

Eye scanpaths during visual imagery reenact those of perception of the same visual scene

Bruno Laeng*, Dinu-Stefan Teodorescu

Department of Psychology, University of Tromsø, Åsgårdveien 9, N-9037 Tromsø, Norway

Received 2 February 2001; received in revised form 1 August 2001; accepted 30 November 2001

Abstract

Eye movements during mental imagery are not epiphenomenal but assist the process of image generation. Commands to the eyes for each fixation are stored along with the visual representation and are used as spatial index in a motor-based coordinate system for the proper arrangement of parts of an image. In two experiments, subjects viewed an irregular checkerboard or color pictures of fish and were subsequently asked to form mental images of these stimuli while keeping their eyes open. During the perceptual phase, a group of subjects was requested to maintain fixation onto the screen's center, whereas another group was free to inspect the stimuli. During the imagery phase, all of these subjects were free to move their eyes. A third group of subjects (in Experiment 2) was free to explore the pattern but was requested to maintain central fixation during imagery. For subjects free to explore the pattern, the percentage of time spent fixating a specific location during perception was highly correlated with the time spent on the same (empty) locations during imagery. The order of scanning of these locations during imagery was correlated to the original order during perception. The strength of relatedness of these scanpaths and the vividness of each image predicted performance accuracy. Subjects who fixed their gaze centrally during perception did the same spontaneously during imagery. Subjects free to explore during perception, but maintaining central fixation during imagery, showed decreased ability to recall the pattern. We conclude that the eye scanpaths during visual imagery reenact those of perception of the same visual scene and that they play a functional role. © 2002 Cognitive Science Society, Inc. All rights reserved.

* Corresponding author. Tel.: +1-47-77646372; fax: +1-47-77645291.

E-mail address: bruno@psyk.uit.no (B. Laeng).

1. Introduction

In a seminal paper concerning imagery, Donald O. Hebb (1968) pointed out that when we build a visual image of a familiar object, for example a rowboat, the different parts of the objects are not clear all at once but successively. Similarly, during perception, we gaze at the actual object by a series of eye movements that bring into focus each of its relevant parts so that our perception is the integration of the results of several eye fixations (cf. Yarbus, 1967). Importantly, Hebb remarked that if the eye movements, observed during perception, are mechanically necessary in scanning the object, the same motor processes could also have an “organizing function” for both perception and imagery. He hypothesized that “if the image is a reinstatement of the perceptual process it should include the eye movements [. . .] and if we can assume that the motor activity, implicit or overt, plays an active part we have an explanation of the way in which the part-images are integrated sequentially” (Hebb, 1968, p. 470). Similarly, Neisser (1967) speculated that the act of constructing an image might require further eye movements like those originally made in perceiving. According to his view, imagery is a process of visual synthesis and construction, much like the one used in perception. Thus, imagery should be considered a coordinate activity between visual memory and eye movement-patterns instead of a revival of stored pictures. Finally, Neisser argues that the more vivid an image the more it is likely to involve some scanning process (Sheehan & Neisser, 1969); but one should not expect a perfect correlation between the experience of vividness of a mental image and the extent and orderliness of eye movements, because even when perceiving real objects shifts of attention without ocular motion can occur.

In the light of the above considerations and the time these ideas were put forward, it is remarkable that only recently Brandt and Stark (1997) have shown that there occur spontaneous eye movements during visual imagery that closely reflect the content and the spatial arrangement of the original visual scene. According to their findings, there is a clear correlation between the eyes’ perceptual analysis of an object or scene and the scanpaths during imagery of the same object or scene. Remarkably, this occurs despite the fact that we would expect that implicit, covert, shifts of attention greatly reduce the need of directing gaze to each different part of an object. Like Hebb, Brandt and Stark (1997) interpret this orderly pattern of eye movements as likely to play a significant, functional, role in the process of visual imagery. Specifically, they propose that the scanpaths are linked to activating and arranging part images of a complex scene into their proper locations; the eyes’ motor system would be a partner to the mechanism of image “generation.”

However, repetitive sequences of movements of bodily sensors during imagery could also be interpreted as epiphenomenal effects of the sequencing of internal commands to shift covertly the attention window. Contemporary research on visual imagery has clarified that images are patterns of activity within a spatial array, a complete image of an object being generated from visual memory by adding iteratively its individual parts (Kosslyn, 1980). This spatial array corresponds to retinotopically mapped cortical areas (Kosslyn, 1994) and a representation within this array could then be scanned in the same way that the representation of an actual object would be (covertly) scanned. However, overt eye movements (saccades and fixations, or eye scanpaths) would seem useless for the purpose of scanning an internal image, because there is no external stimulus to be looked at. Nevertheless, in normal

circumstances, movements of body parts as the hands or the eyes and attentional shifts appear to be functionally linked together. Remarkably, motor areas of the human cortex appear to be engaged during mental transformation tasks of drawings of Shepard-Metzler cubes (e.g., Deutsch et al., 1988). Yet it could be argued that these localized activations of the brain's motor areas represent the preprogramming of movements that either are unexecuted or suppressed or when expressed (like the eyes during imagery) are inherently nonfunctional to the mental transformation process. The motor program would be functional if the subjects were allowed to manipulate the actual 3-D stimuli or look at physically present objects. Indeed, several investigators have suggested that covert shifts of attention may operate and have evolved as preparatory mechanisms for the control of overt shifts in eye fixations (cf. Rizzolatti et al., 1987; Umiltà et al., 1991; Walker, Kentridge, and Findlay, 1995). Shifts of attention precede shifts of the eyes to the same location (cf. Deubel & Schneider, 1996; Henderson, Pollatsek & Rayner, 1989; Rayner, McConkie & Ehrlich, 1978). Thus, we could think that when spontaneous eye movements occur during imagery, these are merely a nonfunctional redundancy, a by-product or "spill over" from the internal process of image "inspection." In other words, we could say that, because attention movements precede eye movements in an obligatory fashion (Irwin & Gordon, 1998), eye movements may also tend to follow in a more or less obligatory fashion the direction of each (or a subset) of the locations visited by attention. Thus, the eyes' motor system would be engaged as a "passive slave" to the visual system while "inspecting" a visual image, regardless of the fact that sometimes there would be no object out there to be viewed.

In summary, we can think of two contrasting accounts for the phenomenon of similar patterns of eye movements during imagery and perception. The functional account can be characterized by the hypothesis that the encoding of each eye fixation during perception participates later, during image generation, as an index to the location of a part in the image. In this view, the efferent commands to the eyes and proprioceptive information are stored along with the visual representation as a form of spatial coding (cf. Roll et al., 1991). Normally, these commands would reflect the order in which parts were inspected during encoding and their locations would be encoded according to a motor-based coordinate system; during retrieval this information could affect the way parts of an image are generated by reenacting the same eye movements in the same order. In support to the idea that the motor system may not be passively slaved to visual system in imagery is the finding that performing movements selectively interferes with mental imagery (Quinn and Ralston, 1986; Wexler et al., 1998). In contrast, the "epiphenomenal" account views movements during imagery as the passive "spill over" from covert shifts of attention during the stage of image inspection, subsequently to its generation. In this phase, neural discharges from the attentional areas program the attention window to move over different parts of an image while these are maintained in the visual buffer. Such an influence of brain areas linking attention and eye movements would result in ocular behavior that reflects the imagery process but it is irrelevant to it. Because of their irrelevance, there would be no need for the visual system to exert active inhibition on these spontaneous eye movements during imagery.

A way to test whether eye movements in imagery play a functional role would be to contrast an experimental situation where a) the subjects' gaze is constrained during the perceptual phase (e.g., by requiring subjects to memorize a pattern or scene while keeping

their gaze on a static fixation point) but free during the imagery phase; with a situation b) in which the eyes are free to move in both the perceptual and imagery phases and another c) in which the eyes are free to move during the perceptual phase but not the imagery phase. We are led to predict that if the functional account is correct, in the condition in which (b) subjects maintained fixation during perception, their gaze should remain fixed to this location during imagery as well. This prediction follows straightforwardly from the idea that the structure of the object is encoded also in reference to the eyes' coordinates, even when there is only one fixed point that is foveated, instead of a series of differentially located points. In contrast, if the epiphenomenal account is correct, it should not make any difference whether the eyes were free to move or constrained during perception. If eye fixations during imagery are only "mirroring" the inspective movements of an attentional window over the target image, then the eyes should spontaneously move towards each of the locations being scanned. Moreover, if the eye movements play a functional role in imagery, then when (c) the eyes are free to explore during perception but are forced to maintain fixation during imagery, imagery should be disrupted and accuracy of recall should consequently suffer. In the following experiments we put to test these ideas. In Experiment 1, we contrasted conditions a) and b) by using a stimulus pattern similar to that used by Brandt and Stark (1997). In Experiment 2, we contrasted conditions a), b) and c) and used pictures of natural objects (i.e., fish) as the objects to be imaged.

2. Experiment 1

Two groups of subjects were instructed to view a 6×6 grid pattern, similar to that used by Brandt and Stark (1997), which was later to be visually imagined. Each diagram contained 5 black filled cells, which would randomly change locations from trial to trial (see Fig. 1). One group, labeled the Free Vision group, was asked to examine freely the pattern for 20 s. We hypothesized that this group would use the mobile attention window, typically limited to a narrow but high-resolution area, to examine serially, both overtly and covertly, the checkerboard. This is a rather complex stimulus but certainly within the visual memory capacity of normal individuals; Irwin (1993) has estimated that between 3 and 6 elements of a visual pattern are maintained in memory across each eye movement. Therefore, in this situation, subjects could either fixate serially on the locations of each of the five blackened squares or fixate onto a few locations on the pattern so as to comprise within the high-resolution area of one fixation a small cluster of elements (e.g., two or three elements). Another group, labeled Central Fixation, was instead requested to maintain gaze on the center of the grid during the whole presentation of the checkerboard. This group of subjects would be characterized by the use of a static gaze and broad focus of the attention window. The main purpose of the experiment was to test whether, when forming a visual image of the same stimulus, the eye movements will reflect the content of the stimulus in both conditions or whether in the Central Fixation condition the eyes would reenact the static position that had been required during perception.

In addition, we collected ratings of the vividness of imagery, both for each mental image as well as vividness as an individual's trait; the former type of vividness was measured at

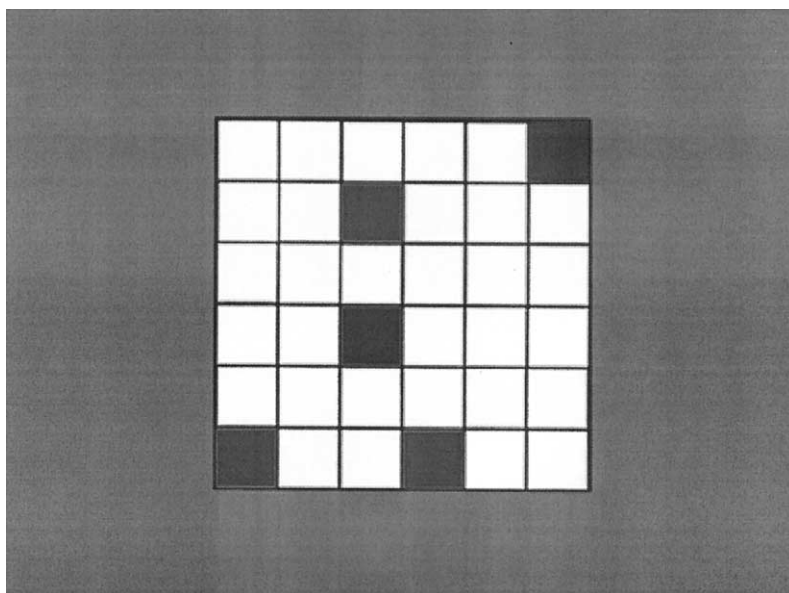


Fig. 1. Experiment 1. An example of the “checkerboard” stimuli that were first perceived and then imagined.

each trial by asking subjects to rate on a discrete scale how vivid each image was, whereas vividness as a trait was measured with the VVIQ Survey questionnaire (Marks, 1973). The so-called Vividness IQ of each participant has been theorized to index ability for imagery that is a stable and general trait of a particular individual. In the experiment, individuals with comparable vividness ratings were assigned in a balanced manner across the groups.

2.1. Methods

2.1.2. Participants

Eight students at the University of Tromsø, 5 females and 3 males (age range 23–31 years), volunteered to participate as paid participants to an experiment on mental imagery. All subjects reported normal vision, or corrected to normal (with contact lenses). Participants were naïve about the hypotheses underlying the experiment, and during the debriefing session at the end of the experiment, it was confirmed that each subject had no specific knowledge or intuitions about the experimenters’ expectations.

2.1.3. Apparatus and stimuli

Eye movements were recorded by means of the Remote Eye Tracking Device, R.E.D., built by SMI-SensoMotoric Instruments from Teltow (Germany). Analyses of recordings were then computed by use of the iView-software, also developed by SMI. The R.E.D.-II can operate at a distance of 0.5–1.5 m and the recording eye tracking sample rate is 50/60 Hz., with resolution better than 0.1 degree. The eye-tracking device operates on the basis of determining the positions of two elements of the eye: The pupil and the corneal reflection. The sensor is an infrared light sensitive video camera typically centered on the left eye of the

subject. Room lighting does not interfere with the recording capabilities of this apparatus. The coordinates of all the boundary points are fed to a computer that, in turn, determines the centroids of the two elements. The vectorial difference between the two centroids is the “raw” computed eye position. The “Vividness of Visual Imagery Questionnaire,” or VVIQ Survey (Marks, 1973) consists of sixteen questions, asking the participant first to image a scene and then to rate the vividness of the mental image on a five-point rating scale. An example of a question from the questionnaire is the following: “Think of some relative or friend whom you frequently see (but who is not with you at present), and consider carefully the picture that comes before your mind’s eye. The exact contour of the face, head, shoulders, and body.” The rating scale was ranging from 1–“no image at all, you only ”know“ that you are thinking of the object”-to 5–“perfectly clear and vivid as normal vision” -; 3 being defined as “moderately clear and vivid.” The visual stimuli used in the experiment were adapted from those used by Brandt and Stark (1997) and consisted of 8 irregularly-checked diagrams, made of 6×6 squares forming a square grid of white empty square cells; however, five of the cells in each grid were filled in black (see Fig. 1). These 5 black squares were randomly located in each diagram and there were no identical or repeated patterns during the experiment. All stimuli were 10×10 cm in size and were presented centered on a 49 cm flat color monitor, surrounded by a uniform blue background. The stimuli presentation was controlled by ACDSee 32v2.4 software. The subject was seated in front of the monitor at a distance of 60 cm, with the head placed in a chin-and-forehead rest to reduce head movements.

2.1.4. Procedure

At the beginning of the experiment, each participant was asked to read the instructions. The instructions for the Free Vision group for the perception phase were: “Look carefully at the figure that shall appear on the monitor and try to remember it as precisely as possible.” The instructions for the Central Fixation group instead were: “Keep your eyes focused in the center of the diagram and try to remember the whole figure as precisely as possible.” The instructions for the imagery phase were common for both groups: “Build a visual image of the figure you just saw while keeping the eyes open.” Subjects in both groups were also specifically told to keep their eyes open at all times during the imagery phase and, importantly, that they were free to look wherever they wanted on the screen during the imagery task. A standard calibration routine was used at the very beginning of each session. The eye position was recorded at nine standard calibration points (appearing as white plus signs on a blue background), corresponding to a regularly spaced 3×3 matrix. The participant was instructed to fixate each location while the eye position was sampled at a rate of 1000Hz for 100 ms near the middle of this interval. The experiment consisted of two main phases, the perceptual phase and the imagery phase, followed by the ratings of vividness for each image and a spatial memory test. Vividness was rated on a five-step scale, each step defined with the same descriptions used in the VVIQ survey. The spatial memory test grid appeared on the screen in the same position of the checkerboard and was identical in size to that used in the previous phases: however, each of its squares was white and contained a number (from 1 to 36, starting from the leftmost and topmost position). The spatial memory test consisted

of naming 5 digits, from the grid, that the subjects judged to correspond to the positions of each of the five black squares seen previously within the checkerboard. Eye movements were recorded for 20 sec both in the perceptual and imagery phase. At the beginning of the experiment a prototype of the stimuli and the empty diagram used in the imagery phase was shown on the screen while the experimenter reread the instructions for the task to the subject. Each test pattern was presented only once, for a total of 8 trials in the experiment. Each trial consisted of the following sequence of events: 1) a fixation cross appeared at the center of the screen for about a second; 2) a diagram was presented for a duration of 20 sec.; 3) an empty screen was presented on the monitor for circa 40 sec, in order to prevent afterimages, while the instructions for the imagery task were repeated to the subject; 4) an empty diagram (i.e., the same grid used as the stimulus but with no black squares) was presented for 20 s while the participant formed and maintained an image of the previously seen stimulus; 5) the participants were then asked to rate the vividness of the image in the imagery task, using the five-step rating scale; 6) the memory test diagram, with a number in each cell, was then presented on the screen and subjects reported the five digits that they thought corresponded to the previously seen and imagined black squares. At the end of each session, a recording was also made while the subject looked in a sequence to each of the 9 calibration points, so as to check the quality of the optical alignment. Finally, a debriefing session was conducted at the end of the experiment to ensure that the subject was naïve about the hypotheses of the experiment.

2.2. Results

The original stimulus' area was subdivided in 3×3 matrix of squares, each square corresponding to an area of interest, called hereafter Regions 1 to 9. Each of these larger squares included 4 squares of the checkerboard actually seen and imaged by the subjects. This coarser grid of 9 areas of interests was preferred to a grid corresponding to the exact grid of 36 squares because 1) subjects may fixate in some circumstances onto "the center of mass" of a cluster of a few elements and 2) if imagery recordings may preserve the overall pattern of Fixations of perceptions at the same time they may introduce some variability and distortion (compressions, expansions, directional bias and drifts) of the original scanpath pattern (cf. Brandt and Stark, 1997). Hence, we believe that adopting a not too conservative measure, where the immediate neighborhood of an original element is taken into account instead of the exact area of each individual squares, may provide a better way of revealing regularities between perception and imagery scanpaths. Moreover, each region of interest was selected by using the I-view software and this software currently constrains such analyses to the upper limit of sixteen separate regions of sampling (i.e., below the number of cells in our checkerboard); nevertheless, I-view software computes precisely the percentage of time the eye spent in each of the defined regions. For those subjects requested to maintain fixation in the center, trials in which the eye-tracker revealed a failure to comply with the instructions were eliminated from the analysis (i.e., 1 trial in 3 subjects).

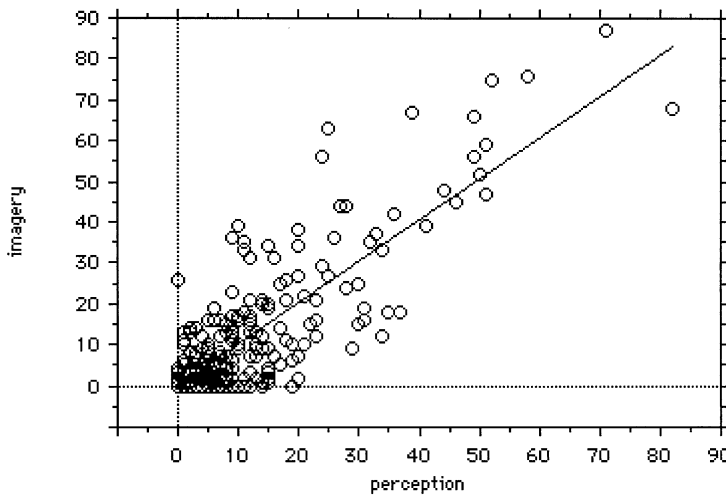


Fig. 2. Experiment 1. Free Vision: Simple regression of % of time in each area of interest during perception and imagery.

2.2.1. Relationship between the percentage of time of fixations on each region during perception and imagery

We first obtained descriptive statistics for each subject. Means of the percentage of time spent in each of the 9 regions of interest were calculated for the perceptual and the imagery phases and each subject's data were pooled over all trials in each phase. The obtained means were then used as variables in separate simple regression analyses, one for each experimental group, with % of time spent in each region in the perceptual phase as the regressor and % of time spent in each region in the imagery phase as the dependent variable.

The regression analysis for the Free Vision condition revealed a strong linear relationship between the percentage of time spent by the eye in each area of interest during the perceptual and the imagery phases; slope coefficient = 1.01, $t(285) = 25.7$, $p < 0.0001$; R-squared = 0.7. The regression plot is shown in Fig. 2.

The same regression analysis was applied to the Central Fixation data. The analysis revealed a significant regression between perception and imagery; slope coefficient = 0.9, $t(286) = 44.1$, $p < 0.0001$; R-squared = 0.9. The regression plot is shown in Fig. 3. As it appears from the graph, the values appear to crowd at the two ends of the measures, indicating that subjects spent nearly the whole recording time looking at the same location. This pattern of results was to be expected if gaze was fixed at the same central location not only during perception but also during imagery. The graph also shows the presence of a few outliers in the two other corners of the quadrant, meaning that there were locations that were not fixated during perception but fixated for a considerably high % of time during imagery and vice versa. These anomalous trials were too rare to affect the regressions between perception and imagery. They reflected occasional lateral drifts of fixations that occurred during imagery into one of the two areas of sampling that were laterally adjacent to the central fixation area; as revealed when each of these outliers was examined individually.

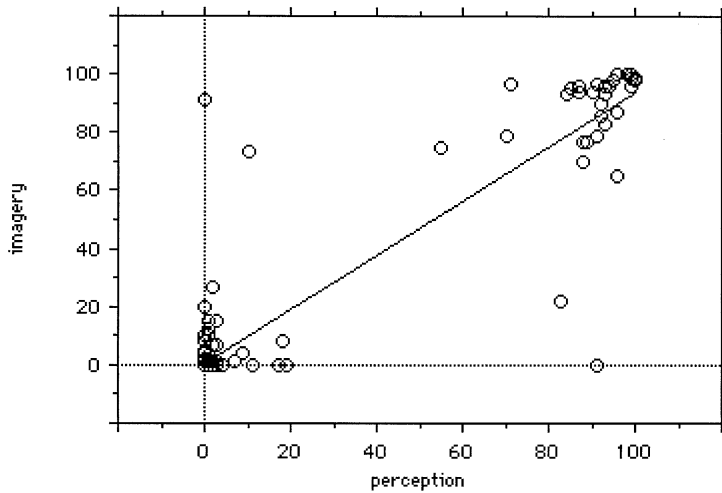


Fig. 3. Experiment 1. Central Fixation: Simple regression of % of time in each area of interest during perception and imagery.

Table 1 shows the mean percentages of time spent in each of the 9 regions of interest for the two experimental groups during the perception and imagery phases.

2.2.2. Relationship between perception and imagery scanpaths

In a separate analysis, the order of scanning of each element was taken into account in order to assess whether imagery reenacts the same order of eye movements used during perception. By necessity, this analysis was performed on the Free Vision data only. A fixation was defined as duration of gaze into one of the specified regions for 300 ms or longer. According to these criteria, subjects in the Free Vision group made a roughly equal number of fixations in each trial during the perceptual phase (average 12 fixations, $SD = 3$)

Table 1

Experiment 1. Mean percentages of time spent in each of the 9 regions of interest for each experimental group (free vision and central fixation) during perception and imagery. Region 5 corresponds to the center of the grid and region 1 to the leftmost and topmost.

Region	Free vision		Central fixation	
	Perception mean %	Imagery mean %	Perception mean %	Imagery mean %
1	5.9	5.8	0.2	0.0
2	6.3	5.8	0.2	0.4
3	3.3	3.9	0.0	0.0
4	8.4	11.1	3.8	1.4
5	35.2	39.4	87.3	88.1
6	5.6	6.4	1.7	3.9
7	8.7	6.5	0.8	0.7
8	13.2	10.5	1.1	2.1
9	4.4	3.7	1.0	0.1

and the imagery phase (average 14 fixations, $SD = 5$). Serial order of scanning was obtained first for the perceptual phase of each trial; namely, the regions of the pattern corresponding to the first to the ninth fixation were identified. Then we coded the imagery condition with respect to the specific regions fixated in the perception condition. To provide an example, if regions 6, 3, and 1 were fixated during perception in fixations 1, 2 and 3, respectively, then we looked at the fixations during imagery and found the first occurring fixation on region 6 and entered the serial number of the fixation as the entry under imagery (e.g., the 4th fixation in imagery may have been the first occurrence of a fixation in region 6). In some cases, a later fixation could also return to a location previously visited (e.g., the 9th fixation during perception was the second occurrence of a fixation on region 1), this would be coded accordingly and paired to the second occurrence of a fixation during imagery on region 1 (e.g., the 8th fixation). Preliminary analyses had shown that limiting the analysis to the first 9 nine fixations during perception (i.e., one standard deviation below the average number of fixations in the perceptual phase) would allow us to always identify the serial position during imagery of a fixation in the same region. Above the limit of nine fixations, it would become increasingly less likely to observe in the imagery data the corresponding n^{th} occurrence of a fixation in the same region.

Clearly, if the scanpaths of imagery reenact those of perception we would expect the two serial ordering of fixations to be highly related. Thus, the obtained serial orders of fixations were used as variables in a simple regression analysis with serial order in the perceptual phase as the regressor and serial order for the same regions in the imagery phase as the dependent variable. This regression analysis indicated a positive linear relationship between the two variables; slope coefficient = 0.7, $t(285) = 6.9$, $p < 0.0001$; $R\text{-squared} = 0.4$.

2.2.3. *Spatial memory test's accuracy and its relationship to the perception and imagery scanpaths*

Means of the correct responses (i.e., the named digit corresponded to the location of a black square in the grid) were calculated and the data were pooled over all the trials. The obtained scores were then entered as cells in repeated-measures ANOVA with Group (Free Vision, Central Fixation) as the between-subjects factor and accuracy of recall as the dependent variable. This analysis showed that there were no differences in terms of accuracy between the Free Vision (Mean = 3.9, $SD = 1.4$) and the Central Fixation groups (Mean = 3.8, $SD = 1.5$). However, a separate analysis on the Free Vision data showed that the degree of relatedness between perception and imagery scanpaths in each individual trial predicted the degree of accuracy of spatial recall for the trial. In this analysis, the obtained serial orders for each trial were used as variables in separate simple regression analyses with serial order in the perceptual phase as the regressor and serial order for the same regions in the imagery phase as the dependent variable. Thus, a total of 32 regression analyses were performed. The obtained R scores for each of these regression analyses were then subsequently used as the regressor in a final regression analysis with accuracy in the corresponding trial as the dependent variable. We reasoned that if scanpaths play a functional role in generating an image then the accuracy of recall during the memory test should vary accordingly to the strength of relatedness of the perceptual and imagery scanpaths. As expected, the analysis indicated a positive linear relationship; slope coefficient = 3.9, $t(31) = 2.8$, $p < 0.001$;

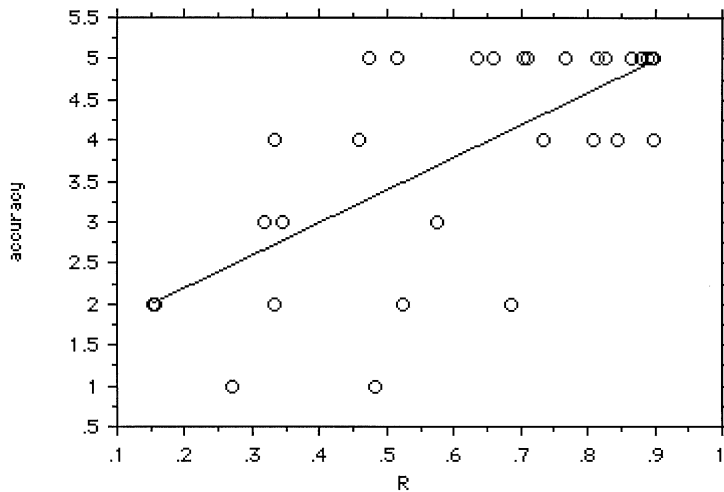


Fig. 4. Experiment 1. Simple regression between the degree of relatedness (R scores) in each trial between perception and imagery scanpaths and accuracy of recall.

R-squared = 0.5. Fig. 4 illustrates this relationship and Table 2 lists the R scores and slope coefficients for each trial and the corresponding accuracy score.

2.2.4. Vividness ratings

Means of each subject's vividness ratings (i.e., a value from 1 to 5) were calculated and data were pooled over all the trials. The obtained means were then entered as cells in repeated-measures ANOVA with Group (Free Vision, Central Fixation) and accuracy of recall as the dependent variable. This analysis revealed no significant difference ($F < 1$) between the groups in this measure of vividness (Free Vision: Mean = 3.0; SD = 1.1; Central Fixation: mean = 2.7; SD = 1.3).

2.2.5. Relationship between the vividness ratings, VVIQ scores, and spatial memory test

Both measures of vividness would be expected to predict accuracy in the visual spatial memory test. A simple regression analysis between each subject's score in the VVIQ Survey and their mean vividness ratings in the imagery phase found no relation (slope coefficients = 0.01, R-Squared = 0.01). However, a simple regression analysis between vividness scores in each trial as the regressor and accuracy of recall in each trial as the dependent variable showed a positive relation; slope coefficient = 3.9, $t(63) = 26.2$, $p < 0.0001$; R-squared = 0.3.

2.3. Discussion

The main conclusion from these findings is that the eye scanpaths during visual perception (i.e., sequences of eye fixations on specific regions of space corresponding to elements of a pattern) are highly correlated to those during visual imagery of the same visual object. These findings replicate those of Brandt and Stark (1997), who used similar stimuli, and support the

Table 2

Experiment 1 R scores, slope coefficients (B score) and accuracy scores in each of the eight trials for each of the four subjects of the Free Vision group

Subject	Trial	R score	B score	Accuracy
1	1	0.71	0.87	5
1	2	0.16	0.04	2
1	3	0.16	0.28	2
1	4	0.52	0.54	5
1	5	0.89	0.9	5
1	6	0.74	0.95	4
1	7	0.32	0.3	3
1	8	0.77	0.72	5
2	1	0.48	0.35	5
2	2	0.67	0.78	5
2	3	0.64	0.75	5
2	4	0.46	0.49	4
2	5	0.58	0.57	3
2	6	0.34	0.67	4
2	7	0.9	1.05	5
2	8	0.82	0.67	5
3	1	0.71	0.62	5
3	2	0.49	0.49	1
3	3	0.87	1.05	5
3	4	0.71	0.87	5
3	5	0.88	0.87	5
3	6	0.9	1.03	5
3	7	0.35	0.4	3
3	8	0.28	0.27	1
4	1	0.53	0.52	2
4	2	0.69	0.74	2
4	3	0.81	0.67	4
4	4	0.83	0.89	5
4	5	0.85	0.95	4
4	6	0.9	0.77	5
4	7	0.34	0.37	2
4	8	0.9	0.8	4

hypothesis that the oculomotor behavior during imagery reenacts that which occurred when perceiving the object (see Fig. 5). Interestingly, when the relatedness of the two serial orders of fixations of perception and imagery (of subjects free to move their eyes during perception) was analyzed on a trial-by-trial basis and the obtained R scores for each of these regression analyses were in turn regressed in relation to each trial's memory accuracy, a positive linear relationship was found. Such a correspondence between scanpaths and recall does not in itself constitute unambiguous evidence for the functional role of eye movements in generating an image. However, if recall did not increase with the strength of relatedness of the perceptual and imagery scanpaths, this would pose a problem for the functional theory.

An important finding was that subjects who were asked to maintain their gaze focused on a central, narrowly defined, location kept their eyes fixed in the same location also during imagery. This behavior supports the idea that eye movement in imagery do not necessarily mirror the programming of covert shifts of attention while "inspecting" or scanning an

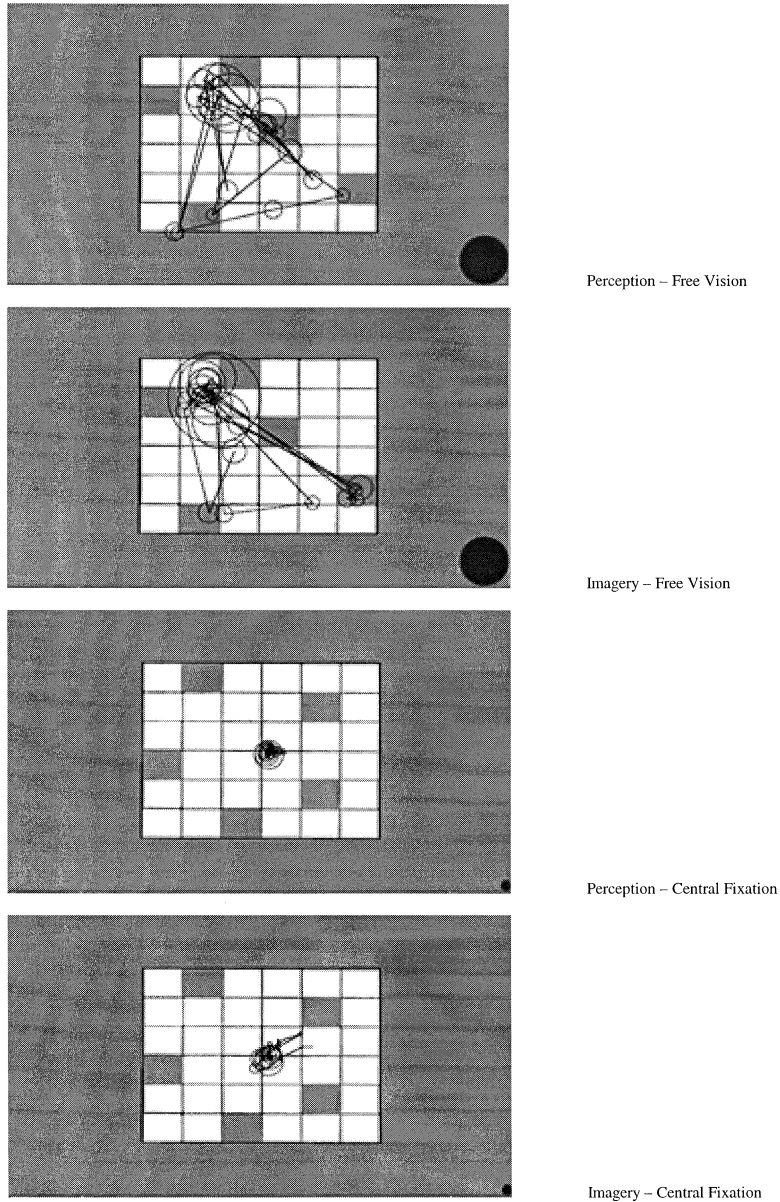


Fig. 5. Experiment 1. Examples of scanpaths from each of the two conditions during perception and imagery (each circle represents a separate fixation and lines indicate the scanpath; the bottom right circle is a time reference area corresponding to 1 s).

internal image. It is reasonable to suppose that subjects asked to maintain gaze in a location do nevertheless covertly scan the image by shifting rapidly the attention window (cf. Posner, 1978; Posner, Nissen & Ogden, 1978). Thus, oculomotor behavior during imagery is more likely to be related to the phase of “image generation” than “image inspection.” In this view,

the original direction (at perception) of the eyes' axis is used as an index for computing the coordinates of different parts of an object or scene while constructing the image. The epiphenomenal theory encounters considerable difficulty in explaining why the Central Fixation subjects kept their gaze centrally fixed also during imagery. Such an account could assume that eye movements, when they occur, always mirror the shifts of an attention window scanning an internal image, but our subjects chose to suppress these spontaneous eye movements. However, such explanation seems obviously posthoc and there is no principled reason for suppressing eye movements. It should be noted that no restriction on eye movements was imposed to subjects in the imagery phase. Above all, the epiphenomenal account fails to explain why a putative suppression of spontaneous eye movements would result with high consistency, across trials and subjects, in fixations onto the same central location. In contrast, the functional theory provides a clear account of the behavior of the Central Fixation group because it predicts the lack of eye shifts or reenactment of static gaze.

The average spatial memory scores of the Free Vision and Central Fixation groups were equally accurate. At a first glance, this is a bit puzzling because Loftus (1972, 1981) has shown that when pictures are viewed for a fixed amount of time (as it was the case in our experiment), memory performance is a positive function of the number of separate fixations in the picture. However, because in our experiment the central fixation point coincided with the center of the grid displays, a null finding may be not so surprising. A large portion of the display fell within the area of central vision and it is likely that the peripheral portions of the pattern subjects could be scanned covertly rather efficiently. Alternatively, subjects in the Central Fixation condition may have chosen to increase the size of their attention window so as to encompass the entire display. Such alternative strategies could have been used either by different individuals or a same individual in different trials; nevertheless, the functional theory would predict that during imagery these subjects would reenact the same behavior of the encoding phase, whether covert or overt. According to the functional theory, covert shifts are encoded in reference to the eyes' coordinates, even when these consist of just one fixed point.

3. Experiment 2

The main goal of the following experiment was to replicate the finding that during imagery the eyes reenact their behavior during perception. The same overall paradigm of the previous experiment was used in the second experiment. However, a novel condition was introduced: A third group of subjects were instructed to view freely the stimuli during the initial perception phase but then requested to maintain central fixation during the imagery phase. We reasoned that when eye movements are free at first to explore the pattern but then, during imagery, prevented or forced into a static oculomotor pattern, the process of image generation should be disrupted to some degree and accuracy of recall should be impaired. Another change introduced in this experiment concerned the type of stimuli used. Previously, we had used stimuli that were both novel, without meaning, and with a regularly defined spatial layout and structure (i.e., grids). As pointed out by Brandt and Stark (1997), it is not clear whether the tight correspondence between scanpaths in perception and imagery would

be also observed with familiar object stimuli. Many classes of objects are also rather variable in shape. Natural kinds (e.g., birds, fish) can be seen as variations around a pattern prototype (Rosch et al., 1976; Edelman & Duvdevani-Bar, 1997), with the exemplars in each class showing even considerable variations in shape, size, texture, and color.

In the following experiment, subjects viewed in each trial a color picture of a different tropical fish, which was always depicted in a side view and positioned in one of the four corners of the computer screen. During the imagery phase, subjects were asked to build an image of the fish just seen, while keeping their eyes open. Once they reported having formed a clear image, the experimenter would question them about whether the fish possessed a certain property or not (e.g., whether the color of the tail was yellow, whether it was swimming leftward, whether there was a round spot on the back). According to Kosslyn (1980, 1994), perhaps the most basic property of imagery is that images make accessible the local geometry and other visual properties of an object or scene by activating a short-term memory representation of stored visual information. This can help to recover properties and relations that are implicitly contained in the image but may have never been encoded explicitly as such (as, for example, when answering a question like: “What shape are a German Shepherd’s ears?”).

Another key difference between the previous and the following experiment was that the object to be imaged appeared in a position (one corner) that never coincided with the location of fixation (the center of the screen). This manipulation would have the effect to reduce the benefit that the Central Fixation group could have had in the previous experiment, when visualizing a pattern that had been seen mainly in central, focal, vision. Moreover, the present manipulation would render unlikely that subjects in the Central Fixation group could solve the task by adjusting the width of the attentional window to encode the pattern as a whole.

To summarize, in this experiment, subjects in one group were allowed to scan freely the picture shown on the screen and no constraints were imposed on gaze in the imagery phase (the Free Perception & Free Imagery group). Given that on the screen there was only one salient pattern, the tropical fish, and the rest was a uniform background, we expected this subjects to move their eyes from the central fixation position onto the area occupied by the fish pattern and scan its details. We surmise that such spontaneous scanpaths would allow the encoding of fine properties of the visual stimuli in long-term visual memory (e.g., the fishes’ global shape, the shape of parts, their color, etc.). In contrast, subjects in a second group (Fixed Perception & Free Imagery, corresponding to the Central Fixation group of the previous experiment) were requested to fixate the center of the monitor during perception. Thus, these subjects would necessarily view the target pattern in peripheral vision. Nevertheless, holding gaze in a fixed location does not prevent the attention window from covertly scanning a peripheral stimulus. Therefore, we would expect these subjects also to be able to recover properties of the visual stimuli from their long-term visual memory when building a visual image. However, for this group the encoding of the object’s parts would be indexed in relation to a single location of gaze, as discussed earlier. Hence, the two groups should reenact during the retrieval phase of visualization the two different ocular behaviors of their respective perception

phases. A third group (Free Perception & Fixed Imagery) viewed freely the stimuli during perception but maintained central fixation during the imagery phase. This should disrupt the image generation process and accuracy of recall should be impaired relative to the other experimental groups.

3.1. *Methods*

3.1.1. *Subjects*

Twelve students at the University of Tromsø, six females and six males, volunteered to participate as paid participants in an experiment on mental imagery. All participants reported having normal or corrected to normal vision (with contact lenses) and their age range was 21–44. None of the subjects had specific knowledge of the tropical fish species shown as stimuli. The participants were randomly divided between three experimental groups, four subjects in each group; namely the Free Perception & Free Imagery group, the Fixed Perception & Free Imagery group, and the Free Perception & Fixed Imagery group.

3.1.2. *Apparatus and stimuli*

Eye movements were recorded by means of the same remote eye-tracking device (built by SMI-SensoMotoric Instruments) used in the previous experiment. A list of 22 questions was prepared, each question probing one visual property for each of the 22 images of fish used during the experiment. The properties could regard the fish's shape, color, or direction of (apparent) movement. Every question required a simple answer of one or two words or a "yes/no" response. For example, one question was: "How many black stripes had the fish?" Correct answer: "Two." The stimuli were color photographs of tropical fish (taken from a "photo safari" guide). All pictures were about of the same size, 2×3 cm, covering an area of 2° of the visual field. The stimuli were first digitized by use of an Agfa scanner and images were edited and formatted by use of Adobe Photoshop software. Each stimulus was presented on a 49 cm flat color monitor. Each fish shape appeared at 10° from the center, in one of the four corners of the screen. Fig. 6 illustrates one of the stimulus presentations. In the middle of the screen there was a white fixation cross covering about 0.5° . The background color of the screen was set to blue; which gave a somewhat natural aspect to each image as if observing a still of a fish swimming in the ocean. The stimuli presentation was done using software ACDSee 32v2.4. The "blank" screen used in the imagery phase had the same blue background color, because Intons-Peterson and Roskos-Ewoldsen (1989) have found that subjects can form images of color objects more easily when the objects are visualized against the same color background.

3.1.3. *Procedures*

At the beginning of the session each subject read the instructions pertaining their experimental group. The Free Perception & Free Imagery group read the following: "Look carefully at the picture of the fish that will appear on the monitor and try to remember it as accurately as possible." Those of the Fixed Perception & Free Imagery group were: "Keep

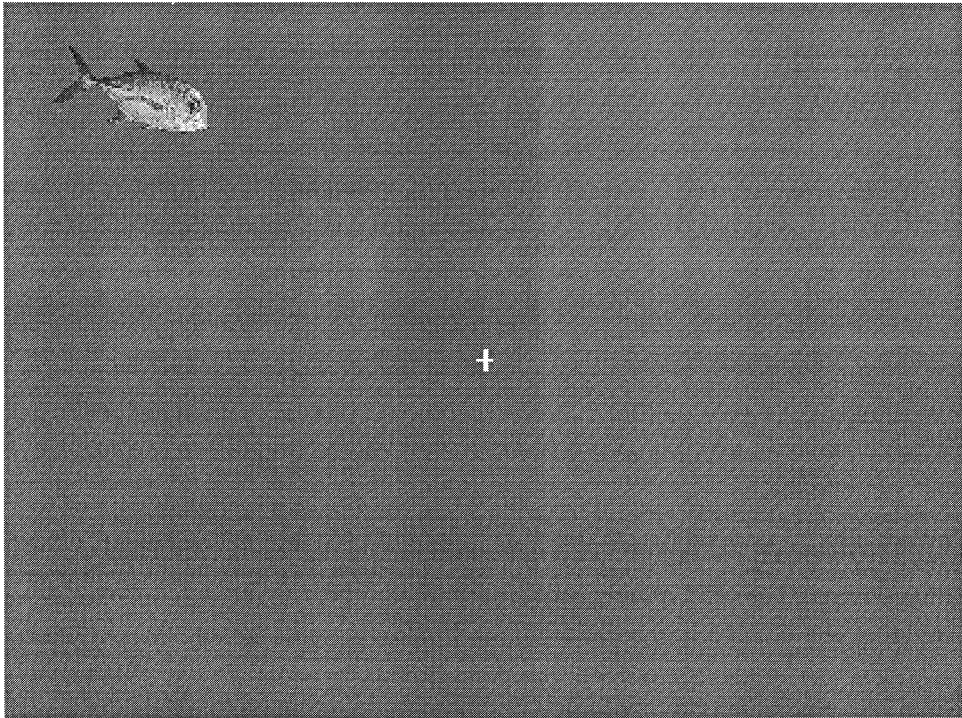


Fig. 6. Experiment 2: Example of one fish stimulus.

your eyes focused on the cross in the middle of the monitor, and try to remember everything that is presented on the monitor.” The instructions for the imagery task were common for both groups: “ Build an image of what you just saw earlier, while keeping your eyes open.” Those of the Free Perception & Fixed Imagery reversed the perception and imagery requirements of the previous group. After reading the experiment’s instructions, each participant was seated comfortably on a chair with the head placed in a chin-and-forehead rest apparatus, centered 50 cm away from the monitor. After the calibration procedure, the experimental session started immediately. The experiment consisted of two practice trials and twenty experimental trials. Eye movements were recorded for 15 sec in the perceptual phase; whereas, in the imagery phase eye movements were recorded from the presentation of the central cross and blue background until the subject gave a verbal response. Each trial consisted of the following sequence of events: 1) a fish appeared for 15 sec, following a semirandom order, in one of the corners of the monitor while a white cross of 0.5° appeared in the middle; 2) a “blank” blue screen was then presented on the monitor. The subject was reminded to image what had been seen previously while keeping the eyes open; 3) the subject indicated that a clear image had been constructed in their minds, at this point the experimenter asked the question probing a property or detail of the fish; 4) the subject answered the question while still holding the image in mind and the experimenter terminated the recording of the imagery phase.

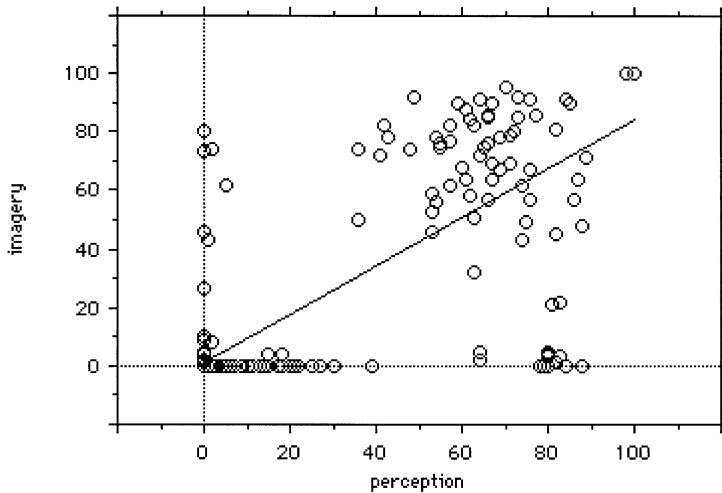


Fig. 7. Experiment 2. Free Perception & Free Imagery group: Simple regression of % of time spent in the regions of interest during perception and imagery.

3.2. Results

3.2.1. Relationship between perception and imagery

We first obtained descriptive statistics for each subject. Means of the % of time spent in each of 5 regions of interest were calculated for both the perceptual and the imagery phases and each subject's data were pooled over all trials in each phase. These 5 regions of interest corresponded to a square area of 5° of diameter surrounding each of the 4 possible locations of the stimuli plus a square area of 5° surrounding the central fixation point. The obtained means were then used as variables in separate simple regression analyses, one for each experimental group, with % of time spent in each region in the perceptual phase as the regressor and the % of time spent in each region in the imagery phase as the dependent variable. The regression analyses for the Free Perception & Free Imagery and Fixed Perception & Free Imagery groups showed a clear linear relationship between fixations during perception and imagery. Subjects in the Free Perception & Free Imagery condition showed a slope coefficient = 0.9, $t(398)=31.8$, $p < 0.0001$; R-squared = 0.7. The regression plot is shown in Fig. 7. We also performed an additional regression analysis on the Free Perception & Free Imagery data, excluding from the data sample all fixations to the quadrant corresponding to the central fixation because including these data may spuriously increase the slope of the regression line between perception and imagery. This regression analysis confirmed a linear relationship between perception and imagery; slope coefficient = 0.9, $t(318)=30.1$, $p < 0.0001$; R-squared = 0.7.

The analysis on the Fixed Perception & Free Imagery group showed a slope coefficient = 0.9, $t(398)=93.5$, $p < 0.0001$; R-squared = 0.9. The regression plot is shown in Fig. 8. Also in this case, an additional regression analysis was performed on the fixations to the central quadrant only, excluding fixations to the 4 peripheral quadrants

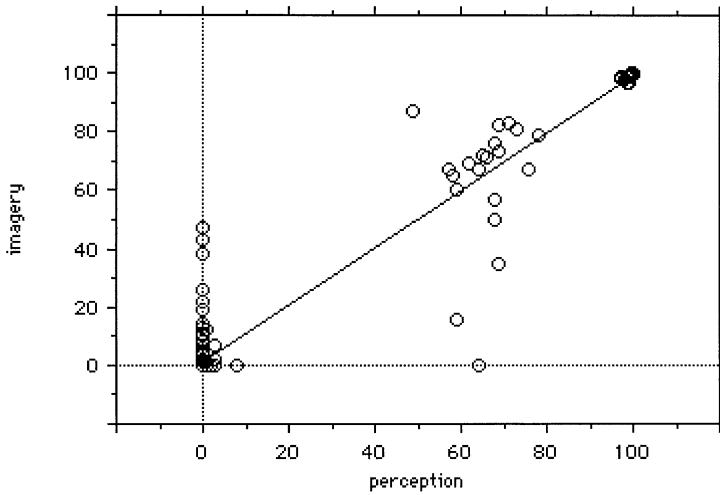


Fig. 8. Experiment 2. Central Fixation: Simple regression of % of time spent in the regions of interest during perception and imagery.

corresponding to the stimuli locations. This regression analysis confirmed the relationship between perception and imagery; slope coefficient = 0.8, $t(78)=14.7$, $p < 0.0001$; $R\text{-squared} = 0.7$. Finally, as expected, the regression analysis on the Free Perception & Fixed Imagery group showed no relation between eye fixations during perception and imagery; $R\text{-squared} = 0.02$. Subjects clearly complied with the instructions and only during 5 trials (2 in one subject and 1 in 3 subjects) did eye movements occurred away from the center. These trials were excluded from the regression analysis. Table 2 lists means and SDs of the percentage of time spent within the 5 regions by each group in the perception and imagery phase.

3.2.2. Accuracy of retrieval of the visual properties

Means of each subject's correct responses to the questions probing a visual property of each fish were calculated and the obtained means were then entered as cells in repeated-measures ANOVA with Condition (Fixed Perception & Free Imagery, Free Perception & Fixed Imagery, Free Perception & Free Imagery) as the between-subjects factor and the number of correct answers as the dependent variable. This analysis showed a significant effect of Condition; $F(2,9)= 6.2$, $p < 0.02$. A posthoc Fisher's LSD test (Critical Difference = 2.26) showed that the Free Perception & Free Imagery group (Mean= 18.1, SD= 1.0) was significantly more accurate than the Free Perception & Fixed Imagery group (Mean= 15.0, SD= 1.4), $p < 0.007$. The difference between the Free Perception & Free Imagery group and the Fixed Perception & Free Imagery group (Mean= 16.5, SD= 1.7) approached significance, $p < 0.07$. Instead, the difference between the Free Perception & Fixed Imagery and the Fixed Perception & Free Imagery groups failed to reach significance, $p < 0.17$. Fig. 9 illustrates the performance of each of the three experimental groups.

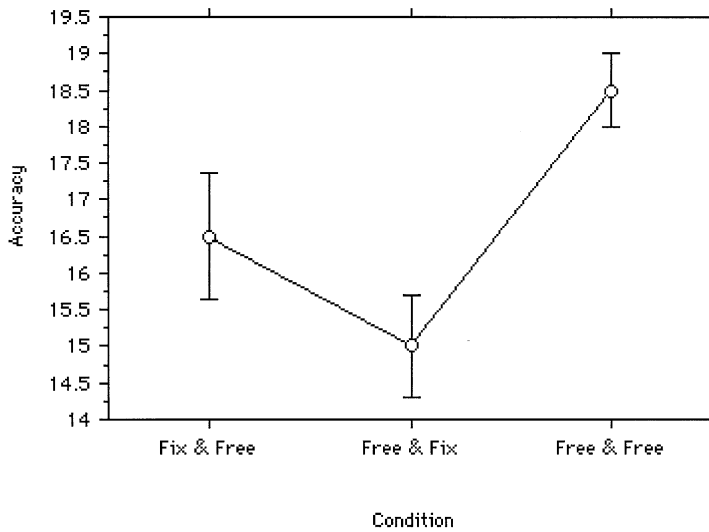


Fig. 9. Experiment 2. Mean % accuracy of recall scores (and standard errors) for subjects who were requested to maintain central fixation during perception but not during imagery (Fix & Free = Fixed Perception & Free Imagery), subjects who were free to view the pattern during perception but requested to maintain central fixation during imagery (Free & Fix = Free Perception & Fixed Imagery), and subjects who were free to move their eyes both during perception and imagery (Free & Free = Free Perception & Free Imagery).

3.3. Discussion

The second experiment replicated successfully the findings of the first experiment. The eye fixations recorded during the imagery phase reoccurred over the same regions of space of the eye fixations during perception in the Free Perception & Free Imagery group. Also, subjects in the Fixed Perception & Free Imagery condition did not show scanpaths in the direction of the probed pattern's position but they reenacted the central fixation behavior that had been required during the perceptual phase. Moreover, subjects in the Free Perception & Fixed Imagery condition demonstrated that constraining eye movements during imagery but not during perception resulted in a loss of accuracy of retrieval from memory of the object's properties. Hence, our conclusion is the same reached earlier: The findings lend support to a functional theory of image generation, according to which the visual system reenacts the same oculomotor behavior that occurred at encoding and this oculomotor behavior assists the construction of the mental image.

4. General discussion

The aim of the present study was to investigate whether oculomotor information encoded together with the visual information at perception is reenacted at retrieval of the visual information or during imagery. Visual encoding was also manipulated by controlling the direction of gaze. Specifically, in one condition the subjects were allowed to make eye

movements freely whereas in the other condition they were not. The main hypothesis was that if oculomotor information is encoded during the visual perception task and it is used as spatial reference in the process of image generation, then the same pattern of oculomotor activity should also be present at retrieval of the visual information. Moreover, we posited that oculomotor information plays a functional role to the image generation process; therefore preventing eye movements during imagery but not perception will disrupt the process of image generation and the ability to retrieve information from the mental image.

The results of both experiments showed highly correlated patterns of oculomotor activity between perception and imagery in both conditions. Specifically, at imagery retrieval, scanpaths occurred over the same regions of space (and in the same order, as shown in Experiment 1) as during the encoding phase of the visual shape. If it is correct to think that the retrieval of a pattern is a “reconstruction” of the original visual image as encoded during oculomotor activity, then according to Hebb’s original proposal (1968), oculomotor activity becomes a relevant aspect of the retrieval component and it is important in assisting the operation of recombining together the pieces of information encoded in several different areas of the brain. Noton and Stark (1971a, b) proposed a “scanpath theory” of perception and imagery, where an image is internally represented as a sequence of sensory and motor activities. Similarly to Hebb’s proposal, the eye movements during imagery would reflect, according to Noton and Stark, the mental process of activating and arranging parts of an imaged object into their proper location. Kosslyn (1994) has distinguished several types of visual imagery, among which he has proposed forms of imagery that are “attention-based.” These types of imagery would require engaging attention at each different location of a multipart image, which would allow each part or component to be added in the correct location onto the image under construction. Because attention-based imagery depends on allocating attention to specific regions of space, this form of imagery may be easier and most effective when visual structure is provided by the environment rather than having to be supplied in imagery. Clearly, within this account, shifts of attention (either overt, like eye movements, or covert) are a direct reflection of the image generation process. Specifically, motor-based information about the eyes’ positions is retrieved and the resulting eye movements follow necessarily an active process of image construction. Moreover, this type of imagery would require the encoding of spatial relation descriptions, which are then used to arrange the parts in an image or scene (Kosslyn, 1987).

The fact that imagery generation can recapitulate a motor or executive program is also exemplified by an experiment of Kosslyn, Cave, Provost and Von Gierke (1988). This study used a task devised by Podgorny and Shepard (1978) where subjects were asked to visualize block letters in grids and decide whether a probe mark (presented on an empty grid) would have fallen on the letter if this were actually present. Kosslyn et al. had observed prior to the experiment how individuals wrote each block letter and selected those that were drawn in a consistent way. Probe marks were presented before the subjects were able to finish visualizing a letter. It was found that “late” letter strokes in drawing a letter’s segment could be probed only by “late” marks, whereas “early” strokes could be accurately probed by most of the probe marks. Indeed, for accurate performance, the timing and location of each probe was highly related to the sequence of hand movements. Thus, this study already showed a

Table 3
Experiment 2. Means and SDs of % of time spent within the 5 sampled regions by each group in the perception and imagery phase

Region	Free perception & free imagery				Fixed perception & free imagery			
	Perception		Imagery		Perception		Imagery	
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
Center	5.8	8.9	3.1	12.7	90.8	15.1	90.1	19.3
Lower left	21.7	33.3	22.5	36.0	0.0	0.0	0.9	4.6
Lower right	14.1	26.1	13.7	26.9	0.0	0.0	1.3	4.7
Upper left	36.2	51.5	30.8	52.4	0.0	0.0	2.3	7.7
Upper right	18.4	31.7	14.9	29.7	0.0	0.0	0.4	1.7

Region	Free perception & fixed imagery			
	Perception		Imagery	
	MEAN	SD	MEAN	SD
Center	4.9	8.9	95.1	1.8
Lower left	26.9	33.3	0.0	0.0
Lower right	16.1	26.1	0.0	0.0
Upper left	33.4	51.2	0.0	0.0
Upper right	16.8	31.7	0.0	0.0

relationship between the image generation process and properties of a motor program (in this case, the writing of each letter of the alphabet).

A “classic” objection (Pylyshyn, 1981; Intons-Peterson, 1983), often leveled to the whole paradigm of imagery experiments, could be raised for the present study; namely, that subjects are complying with “demand characteristics” of the experiment and recapitulate the behavior of perception simply on the basis of their intuition of what the experimenters are expecting them to do or by “simulating” their past perceptual behavior. Many responses have been given in the past to this type of criticism (e.g., Finke & Pinker, 1982, 1983; Jolicoeur & Kosslyn, 1985). In the specific case of the present study, we are skeptical that such explanation can account for both the character and the complexity of the findings. First, subjects in the Central Fixation group reenacted the previous fixation behavior; despite the fact that they were given clear imagery instructions (i.e., no constraints on gaze) that run contrary to those used for the perceptual phase. These instructions should have either introduced the opposite “demand characteristic” or released the subject from compliance to the previous instructions. Second, as pointed out by Brandt and Stark (1997), to recreate voluntarily a complex scanpath, like that seen in the Free Vision condition (see Fig. 5), seems an unlikely motoric memory feat; in fact, conscious proprioceptive information provided by the eye muscle spindles is extremely weak (Bridgeman & Stark, 1991).

Alternative accounts, somewhere in between the functional and the epiphenomenal ones, could also be taken into consideration. For instance, one could state that eye movements’ programming and execution still play no informational role about the structure of the image but eye movements interfere with the activity of the visual cortex by triggering a mechanism of “saccadic suppression” (i.e., “erasing” input in the visual cortex while the eye is in transit).

Thus, in this view, if images are pattern of activation in the visual buffer (or the visual cortex; see Kosslyn, 1994), eye movements could actually assist the process of suppressing a mental image. This account would actually predict that, when subjects engage in imagery tasks, eye movements are inhibited. This prediction is clearly at variance with the current findings as well as those of previous studies but, apparently supporting this account, Weiner and Ehrlichman (1976) found that spatial questions (e.g., “how many windows are there in your house or apartment?”) elicited fewer eye movements than did verbal questions (e.g., “define these words: charity, liquorice”). They concluded that eye movements do not reflect a process of image scanning and, instead, proposed a limited-capacity account (cf. Antrobus, 1973; Singer et al., 1971; Glenberg et al., 1998) in which eye movements are inhibited so as to “gate out” visual and motoric input during highly demanding processing task (like visuo-spatial imagery). However, a close look at several of the questions used by Weiner and Ehrlichman casts doubts on their assumption that their “spatial” questions required visuo-spatial imagery to a greater extent than their “verbal” questions. For example, “define a narrow cylinder with lenses at both ends” from the “verbal” set of questions may actually prompt subjects to build a multipart image of a telescope; and there is little doubt that three-term series problems like “Jim is better than Ted; Ted is better than Bob; who is best?” are often solved by building a mental model of the terms (cf. Huttenlocher, 1968; Johnson-Laird, 1983; Kaufmann, 1984). The cost in memory retrieval for the Free Perception & Fixed Imagery subjects of Experiment 2 also argues against the above idea that quiescence of gaze (due to inhibition or “gate out”) helps maintaining the current mental image. Moreover, such an account fails to explain why subjects in both the Central Fixation group of Experiment 1 or Fixed Perception & Free Imagery group of Experiment 2 consistently showed quiescence of gaze whereas those in the Free Vision or Free Perception & Free Imagery groups consistently chose scanpaths towards the targets’ locations.

To conclude, Hebb (1968) may have been the first psychologist to propose explicitly an inherent oculomotor component in visual perception and visual imagery, claiming that if imagery is the reinstatement of a perceptual process, then this process should include eye movements. On this account, oculomotor patterns in imagery are not irrelevant but essential (Hebb, 1968). The present study brings some support to this idea and, generally, to the view that eye movements in visual imagery play a functional, computational role during image generation, against the view that these eye movements simply mirror the internal scanning of the image in an epiphenomenal way. Recently, several researchers and theorists have remarked that perception is “active” and some have claimed that a new paradigm in the study of perception may be emerging; namely, “perceptual activity theory” or, in robotics, “active vision” (Thomas, 1999). According to this account, influenced by the writings of Hebb (1968), Hochberg (1968), Neisser (1967), Gibson (1966, 1979), and of the philosopher Merleau-Ponty, perception is not only about storing descriptions or pictures but is deeply (essentially, in the more radical views) “procedural” in kind; that is, we encode also how to direct attention, examine, and explore an object. Imagery consists in going through the motions of the equivalent perceptual process. Information about the motor control of sensors and transducers (e.g., the hand, the eyes) constitutes therefore a fundamental part of the perceptual experience. More traditional contemporary information processing accounts also view the flow of information as running top-down and bottom-up in an integrated manner

and posit that the control of attention may be essential to several perceptual processes (cf. Kosslyn, 1994; Ullman, 1996; Laeng, Shah & Kosslyn, 1999).

References

- Antrobus, J. S. (1973). Eye movements and nonvisual cognitive tasks. In V. Zikmund (Ed.), *The oculomotor system and brain functions*. London: Butterworth.
- Brandt, S. A., & Stark, L. W. (1997). Spontaneous eye movements during visual imagery reflect the content of the visual scene. *Journal of Cognitive Neuroscience*, 9, 27–38.
- Bridgeman, B., & Stark, L. (1991). Ocular proprioception and efference copy in registering visual direction. *Vision Research*, 31, 1903–1913.
- Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, 36, 1827–1837.
- Deutsch, G., Bourbon, W. T., Papanicolaou, A. C., & Eisenberg, H. M. (1988). Visuospatial experiments compared with activation of regional cerebral blood flow. *Neuropsychologia*, 26, 445–452.
- Edelman, S., & Duvdevani-Bar, S. (1997). A model of visual recognition and categorization. *Philosophical Transactions of the Royal Society of London, Series B*, 352, 1191–1202.
- Finke, R. A., & Pinker, S. (1982). Spontaneous imagery scanning in mental extrapolation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8, 142–147.
- Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston, MA: Houghton Mifflin.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.
- Glenberg, A. M., Schroeder, J. L., & Robertson, D. A. (1998). Averting gaze disengages the environment and facilitates remembering. *Memory & Cognition*, 26, 651–658.
- Hebb, D. O. (1968). Concerning imagery. *Psychological Review*, 75, 466–477.
- Henderson, J., Pollatsek, A., & Rayner, K. (1989). Covert visual attention and extrafoveal information use during object identification. *Perception and Psychophysics*, 57, 787–795.
- Hochberg, J. (1968). In the mind's eye. In R. N. Haber (Ed.), *Contemporary theory and research in visual perception* (pp. 309–331). New York: Holt Rinehart & Winston.
- Huttenlocher, J. (1968). Constructing spatial images: A strategy in reasoning. *Psychological Review*, 75, 556–560.
- Intons-Peterson, M. J. (1983). Imagery paradigms: How vulnerable are they to experimenters' expectations? *Journal of Experimental Psychology: Human Perception and Performance*, 18, 307–317.
- Intons-Peterson, M. J., & Roskos-Ewoldsen, B. B. (1989). Sensory-perceptual qualities of images. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 188–199.
- Irwin, D. E. (1996). Integrating information across saccadic eye movements. *Current Directions in Psychological Science*, 5, 94–100.
- Irwin, D. E., & Gordon, R. D. (1998). Eye movements, attention and trans-saccadic memory. *Visual Cognition*, 5, 127–155.
- Johnson-Laird, P. N. (1983). *Mental models*. Cambridge: Harvard University Press.
- Jolicoeur, P., & Kosslyn, S. M. (1985). Is time to scan visual images due to demand characteristics? *Memory and Cognition*, 13, 320–332.
- Kaufmann, G. (1984). Mental imagery in problem solving. *International Review of Mental Imagery*, 1, 23–25.
- Kosslyn, S. M. (1980). *Image and mind*. Cambridge, Mass.: Harvard University Press.
- Kosslyn, S. M. (1987). Seeing and imagining in the cerebral hemispheres: A computational approach. *Psychological Review*, 94, 148–175.
- Kosslyn, S. M. (1994). *Image and brain*. Cambridge, Mass.: The MIT Press.
- Kosslyn, S. M., Cave, C. B., Provost, D., & Von Gierke, S. (1988). Sequential processes in image generation. *Cognitive Psychology*, 20, 319–343.
- Laeng, B., Shah, J., & Kosslyn, S. M. (1999). Identifying objects in conventional and contorted poses: Contributions of hemisphere-specific mechanisms. *Cognition*, 70, 53–85.

- Loftus, G. R. (1972). Eye fixations and recognition memory for pictures. *Cognitive Psychology*, 3, 525–551.
- Loftus, G. R. (1981). Tachistoscopic simulations of eye fixations on pictures. *Journal of Experimental Psychology: Human Learning and Memory*, 7, 369–376.
- Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology*, 64, 17–24.
- Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.
- Noton, D., & Stark, L. (1971a). Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vision Research*, 11, 929–942.
- Noton, D., & Stark, L. W. (1971b). Scanpaths in eye movements during perception. *Science*, 171, 308–311.
- Podgorny, P., & Shepard, R. N. (1978). Functional representations common to visual perception and imagination. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 380–393.
- Posner, M. I. (1978). *Chronometric explorations of mind*. New York: Oxford University Press.
- Posner, M. I., Nissen, M., & Ogden, W. (1978). Attended and unattended processing modes: The role of set for spatial location. In H. Pick and I. Saltzman (Eds.), *Modes of perceiving and processing information* (pp. 137–157). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Pylyshyn, Z. (1981). The imagery debate: Analogue media versus tacit knowledge. *Psychological Review*, 87, 16–45.
- Quinn, J. G., & Ralston, G. E. (1986). Movement and attention in visual working memory. *Quarterly Journal of Experimental Psychology*, 38A:689–703.
- Rayner, K., McConkie, G. W., & Ehrlich, S. (1978). Eye movements and integrating information across fixations. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 529–545.
- Rizzolatti, G., Riggio, L., Dascola, I., & Umiltà, C. (1987). Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention. *Neuropsychologia*, 25, 31–40.
- Roll, R., Velay, J. L., & Roll, J. P. (1991). Eye and neck proprioceptive messages contribute to the spatial coding of retinal input in visually oriented activities. *Experimental Brain Research*, 85, 423–431.
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8, 382–439.
- Sheehan, P. W., & Neisser, U. (1969). Some variables affecting the vividness of imagery recall. *British Journal of Psychology*, 60, 71–80.
- Singer, J. L., Greenberg, S., & Antrobus, J. S. (1971). Looking with the mind's eye: Experimental studies of ocular motility during daydreaming and mental arithmetic. *Transactions of the New York Academy of Sciences*, 3, 694–709.
- Thomas, N. J. T. (1999). Are theories of imagery theories of imagination? *An active perception approach to conscious mental content*. *Cognitive Science*, 23, 207–245.
- Ullman, S. (1996). *High-level vision. Object recognition and visual cognition*. Cambridge, MA: The MIT Press.
- Umiltà, C., Riggio, L., Dascola, I., & Rizzolatti, G. (1991). Differential effects of central and peripheral cues on the reorienting of spatial attention. *European Journal of Cognitive Psychology*, 3, 247–267.
- Walker, R., Kentridge, R., & Findlay, J. (1995). Independent contributions of the orienting of attention, fixation offset, and bilateral stimulation on human saccadic latencies. *Experimental Brain Research*, 103, 294–310.
- Weiner, S. L., & Ehrlichman, H. (1976). Ocular motility and cognitive process. *Cognition*, 4, 31–43.
- Wexler, M., Kosslyn, S. M., & Berthoz, A. (1998). Motor processes in mental rotation. *Cognition*, 68, 77–94.
- Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum Press.