ± 0.02 in the perception condition and with a threshold elevation of 0.04 ± 0.01 in the control condition (mean \pm SE, $n = 7$ observers)].

Further analysis revealed similarity between the facilitation seen in both imagery and perception. The area in the enhancement zone, from 2λ to 12λ , was computed for all conditions in each session. The differences between the enhancement area in the imagery and perception conditions were statistically not significant, but the differences between the control and perception conditions were significant, as well as the differences

between the control and imagery conditions (repeated measures analysis of variance followed by Scheffe multiple comparison, $P < 0.001$). Another criterion was the distribution of minima points: Although in the control condition a minimum was observed almost equally likely at all target-to-mask distances, in the perception and the imagery conditions the minima were in 75 and 50% of the sessions, respectively, at 2λ and 3λ (Fig. 3). The differences between perception and imagery distributions were statistically not significant, whereas the differences between perception and control conditions were significant (Kolmogorov-Smirnov test, $P < 0.005$), as well as the differences between imagery and control conditions (Kolmogorov-Smirnov test, $P < 0.05$) (10).

The imagery facilitation was found to be orientation-specific. When target and masks were orthogonal (vertical target flanked by horizontal masks), no facilitatory effect was observed in either perception or imagery [at 3λ , a threshold reduction of 0.04 ± 0.03 logarithmic unit in the perception condition, as compared with 0.02 ± 0.01 in the imagery condition and with a threshold elevation of 0.01 ± 0.04 in the control condition (mean \pm SE, $n = 3$ observers)]. When observers were instructed to perform a mental horizontal task (for example, imagining horizontal masks after perceiving vertical ones), the imagery facilitation was significantly reduced. Moreover, when the horizontal masks in the perception task were followed by mental rotation to vertical ones, no imagery facilitation was obtained, indicating the necessity of previous sensory activation (11).

Additional experiments showed that the imagery facilitation was monocular. When the perception task was performed with one eye covered and the following imagery task with the other eye covered, there was no imagery-induced facilitation [at 3λ , a threshold reduction of 0.2 ± 0.03 logarithmic unit in the perception condition, as compared with a threshold elevation of 0.03 ± 0.02 in the imagery condi-

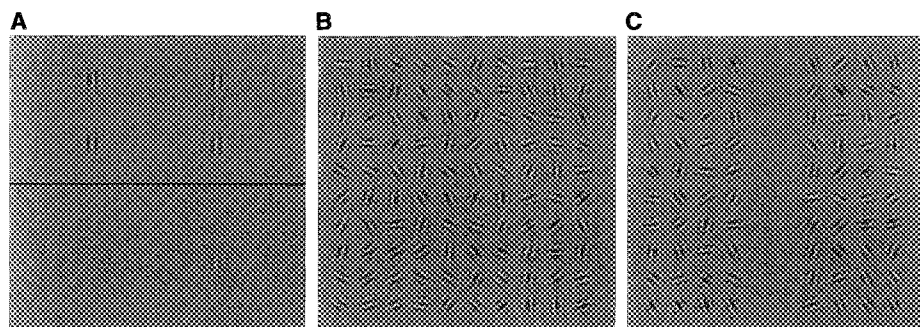


Fig. 1. Computer-generated Gabor stimuli. (A) A foveal Gabor target flanked by two high-contrast Gabor masks, at a distance of three times the target wavelength (3λ), used in the perception condition (upper panel) and an isolated target without the masks, used in both control and imagery conditions (lower panel). Gabor target and masks were vertical, arranged along the vertical meridian. In both examples, target appeared on the right (that is, on the second presentation of stimuli). The display had an area of 9.6° by 9.6° with a mean display luminance of 50 cd/m^2 , viewed binocularly from a distance of 150 cm in a dark environment. (B) Visual noise composed of Gabor patches of random orientations and phases. (C) Surround visual noise with a gap at the location of target and masks.

Fig. 2. Imagery-induced facilitation. The figure shows the contrast detection threshold as a function of target-to-mask distance, for vertical target and masks, averaged across all observers. Error bars in this and subsequent graphs indicate the standard error of the mean.

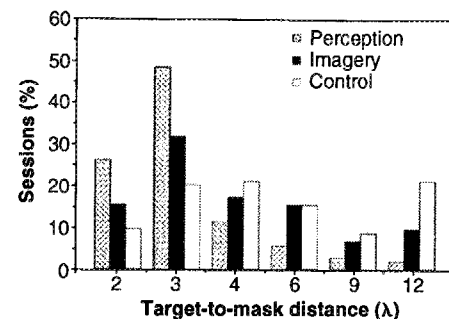
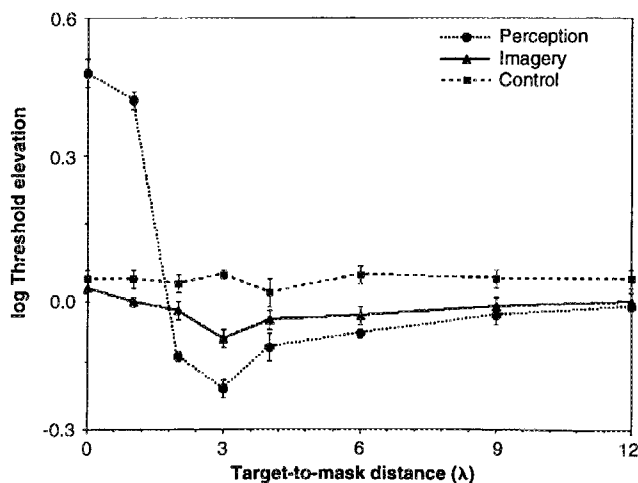


Fig. 3. Distribution of minima points in perception, imagery, and control conditions, averaged across all observers.

tion and 0.04 ± 0.01 in the control condition (mean \pm SE, $n = 3$ observers) (11). The necessity of matching between the sensory and the recalled stimuli for facilitation to occur implicates an iconic storage of the former. Reducing the number of trials in the perception blocks to 10 abolished the imagery effect, which suggests that the imagery facilitation was not merely due to an effect of attention and indicates the need for repetition to create a memory trace (11).

To quantify the temporal aspects of the memory involved in the imagery-induced facilitation, we tested the effect of various delay periods between the perception and imagery conditions. With short delay periods (up to 5 min), imagery-induced facilitation was obtained (Fig. 4A). However, with a delay of 10 min, no facilitation was seen. These results suggest the involvement of a memory system that stores the sensory trace for several minutes and that can be reactivated by higher level processes.

To test whether "visual noise" could affect the sensory trace, we introduced a display composed of Gabor patches of random orientations and phases at the end of each block or at the end of each trial (Fig. 4B). Although imagery facilitation still was detected when the visual noise was added at the end of each block (Fig. 1B),

no imagery effect was obtained when the visual noise was presented after each trial. When the visual noise covered only the surrounding area after each trial, leaving the location of target and masks intact (see Fig. 1C), imagery facilitation was seen. In all cases, the addition of the visual noise did not affect the perception enhancement. These results suggest that the sensory storage is created during the perceptual task. The addition of visual noise after each trial interfered with the accumulation of the sensory trace and, hence, affected the observer's ability to recall the relevant Gabor patches during the imagery task.

The main effect uncovered by use of the lateral masking paradigm is the imagery facilitation (12), which shared the same characteristics with the perception facilitation (Figs. 2 and 3). This paradigm has revealed the existence of facilitatory interactions between spatial channels, which were found to be monocular, orientation-specific, and spatial frequency-specific (7, 8). We believe, therefore, that mental images can be interfaced with perceptual representations at early stages of visual information processing. Our data provide indirect support for the hypothesis that visual imagery activates the primary visual cortex (2, 4, 13). As the stimuli we used had no semantic significance, it is reasonable to infer that the facilitatory effect of visual imagery on perception was due to activation of early representations. Neuropsychological case studies have shown dissociation of visual imagery and visual perception (14), which suggests impairment at the level of activation of "internal representations." Our results indicate, therefore, the existence of common representational structures that can subserve both perception (for instance, matching visual input to stored information for recognition) and image generation.

The time course of the imagery facilitation and visual noise experiments (Fig. 4) exposed a memory system that is capable of storing the sensory trace for several minutes. The stored information is accessible to higher level processes. Reactivation by visual imagery can induce facilitation when the recalled images include features of the stored images, such as orientation and location, as well as the eye used. This stimulus-specific memory suggests that cortical cells that process the stimulus serve also as memory cells. Earlier studies have shown the involvement of monocular and orientation-specific cells in perceptual learning (15). It is possible that the memory system involved in imagery subserves learning by enabling spatiotemporal associations across a time window of few minutes (8).

REFERENCES AND NOTES

1. P. Roland, L. Eriksson, S. Stone-Elander, L. Widen, *J. Neurosci.* **7**, 2373 (1987); M. Farah, F. Peronnet, M. Gannon, *J. Exp. Psychol. Gen.* **117**, 248 (1988); G. Goldenberg *et al.*, *Neuropsychologia* **27**, 641 (1989).
2. D. L. Bihan *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 1182 (1993); S. Kosslyn *et al.*, *J. Cogn. Neurosci.* **5**, 263 (1993).
3. M. Farah, *Trends Neurosci.* **12**, 395 (1989).
4. P. Roland and B. Gulyas, *ibid.* **17**, 281 (1994); S. Kosslyn and K. Ochsner, *ibid.*, p. 290; K. Sakai and Y. Miyashita, *ibid.*, p. 287.
5. S. Segal and V. Fussella, *J. Exp. Psychol.* **83**, 458 (1970); C. Craver-Lemley and A. Reeves, *Psychol. Rev.* **99**, 6633 (1992).
6. M. Farah, *J. Exp. Psychol. Gen.* **114**, 91 (1985).
7. U. Polat and D. Sagi, *Vision Res.* **33**, 993 (1993).
8. ———, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 1206 (1994).
9. A two-alternative forced-choice procedure was used. Each trial consisted of two stimuli presented sequentially, only one of which had a target. Before each trial, a small fixation cross was presented at the center of the screen. When ready, the observer pressed a key activating the trial sequence: a no-stimulus interval (0.5 s), a first-stimulus presentation (90 ms), a no-stimulus interval (1 s), and a second-stimulus presentation (90 ms). In visual noise experiments, after another interval (0.5 s), the noise was presented for 0.5 s. The observer was asked to perform a detection task, that is, to determine which of the stimuli contained the target. Each block consisted of 50 trials on average, across which the distance between the Gabor target and masks was kept constant. Auditory feedback, by means of a keyboard bell, was given immediately after an erroneous response. Each session included eight alternating blocks of either perception followed by control or perception followed by imagery, with one of two sets of target-to-mask distances (either 0, 2, 4, 9 λ or 1, 3, 6, 12 λ). Mask amplitude was 40% of mean luminance, with the wavelength (λ), and the Gaussian envelope size (σ) equal to 0.15° [as in (7, 8)]. Target threshold contrast (which ranged from 5 to 15%) was determined by a staircase method. Threshold elevation was computed relative to detection of the isolated target in the presence of two peripheral high-contrast crosses.
10. Each daily session included either perception and control conditions or perception and imagery conditions. The comparison between imagery and control is, therefore, a comparison between different sessions.
11. A. Ishai and D. Sagi, *Tech. Rep. GC-DS/95-1* (Weizmann Institute of Science, Rehovot, Israel, 1995).
12. We did not observe suppression at short distances (0 and 1 λ), where target and masks were overlapped, as may be expected by the classical Perky effect. This might be due to the fact that we used a simple detection task and stimuli that have no meaning, as opposed to common objects (such as lines and circles) used in the Perky effect (5).
13. Because the Perky effect is not affected by orientation, it was suggested that the interference may occur at an earlier level of visual processing, before the extraction of features such as orientation [C. Craver-Lemley and A. Reeves, *Perception* **16**, 599 (1987)].
14. Z. Mehta, F. Newcombe, E. D. Haan, *Neuropsychologia* **30**, 645 (1992); M. Behrmann, C. Winocur, M. Moscovitch, *Nature* **359**, 636 (1992); J. Jankowiak, M. Kinsbourne, R. Shalev, D. Bachman, *J. Cogn. Neurosci.* **4**, 119 (1992).
15. A. Karni and D. Sagi, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 4966 (1991).
16. We thank Y. Dudai, S. Edelman, and R. Malach for helpful comments on an early version of the manuscript, A. Reeves for thoughtful suggestions, and E. Schechtman for help with statistics. Supported by the Science Foundation administered by the Israel Academy of Science and Humanities, by the Charles H. Revson Foundation, and by Foundation Mordoch Mijan de Salonique.

10 February 1995; accepted 3 May 1995

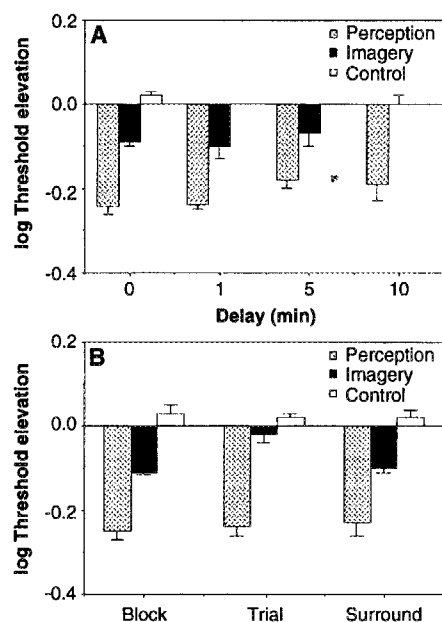


Fig. 4. Time course and memory masking. (A) Relative enhancement (at 3λ), as a function of delay periods between perception and imagery tasks, averaged across three observers. (B) Relative enhancement (at 3λ), in the presence of visual noise, averaged across three observers. The visual noise was introduced at the end of each block for 1 min, or at the end of each trial for 0.5 s, covering the entire display (see Fig. 1B) or the surrounding area (see Fig. 1C).