

## Scrutinizing visual images: The role of gaze in mental imagery and memory



Bruno Laeng<sup>a,\*</sup>, Ilona M. Bloem<sup>b</sup>, Stefania D'Ascenzo<sup>c</sup>, Luca Tommasi<sup>d</sup>

<sup>a</sup> Department of Psychology, University of Oslo, Oslo, Norway

<sup>b</sup> Faculty of Psychology and Neuroscience, Maastricht University, The Netherlands

<sup>c</sup> Department of Communication and Economics, University of Modena and Reggio Emilia, Italy

<sup>d</sup> Department of Psychological Sciences, Humanities and Territory, G. d'Annunzio University, Chieti, Italy

### ARTICLE INFO

#### Article history:

Received 28 August 2012

Revised 13 January 2014

Accepted 16 January 2014

#### Keywords:

Imagery

Memory

Eye-tracking

Enactment

### ABSTRACT

Gaze was monitored by use of an infrared remote eye-tracker during perception and imagery of geometric forms and figures of animals. Based on the idea that gaze prioritizes locations where features with high information content are visible, we hypothesized that eye fixations should focus on regions that contain one or more local features that are relevant for object recognition. Most importantly, we predicted that when observers looked at an empty screen and at the same time generated a detailed visual image of what they had previously seen, their gaze would probabilistically dwell within regions corresponding to the original positions of salient features or parts. Correlation analyses showed positive relations between gaze's dwell time within locations visited during perception and those in which gaze dwelled during the imagery generation task. Moreover, the more faithful an observer's gaze enactment, the more accurate was the observer's memory, in a separate test, of the dimension or size in which the forms had been perceived. In another experiment, observers saw a series of pictures of animals and were requested to memorize them. They were then asked later, in a recall phase, to answer a question about a property of one of the encoded forms; it was found that, when retrieving from long-term memory a previously seen picture, gaze returned to the location of the part probed by the question. In another experimental condition, the observers were asked to maintain fixation away from the original location of the shape while thinking about the answer, so as to interfere with the gaze enactment process; such a manipulation resulted in measurable costs in the quality of memory. We conclude that the generation of mental images relies upon a process of enactment of gaze that can be beneficial to visual memory.

© 2014 Elsevier B.V. All rights reserved.

### 1. Introduction

In his book *Inquiries into Human Faculty and its Development* (1883), Sir Francis Galton discussed mental imagery as a special ability of human visual memory. Specifically, he wondered whether mental images could be “so clear and sharp as [...] to be scrutinized with nearly as much

ease and prolonged attention as if they were real objects.” Galton prompted his informants to “think of some definite object—suppose it is your breakfast-table as you sat down to it this morning—and consider carefully the picture that rises before your mind's eye [...] Is the image dim or fairly clear? [...] Are all the objects pretty well defined at the same time, or is the place of sharpest definition at any one moment more contracted than it is in a real scene?” Reports about the “definition” of the imagined breakfast items varied very much across individuals; however, a

\* Corresponding author. Address: Department of Psychology, University of Oslo, 1094 Blindern, 0317 Oslo, Norway. Tel.: +47 22 84 51 16.

E-mail address: [bruno.laeng@psykologi.uio.no](mailto:bruno.laeng@psykologi.uio.no) (B. Laeng).

common report was that one or two objects would appear much more distinct than the others but these could come out clearly if attention be paid to them. Thus, different objects were not clear all at once but only successively, by focusing attention on them at different time points.

About a century later, although accounts of imagery did not rely any longer exclusively on introspective reports, the modern cognitive psychologists also concluded that whenever we generate a visual image of an object, the different parts of the object are not clear all at once but only successively (e.g., Hebb, 1968; Neisser, 1976). Kosslyn (1980), Kosslyn (1994) has also put forward an influential computational model for visual imagery, according to which each part of an image is added in successive steps (Kosslyn, Cave, Provost, & Von Gierke, 1988; Kosslyn, Reiser, Farah, & Fliegel, 1983). Visual images take time both to generate and to inspect and, in many respects, they strongly resemble the normal perception of objects at close range, where a high-resolution perceptual representation of the object cannot be achieved in a single glance but a series of eye movements must bring into ‘foveal’ focus the different parts of the object.

One remarkable finding of several studies of imagery is that while imagining something there appears to be a lot of motor activity, which resembles the exploratory movements typically made during perceptual scrutiny of an object or scene. Jacobson (1932; see also Totten, 1935) had originally observed with a galvanometer that engaging in imagery (e.g., recollection) resulted in the measurement of action potentials in muscle groups that were specific to the body part which was imaginatively moved (e.g., during visual imagination, movements of the eye-balls was registered, while when thinking, one could register brief contractions in muscles of tongue). Moreover, several researchers have noticed a remarkable similarity in the duration of imagined actions compared to the time it takes to perform them (e.g., Decety, 1996; Decety, Jeannerod, & Prablanc, 1989; Jeannerod, 1994; Parsons, 1987). These findings clearly implicate the presence of motor processing during imagery, although the motor processes would often seem to constitute only a subset of those activated during overt movement (Ellis, 1995).

According to recent studies, gaze patterns (i.e., fixations and/or direction of saccades) that are measured in real time during recollection of a previous event look remarkably similar to the scanpaths during a perceptual recognition test of the same scene, despite the fact that when thinking about the episode there is nothing at all to look at on a blank computer screen. This phenomenon has been repeatedly observed in a variety of studies (e.g., Moore, 1903; Altmann, 2004; Brandt & Stark, 1997; Brandt, Stark, Hacısalihzade, Allen, & Tharp, 1989; de’Sperati, 2003; Gbadamosi & Zangemeister, 2001; Hollingworth, 2005; Humphrey & Underwood, 2008; Jeannerod & Mouret, 1962; Johansson, Holsanova, & Holmqvist, 2006; Laeng & Teodorescu, 2002; Laeng et al., 2007; Martarelli & Mast, 2013; Renkewitz & Jahn, 2012; Spivey & Geng, 2001). It would seem that, when retrieving a visual image or episode, not only there occur spontaneous eye movements but these tend to reflect the content of the original scene. Deckert (1964) had observed that participants instructed

to imagine a beating pendulum developed pursuit ocular movements of a frequency comparable to the frequency of a previously seen real pendulum. Intriguingly, studies of rapid eye movements or REM during sleep also would seem to show some relationship between the types of eye movements and the content of dreams (e.g., Aserinsky & Kleitman, 1953; Dement & Kleitman, 1957; Doricchi, Iaria, Silveti, Figliozzi, & Siegler, 2007; Hong et al., 1997; Hong et al., 2009; Roffwarg, Dement, Muzio, & Fisher, 1962) as well as time-locked activity within primary visual cortex (Miyauchi, Misaki, Kan, Fukunaga, & Koike, 2009).

At a first glance, the above phenomena are puzzling because it seems a meaningless expenditure of bodily energy and cognitive effort to move about the eyes when there is nothing to be seen. Purposeful saccades that cannot garner any visual input appear completely paradoxical in relation to normal visual processing, since the pattern of saccadic movements during perception seems to be purposefully guided towards visual information or ‘objects’ that are relevant for the cognitive system at that particular time (e.g., Einhäuser, Spain, & Perona, 2008; Findlay & Gilchrist, 2003; Hayhoe & Ballard, 2005; Noton & Stark, 1971a; Noton & Stark, 1971b; Rothkopf, Ballard, & Hayhoe, 2007; Rucci, Iovin, Poletti, & Santini, 2007; Schütz, Trommershäuser, & Gegenfurtner, 2012; Stark & Ellis, 1981; Trommershäuser, Maloney, & Landy, 2009; Yarbus, 1967). Importantly, eye movements indicate the occurrence of shifts in spatial attention (Craighero, Nascimben, & Fadiga, 2004; Deubel & Schneider, 1996; Henderson, 1992; Moore & Fallah, 2001; Rolfs, Jonikaitis, Deubel, & Cavanagh, 2011; Shepherd, Findlay, & Hockey, 1986) and covert visual attention may consist in the motor preparation of an eye movement (Rizzolatti, Matelli, & Pavesi, 1983; Rizzolatti, Riggio, Dascola, & Umiltà, 1987). Hence, oculomotor activity could overload the cognitive system and/or interfere with other processes (cf. Loftus, 1972). Since the early days of research on mental imagery, both Francis Galton and Alfred Binet (Hadamard, 1945, pp. 72–73) had suggested that there may be an antagonism between the vividness or detail of a visual image and the presence of other activities.

A solution to the above puzzle is to assume that, contrary to the idea that such “empty” looks during recollection and imagination are either deleterious or irrelevant to cognition, they may actually serve some useful function. There is growing evidence for shared mechanisms of perception and imagery (e.g., Kan, Barsalou, Solomon, Minor, & Thompson-Schill, 2003; Kosslyn & Thompson, 2000). In addition, the idea that perception is “active” or “embodied” has been gaining strength over the years within the cognitive sciences and neurosciences (Barsalou, 1999; Ellis, 1995; Findlay & Gilchrist, 2003; Gibbs, 2006; Gibson, 1979; Pezzulo et al., 2001; Pulvermüller & Fadiga, 2010). This perspective stresses the idea that the visual system does not merely register its environment but explores it and poses questions by “grasping” objects with the eyes and/or hands (Ballard, Hayhoe, Pook, & Rao, 1997; Castelano, Mack, & Henderson, 2009; Karn & Hayhoe, 2000; Land et al., 1999). If perception and imagery share processing mechanisms, then also imagery may be “active” in the sense that adjustments of the body organs, even in a vacuum, could play a significant role in the retrieval of

internally stored information. A straightforward hypothesis, already well-formulated by Hebb (1968), is that such an empty gaze serves the function of assisting the mental re-construction of a representation. According to Hebb (1968, p. 470), “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”. Neisser (1976) also speculated that the act of constructing an image would require eye movements like those originally made in perceiving, because imagery is a process of visual synthesis and construction, much like perception.

The fundamental “Hebbian” idea behind the present study is that eye fixations can provide a sort of “scaffolding structure” for generating a visual image part-by-part. As put by Mast and Kosslyn (2002), eye movements could play an important role in allowing one to visualize a montage, a composite created on the basis of memories of multiple fixations. In other words, a single object’s image may be constructed in a manner that is not that different from imagining a scene; since an object has a categorical spatial structure between its parts (Laeng, Shah, & Kosslyn, 1999), these can be treated as separate units or “objects”. Thus, gaze could trigger sequences of memories and could also help to position correctly each image of a part relative to other parts. For example, we may have vivid imagery of, say, a cat, when we go through (some of) the motions of looking at something and determining that it is a cat, even though there is actually no cat (Thomas, 2011). Thus, contrary to the idea that motoric activity during imagery may be an epiphenomenon, a meaningless spill-over of mental activity while thinking to be back in a previously encountered situation, which in itself could bear no meaningful effect on cognitive processing (e.g., Marks, 1973; Teichner, LeMaster, & Kinney, 1978, p. 278), we believe that the present phenomena actually reflect something very important about the nature of mental representations.

Most current models of episodic memory do posit that one of the key functions of imagery is to allow reconstructing the past and, in particular, to generate specific predictions based on past experience (Addis, Wong, & Schacter, 2007; Hassabis, Kumaran, & Maguire, 2007; Moulton & Kosslyn, 2009; Schacter, Addis, & Buckner, 2007; Schacter, Addis, & Buckner, 2008). That is, imagery allows making explicit and accessible aspects of a specific situation. If someone’s gaze is engaged during recollection, despite being actually “looking at nothing”, this might actually tell us a great deal about mechanisms involved in memory recall (Ferreira et al., 2008; Ryan, Althoff, Whitlow, & Cohen, 2000). Specifically, memory representations are based on integrating input from various sources with spatial information, which would seem to be registered by default in working memory as part of a dynamic motor system (Altmann, 2004; Altmann & John, 1999; Ballard et al., 1997; Hodgson, Bajwa, Owen, & Kennard, 2002; Logie, 1995; Richardson, Altmann, Spivey, & Hoover, 2009; Richardson & Kirkham, 2004). Thus, the visual system automatically registers a spatial index or pointer to a position in the visual field as a core element of an episodic trace, also in circumstances in which actions are not required, the location

information is not relevant for solving the task, and there is no intention or demand to learn the spatial information (Laeng et al., 2007; Richardson & Spivey, 2000). Kent and Lamberts (2008) have proposed that memory retrieval is generally elicited by “mental simulation” (Barsalou, 1999); supposedly, when the integrated memory episode is reactivated at a later time, the spatial index relating to an object or part will also be automatically retrieved (Bourlon, Oliviero, Wattiez, Pouget, & Bartolomeo, 2011; Hoover & Richardson, 2008), which in turn triggers the eyes to move to the indexed location in which the part originally appeared. As Ballard et al. (1997, p. 724) point out: “Because humans can fixate on an environmental point, their visual system can directly sample portions of three dimensional space [...] and as a consequence, the brain’s internal representations are implicitly referred to an external point.” Thus, gaze direction may indicate a retrieval attempt for a specific item of information (Renkewitz & Jahn, 2012). In fact, outside of on-screen laboratory experiments, locations in the environment are rarely completely empty. Therefore, gaze might garner useful contextual visual cues (like noticing an empty chair) when attempting to recall visual information. Finally, returning the eyes to the former location of an object could also improve memory for information associated with that object (e.g., Hollingworth, 2006; Johansson & Johansson, *in press*), especially if spatial information contributes to maintaining the continuity and integrity of the “object file” or event (Hommel, 2004; Hommel, Musseler, Aschersleben, & Prinz, 2001; Kahneman, Treisman, & Gibbs, 1992).

In support of the above idea that the motoric activity during imagery plays a functional role, there exists some evidence that the accuracy of memory retrieval can be disrupted when someone who is holding an image in mind is restrained from making an eye movement or deliberately moves in an image-irrelevant way (e.g., Andrade, Kavanagh, & Baddeley, 1997; Antrobus, Antrobus, & Singer, 1964; Barrowcliff, Gray, Freeman, & Macculloch, 2004; Gunter & Bodner, 2008; Postle, Idzikowski, Della Sala, Logie, & Baddeley, 2006; Ruggieri, 1999; Singer & Antrobus, 1965). Several studies have shown the same phenomenon with other movement types; e.g., the recall of an imagined path can be disrupted by a concurrent movement of the arm (Quinn, 1994). A counterclockwise manual rotation hinders the concomitant clockwise mental rotation of a visual object and vice versa; however, a counterclockwise mental rotation of a visual object does facilitate a clockwise mental rotation (Wexler, Kosslyn, & Berthoz, 1998). Demarais and Cohen (1998) observed that, while solving transitive inference problems with the terms left/right or above/below, participants spontaneously made more horizontal than vertical saccades during the former task but they showed the reverse pattern for the latter. Glenberg and Kaschak (2002) found that, when judging whether a sentence was sensible (e.g., “close the drawer”), participants had difficulty making such a judgment if required to make a response in the opposite direction. Dijkstra, Kaschak, and Zwaan (2008) have found that participants could retrieve more efficiently autobiographical information when their body positions while being queried were similar to the body position they had during the original event.

In eye-tracking studies, when participants perform a problem solving tasks and simultaneously their eye movements are “guided” either according to a scanpath related to the problem’s solution or in an irrelevant way, the former gaze patterns lead to successful problem-solving than the latter ones (e.g., Grant & Spivey, 2003; Thomas & Lleras, 2007). Laeng and Teodorescu (2002) specifically found that memory suffered when spontaneous fixations during recall were prevented by enforcing fixation on a central cross at the time the participant attempted to answer a question regarding a previously seen object, which strongly suggests that the eye movements occurring during image generation are not epiphenomenal or a consequence of the experiment’s task demands (Jolicoeur & Kosslyn, 1985). Instead, they strongly suggest that, by disrupting a spontaneous action pattern, the memory system may be hindered in the retrieval of the details of a mental representation and that they play a functional role in the process of recollecting and re-constructing a previous perception. Consistently with the findings of Laeng and Teodorescu’s (2002; Experiment 2), successive studies have found evidence, by forcing fixation during retrieval, that eye movements played a functional role for memory, since this procedure reduced episodic memory performance (Johansson, Holsanova, Dewhurst, & Holmqvist, 2012; Johansson & Johansson, 2013; Mäntylä & Holm, 2006).

In the specific case of visual imagery, eye-tracking studies support the idea that the original locations get automatically stored as a part of the scene’s representation and, to some extent, a trace of the whole oculomotor sequence may also be kept, as originally suggested by Noton and Stark (1971a), Noton and Stark (1971b), although the evidence for a “scanpath” memory remains weak (cf. Johansson, Holsanova, & Holmqvist, 2006). Moreover, a key aspect that is still unclear is whether these eye movements during imagery simply return to a generic position occupied by an object (e.g., its center of mass) or they actually “mirror” to some degree the details or parts of a single object, as classic accounts of imagery would appear to imply. The few extant studies in which observers were asked to visualize single pattern stimuli remains ambiguous in this respect (i.e., Brandt & Stark, 1997; Laeng & Teodorescu, 2002; Martarelli & Mast, 2013; Noton & Stark, 1971a; Noton & Stark, 1971b). Therefore one aim of the present study is to directly address the question of whether fixations during imagery do not simply occur over a generic, center-of-mass, position of the object but also on specific locations corresponding to an object’s features. If imagery of a single object can be based on a global encoding of the shape as a single unit (cf. Kozhevnikov, Kosslyn, & Shepard, 2005), then it might be possible than no more than a generic gravitation of gaze over the region previously occupied by the object would be observed. However, a large literature on looking at patterns has revealed that the eye’s “dwell time” is a function of the information value of specific parts or features of an object (e.g., Buswell, 1935; Deco & Schürmann, 2000; Kaufman & Richards, 1969; Leek et al., 2012; Mackworth & Morandi, 1967; Renninger, Vergheese, & Coughlan, 2007; Yarbus, 1967; Zusne & Michels, 1964). Thus, we hypothesize that when reconstructing the image of a single object, like the figure of an animal or a

geometrical shape, fixations during imagery will concentrate in the locations of those ‘parts-rich’ regions of the shape where gaze mainly dwells during perception.

In the present study, we monitored gaze by use of an infrared remote eye-tracker during perception and imagery of geometric forms (Experiment 1) and figures of animals (Experiment 2 and 3). In the first experiment, we provide evidence that when observers look at an empty screen and at the same time they generate a detailed visual image of a simple geometrical form that was previously seen, their gaze probabilistically dwells within a region that corresponds to the original positions and shape of the imagined object. In Experiments 2 and 3, we show that the more faithful an observer’s gaze enactment between perception and imagery over specific parts of figures of animals, the more accurate is the observer’s memory. Moreover, when participants were queried in a recall phase about properties of each one of the animal pictures, gaze not only returned to the location of the part probed by the question, but interfering with this process (by asking them to maintain fixation away from the original location of the shape while thinking about the answer), resulted in measurable costs in the quality of memory.

## 2. Experiment 1

In the first experiment, we showed pictures of equilateral triangles on a computer monitor, while the participants’ eye fixations were monitored by an infrared eye-tracker. The triangles were always shown centered over the same gray background (see Fig. 1) but their orientation was in half of the trials upright (i.e. with one corner pointing up) and in the other half upside-down (i.e. one side was on the top and one corner pointed down). To introduce some variety in the stimuli, the internal area of each triangle changed luminance from trial to trial. During imagery generation, the screen was devoid of stimuli and it had the same background color over which the triangle shapes had previously appeared, so that the eyes could be monitored while kept open and without a need to avert gaze away from the screen while imagining the forms (cf., Glenberg, Schroeder, & Robertson, 1998).

Studies of eye fixations on pictures of shapes have found that, although the initial landing position of gaze on an object tends to occur over the center-of-gravity or mass of a shape (e.g., Kaufman & Richards, 1969; Melcher & Kowler, 1999; Vishwanath & Kowler, 2004), during a prolonged inspection of an object or scene, gaze concentrates on information-rich local areas of the shape, like borders, depth gradients, corners and junctions (Chelnokova & Laeng, 2011; Deco & Schürmann, 2000; Leek et al., 2012; Renninger et al., 2007; Sæther, Van Belle, Laeng, Brennen, & Øvervoll, 2009). Thus, we expected that most of the fixations would occur within the regions of the triangular shapes and also in the vicinity of their corners and borders. Most importantly, we expected that our observers would do the same when generating mental images of polygons. Given that half of the triangles were oriented up and the other half down, a simple manner to test whether fixations during the imagery condition resembles

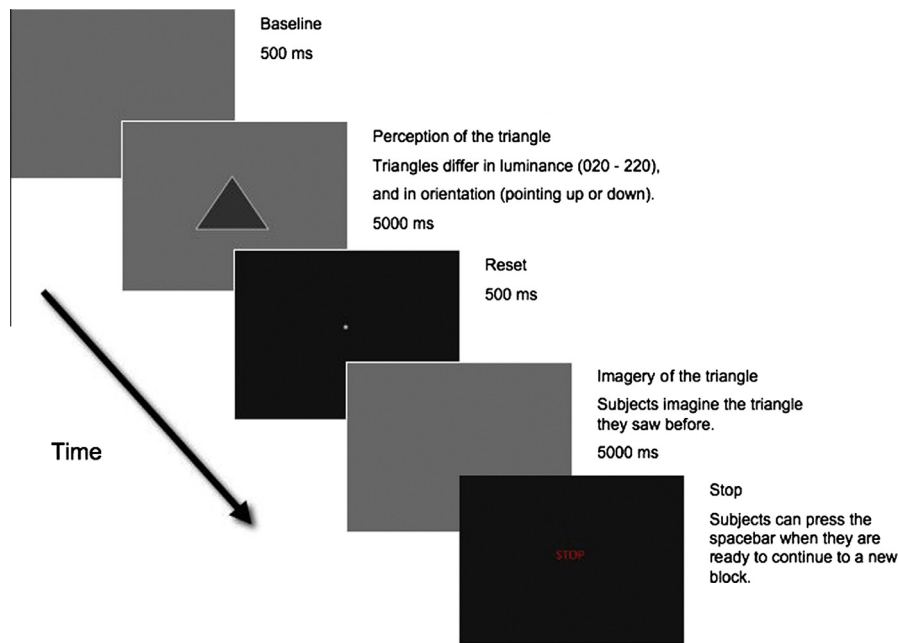


Fig. 1. Illustration of events within a trial of Experiment 1.

those during perception consists in computing the average Y position of gaze for the two conditions. One would therefore expect that fixations should be on average higher for the down-pointing triangles than for the upward pointing triangles, both during perception and imagery conditions. In contrast, the average X positions of fixations should not significantly differ between the two shape orientations.

## 2.1. Methods

### 2.1.1. Participants

Thirty participants (21 female; mean age = 23.7;  $SD = 7.6$ ) were recruited from the psychology department at the University of Oslo. All participants had normal or corrected to normal vision (with contact lenses). They were rewarded for their participation by means of a gift voucher of 100 Norwegian Crowns.

### 2.1.2. Apparatus

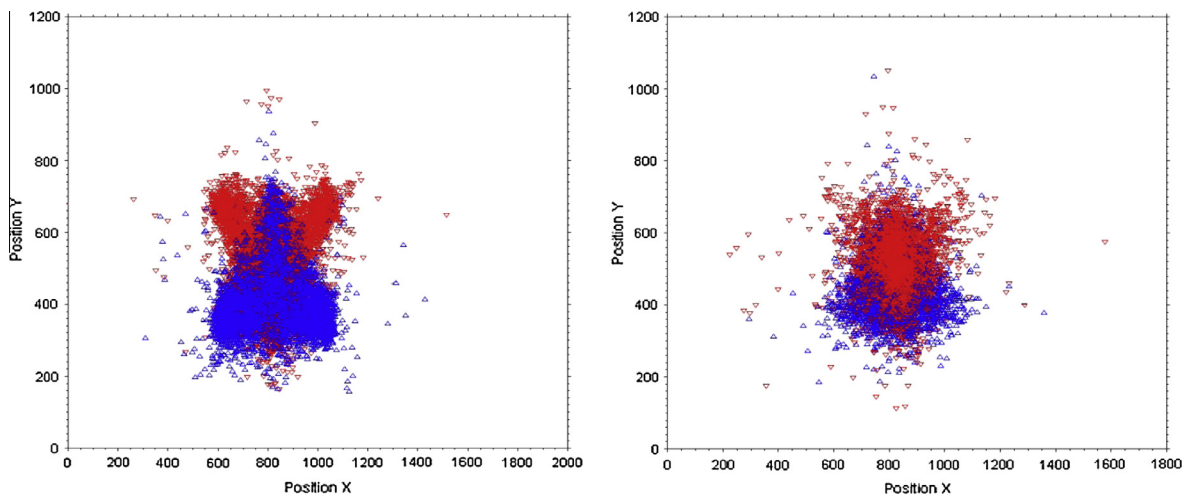
A Remote Eye Tracking Device from Senso-Motoric Instruments (SMI, Berlin, Germany) was employed to monitor eye fixations. Samples of eye positions were taken at 60 Hz. The spatial resolution of the system is accurate within  $0.03^\circ$  of visual angle. Room lighting does not interfere with the recording capabilities of the eye tracker but the illumination in the room was kept constant. The eye tracker is developed for a contact-free gaze measurement with automatic head-movement compensation, in a range of  $40 \times 20$  cm at a 70 cm distance. Fixations are automatically detected according to an algorithm based on minimum gaze duration of 80 ms within a circular region with maximal dispersion of 100 pixels.

### 2.1.3. Stimuli and procedure

We showed on a Dell LCD monitor pictures of 22 equilateral triangles (half of them pointing up and half down; see Fig. 1). Each picture had a resolution of  $1680 \times 1050$  pixels. The triangles (5 cm each side) were always shown for 5 s and centered over a gray background but the internal area of the triangle could have different luminance levels (i.e., one of 11 steps, ranging from 20 to 220 RGB units, with the middle step corresponding to the luminance of the background and baseline picture). A thin white line was used as border to outline the triangles. A fixed, randomized, presentation order of the trials was used for all participants.

The stimuli were presented by use of iView 3.0® ‘Experiment Center’ software. A standard eye tracker calibration routine was used at the beginning of the experiment, when eye position was recorded at four standard calibration points (appearing as whitish disks with a small red dot in the center). Participants were instructed to fixate on the red dot, which moved at regular intervals and stopped in the four positions forming a regular  $2 \times 2$  matrix, where eye position was sampled. No chinrest was used but participants were instructed to keep their head as stable as possible and to keep their eyes open during both the perception and imagery phases of the experiment. Participants were told that the experiment’s goal was to measure pupil diameters during a cognitive task and to stimuli varying in brightness. No mention was made about recording eye movements during the task.

The experimental task consisted of 22 trials, one for each stimulus type. Observers could freely view the stimuli. Each trial started with a 500 ms baseline, consisting of a uniform gray picture, followed for 5 s by a triangle



**Fig. 2.** Cumulative fixations of all participants (color triangles) plotted for both *X* and *Y* positions and for all trials of Experiment 1. Different distributions of fixations are apparent for upright triangles (in blue) versus upside-down triangles (in red) and can be seen both for the perception condition (left panel) and the imagery condition (right panel). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

positioned in the center of the screen (see Fig. 1). Subsequently a black screen with a white central fixation dot ( $1^\circ$  visual angle) was presented for 500 ms. After this picture, an empty gray screen (with the same color as the background of the perception stimuli) appeared for 5 s and participants were requested to generate a visual image (keeping the eyes open) of the triangle they had just seen. The last event of each trial was a black screen with the word STOP written in red<sup>1</sup> in the center of the screen, indicating that it was no longer needed to keep the mental image of the stimulus and that a new trial could be started (by pressing, when ready, the spacebar on the computer keyboard). During the experimental conditions, no explicit response (verbal or key press) was required.

## 2.2. Results

We extracted by means of SMI's BeGaze<sup>®</sup> analysis software, the *X* and *Y* coordinates on the screen of all fixations (see Fig. 2). Mean *X* and *Y* positions were then computed for each individual and for Up versus Down triangle orientations in either the perception and imagery conditions.

Two repeated-measures ANOVA were performed with Condition (perception, imagery), Orientation (Up, Down) for the mean *X* or *Y* fixations' positions as dependent variables. As expected for the *Y* coordinates, there was a significant interactive effect of Orientation,  $F(1,29) = 277.7$ ,  $p < 0.0001$ , showing that, both during perception and imagery, fixations gravitated to a lower location when the triangle was oriented upwards (mean *Y* position = 459.5;  $SD = 34.8$ ) and to a higher location when the triangle was oriented downward (mean *Y* position = 516.4;  $SD = 346.6$ ), consistent with a distribution of fixations reflecting the shape of the perceived but also imagined

stimulus. A significant interaction of Condition and Orientation,  $F(1,29) = 22.2$ ,  $p < 0.0001$ , was due to a tendency to fixate slightly more upward during imagery compared to perception and especially when imagining upward pointing triangles (Perception: mean *Y* = 448.8;  $SD = 33.7$ ; Imagery: mean *Y* = 471.5;  $SD = 36.5$ ) than downward ones (Perception: mean *Y* = 517.8;  $SD = 32.6$ ; Imagery: mean *Y* = 515.5;  $SD = 37.6$ ). This effect also caused a significant main effect in average *Y* positions between perception and imagery.

In contrast, but also consistently to our expectation, the ANOVA on the mean *X* fixations' positions showed no significant differences in eye positions between perception and imagery,  $F(1,29) = 2.2$ ,  $p = 0.15$ , or between the triangles' orientation,  $F(1,29) = 2.2$ ,  $p = 0.15$ , or the interaction of the two factors,  $F(1,29) = 0.001$ ,  $p = 0.98$ . See Fig. 2 for an illustration of the distributions of fixations in the various conditions.

## 2.3. Discussion

During imagery participants clearly fixated a region of the screen that corresponded to that of the triangles they had just seen, despite the screen was devoid of any form. Most important, the distribution of fixations closely corresponded to the form and orientation of the shapes both during perception and imagery. Indeed, gaze gravitated in both the perception and imagery situations to different average positions that reflected the difference in orientation of the shapes. These findings strongly support the idea that eye fixations during imagery enact a similar behavior shown during perception of the imagined object, like "what one would be doing if actually perceiving that thing" (Thomas, 2011).

As illustrated in Fig. 2, the cumulative pattern of fixations appeared to cover the entire shape, occasionally falling on outer regions both during perception and imagery.

<sup>1</sup> For interpretation of color in Fig. 1, the reader is referred to the web version of this article.

Remarkably, the cumulative plots shown in Fig. 2 reveal distinctively different patterns for upright versus upside-down shapes, not only during perception but also during imagery. Indeed, it was possible to recognize the original orientation of a triangle in a particular trial by simply inspecting the cumulative plot during imagery.

### 3. Experiment 2

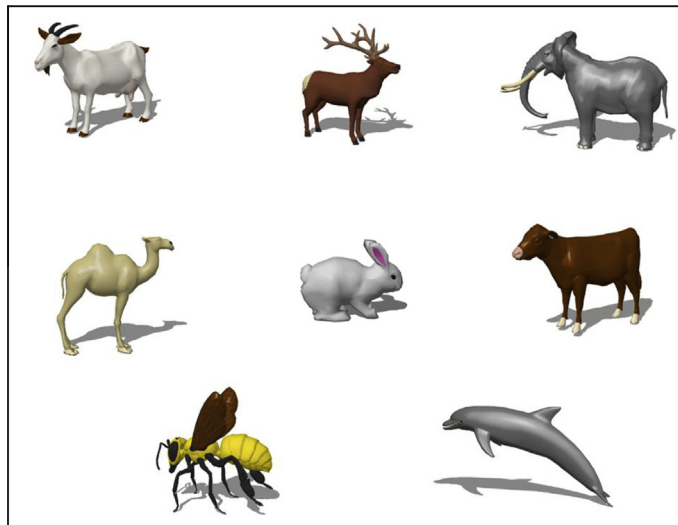
The previous experiment used simple geometric figures that the observers visualized. In these shapes, the key features are evenly or symmetrically distributed. In contrast, living organisms and many man-made objects (e.g., automobiles) tend to have a hierarchical distribution of features, also rather asymmetric, so that several of the relevant features may be densely crowded within a same section of the object's body (e.g., the head of the animal or the front of the automobile). Moreover, many of these relevant features can be based on rather 'local' information and may require the high resolution of the fovea in order to be scrutinized.

As Leek et al. (2012) point out, some kinds of visual features that are useful in recognition are likely to be detected only at a relatively coarse spatial scale (e.g., edge co-linearity or parallelism, elongation, symmetry, aspect ratio, and global outline; see also Biederman, 1987; Hayward, Tarr, & Corderoy, 1999). However, other useful shape features may (and in some cases, must) be encoded locally at a relatively finer spatial scale (e.g., edge boundaries, corners, vertices, surface depth gradients and curvature, as well as color and texture; see Lowe, 2004; Ullman, Vidal-Naquet, & Sali, 2002). Thus, one would expect that perceptual scrutiny would concentrate on these local features and that, in terms of gaze scanpaths, eye fixations and movements would be more likely to occur over these key features or interest points. In other words, gaze should disperse asymmetrically over the regions of space occupied by a natural body or artifact depending on the location within the visual field of those local features that require close scrutiny. Yarbus (1967) had already observed that the eyes "fixate on those elements of an object which carry or may carry essential or useful information" (p. 211). Moreover, he observed that once these elements had been fixated a first time, they would be fixated again and again in cycles (cf. DeAngelus & Pelz, 2009; Greene, Liu, & Wolfe, 2012; Jacob & Hochstein, 2011; Mannan, Ruddock, & Wooding, 1997; Zelinsky, Loschky, & Dickinson, 2011). A few years later, Noton and Stark (1971a), Noton and Stark (1971b) proposed and found supportive evidence that similar cycles of fixations would be repeated at recognition (i.e., when viewing the same shape a second time). They argued that recognition would benefit from a close match of the sensory-motor patterns elicited in the two encounters with the same object (i.e., in line with the 'encoding specificity principle' by Tulving, 1983, and the 'remembering operations' or 'transfer appropriate processing account' by Kolers, 1973; Roediger, Weldon, & Challis, 1989). Consequently, the hypothesis that an enactment of fixation cycle may be functional to successful recognition has been entertained by several researchers (Mäntylä & Holm, 2006).

Some imagery experiments that monitored gaze have given some support to the idea that an enactment of fixations, both in terms of their position or sequence (e.g., Laeng & Teodorescu, 2002), could give a memory advantage. However, these studies did not directly address the question of whether the distribution of gaze actually concentrated over places occupied by local salient features. For example, one experiment in the study by Laeng and Teodorescu (2002) reported gaze behavior while participants responded to queries about visual properties of previously seen pictures of fish (e.g., whether the previously seen fish possessed a particular trait, like a colored spot on the tail or on the head). Such an experimental paradigm clearly intended to prompt the use of visual imagery as a strategy for recalling details of memory episodes and, in particular, features that may be only implicitly encoded (as implied by Galton's, 1883, famous "breakfast-table questionnaire" or by Kosslyn's well-known query: "what shape are a German Shepherd's ears?"; Kosslyn, 2002; Kosslyn, 2007; Thompson, Kosslyn, Hoffman, & van der Kooij, 2008).). As observed by Laeng and Teodorescu's (2002), the participants re-fixated, while thinking about the answer, the region of the (empty) screen where the fish had originally appeared. Gaze also tended to be distributed over a region of about the same size at which the original object had appeared, which would suggest that eye movements over several features of the original shape had been stored and re-enacted during recollection. However, detailed analyses of fixations' position were not performed and it remains unclear whether the pattern of fixations could also somehow "mirror" the body structure of the animal.

Hence in the next experiment, we decided to directly assess whether (a) eye movements during imagery reflect the manner gaze scrutinized the object during perception and (b) if such a repetition of eye movements can predict memory performance or, in other words, whether a strong similarity of eye fixations during perception and imagery indicates the presence of robust episodic memory representations. Sheehan and Neisser (1969) proposed that the more vivid an image the more likely it is to involve some scanning process (Goldthwait, 1933; Sima, Lindner, Schultheis, & Barkowsky, 2010). On the basis of these ideas, one would expect that individual differences in spatial memory accuracy should be related to the degree of resemblance in gaze as performed during perception and imagery.

We chose to use several drawings of animals (e.g., a dog, a dolphin, an elephant; see Fig. 3) as the to-be-visualized stimuli. Clearly, these pictures contain many details but the most salient and defining features of each animal appear to be contained within the head region of the body (e.g., eyes, mouth, nose and ears). This is also a region typically containing changes in color/texture and curvature gradients (which are known to attract gaze; Leek et al., 2012; Wexler & Ouarti, 2008). Consequently, we expected that a larger proportion of gaze dwell time would be employed over the region that contained the animal's head than over regions containing other body parts (e.g., the legs, the trunk), both during perception and imagery. Moreover, we tested whether the strength of similarity between the mean% dwell times during perception and imag-



**Fig. 3.** Pictures used in Experiment 2 depicting 8 animals. In different trials, the same animal was shown turned either leftwards or rightwards.

ery could predict the quality of memory for the pictures. Therefore, right after the imagery task, we asked the same participants to judge whether a new series of pictures of the same animals depicted these at either the same size as seen earlier or if there had been a change (increase or decrease) in the total area size. We reasoned that eye shifts between points of fixations provide information not only about an object's spatial structure but also about its size. Thus, we specifically predicted that a participant's accuracy in this spatial judgment task would be positively related to the degree in which the participant's gaze during imagery re-enacted that of perceptual encoding.

### 3.1. Methods

#### 3.1.1. Participants

Forty participants (27 female; mean age = 24.5;  $SD = 5.2$ ) were recruited from the psychology department at the University of Oslo. All participants had normal or corrected to normal vision (with contact lenses). They were rewarded for their participation by means of a gift voucher of 100 Norwegian Crowns.

#### 3.1.2. Apparatus

The same Remote Eye Tracking Device (SMI, Berlin, Germany), computers and screen, employed in Experiment 1 were used in this experiment.

#### 3.1.3. Stimuli and procedure

We showed pictures of 8 animals (Fig. 3) twice (one version showing the animal turned leftwards and one version with the same picture flipped horizontally so that it looked rightwards). Each picture had a resolution of  $1680 \times 1050$  pixels. Each animal picture was always shown for 5 s within a white rectangle of size  $12 \times 9$  cm and centered over a gray background. A fixed, randomized, presentation order of the trials was used for all participants. The stimuli were presented by use of iView 3.0®

Experiment Center software. The same standard calibration routine used in Experiment 1 was employed. The experimental procedure for the imagery phase was practically identical to that used in Experiment 1 (i.e., a 500 ms baseline, a 5 s presentation of a picture of an animal, a black screen for 500 ms, and an empty gray screen for 5 s that participants were requested to look at while generating a visual image of the animal they had just seen previously). During the experimental conditions, no explicit response (verbal or key press) was required.

A final phase of the experiment was performed immediately after the imagery task. The same participants were asked to judge whether a new series of 16 pictures of the same animals showed these at either exactly the same size as seen earlier or if there had been change (either an increase or decrease in their total area size). The differently sized pictures were generated by use of Adobe Photoshop by use of the Image/Resize function (in percent size).

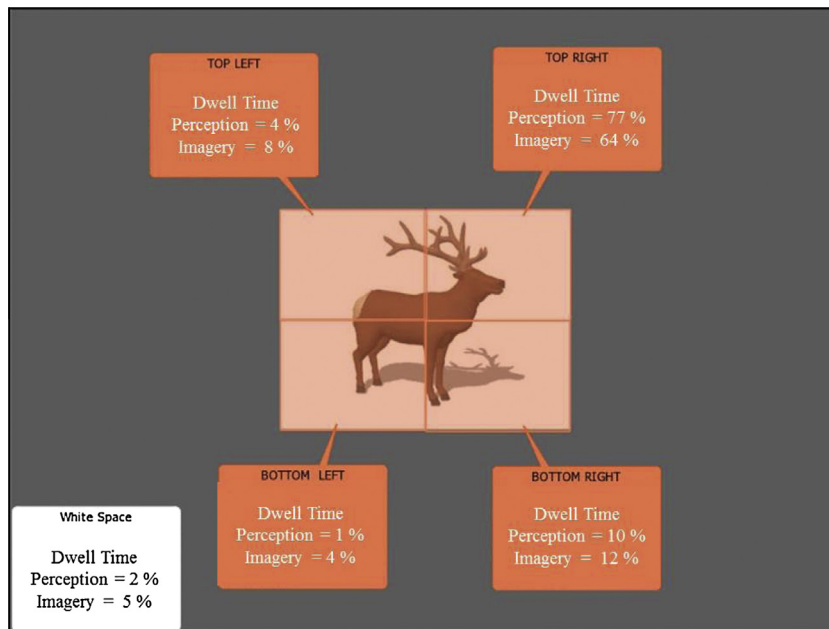
### 3.2. Results

We defined, by means of SMI's BeGaze® analysis software, 4 Areas of Interest (AOI) corresponding to the four quadrants of the white rectangle within which each animal shape appeared (see Fig. 4). Mean% dwell time within each AOI was then computed for both perception and imagery conditions in each trial and within the four AOI.

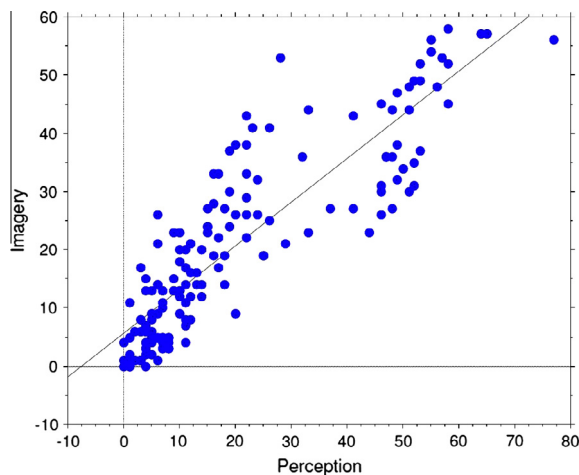
We performed a simple regression analysis using mean percent dwell time spent within each of the AOI during perception as the regressor and mean percent dwell time during imagery as the dependent variable. The analysis revealed a highly significant correlation;  $y = -5.62 + 0.75x$  ( $r = 0.88$ ;  $p < 0.0001$ ). Fig. 5 shows a scatterplot of these data with the interpolating line showing the regression of dwell time during imagery over perception.

Moreover, as expected, a significantly larger portion of dwell time occurred within the AOI or quadrant containing the animals' head than in the other quadrants combined,





**Fig. 4.** Mean% Dwell Time for the perception and imagery conditions of one trial, showing a deer facing right. During imagery only the empty white rectangle was visible. The division in quadrants was not shown during the experiment and it illustrates here the four Areas of Interest (AOI) used in the analyses. In this example, the top right AOI, containing the head of the animal, yielded a larger proportion of gaze dwell time, both during perception and imagery. Dwell time outside the AOI is accounted as ‘white space’ and time taken by saccades or eye blinks is not displayed.

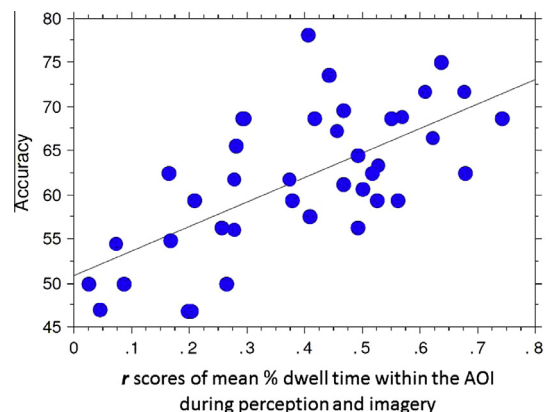


**Fig. 5.** Experiment 2: Regression between mean% Dwell Time spent by gaze within the same AOI during perception and imagery.

not only during the perception (Head-AOI: mean% dwell time = 43.1; *SD* = 14.1; Other-AOI: mean% dwell time = 11.3; *SD* = 9.9;  $t(158) = 15.7, p < .0001$ ) but also during imagery (Head-AOI: mean% dwell time = 35.8; *SD* = 12.9; Other-AOI: mean% dwell time = 14.9; *SD* = 12.5;  $t(158) = 8.9, p < .0001$ ).

Finally, means of correct responses in the spatial task (i.e., whether the animal picture appeared the same size or not) were calculated and the data were pooled over all the trials for each participant. The mean percent dwell time data within each of the 4 regions of interest for each

participant was then used to calculate a correlation score for each individual. The obtained *r* scores were taken as estimates of the strength or fidelity of enactment and used to predict in another simple regression analysis the accuracy scores in the spatial memory task. As shown in Fig. 6, this analysis revealed a significant positive correlation;  $y = 51.6 + 21.35x$  ( $r = 0.74; p = 0.0002$ ). An ANOVA on mean fixation duration in the spatial task revealed that participants on average showed longer fixations when looking at the differently sized pictures (Mean = 335 ms,



**Fig. 6.** Experiment 2: Regression between each participant's mean% Accuracy score in size judgments of the animals (as previously seen) with each participant's correlation coefficient *r* between mean% Dwell Times within the four AOI during perception and imagery (i.e., enactment fidelity).

$SD = 235$ ) than when viewing the same picture a second time (Mean = 311 ms,  $SD = 245$ ),  $F(1, 38) = 6.7$ ,  $p = 0.008$ .

### 3.3. Discussion

We predicted that gaze during imagery should probabilistically dwell within empty regions corresponding to the original positions of the salient features and especially those regions where several local features are naturally crowded (e.g., the rabbit's head). Instead, we did not expect gaze to be randomly distributed or to dwell on the shape's center-of-mass and/or to simply remain still. As expected, eye fixations within the AOs (or quadrants) at encoding and image generation were highly correlated. It was also clear that the quadrant containing the head region received much more scrutiny than the other parts of the animal shapes during perception and, importantly, the same behavior was observed for gaze during imagery. Such a result cannot be attributed to a bias of gaze towards a certain position on the blank screen, since the positions of the head and body were swapped for each animal picture in half of the trials.

We also reasoned that eye shifts between points of the image could provide information not only about an object's spatial structure but also its dimension and aspect ratio as experienced during the reconstruction of the episode. Charles Judd (1905) had observed that, when looking at the Müller-Lyer illusion, successive movements of the eye did not end at exactly the same relative points and that observers showed restricted movements in looking across the underestimated Müller-Lyer figure while showing wider movements when looking across the overestimated figure (see also de Grave & Bruno, 2010). In other words, the oculomotor commands contain additional spatial information that can be absent from retinal signals (Hafed & Krauzlis, 2006). In a study by Ryan and Villate (2009), gaze was monitored while participants judged whether the spatial relations among a few objects had been changed or not in a successive presentation of the display; their gaze clearly reflected the knowledge of positions previously occupied by objects, since the eyes transitioned between the locations of the presented objects and the locations that were previously occupied. According to Ryan and Villate (2009), eye position could be a conduit by which visuo-spatial information is integrated into a lasting memory representation (see also Olsen, Chiew, Buchsbaum, & Ryan, 2014).

Based on the above findings, we believe that eye positions and distance traveled by the eyes can provide pointers used by the spatial system that assist the computation of coordinate spatial relations between parts of an object or between objects in a scene (cf. Jones & Henriques, 2010). These coordinates can also provide information about the specific aspect ratio or viewing distance of the object when compared to knowledge in memory about spatial attributes of the object. Hence, if the original eye coordinates can be retrieved when recalling the object, this information should result in generating a visual image of the object as seen at approximately the same distance/size as it was experienced during the perceptual episode. Consequently, we expected that individual differences in the

similarity between eye fixations during perception and imagery predict the quality of the memory; i.e., the higher the 'fidelity' of the enactment, the more accurate the participant should be in judging the dimension/distance at which objects were originally perceived. Indeed, we found a strong correlation between a participant's enactment fidelity and accuracy of spatial memory.

## 4. Experiment 3

Although the findings of the previous two experiments are highly consistent with a theoretical account positing that imagery emulates perception not only phenomenally or subjectively but also in term of the oculomotor operations that occur during perception, one could remain skeptical that we really tested imagery in the traditional sense of the concept. That is, imagery often refers to the internal re-creation of a visual stimulus on the basis of top-down knowledge, as in "please imagine an elephant". In contrast, the previous tasks can be interpreted as requiring visual working memory and not necessarily retrieving a trace from long-term memory. In fact, a stimulus was presented and then withdrawn, so that the participant might have simply maintained that image "on-line." Thus, image 'maintenance' rather than 'generation' may account for the similarity in eye movements and fixations, if one would also posit that eye movements might assist in maintaining an image (cf. Brockmole & Irwin, 2005; Theeuwes, Belopolski, & Olivers, 2009). In other words, it is possible that repeating (rehearsing) a series of eye-movements may help keeping the size and position of a stimulus' key features in one's mind.

In our opinion, the above objection seems rather weak, since after each imagery phase the participants were explicitly instructed to "stop" imaging the object and, in addition, a 500 ms black-screen interval occurred before they received the imagery instruction. Nevertheless, we have no direct evidence that either the participants would comply with the 'interrupt' request or that a black screen was sufficient to "flush" the participant's visual working memory. Hence the motivation for a new experiment, where the same stimuli are used in an imagery task based on recall from long-term memory. Specifically, stimuli can be initially presented one by one as a slide show. In such an encoding phase, working memory would be necessarily updated for each new picture. In a later phase, a verbal instruction can be used to recall and visualize one of the previously seen objects in the series (e.g., "please imagine the \*elephant\* you saw earlier").

If our hypothesis about the usefulness of retaining oculomotor memory traces is correct, we would expect results in this experiment that are very similar to those of the previous experiments. Such a result would also extend that based on size discrimination performance in the previous experiment that may be considered suggestive but not conclusive. In fact, because there is variability in the precision of the memory representation formed when viewing each picture, the more precise the memory representation, the more similar one would expect the gaze patterns during perception and imagery to be. Within such an alterna-

tive account, size discrimination performance may just reflect the precision of the memory representation and the observed relationship between similarity in gaze and size discrimination performance in the previous experiment may occur without any functional role for eye movements (either during imagery formation or retrieval of the memory representation). Hence, a critical test for the hypothesis of a function role of gaze enactment is showing that controlling gaze during imagery can disrupt memory performance (see also Johansson et al., 2012).

Therefore in the third experiment, one condition (Experiment 3A) examined again the degree in which fixations during image generation were related to positions of salient body parts as seen in perception. In two other conditions (Experiment 3B and Experiment 3C), we gathered direct evidence that fixating on a position previously occupied by a part of the object provided a memory advantage. In the ‘part focus’ condition we questioned the participants about a part’s property (e.g., “what was the color of the bee’s wings?”) and also checked whether gaze would return in the specific region of the screen that was previously occupied by the probed feature (Experiment 3B). In a control ‘fixed focus’ condition, we checked whether the accuracy of recall would be lower if gaze was forced to remain on a fixation point away from the original location of the stimulus while the participants thought about the answer (Experiment 3C). Given that the goal is to capture fixations on specific body parts, when analyzing results from all three experiments, we used stricter AOI that fit tightly around the boundaries of the salient body parts of the animals (e.g., head, feet, tail; see Fig. 7 for examples). Naturally, we expected that using AOI that overlap precisely to the shape and position of the original object would yield lower correlations than those observed earlier, given that distortions in spatial memory and a loss of spatial precision in long-term memory are common (cf. Giudice, Klatzky, Bennett, & Loomis, 2013) and may result in drifts of the patterns of fixations even in the occurrence of “minimization” (Gbadamosi & Zangemeister, 2001) or

“scaling down” (Johansson, Holsanova, & Holmqvist, 2011) of the viewing gaze patterns during recall.

All questions probing specific features of body parts were chosen to reflect properties that could be answered correctly only by referring to the specific episode and not on the basis of general knowledge about the animal (e.g., “what was the color of the beak of the toucan bird?”; “Were the ears of the dog up or down?”; note that for each of these cases, there is no correct answer based on general or encyclopedic knowledge, given that the beak of a toucan can be of various colors, e.g. red or black, and different species of dogs have ears that typically either droop or stand up).

#### 4.1. Participants

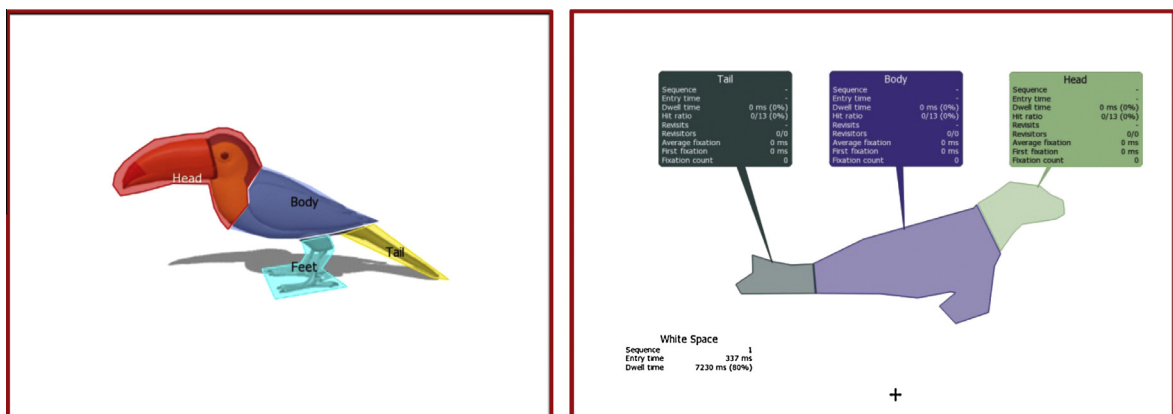
Forty new participants (24 females; mean age = 23.7;  $SD = 3.9$ ) were recruited from the psychology department at the University of Oslo. All participants had normal or corrected to normal vision (with contact lenses). They were rewarded for their participation by means of a gift voucher of 100 Norwegian Crowns.

#### 4.2. Apparatus

The same Remote Eye Tracking Device (SMI, Berlin, Germany), computers and screen, employed in the previous experiments were used.

#### 4.3. Stimuli and procedure

We showed pictures of 12 animals to all participants. Five of the pictures were the same used in the previous experiment (bee, camel, deer, elephant, goat) and the other seven (dog, giraffe, mantis, ostrich, rat, seal, toucan) were taken from the same picture database as the others. The pictures were roughly the same size and their centers-of-mass were positioned in slightly different position around the center of screen so that none of the probed parts (in Experiment 3B and 3C) would occupy the same region of



**Fig. 7.** Experiment 3: Examples of the Areas of Interest (AOI) used in the analyses. The left panel shows the picture of the ‘toucan’ overlaid to the picture as seen at encoding. The right panel shows the AOI overlaid to the picture of the ‘seal’ overlaid to the blank screen as seen during the “fixed focus” condition (in Experiment 3C); the only visible item in this condition was the small fixation cross at the bottom (the boxes represents dwell time data in relation to a particular trial (as expected, no time was spent within the AOI in this condition, since fixation was forced on the cross).

space. In the encoding phase, each animal picture was shown for 10 s on a white screen. All participants, regardless of whether they would be subsequently assigned to one of three different recall conditions, viewed during this encoding phase the 12 pictures as a slideshow; each image presentation was triggered by the participant by pressing the spacebar on the PC keyboard. A fixed, randomized, presentation order of the trials was used for all participants. The stimuli were presented by use of iView 3.0<sup>®</sup> Experiment Center software. The same standard calibration routine used in the previous experiments was employed.

After a break of 15 min, a subgroup of participants ( $N = 14$ ; females = 8) was subsequently tested with the imagery generation condition (or Experiment 3A), which consisted in presenting a blank screen for 10 s, while the participant was prompted to generate a mental image of a specific animal that was seen previously. The participant initiated a trial by pressing the spacebar, which triggered the computer to play an audio file naming one of the 12 animals seen in the encoding phase (e.g., 'rat'). The order of stimuli to be imagined was fixed across participants and it was different from the order in which they were perceived at encoding. Participants were instructed at the beginning of the task about building mental images of each named animal while looking at the blank screen, without blinking. During this experimental condition, no explicit response (verbal or key press) was required.

In contrast, those participants ( $N = 13$ ; females = 7) that were tested in the "parts' focus" condition (or Experiment 3B), when pressing the spacebar they heard an audio file posing a question about a specific animal and a specific body part (e.g., "was the tip of the beak of the 'toucan bird' orange or red in color?"). Participants were allowed to look freely at the blank screen while mentally looking for an answer. No feedback was given about the correctness of the response. Responses were recorded manually by the experimenter on a sheet of paper and later coded as 'correct' versus 'incorrect.' Finally, those participants ( $N = 13$ ; females = 8) that were tested in the "fixed focus" condition (or Experiment 3C), heard the same questions posed to the participants in the previous group but they were requested to maintain fixation, while mentally looking for an answer, on a small cross centrally located but close to the bottom border of the screen (i.e., more than 9° away from the original location of any of the probed features).

### 4.3. Results

We defined, by means of SMI's BeGaze<sup>®</sup> analysis software, AOI corresponding to the major parts of each animal's body. In most cases these were the head, body, legs and tail. However, one animal had no legs (i.e., seal) and in some pictures the animal's tail was not visible (e.g., elephant). Some animals had additional but salient body parts, so that separate AOI were designed for these (e.g., the wings of the bee or the trunk of the elephant). Differently from the previous experiment, AOI followed rather closely the border of the body and parts' boundaries (see Fig. 7).

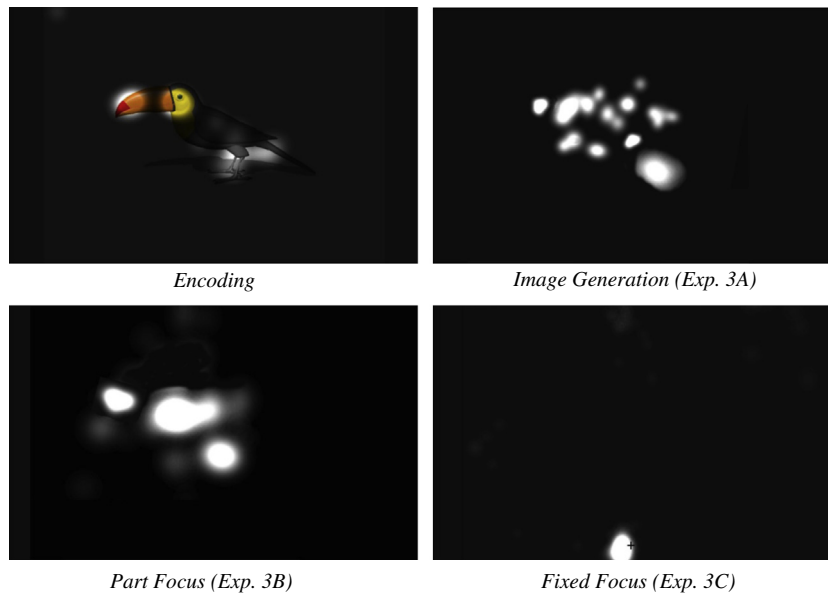
Mean% dwell time within each AOI was then computed by means of BeGaze<sup>®</sup> for both the 'encoding' (or percep-

tion) condition and the three 'recall' (or imagery) conditions in each trial and within each of the AOI. We then performed simple regression analyses using mean percent dwell time spent within each of the AOI during encoding (perception) as the regressor and mean percent dwell time during recall (imagery) as the dependent variable.

The analysis between dwell time at encoding and image generation (Experiment 3A) revealed a significant correlation;  $y = 4.46 + 0.24x$  ( $r = 0.15$ ;  $p < 0.0002$ ). This finding confirms the hypothesized enactment of fixations during image generation, replicating the results of the previous experiment, though in this case the strength of the relation (or ' $r$ ') was weaker, probably due to the highly conservative AOI (tightly fitting the original parts' locations) used in the present analyses. As one would expect, the regression analysis when applied to the 'part focus' (Experiment 3B) condition yielded no significant relation,  $y = 23.83 + -0.13x$  ( $r = 0.06$ ;  $p = 0.12$ ). Given that we hypothesized that fixations would focus on a specific part, then all the other AOI for incorrect parts should receive less attention, which in turn should reduce or nullify the correlation between encoding and recall. Thus, in order to reveal the hypothesized change, we labeled each AOI corresponding to a part probed as a question (e.g., the trunk for the elephant or the legs for the ostrich) as 'correct' and instead we labeled any of the other parts as 'incorrect' (e.g., the body for the elephant or the head for the ostrich). We then run two separate ANOVAs for correct versus incorrect body parts as the within-subject factor and% dwell time as the dependent variable. During perception, the so-called 'incorrect' parts (i.e., not probed later) were looked more on average (mean% dwell time = 23.3;  $SD = 17$ ) than the 'correct' part (mean% dwell time = 19.1;  $SD = 15$ ),  $F(1,583) = 8.6$ ,  $p = 0.004$ , confirming that the part chosen for the question did not receive any privileged scrutiny at encoding. However, the analysis of dwell time during imagery showed an advantage for the AOI corresponding to the 'correct' part (mean% dwell time = 17.9;  $SD = 34.6$ ) compared to 'incorrect' parts (mean% dwell time = 10.7;  $SD = 26.7$ ),  $F(1,583) = 7.0$ ,  $p = 0.008$ . Fig. 8 illustrates cumulative fixations in each condition.

The regression analysis between dwell time at encoding and image generation for the 'fixed focus' condition (Experiment 3C) showed no significant relation,  $y = .95 + 0.01x$  ( $r = 0.002$ ;  $p = 0.91$ ), as it should be if participants did not look at all at the AOI during recall and maintained fixation as requested.

Finally, means of correct responses in Experiments 3B and 3C (i.e., whether the answer was given in a condition where gaze was free to move about spontaneously versus one in which it was forced to remain on a fixation point) were calculated and the data were pooled over all the trials for each participant. Accuracy was typically high (mean number of errors = 1.94) no participant made more than five errors (out of twelve questions). A  $t$ -test on mean number of errors in each condition revealed that, as expected, participants made fewer errors when they were free to look at the blank screen (Mean error = 1.39,  $SD = 0.5$ ) than when they maintained fixation (Mean error = 2.5,  $SD = 1.8$ ),  $t(23) = 2.1$ ,  $p = 0.05$ .



**Fig. 8.** “Focus maps” indicating the locations on screen where fixations accumulated during the 10 s period (light is proportional to the sum of fixation time and dark areas received less than 1% of fixation time). All maps relate to the same stimulus (the toucan) in each of the four conditions and are based on data from all of the participants in each condition.

#### 4.4. Discussion

When taking together the findings from the three recall conditions (generating an image versus answering a question about a detail of a previously seen picture with either free or fixed gaze), there was clear support to the conclusion that enactment of fixations occurs also when recalling a long-term memory trace (Humphrey & Underwood, 2008) and not just in a “working memory-like” situation. These findings are also highly consistent with a study by Martarelli and Mast (2013) that used a similar paradigm to that used here. In their study, pictures of animals (e.g., a parrot, a dog) or imaginary creatures (e.g., an angel, a four-legged tennis player) were first learned. When later probing participants with specific questions about the stimuli (e.g., “did the parrot have blue wings?”), either immediately after encoding or one week later, they found that participants looked longer at the areas where the stimuli were originally encoded.

Additionally, the present results support the idea that having the ability to fixate on a previously occupied location can be beneficial to the part’s memory, compared to a condition in which such a spontaneous expression of gaze movement is voluntarily suppressed. Thus the present results both generalize the role of enactment to situations beyond immediate recall or image maintenance and they replicate the findings of Laeng and Teodorescu (2002), consistent with idea that the eye movements during imagery or recall can play a functional role for memory. However, the memory advantage accrued by freely moving the eyes was rather small. Yet, such a small advantage may simply be explained by the fact that the memory task was not particularly challenging (only 12 items had to be remembered and rather simplified pictures of distinctive

animals were used). Indeed, accuracy of recall was generally high. A more challenging memory task, with more items and more similar with one another, could have yielded a greater advantage of the ‘free gaze’ (part focus) over the ‘forced gaze’ (fixed focus) conditions. Moreover, as mentioned earlier, whether the spatial information is relevant or not for solving the task may play a crucial role in whether it can influence or not memory performance. We surmise that some of the studies that failed to show that eye movements during recollection helped memory (e.g., Hoover & Richardson, 2008; Richardson & Spivey, 2000) may have reflected the fact that the task did not need generating a detailed image (e.g., a “token” level representation; cf. Hollingworth & Henderson, 2002) and/or spatial structure of the original scene. In fact, in Richardson and colleagues’ experiments, semantic memory for non-visual (auditory) information was assessed instead of visual information and the tasks could have in principle be performed by participants with their eyes closed. Thus, in their experiments, gaze may have spontaneously tended to resemble the generic structure of the imagined object without being a faithful repetition of what gaze actually did during the original encoding episode.

Nevertheless, in some conditions of their study, Martarelli and Mast (2012) manipulated gaze position during retrieval. Differently from Laeng and Teodorescu’s participants or those of Johansson et al. (2012) who were requested to maintain fixation on a cross during recall, Martarelli and Mast’s participants were free to move their eyes when searching an answer to questions about the objects (e.g. “was the dog sitting?”) but only within predefined areas or quadrants that could have included the object’s original position or not. Allowing eye movements within a quadrant would, according to Martarelli and Mast,

minimize the difference in working memory load between conditions; that is, maintaining fixation while imagining a previously stored image could by itself act like a distraction and decrease performance (Mast & Kosslyn, 2002). Crucially, when participants were probed with specific questions about the stimuli, but they were free to move their eyes within a different quadrant than that of the original position of the stimuli, such a manipulation had no deleterious effect on memory compared to the condition in which they could move the eyes within the correct quadrant.

The above results suggest that the absolute spatial position per se may not be crucial for accurate recall. In fact, one of the fundamental properties of mental imagery is of being dynamic, so that imagined elements can be imagined either translated or rotated or moving in space (Kosslyn, 1980). Moreover, we also seem to be able to re-map eye movements made on a map (or “oculomotor navigation”) to perform a locomotor memory task or navigation in real space (Demichelis, Olivier, & Berthoz, 2012), requiring translation and expansion of the spatial reference frame originally based on the effector-specific coordinates. Such functional properties give imagery its “imaginative” character and can serve a clearly adaptive role, since one is able to conjure up novel combinations of known elements (e.g., a sphynx) or predict future or possible spatial arrangements (e.g., “will this sofa that I see now in the furniture shop fit against the wall and under the window of my living room?”). Hence, it may not be surprising that there was no memory loss in the ‘incorrect quadrant’ condition of Martarelli and Mast’s experiment, since participants should possess the ability to enact a similar pattern of eye movements at encoding by displacing it to another quadrant. Finally, it is questionable that one should expect the memory to be disrupted when forcing gaze to fixate and assume that such a requirement taxes working memory more than moving freely the eye (Johansson et al., 2012). In fact, it seems more likely that unnecessary oculomotor activity would overload the cognitive system and/or interfere with other processes (Weiner & Ehrlichman, 1976). In contrast, fixating gaze on an object is not only a natural behavior in humans but it has been an essential requirement in numerous psychological experiments without these suggesting any evidence that fixation per se would cause a disruption of cognitive processing or a detectable increase in mental effort (e.g., Micic, Ehrlichman, & Chen, 2010; Postle et al., 2006). It is clear within the present account that the act of maintaining fixation should suppress the spontaneous unfolding of gaze and interfere directly with the enactment process and, consequently, with the quality of the memory representation.

In fact, one should not expect significant memory advantages based on gaze enactment when the memory task is little challenging (e.g., recalling a small set of highly distinctive items) and it could be solved on the basis of recalling a shape pattern, color, or semantically (verbally) coded information. In such cases, the solution to the problem may come to mind before reconstruction of the spatial structure of the remembered object or scene is completed, and despite forcing inconsistent fixations during the act of recall. In fact, Johansson and Johansson (2013) have also

used manipulations similar to that of Martarelli and Mast (2013), where participants could view freely on a blank screen or maintain central fixation but also look inside a square area congruent with the location of the to-be-recalled objects or inside a square that was incongruent with their locations. However, they probed their participants in the recall phase about spatial properties of the original scene (e.g., “the car was facing left” or “the train was located to the right of the car”). Interestingly, Johansson and Johansson did observe that looking at a congruent (empty) region facilitated memory retrieval compared to an incongruent region, a finding that is highly consistent with those of the present study as well as of several previous studies (Johansson et al., 2012; Laeng & Teodorescu, 2002; Mäntylä & Holm, 2006).

Nevertheless, the re-constructive process expressed by the eye movements occurring during imagery may not simply be a “replay” of the previous gaze behavior. In fact, if gaze during imagery is an enactive process and the eye movements assist as a scaffolding structure for generating a detailed image, gaze should resemble the structure of the imagined object rather than be a literal re-capitulation of what gaze did during the original perceptual episode or encoding. In fact, according to Kosslyn, Thompson, Sukel, and Alpert (2005), imagery does not necessarily recapitulate all processing that occurred during encoding. In their experiment, PET imaging was used to reveal patterns of brain activation when participants were asked to recall mental images of simple geometrical arrangements. Crucially, while all participants formed identical images of the same figures, some participants were originally given verbal descriptions of what to imagine, whereas others viewed parts of the whole arrangement and then were asked to visualize the complete figure. Given that the brain activity was the same during imagery, despite the patterns were originally encoded differently, these findings indicate that forming a mental image does not re-enact all aspects of what the participants did at the time of encoding.

Richardson and Spivey (2000; Experiment 3) forced participant to keep their eyes still during encoding and to fixate a central cross; in contrast to Laeng and Teodorescu (2002), they failed to observe an enactment of ‘fixation’ on the cross during recall and instead they observed eye movements during retrieval which were based on the spatial representation of the scene and clearly not on the basis of the previous oculomotor behavior (see also Johansson et al., 2012). However, enactive theories would predict that visual images are constituted by (partial) enactment of the perceptual acts that would be carried out if one were actually perceiving whatever is being imagined (Thomas, 2011). According to an enactive account, if imagery uses eye positions as a scaffolding structure for generating a detailed image, then the pattern of gaze during imagery can be based on a subset of the original movements (i.e., relevant gaze pointers; cf. Ballard et al., 1997) and it should actually tend to resemble the structure of the imagined object more than being a faithful repetition, in its minute details, of what gaze actually did during the original perceptual episode or encoding. One speculation based on the above findings and considerations is that the enactment of fixations that occurred at encoding may be an “op-

tional” strategy, most likely put to use in situations where retrieving the original anchor points could make a difference for solving the task. In Richardson and Spivey’s (2000) task, the location of the stimulus was entirely irrelevant to the task itself (i.e., remembering verbal information presented auditorily), therefore fidelity to the episode as encoded in spatial terms would not have been a useful strategy.

## 5. General discussion

The present experiments showed highly correlated patterns of gaze between perception and imagery of a same visual object. Specifically, at imagery retrieval, gaze fixations were likely to re-occur over the same regions of space as those scrutinized during the encoding or perceptual scrutiny of the shape. In the first experiment, we observed that imagining upright versus upside-down triangles resulted in a pattern of fixation that mirrored the shape and orientation of each form over the same region of the screen where they had originally appeared. In the second experiment, more complex shapes consisting in drawings of animals were used. As predicted, during imagery, gaze dwelled mainly within regions of the visual field that during perception had contained cluster of salient features (i.e., most of the total fixation time was spent over the head region). Crucially, those individuals who showed a high fidelity in enacting during imagery the pattern of fixation shown during perception also showed higher scores in a spatial memory task (Experiment 2) and when gaze was forced to remain on a fixation point distant from the original fixations (Experiment 3C), memory accuracy was slightly but significantly lower than when gaze was allowed to roam freely during recall (Experiment 3B). These findings strengthens the idea that fixations during the process of recall may index a process of image generation, so that the more faithfully an oculomotor pattern is retrieved and enacted, the better the quality of the memory of the episode. Specifically, the original size of the previously seen shape was remembered better, when the resemblance between oculomotor patterns during perception and imagery was higher (see also: Johansson & Johansson, *in press*; Johansson et al., 2012; Laeng & Teodorescu, 2002; Mäntylä & Holm, 2006).

A few previous studies have failed to reveal positive correlations between looking at the correct location and correct recall (e.g., Hoover & Richardson, 2008; Richardson & Kirkham, 2004; Richardson & Spivey, 2000; Martarelli & Mast, 2011). This may suggest the tentative conclusion that gaze enactment during recall may be used as an “optional” resource when it can improve memory performance. That is, re-constructing the spatial structure is a time-consuming process, given that it is based on a serial process of generating the elements of an image one-by-one and placing them in their correct positions in relation to one another. A host of other retrieval processes, like activating a particular pattern representation, color, or semantic (verbal) information may, in many cases, reach the solution before the imagery or enactment process is completed. Indeed, a similar account relatively to the relevance

of spatial information in object recognition was proposed by Warrington and James (1986). As originally shown in a study by Warrington and Taylor (1973), patients with right parietal lesions had difficulties in recognizing objects when viewed at unconventional perspectives. According to Marr (1982), these findings suggested that the patients’ difficulties reflected an inability to transform or align an internal spatial frame of reference centered on the object’s intrinsic coordinates (i.e., its axes of elongation) so as to match the perceived image. However, it is well known that damage to the dorsal system in the human brain does not typically result in object recognition problems or agnosia (Farah, 1990). The bottom line is that a spatial analysis of an object may come useful in challenging object recognition (or recall) situations (Laeng, Carlesimo, Caltagirone, Capasso, & Miceli, 2000; Laeng et al., 1999).

The finding that gaze during imagery or memory of an absent stimulus re-enacts perception is also consistent with the idea, originally proposed by the philosopher David Hume (in *An Enquiry Concerning Human Understanding*), that imagining seeing an object is very much like actually seeing it (cf. Kosslyn, Thompson, & Ganis, 2006; Kosslyn et al., 1999; Laeng & Sulutvedt, 2014). Interestingly, neuro-imaging studies also strongly support this account, since the patterns of activity within visual cortex are nearly identical when perceiving something and later imagining it. A logical consequence of this state of things is that such a neural “re-presentation” (Kosslyn, 1999) of the imagined form should be expected to provoke the same type of cognitive control processes in the brain that a real perceptual stimulus would provoke (Ganis, Thompson, & Kosslyn, 2004). Specifically, one would expect that imagining a form would trigger a similar pattern of eye movements (and perhaps even the sequence or “scanpath”) over the positions of the parts of an imagined shape despite their absence within the visual field. Consequently, eye movements will occur that closely resemble those shown during the original perceptual episode (Laeng & Teodorescu, 2002).

Alternatively, from a Hebbian perspective of mental imagery, the mental retrieval of a pattern is a “reconstruction” of the original visual image as encoded during perception (Hebb, 1968) and oculomotor activity that occurred during perception constitutes a relevant aspect of the retrieval component which assists the operation of recombining together the pieces of information of the memory episode. Hence, gaze shifts are not only mechanically necessary during perception, for scanning an object and extracting detailed information from it, but they may also have, at a later stage of recall, an “organizing function”, especially when imagery is used as a mnemonic retrieval mechanism of information related to the absent object. Kosslyn (1987, 1994) distinguished several types of visual imagery, among which he proposed forms of imagery that are “attention-based” (see also Thomas, 1999). These types of imagery would require engaging attention at the different locations of a multipart image, so that each part or component gets added in the correct location onto the image under construction. Importantly, eye movements index in an explicit or overt manner the occurrence of shifts in spatial attention (Henderson,

1992; Rizzolatti et al., 1987). Thus, within an attention-based imagery account, overt shifts of attention can be interpreted as a direct reflection of the image generation process. Specifically, motor-based information about the eyes' positions is automatically retrieved and the resulting eye movements reflect the active process of image construction. In other words, spatial pointers are included in the episodic trace associated with each learned item and when the trace is activated then also the encoded location of the item is necessarily activated and this component can automatically drive the eyes towards that location (Altmann, 2004).

Thus, there seem to be two alternative ways for interpreting the occurrence of eye movements and fixations during visual imagery: (a) as a reflection of the formation of a quasi-perceptual state in the visual cortex (i.e., the oculomotor system may not distinguish between a bottom-up driven percept in the cortex and a similar neuronal state that is generated top-down and equally react to both); versus (b) eye fixations reflect the serial activation of oculomotor memory traces that, when actively retrieved, serve as a spatial code for re-constructing the image. Although these two accounts could be viewed as incompatible alternatives, in our opinion, the two accounts are not mutually exclusive and it may be more useful to regard them as complementary instead of competing. That is, "empty" movements during recollection may be both a consequence of activating the integrated episodic representation and a strategy for facilitating the retrieval of further information (Ferreira et al., 2008). In other words, gaze programming and execution may not only assist image generation but, at the same time, react to parts of an already generated image to guide further attentional scrutiny and possibly discover implicitly stored features (cf. Rouw, Kosslyn, & Hamel, 1997; Thompson et al., 2008). Neisser (1976), Neisser (1978) had already envisaged such an interactive relationship when positing a common process of analysis-by-synthesis between perception and imagery (see also Bar et al., 2006; Bever & Poeppel, 2010; Yuille & Kersten, 2006). According to Neisser (1976, pp. 130–131), "the experience of having an image is just the inner aspect of a readiness to perceive the imagined object" and so that imagining and seeing are "only parts of a perceptual cycle" and under the control of "plans for obtaining information from potential environments." Within this account, imagery could be an anticipatory phase of perception like a "disposition to see" (see also Freyd, 1987; Grush, 2004; Kosslyn & Sussman, 1994; Ryle, 1949) that takes place all the time; only when the perceptual pickup of information is either interrupted or delayed, imagery becomes subjectively experienced. Thus, in Neisser's account, the role of eye fixations during imagery seems particularly relevant, since anticipating visual information can guide gaze to the likely locations where this information will be found (cf. Vickers, 2007). An image could represent something partially illuminated or occluded or concealed, but about to appear in a specific location, and it could therefore prepare the visual system to its occurrence and even constitute the basis of humans' "amodal perception" ability (Nanay, 2010; Plomp, Nakatani, Bonnardel, & van Leeuwen, 2004; Strawson, 1974). Never-

theless, images can also be deliberately used as a way of remembering things, in which case the sensorimotor information is activated off-line with no support from the visual input (Cuthbert, Vrana, & Bradley, 1991; Glenberg, 1997; Grush, 2004).

The above account seems also highly consistent with current cognitive and philosophical perspectives that stress that perception, or more generally cognition, is "active" and/or "embodied" (Gibbs, 2006; Thomas, 1999; Varela, Thompson, & Rosch, 1991). Intuitively, gaze may seem to differ from most bodily actions in that, while these can have direct consequences in the environment as well as on the physical state of the agent (e.g., hand movements), the motions of the eyes seem to have as their main purpose to obtain information. But this view is clearly insufficient and, according to active/embodied accounts that are influenced by the writings of philosophers like Dewey or Merleau-Ponty, perception is not only about storing descriptions or pictures but is "procedural" in kind; that is, the visual brain encodes also how to direct attention, examine, and explore an object. As John Dewey put it in a seminal article in the *Psychological Review* (1896, p. 358), "Upon analysis, we find that we begin not with a sensory stimulus, but with a sensori-motor coordination, the optical-ocular, and that in a certain sense it is the movement which is primary, and the sensation which is secondary, the movement of body, head and eye muscles determining the quality of what is experienced. In other words, the real beginning is with the act of seeing; it is looking, and not a sensation of light." Indeed, eye movements may constitute the majority of human behaviors in a lifetime and, if our eyes did not flick about the environment, one would be practically unable to see (Bridgeman, 1992; cf. Findlay & Gilchrist, 2003).

Within the active/embodied perspective, imagery quintessentially 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 (i.e., the embodiment). Lakoff and Johnson (1999) suggested that there exist at least three levels of embodiment: the neural level (i.e., cognition is what the brain does), the phenomenological level (i.e., we experience or subjectively "feel" our actions and perceptions), and the cognitive unconscious level (i.e., most cognitive processes are "unfelt" and the mind operates on the basis of sensori-motor information that goes beyond the classic five senses and which to a great extent is not accessible to consciousness). Within the present account, we believe that the latter (cognitive unconscious) level of embodiment better captures the role of oculomotor mechanisms in imagery. Although we may have a rough feel of where our eyeballs may be at a specific time within the eye sockets, the conscious proprioceptive information provided by the eye muscle spindles during the rapid movements and brief fixations is extremely weak (Bridgeman & Stark, 1991). Given that the phenomenological level is unlikely to play a significant role in accounting for these phenomena, the present findings avoid a classic criticism of mental imagery research, where behavior during imagery is explained away as compliance to task demands on the basis of tacit knowledge



(e.g., [Intons-Peterson, 1983](#)). That is, participants in mental imagery experiments recapitulate the behavior of perception simply on the basis of their intuition of what the experimenters are expecting them to do and “simulate” their past perceptual behavior on the basis of their memory of what one did in a similar situation. However, as pointed out by [Brandt and Stark \(1997\)](#), the voluntary recreation of complex scanpaths, like those recorded in this and previous experiments, on the basis of extremely weak proprioceptive signals, would seem an unlikely memory feat.

Because of such unconscious embodied knowledge, perceptual experience and the sensori-motor event can be “simulated” or “emulated” in a very different sense than that intended by a compliance to task demands ([Grush, 2004](#); [Moulton & Kosslyn, 2009](#)). Specifically, the sensory and motor systems can be spontaneously active without the presence of an external stimulus and without necessarily executing the movement (or by activating only a subset of the movements or a weaker version of the original action). By doing so, the process of recall or recognition or more generally our mental representations could be established more efficiently (e.g., [Barsalou et al., 2003](#); [Pezzulo et al., 2001](#); [Pulvermüller & Fadiga, 2010](#); [Zwaan & Taylor, 2006](#)). Indeed, in the present study, eye movements that during imagery showed a very similar activity pattern to those observed during perceptual encoding were also associated to more accurate and robust episodic representations.

In general, eye movements can be seen to provide a coordinate frame or “pointers” that can be used in the memory encoding of information about the external world (cf. [O’Regan, 1992](#)). If the object is still present in the visual field, like during recognition or controlled action, the details of the shape will be efficiently searched by using eye positions as “deictic” markers or pointers ([Ballard et al., 1997](#); [Xu & Chun, 2006](#)). The ability to use an external frame of reference centered at the fixation point that can be rapidly moved to different locations leads to great simplifications in algorithmic complexity ([Ballard, 1991](#)) and vision can be modeled as composed by mechanical pointing devices (eye fixations and grasping) and localization by attention as a neural pointing device ([Rolfs et al., 2011](#)).

If the object is no longer present, as in imagery recall, then the oculo-motor based “pointers” could help re-constructing the visual representation by regenerating a spatially similar pattern of activity within the visual cortex ([Ganis et al., 2004](#); [Kosslyn, 1993, 1994](#); [Kosslyn, Thompson, Kim, & Alpert, 1995](#); [Kosslyn, Thompson, Kim, Rauch, & Alpert, 1996](#)). Such a neural reconstruction of a pattern seems possible especially in light of the fact that the visual areas of the brain are topographically organized in retinotopic “visual field maps” ([Kaas, 1997](#); [Tootell et al., 1998](#); [Wandell, Brewer, & Dougherty, 2005](#)). These maps preserve, to some extent, the geometric structure of the retina, which in turn, by the laws of optic refraction, reflects the geometric structure of the external visual world as a planar projection onto a two dimensional surface. A topographically organized structure can depict visual information as “points” organized by their relative locations in space.

Points near each other in the represented space are represented by points near each other in the representing substrate; this (internal) space can be used to represent (external) space ([Markman, 1999](#)). Thus, one fundamental property of the brain’s representation of space is that “the brain uses space on the cortex to represent space in the world” ([Kosslyn et al., 2006](#)). Another advantage of a topographical organization of the visual brain is in guiding ocular movements by maintaining a faithful representation of the position of the target of a saccade ([Optican, 2005](#)). Primate studies also show that the parietal lobes contain neurons that encode the shape of objects ([Sereno & Maunsell, 1998](#); [Taira, Mine, Georgopoulos, Murata, & Sakata, 1990](#)). Thus, human cortical structures contain motor-relevant or effector-based information about the shapes of some objects, information that would seem necessary in order to control efficiently specific actions, including directing eye movements towards relevant shape features ([Moore, 1999](#)). Interestingly, such gaze control mechanisms can operate extremely fast despite being dependent on object processing within cortical pathways ([Miles, 1998](#)).

To conclude, eye fixations while observers look at an empty screen and at the same time generate a visual image by recalling something they had just seen tend to dwell within those empty regions corresponding to the original positions of the object’s salient features (e.g., regions where several local features are naturally crowded, like an animal’s head). Thus, observers enact a similar oculomotor behavior during imagery to that they displayed during perception. The fidelity of such an enactment also predicts the level of accuracy in remembering a visuo-spatial property of the object (i.e., dimension or size) as experienced in the original episode and interfering with the enactment process can result in measurable costs in the quality of memory. These findings support the conclusion that fixations reflect the feature complexity of the seen form and do not simply return to a generic position originally occupied by the remembered object.

## References

- Addis, D. R., Wong, A. T., & Schacter, D. L. (2007). Remembering the past and imagining the future: Common and distinct neural substrates during event construction and elaboration. *Neuropsychologia*, *45*, 1363–1377.
- Altmann, G. T. M. (2004). Language-mediated eye movements in the absence of a visual world: The ‘blank screen paradigm’. *Cognition*, *93*, 79–87.
- Altmann, E. M., & John, B. E. (1999). Episodic indexing: A model of memory of attention events. *Cognitive Science*, *23*, 117–156.
- Andrade, J., Kavanagh, D., & Baddeley, A. (1997). Eye-movements and visual imagery: A working memory approach to the treatment of post-traumatic stress disorder. *British Journal of Clinical Psychology*, *36*, 209–223.
- Antrobus, J. S., Antrobus, J. S., & Singer, J. L. (1964). Eye movements accompanying daydreaming, visual imagery, and thought suppression. *Journal of Abnormal and Social Psychology*, *69*, 244–252.
- Aserinsky, E., & Kleitman, N. (1953). Two types of ocular motility occurring in sleep. *Journal of Applied Psychology*, *8*, 1–10.
- Ballard, D. H. (1991). Animate vision. *Artificial Intelligence*, *48*, 57–86.
- Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. N. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, *20*, 723–767.
- Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., et al. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences*, *103*, 449–454.

- Barrowcliff, A. L., Gray, N. S., Freeman, T. C. A., & Macculloch, M. J. (2004). Eye-movements reduce the vividness, emotional valence and electrodermal arousal associated with negative autobiographical memories. *The Journal of Forensic Psychiatry & Psychology*, *15*, 325–345.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*, 577–600.
- Barsalou, L. W., Simmons, W. K., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*, *7*, 84–91.
- Bever, T. G., & Poeppel, D. (2010). Analysis by synthesis: A (re)-emerging program of research for language and vision. *Biolinguistics*, *4*, 174–200.
- Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*, 115–147.
- Bourlon, C., Oliviero, N., Wattiez, P., Pouget, P., & Bartolomeo, P. (2011). Visual mental imagery: What the head's eye tells the mind's eye. *Brain Research*, *1367*, 287–297.
- Brandt, S. A., Stark, L. W., Hacısalihzade, S., Allen, J., & Tharp, G. (1989). Experimental evidence for scanpath eye movements during visual imagery. In *Proceedings of the 11th IEEE conference on engineering, medicine, and biology*.
- 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. (1992). Conscious vs. unconscious processes: The case of vision. *Theory and Psychology*, *2*, 73–88.
- Bridgeman, B., & Stark, L. (1991). Ocular proprioception and efference copy in registering visual direction. *Vision Research*, *31*, 1903–1913.
- Brockmole, J. R., & Irwin, D. E. (2005). Eye movements and the integration of visual memory and visual perception. *Perception & Psychophysics*, *67*, 495–512.
- Buswell, G. T. (1935). *How people look at pictures*. Chicago: University of Chicago Press.
- Castelhano, M. S., Mack, M. L., & Henderson, J. M. (2009). Viewing task influences eye movement control during active scene perception. *Journal of Vision*, *9*(3), 6, 1–15.
- Chelnokova, O., & Laeng, B. (2011). Three-dimensional information in face recognition: An eye-tracking study. *Journal of Vision*, *11*(13), 1–15.
- Craighero, L., Nascimben, M., & Fadiga, L. (2004). Eye position affects orienting of visuospatial attention. *Current Biology*, *14*, 331–333.
- Cuthbert, B. N., Vrana, S. R., & Bradley, M. M. (1991). Imagery: Function and physiology. In P. K. Ackles, J. R. Jennings, & M. G. A. Coles (Eds.). *Advances in psychophysiology* (vol. 4). JAI Press.
- de Grave, D. D. J., & Bruno, N. (2010). The effect of the Müller-Lyer illusion on saccades is modulated by spatial predictability and saccadic latency. *Experimental Brain Research*, *203*, 671–679.
- DeAngelus, M., & Pelz, J. B. (2009). Top-down control of eye movements: Yarusus revisited. *Visual Cognition*, *17*(6), 790–811.
- Decety, J. (1996). Do imagined and executed actions share the same neural substrate? *Cognitive Brain Research*, *3*, 87–93.
- Decety, J., Jeannerod, M., & Prablanc, C. (1989). The timing of mentally represented actions. *Behavioral Brain Research*, *34*, 35–42.
- Deckert, G. H. (1964). Pursuit eye movements in the absence of a moving visual stimulus. *Science*, *143*, 1192–1193.
- Deco, G., & Schürmann, B. (2000). A neuro-cognitive visual system for object recognition based on testing of interactive attentional top-down hypotheses. *Perception*, *29*, 1249–1264.
- Demarais, A. M., & Cohen, B. H. (1998). Evidence for image-scanning eye movements during transitive inference. *Biological Psychology*, *49*, 229–247.
- Dement, W., & Kleitman, N. (1957). The relation of eye movements during sleep to dream activity: An objective method for the study of dreaming. *Journal of Experimental Psychology: General*, *55*, 339–346.
- Demicheli, A., Olivier, G., & Berthoz, A. (2012). Motor transfer from map ocular exploration to locomotion during spatial navigation from memory. *Experimental Brain Research*, *224*, 605–611.
- deSperati, C. (2003). Precise oculomotor correlates of visuospatial mental rotation and circular motion imagery. *Journal of Cognitive Neuroscience*, *15*, 1244–1259.
- Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, *36*, 1827–1837.
- Dijkstra, K., Kaschak, M. P., & Zwaan, R. A. (2008). Body posture facilitates retrieval of autobiographical memories. *Cognition*, *102*, 139–149.
- Doricchi, F., Iaria, G., Silvestri, M., Figliozzi, F., & Siegler, I. (2007). The "ways" we look at dreams: Evidence from unilateral spatial neglect (with an evolutionary account of dream bizarreness). *Experimental Brain Research*, *178*, 450–461.
- Einhäuser, W., Spain, M., & Perona, P. (2008). Objects predict fixations better than early saliency. *Journal of Vision*, *8*(14), 1–26.
- Ellis, R. D. (1995). *Questioning consciousness: The interplay of imagery, cognition, and emotion in the human brain*. Amsterdam: John Benjamins.
- Farah, M. J. (1990). *Visual agnosia: Disorders of object recognition and what they tell us about normal vision*. Cambridge, MA: The MIT Press.
- Ferreira, F., Apel, A., & Henderson, J. M. (2008). Taking a new look at looking at nothing. *Trends in Cognitive Science*, *12*(11), 405–410.
- Findlay, J. M., & Gilchrist, I. D. (2003). *Active vision: The psychology of looking and seeing*. Oxford University Press.
- Freyd, J. J. (1987). Dynamic mental representations. *Psychological Review*, *94*, 427–438.
- Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2004). Brain areas underlying visual mental imagery and visual perception. *Cognitive Brain Research*, *20*, 226–241.
- Gbadamosi, J., & Zangemeister, W. H. (2001). Visual imagery in hemianopic patients. *Journal of Cognitive Neuroscience*, *13*(7), 855–866.
- Gibbs, R. W. (2006). *Embodiment and cognitive science*. New York: Cambridge University Press.
- Gibson, J. (1979). *The ecological approach to visual perception*. Lawrence Erlbaum.
- Giudice, N. A., Klatzky, R. L., Bennett, C. R., & Loomis, J. M. (2013). Combining locations from working memory and long-term memory into a common spatial image. *Spatial Cognition and Computation*, *13*, 103–128.
- Glenberg, A. (1997). What memory is for. *Behavioral and Brain Sciences*, *20*, 1–55.
- Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin and Review*, *9*, 558–565.
- Glenberg, A. M., Schroeder, J. L., & Robertson, D. A. (1998). Averting gaze disengages the environment and facilitates remembering. *Memory & Cognition*, *26*, 651–658.
- Goldthwait, C. (1933). Relation of eye movements to visual imagery. *American Journal of Psychology*, *45*, 106–110.
- Grant, E. R., & Spivey, M. J. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science*, *14*, 462–466.
- Greene, M. R., Liu, T., & Wolfe, J. M. (2012). Reconsidering Yarbus: A failure to predict observers' task from eye movement patterns. *Vision Research*, *62*, 1–8.
- Grush, R. (2004). The emulation theory of representation: Motor control, imagery, and perception. *Behavioral and Brain Sciences*, *27*, 377–442.
- Gunter, R. W., & Bodner, G. E. (2008). How eye movements affect unpleasant memories: Support for a working-memory account. *Behaviour Research and Therapy*, *46*, 913–931.
- Hadamard, J. (1945). *The psychology of invention in the mathematical field*. New York: Dover Publications.
- Hafed, Z. M., & Krauzlis, R. J. (2006). Ongoing eye movements constrain visual perception. *Nature Neuroscience*, *9*, 1449–1457.
- Hassabis, D., Kumaran, D., & Maguire, E. A. (2007). Using imagination to understand the neural basis of episodic memory. *Journal of Neuroscience*, *27*, 14365–14374.
- Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences*, *9*, 188–194.
- Hayward, W. G., Tarr, M. J., & Corderoy, A. K. (1999). Recognizing silhouettes and shaded images across depth rotation. *Perception*, *28*, 1197–1215.
- Hebb, D. O. (1968). Concerning imagery. *Psychological Review*, *75*, 466–477.
- Henderson, J. M. (1992). Visual attention and eye movement control during reading and picture viewing. In K. Rayner (Ed.), *Eye movements and visual cognition* (pp. 260–283). Berlin: Springer Verlag.
- Hodgson, T. L., Bajwa, A., Owen, A. M., & Kennard, C. (2002). The strategic control of gaze direction in the tower of London task. *Journal of Cognitive Neuroscience*, *12*, 894–907.
- Hollingworth, A. (2005). The relationship between online visual representation of a scene and long-term scene memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(3), 396–411.
- Hollingworth, A. (2006). Scene and position specificity in visual memory for objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 58–69.
- Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 113–136.
- Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Science*, *8*, 494–500.

- Hommel, B., Musseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, 24, 849–878.
- Hong, C. C.-H., Harris, J. C., Pearson, G. D., Kim, J.-S., Calhoun, V. D., Fallon, J. H., et al. (2009). fMRI evidence for multisensory recruitment associated with rapid eye movements during sleep. *Human Brain Mapping*, 30, 1705–1722.
- Hong, C. C.-H., Potkin, S. G., Antrobus, J. S., Dow, B. M., Callaghan, G. M., & Gillin, J. C. (1997). REM sleep eye movement counts correlate with visual imagery in dreaming: A pilot study. *Psychophysiology*, 34, 377–381.
- Hoover, M. A., & Richardson, D. C. (2008). When facts go down the rabbit hole: Contrasting features and objecthood as indexes to memory. *Cognition*, 108, 533–542.
- Humphrey, K., & Underwood, G. (2008). Fixation sequences in imagery and in recognition during the processing of pictures and real-world scenes. *Journal of Eye Movement Research*, 2, 1–15.
- 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.
- Jacob, M., & Hochstein, S. (2011). Gathering and retaining visual information over recurring fixations: A model. *Cognitive Computation*, 3, 105–123.
- Jacobson, E. (1932). Electrophysiology of mental activities. *American Journal of Psychology*, 44, 677–694.
- Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences*, 17, 187–245.
- Jeannerod, M., & Moutet, J. (1962). Etude des mouvements oculaires observés chez l'homme au cours de la veille et du sommeil. *Comptes Rendus de la Société de Biologie*, 156, 1407–1410.
- Johansson, R., & Johansson, M. (2013). Look here, eye movements play a functional role in memory retrieval. *Psychological Science*, 25(1), 236–242.
- Johansson, R., Holsanova, J., Dewhurst, R., & Holmqvist, K. (2012). Eye movements during scene recollection have a functional role, but they are not reinstatements of those produced during encoding. *Journal of Experimental Psychology: Human Perception and Performance*, 38, 1289–1314.
- Johansson, R., Holsanova, J., & Holmqvist, K. (2006). Pictures and spoken descriptions elicit similar eye movements during mental imagery, both in light and in complete darkness. *Cognitive Science*, 30, 1053–1079.
- Johansson, R., Holsanova, J., & Holmqvist, K. (2011). The dispersion of eye movements during visual imagery is related to individual differences in spatial imagery ability. In L. Carlson, C. Holscher, & T. Shipley (Eds.), *Proceedings of the 33rd annual meeting of the cognitive science society* (pp. 1200–1205). Austin, TX: Cognitive Science Society.
- Jolicoeur, P., & Kosslyn, S. M. (1985). Is time to scan visual images due to demand characteristics? *Memory and Cognition*, 13, 320–332.
- Jones, S. A. H., & Henriques, D. Y. P. (2010). Memory for proprioceptive and multisensory targets is partially coded relative to gaze. *Neuropsychologia*, 48, 3782–3792.
- Judd, C. H. (1905). The Müller-Lyer illusion. *Psychological Monographs*, 7(1), 55–81.
- Kaas, J. H. (1997). Topographic maps are fundamental to sensory processing. *Brain Research Bulletin*, 44, 107–112.
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24, 175–219.
- Kan, I. P., Barsalou, L. W., Solomon, K. O., Minor, J. K., & Thompson-Schill, S. L. (2003). Role of mental imagery in a property verification task: fMRI evidence for perceptual representations of conceptual knowledge. *Cognitive Neuropsychology*, 20, 525–540.
- Karn, K., & Hayhoe, M. (2000). Memory representations guide targeting eye movements in a natural task. *Visual Cognition*, 7, 673–703.
- Kaufman, L., & Richards, W. (1969). Spontaneous fixation tendencies for visual forms. *Perception and Psychophysics*, 5, 85–88.
- Kent, C., & Lamberts, K. (2008). The encoding–retrieval relationship: Retrieval as mental simulation. *Trends in Cognitive Sciences*, 12, 92–98.
- Kolers, P. (1973). Remembering operations. *Memory and Cognition*, 3, 347–355.
- Kosslyn, S. M. (1987). Seeing and imagining in the cerebral hemispheres: A computational approach. *Psychological Review*, 94, 148–175.
- Kosslyn, S. M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.
- Kosslyn, S. M. (1994). *Image and brain*. Cambridge, MA: The MIT Press.
- Kosslyn, S. M. (2002). What shape are a German Shepherd's ears? A talk with Stephen M. Kosslyn. In John Brockman (Ed.), *The reality club*. <[http://www.edge.org/3rd\\_culture/kosslyn/kosslyn\\_print.html](http://www.edge.org/3rd_culture/kosslyn/kosslyn_print.html)>.
- Kosslyn, S. M., Alpert, N. M., Thompson, W. L., Maljkovic, V., Weise, S. B., Chabris, C. F., et al. (1993). Visual imagery activates topographically organized visual cortex: PET investigations. *Journal of Cognitive Neuroscience*, 5, 263–287.
- Kosslyn, S. M., Cave, C. B., Provost, D. A., & Von Gierke, S. (1988). Sequential processes in image generation. *Cognitive Psychology*, 20, 319–343.
- Kosslyn, S. M. (1999). Visual mental images as re-presentations of the world. In J. S. Gero & B. Tversky (Eds.), *Visual and spatial reasoning in design* (pp. 83–92). Sydney, Australia: University of Sydney Press.
- Kosslyn, S. M. (2007). Remembering images. In M. A. Gluck, J. R. Anderson, & S. M. Kosslyn (Eds.), *Memory and mind: A festschrift for Gordon H. Bower*. Hillsdale, NJ: Erlbaum Associates.
- Kosslyn, S. M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J. P., Thompson, W. L., et al. (1999). The role of area 17 in visual imagery: Convergent evidence from PET and rTMS. *Science*, 284, 167–170.
- Kosslyn, S. M., Reiser, B. J., Farah, M. J., & Fliegel, S. L. (1983). Generating visual images: Units and relations. *Journal of Experimental Psychology: General*, 112, 278–303.
- Kosslyn, S. M., & Sussman, A. L. (1994). Roles of imagery in perception: Or, there is no such thing as immaculate perception. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 1035–1042). Cambridge, MA: MIT Press.
- Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). *The case for mental imagery*. Neurosciences. Oxford, UK: Oxford University Press.
- Kosslyn, S. M., & Thompson, W. L. (2000). Shared mechanisms in visual imagery and visual perception: Insights from cognitive neuroscience. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences*. Cambridge, MA: MIT Press.
- Kosslyn, S. M., Thompson, W. L., Kim, I. J., & Alpert, N. M. (1995). Topographical representations of mental images in primary visual cortex. *Nature*, 378, 496–498.
- Kosslyn, S. M., Thompson, W. L., Kim, I. J., Rauch, S. L., & Alpert, N. M. (1996). Individual differences in cerebral blood flow in area 17 predict the time to evaluate visualized letters. *Journal of Cognitive Neuroscience*, 8, 78–82.
- Kosslyn, S. M., Thompson, W. L., Sukel, K. E., & Alpert, N. (2005). Two types of image generation: Evidence from PET. *Cognitive, Affective, & Behavioral Neuroscience*, 5, 41–53.
- Kozhevnikov, M., Kosslyn, S. M., & Shepard, J. (2005). Spatial versus object visualizers: A new characterization of visual cognitive style. *Memory & Cognition*, 33, 710–726.
- Laeng, B., Carlesimo, G. A., Caltagirone, C., Capasso, R., & Miceli, G. (2000). Rigid and non-rigid objects in canonical and non-canonical views: Effects of unilateral stroke on object identification. *Cognitive Neuropsychology*, 19, 697–720.
- Laeng, B., Shah, J., & Kosslyn, S. M. (1999). Identifying objects in conventional and contorted poses: Contributions of hemisphere-specific mechanisms. *Cognition*, 70(1), 53–85.
- Laeng, B., & Sulutvedt, U. (2014). The eye pupil adjusts to imaginary light. *Psychological Science*, 25(1), 188–197.
- Laeng, B., & Teodorescu, D.-S. (2002). Eye scanpaths during visual imagery re-enact those of perception of the same visual scene. *Cognitive Science*, 26, 207–231.
- Laeng, B., Waterloo, K., Johnsen, S. H., Bakke, S. J., Låg, T., Simonsen, S. S., et al. (2007). The eyes remember it: Oculography and pupillometry during recollection in three amnesic patients. *Journal of Cognitive Neuroscience*, 19(11), 1888–1904.
- Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to Western thought*. New York: Basic Books.
- Land, M., Mennie, N., & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception*, 28, 1311–1328.
- Leek, E. C., Cristino, F., Conlan, L. I., Patterson, C., Rodriguez, E., & Johnston, S. J. (2012). Eye movement patterns during the recognition of three-dimensional objects: Preferential fixation of concave surface curvature minima. *Journal of Vision*, 12(1), 1–15.
- Loftus, G. R. (1972). Eye fixations and recognition memory for pictures. *Cognitive Psychology*, 3, 525–551.
- Logie, R. (1995). *Visuo-spatial working memory*. Hillsdale, NJ: Erlbaum.
- Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60, 91–110.
- Mackworth, N. H., & Morandi, A. J. (1967). The gaze selects informative details within pictures. *Perception and Psychophysics*, 2, 547–552.
- Mannan, S., Ruddock, K., & Wooding, D. (1997). Fixation sequences made during visual examination of briefly presented 2D images. *Spatial Vision*, 11, 157–178.
- Mäntylä, T., & Holm, L. (2006). Gaze control and recollective experience in face recognition. *Visual Cognition*, 13(3), 365–386.

- Markman, A. B. (1999). *Knowledge representation*. Psychology Press.
- Marks, D. E. (1973). Visual imagery differences and eye movements in the recall of pictures. *Perception & Psychophysics*, *14*, 407–412.
- Marr, D. (1982). *Vision*. San Francisco, CA: Freeman and Co..
- Martarelli, C., & Mast, F. (2013). Eye movements during long-term pictorial recall. *Psychological Research*, *77*, 303–309.
- Mast, F. W., & Kosslyn, S. M. (2002). Eye movements during visual mental imagery. *Trends in Cognitive Sciences*, *6*, 271–272.
- Melcher, D., & Kowler, E. (1999). Shapes, surfaces and saccades. *Vision Research*, *39*, 2929–2946.
- Micic, D., Ehrlichman, H., & Chen, R. (2010). Why do we move our eyes while trying to remember? The relationship between non-visual gaze patterns and memory. *Brain and Cognition*, *74*, 210–224.
- Miles, F. A. (1998). The neural processing of 3-D information: Evidence from eye movements. *European Journal of Neuroscience*, *10*, 811–822.
- Miyauchi, S., Misaki, M., Kan, S., Fukunaga, T., & Koike, T. (2009). Human brain activity time-locked to rapid eye movements during REM sleep. *Experimental Brain Research*, *192*, 657–667.
- Moore, C. S. (1903). Control of the memory image. *Psychological Review*, *4*, 277–306.
- Moore, T. (1999). Shape representations and visual guidance of saccadic eye movements. *Science*, *285*, 1914–1917.
- Moore, T., & Fallah, M. (2001). Control of eye movements and spatial attention. *PNAS*, *98*, 1273–1276.
- Moulton, S. T., & Kosslyn, S. M. (2009). Imagining predictions: Mental imagery as mental emulation. *Philosophical Transactions of the Royal Society, B*, *364*, 1273–1280.
- Nanay, B. (2010). Perception and imagination: Amodal perception as mental imagery. *Philosophical Studies*, *150*, 239–254.
- Neisser, U. (1976). *Cognition and reality*. San Francisco, CA: W.H. Freeman.
- Neisser, U. (1978). Anticipations, images, and introspection. *Cognition*, *6*, 169–174.
- 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.
- Olsen, R. K., Chiew, M., Buchsbaum, B., & Ryan, J. D. (2014). The relationship between delay period eye movements and visuospatial memory. *Journal of Vision*, *14*(1), 1–11.
- Optican, L. M. (2005). Sensorimotor transformation for visually guided saccades. *Annals of the New York Academy of Sciences*, *1039*, 132–148.
- O'Regan, J. K. (1992). Solving the “real” mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology*, *46*, 461–488.
- Parsons, L. M. (1987). Image spatial transformations of one's body. *Journal of Experimental Psychology: General*, *116*, 172–191.
- Pezzulo, G., Barsalou, L. W., Cangelosi, A., Fische, M. H., McRae, K., & Spivey, M. J. (2001). The mechanics of embodiment: A dialog on embodiment and computational modeling. *Frontiers in Psychology*, *2*(5), 1–21.
- Plomp, G., Nakatani, C., Bonnardel, V., & van Leeuwen, C. (2004). Amodal completion as reflected by gaze durations. *Perception*, *33*(10), 1185–1200.
- Postle, B. R., Idzikowski, C., Della Sala, S., Logie, R. H., & Baddeley, A. D. (2006). The selective disruption of spatial working memory by eye movements. *Quarterly Journal of Experimental Psychology*, *59*, 100–120.
- Pulvermüller, F., & Fadiga, L. (2010). Active perception: Sensorimotor circuits as a cortical basis for language. *Nature Reviews: Neuroscience*, *11*, 351–360.
- Quinn, J. (1994). Toward a clarification of spatial processing. *Quarterly Journal of Experimental Psychology*, *47A*, 465–480.
- Renkewitz, F., & Jahn, G. (2012). Memory indexing: A novel method for tracing memory processes in complex cognitive tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 1622–1639.
- Renninger, L. W., Verghese, P., & Coughlan, J. (2007). Where to look next? Eye movements reduce local uncertainty. *Journal of Vision*, *7*, 1–17.
- Richardson, D. C., Altmann, G. T. M., Spivey, M. J., & Hoover, M. A. (2009). Much ado about eye movements to nothing: A reply to “Taking a new look at looking at nothing”. *Trends in Cognitive Sciences*, *13*(6), 235–236.
- Richardson, D. C., & Kirkham, N. Z. (2004). Multimodal events and moving locations: Eye movements of adults and 6-month-olds reveal dynamic spatial indexing. *Journal of Experimental Psychology: General*, *133*, 46–62.
- Richardson, D. C., & Spivey, M. J. (2000). Representation, space and Hollywood squares: Looking at things that aren't there anymore. *Cognition*, *76*, 269–295.
- Rizzolatti, G., Matelli, M., & Pavesi, G. (1983). Deficits in attention and movements following the removal of postarcuate (area 6) and prearcuate (area 8) cortex in macaque monkey. *Brain*, *106*, 655–673.
- 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.
- Roediger, H. L., Weldon, M. S., & Challis, B. H. (1989). Explaining dissociations between implicit and explicit measures of retention: A processing account. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 3–41). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Roffwarg, H. P., Dement, W. C., Muzio, J. N., & Fisher, C. (1962). Dream imagery: Relation to rapid eye movements of sleep. *Archives of General Psychiatry*, *7*, 235–258.
- Rolfs, M., Jonikaitis, D., Deubel, H., & Cavanagh, P. (2011). Predictive remapping of attention across eye movements. *Nature Neuroscience*, *14*(2), 252–256.
- Rothkopf, C. A., Ballard, D. H., & Hayhoe, M. M. (2007). Task and context determine where you look. *Journal of Vision*, *7*, 1–20.
- Rouw, R., Kosslyn, S. M., & Hamel, R. (1997). Detecting high-level and low-level properties in visual images and visual percepts. *Cognition*, *63*, 209–226.
- Rucci, M., Iovin, R., Poletti, M., & Santini, F. (2007). Miniature eye movements enhance fine spatial detail. *Nature*, *447*, 851–854.
- Ruggieri, V. (1999). The running horse stops: The hypothetical role of the eyes in imagery of movement. *Perceptual and Motor Skills*, *89*, 1088–1092.
- Ryan, J. D., Althoff, R. R., Whitlow, S., & Cohen, N. J. (2000). Amnesia is a deficit in relational memory. *Psychological Science*, *11*, 454–461.
- Ryan, J. D., & Villate, C. (2009). Building visual representations: The binding of relative spatial relations over time. *Visual Cognition*, *17*, 254–272.
- Ryle, G. (1949). *The concept of mind*. London: Hutchinson.
- Sæther, L., Van Belle, W., Laeng, B., Brennen, T., & Øvervoll, M. (2009). Anchoring gaze when categorizing faces' sex: Evidence from eye-tracking data. *Vision Research*, *49*, 2870–2880.
- Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: The prospective brain. *Nature Reviews: Neuroscience*, *8*, 657–661.
- Schacter, D. L., Addis, D. R., & Buckner, R. L. (2008). Episodic simulation of future events: Concepts, data, and applications. *Annals of the New York Academy of Science*, *1124*, 39–60.
- Schütz, A. C., Trommershäuser, J., & Gegenfurtner, K. R. (2012). Dynamic integration of information about salience and value for saccadic eye movements. *Proceedings of the National Academy of Sciences*, *109*, 7547–7552.
- Sereno, A. B., & Maunsell, J. H. R. (1998). Shape selectivities in primate lateral intraparietal cortex. *Nature*, *395*, 500–503.
- Sheehan, P. W., & Neisser, U. (1969). Some variables affecting the vividness of imagery recall. *British Journal of Psychology*, *60*, 71–80.
- Shepherd, M., Findlay, J. M., & Hockey, R. J. (1986). The relationship between eye movements and spatial attention. *Quarterly Journal of Experimental Psychology*, *38A*, 475–491.
- Sima, J. F., Lindner, M., Schultheis, H., & Barkowsky, T. (2010). Eye movements reflect reasoning with mental images but not with mental models in orientation knowledge tasks. *Spatial Cognition*, *VII*, 248–261.
- Singer, J. L., & Antrobus, J. S. (1965). Eye movements during fantasies. *Archives of General Psychiatry*, *12*, 71–76.
- Spivey, M. J., & Geng, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research*, *65*, 235–241.
- Stark, L. W., & Ellis, S. R. (1981). Scanpath revisited: Cognitive models direct active looking. In D. E. Fisher, R. A. Monty, & J. W. Senders (Eds.), *Eye movements: Cognition and visual perception* (pp. 193–226). Hillsdale, NJ: Lawrence Erlbaum.
- Strawson, P. (1974). Imagination and perception. In P. Strawson (Ed.), *Freedom and resentment* (pp. 45–65). London: Methuen.
- Taira, M., Mine, S., Georgopoulos, A. P., Murata, A., & Sakata, H. (1990). Parietal cortex neurons of the monkey related to the visual guidance of hand movement. *Experimental Brain Research*, *83*, 29–36.
- Teichner, W. H., LeMaster, D., & Kinney, P. A. (1978). Eye movements during inspection and recall. In J. W. Senders, D. E. Fisher, & R. A. Monty (Eds.), *Eye movements and the higher psychological functions* (pp. 259–278). Hillsdale, NJ: Lawrence Erlbaum.
- Theeuwes, J., Belopolski, A., & Olivers, C. N. L. (2009). Interactions between working memory, attention and eye movements. *Acta Psychologica*, *132*, 106–114.

- 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.
- Thomas, N. J. T. (2011). Mental imagery. In Edward N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. <<http://plato.stanford.edu/archives/win2011/entries/mental-imagery/>>.
- Thomas, L., & Lleras, A. (2007). Moving eyes and moving thought: On the spatial compatibility between eye movements and cognition. *Psychonomic Bulletin and Review*, 14, 663–668.
- Thompson, W. L., Kosslyn, S. M., Hoffman, M. S., & van der Kooij, K. (2008). Inspecting visual mental images: Can people "see" implicit properties as easily in imagery and perception? *Memory & Cognition*, 36, 1024–1032.
- Tootell, R. B., Hadjikhani, N., Hall, E. K., Marrett, S., Vanduffel, W., Vaughan, J. T., et al. (1998). The retinotopy of visual spatial attention. *Neuron*, 21, 1409–1422.
- Totten, E. (1935). Eye movements during visual imagery. *Comparative Psychology Monographs*, 11, 1–46.
- Trommershäuser, J., Maloney, L., & Landy, M. (2009). The expected utility of movement. In P. Glimcher, C. Camerer, R. Poldrack, & E. Fehr (Eds.), *Neuroeconomics: Decision making and the brain* (pp. 95–111). Amsterdam: Elsevier Inc.
- Tulving, E. (1983). *Elements of episodic memory*. New York: Oxford University Press.
- Ullman, S., Vidal-Naquet, M., & Sali, E. (2002). Visual features of intermediate complexity and their use in classification. *Nature Neuroscience*, 5, 682–687.
- Varela, F., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: MIT Press.
- Vickers, J. N. (2007). *Perception, cognition, and decision training: The quiet eye in action*. Windsor: Human Kinetics.
- Vishwanath, D., & Kowler, E. (2004). Saccadic localization in the presence of cues to three-dimensional shape. *Journal of Vision*, 4, 445–458.
- Wandell, B. A., Brewer, A. A., & Dougherty, R. F. (2005). Visual field map clusters in human cortex. *Philosophical Transactions of the Royal Society, B*, 360, 693–707.
- Warrington, E. K., & James, A. M. (1986). Visual object recognition in patients with right-hemisphere lesions: Axes or features. *Perception*, 15, 355–366.
- Warrington, E. K., & Taylor, A. M. (1973). The contribution of the right parietal lobe to object recognition. *Cortex*, 9, 152–164.
- 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.
- Wexler, M., & Ouarti, N. (2008). Depth affects where we look. *Current Biology*, 18, 1872–1876.
- Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, 440, 91–95.
- Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum Press.
- Yuille, A., & Kersten, D. (2006). Vision as Bayesian inference: Analysis by synthesis? *Trends in Cognitive Sciences*, 10, 301–308.
- Zelinsky, G. J., Loschky, L. C., & Dickinson, C. A. (2011). Do object refixations during scene viewing indicate rehearsal in visual working memory? *Memory and Cognition*, 39, 600–613.
- Zusne, L., & Michels, K. M. (1964). Nonrepresentational shapes and eye movements. *Perceptual and Motor Skills*, 18(1), 1–20.
- Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General*, 135, 1–11.