Running Head: Mental Imagery and Visual Memory

Using Space to Represent Categories: Insights from Gaze Position

Corinna S. Martarelli[1,2]*, Sandra Chiquet[1,2], Bruno Laeng[3], and Fred W. Mast[1,2]

[1]Department of Psychology, University of Bern, Switzerland

[2]Center for Cognition, Learning and Memory, University of Bern, Switzerland

[3] Department of Psychology, University of Oslo, Norway

*Corresponding author: Corinna Martarelli, Department of Psychology, University of Bern,

Fabrikstrasse 8, CH-3012 Bern, Switzerland

Email: corinna.martarelli@psy.unibe.ch; phone: +41 31 631 40 31; fax: +41 31 631 82 12

Using Space to Represent Categories: Insights from Gaze Position

Abstract

We investigated the boundaries between imagery, memory, and perception by measuring gaze during retrieved versus imagined visual information. Eye fixations during recall were bound to the location at which a specific stimulus was encoded. However, eye position information generalized to novel objects of the same category that had not been seen before. For example, encoding an image of a dog in a specific location enhanced the likelihood of looking at the same location during subsequent mental imagery of other mammals. The results suggest that eye movements can also be launched by abstract representations of categories and not exclusively by a single episode or a specific visual exemplar.

Keywords: mental imagery, visual memory, eye gaze, embodied cognition, prediction

Using Space to Represent Categories: Insights from Gaze Position

Introduction

The functional role of eye movements does not seem to be limited to the processing of visual information; eye movements are also relevant in cognitive tasks when there would seem to be no obvious reason to move one's eyes. The so-called "blank screen paradigm" illustrates that empty areas visited during imagery and memory tasks correspond to locations that were inspected during perception (Altmann, 2004; Brandt & Stark, 1997; Foerster, Carbone, Koesling, & Schneider, 2012; Fourtassi et al., 2013; Johansson, Holsanova, Dewhurst, & Holmqvist, 2012; Johansson, Holsanova, & Holmqvist, 2006; Johansson & Johansson, 2014; Laeng, Bloem, D'Ascenzo, & Tommasi, 2014; Laeng & Teodorescu, 2002; Martarelli & Mast, 2011, 2013; Richardson & Spivey, 2000; Scholz, Mehlborn, & Krems, 2016; Spivey & Geng, 2001; Wantz, Martarelli, & Mast, 2015; Wantz, Martarelli, Cazzoli, Kalla, Müri, & Mast, 2016). In a seminal article, Noton and Stark (1971) proposed a scanpath theory, according to which, while scanning a visual scene, the brain stores the sequence of fixations in memory and reactivates it when seeing the image again or when visualizing it later in the absence of any perceptual information (Brandt & Stark, 1997). Foulsham and Kingstone (2013) revisited the scanpath theory and illustrated that areas seen at encoding were indeed more often looked at during recognition. However, their results showed that the order in which the areas were scanned did not correspond to the encoding order as suggested by the scanpath theory. It is noteworthy that a recent study (Bochynska & Laeng, 2015) found some evidence in favor for retention of the scanpath's sequence, thus this issue remains controversial.

In previous studies while recalling from memory the image of an animal (e.g., a dog) that was encoded in a defined area (e.g., upper-left area of the screen), participants spent more time in the same area of interest although the screen was blank (e.g., Laeng & Teodorescu, 2002). Despite many studies have replicated the above finding, it remains unclear whether eye movements during memory/imagery are only related to the recollection of a previous episode with its specific element (e.g., a dog) or can also be generalized to other items in the same category (e.g., other mammals or four-legged animals). Thus, in the present study, we assessed whether the "corresponding area effect" also occurs during visualization of semantically related items (e.g., a cat) that had not been seen or associated with a specific test episode before.

By finding out whether eye fixations transfer to other categories (i.e., *semantic eye fixations*) and not only to specific exemplars or episodes (i.e., *episodic eye fixations*), we will gain considerable insight into the nature of eye movements and the underlying format of mental images. In fact, the existence of *semantic eye fixations* would support a view of mental imagery that is intrinsically flexible and creative in kind, since it would show to be a process that, although grounded on specific past experiences, is able to generalize the past information to novel images (e.g., its generation) by selecting a past episode (e.g., dog) that shares some features with the novel item (e.g., cat).

Another issue that researchers have debated concerns the reference frame used during imagery by which episodic visual memory is encoded and used to trigger eye movements: Some believe that the reference frame is a retinotopic coordinate, others believe it is a location in absolute space, while others believe it is the object's structure. Hoover and Richardson (2008) supported the notion that the object's structure serves as the reference frame because re-fixations seem to follow the new locations of moving objects.

However, the results of other experiments in which eye position was manipulated (e.g., Johansson & Johansson, 2014; Laeng & Teodorescu, 2002; Scholz et al., 2016) suggest that the location in space may be encoded by default. If locations are stored and integrated into the memory trace, then eye movements may play a critical role in many cognitive tasks.

Little is known about the role of eye movements in the representation and organization of categories. The work of Zelinsky and colleagues (e.g., Maxfield, Stalder, & Zelinsky, 2014; Zelinsky, Peng, & Samaras, 2013) highlights the role of eye movements in categorical search tasks (tasks involving finding an object from a target category) by showing that target typicality affects eye behavior. To our knowledge, no study has investigated eye behavior and categories using the *blank screen paradigm*. It seems possible that, in the absence of a specific motoric or spatial component that has been encoded, a new mental image could still launch eye fixations to specific locations. Hoover and Richardson (2008) suggested that the object-based effect they identified plays a role in imagining possible future events. Our working hypothesis is that, when an object's location is encoded, visual and semantic information about that object will generalize or transfer to items of the same or neighboring categories and thus trigger eye movements to the relevant areas. This will occur during visual recollection of old items as well as during mental imagery of new items as long as the new items belong to the same category or are semantically related to old items.

Method

Participants

Twenty-five students (24 female, 1 male) ranging in age from 18 to 40 (M = 22.3, SD = 4.9) took part in this study. The data from one participant had to be excluded because of

technical problems (the tracking ratio was 24.6%). Participants were naïve to the purpose of the experiment and received course credit for participation. They had corrected-to-normal visual acuity.

Apparatus

Eye movements were recorded using an SMI RED tracking system (SensoMotoric Instruments, Teltow, Germany). Data were recorded with a sampling rate of 50 Hz, a spatial resolution of 0.1° and a gaze position accuracy of 0.5°. The eye-tracking device was contact-free and determined the direction of gaze by combining the cornea reflex with the pupil location via an infrared light-sensitive video camera. The stimuli were presented on a 17-inch screen (1280 x 1024 pixel) using SMI Experiment Center Software and eye data were recorded with I-View X Software, both developed by SensoMotoric Instruments (Teltow, Germany).

Stimuli

The items were 64 three-dimensional color images presented as two-dimensional projections taken from an online database (dennisharoldsen.com). Each image belonged to one of four categories (mammals, birds, machines, home furniture). The spatial orientation of the three-dimensional objects was kept constant across categories (right/left). Images belonging to the bird category always appeared in the upper left area, furniture images in the upper right area, machine images in the lower left area, and mammal images in the lower right area. We had two versions of the experiment (Versions A and B). We presented 32 images randomly selected from the initial 64 images to half of the participants and the remaining 32 images to the other half.

Procedure

The experiment was divided into three phases: perceptual encoding phase, distraction phase, and recall phase (see Figure 1 for a schematic representation of the different phases of the experiment). Participants were seated in front of the computer screen. The distance between participant and screen was approximately 70 cm. We used a 5-point calibration and validation procedure (only error values below 0.8° were accepted).

In the perceptual encoding phase, participants were presented with 32 images from the four categories mammals, birds, machines, and home furniture (eight per category). The stimuli appeared for 6 s (preceded by a fixation cross presented for 3 s). Simultaneously, participants heard the name of the presented object. All audio files were created using Audacity (http://audacity.sourceforge.net) and presented via loudspeakers. The stimuli were presented in random order.

In the recall phase, participants were given three tasks: to visualize 32 old and 32 new items (image generation task); to evaluate (true/false) a pre-recorded statement about the visual details of the object such as "the flamingo is standing on one leg" (image inspection task); and to judge whether they had seen the item previously (old/new recognition task). Each of the 32 new items also belonged to one of the four categories (mammals, birds, machines, home furniture; 8 per category). The procedure was similar to procedures used by Kosslyn and colleagues (e.g., Thompson, Kosslyn, Sukel, & Alpert, 2001), who investigated the stages of mental imagery (image generation and image inspection). We added an old/new recognition task to ensure that the participants did not confound the items.

During the recall phase, the screen was blank white (and participants were free to move their eyes). To facilitate spontaneous eye movements, we explicitly avoided using

fixation crosses and the three tasks (image generation, image inspection, old/new recognition) were self-paced (see Fig. 1). After hearing the pre-recorded cue, participants generated the mental image (image generation) and informed the experimenter that they had done so by saying "ok." Then they heard a specific question (auditory file) and gave their response verbally (image inspection). The experimenter instantly pressed a button on the keyboard. Finally, the participants judged whether they had seen the item previously (old/new recognition task); this response was recorded by the experimenter via keyboard. Key presses initiated and terminated the recording of the eye-tracking sequence.

The 64 trials of the recall phase were presented in random order. A distraction phase (involving additions and subtractions for the duration of five minutes) was presented between the perceptual encoding and the recall phase to prevent active rehearsal. At the end of the experiment, the participants were presented with the 64 visual images (centrally on the screen) and the corresponding audio file. They were to evaluate the images with respect to category typicality on a 5-point Likert scale (1 = not at all typical, 2 = not typical, 3 = neutral, 4 = typical, 5 = very typical). The 16 images from each category (mammals, birds, machines, and home furniture) were presented blockwise. Each block was preceded by task instructions (i.e., to rate the typicality of the image for the given category), which appeared on the screen.

Insert Figure 1 about here

A) Encoding Task
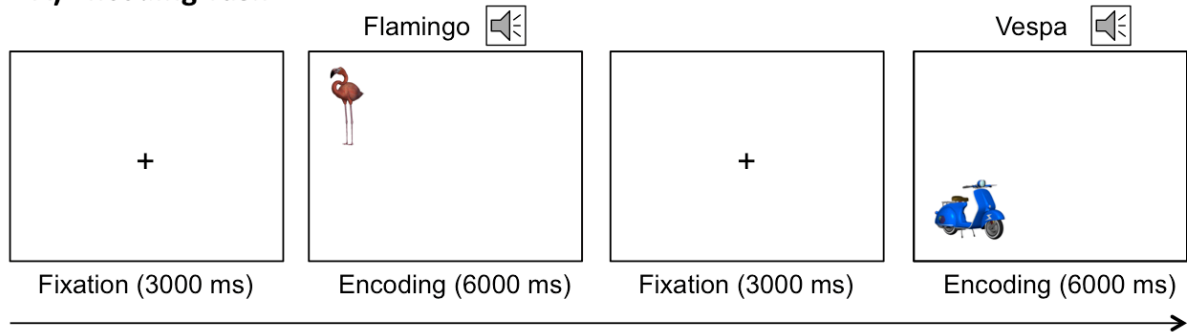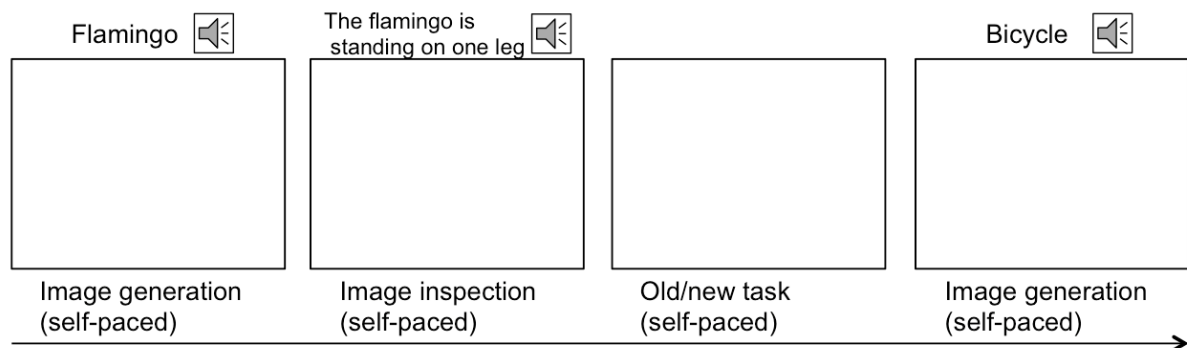

B) Distraction Task (5 min)

C) Recall Task

Figure 1. A schematic representation of the temporal order of events in the experiment including the stimuli used during the different tasks.

Results

The eye data analyses were based on fixations extracted using BeGaze^{TM} software (SensoMotoric Instruments, Teltow, Germany). Fixations were detected when the sum of the gaze stream on the x- and y-axes was within an area of 100 pixels and when the fixation duration exceeded 80 ms. Blink events were automatically subtracted from the original gaze stream by the software and treated as missing data.

The screen was divided at the vertical and horizontal midlines into four equally sized areas of interest (AOIs). The eye data from the perceptual encoding phase and the image generation task, the image inspection task, and the old/new recognition task of the recall

phase were analyzed. The main analyses involved repeated measures ANOVAs in order to compare the time spent in the AOIs in which the pictures were displayed with the mean time spent fixating one of the non-corresponding areas both for old and new items. When Mauchly's test indicated that the sphericity assumption was violated ($p < .05$), we used the Huynh-Feldt correction to adjust the degrees of freedom. We report partial η^2 and Cohen's *d* as measures of effect size.

Validity of stimulus material

We tested whether the four categories differed with respect to perceived typicality and whether the participants' ratings in Version A of the experiment (*n* = 12) differed from the participants' ratings in Version B of the experiment (*n* = 12)[1]. These analyses are reported in Appendix A. The inclusion of experiment version (A, B) in the eye data analyses did not change the results and the factor turned out to be non-significant. Thus, we report the results without experiment version included in the model. The category typicality ratings are reported in Appendix A: They were relatively high with an overall mean of 4.01. We concluded that the objects were valid stimuli.

Behavioral data

Accuracy. Participants were correct in 70.7% (*SD* = 6.8) of the trials with old stimuli in the image inspection task. Participants were correct in 84.5% (*SD* = 9.3) of the trials with old

---

[1] In Version A of the experiment, participants were presented 32 images that had been randomly selected from the initial 64 images; in Version B they were presented the remaining 32 images. The images that were presented in the encoding phase constituted the old items in the recall phase. The remaining 32 images were used as new items in the recall phase.

stimuli in the old/new recognition task and in 89.6% (SD = 7.1) of the trials with new stimuli in the old/new recognition task.

Response times (RTs). After we had removed the outliers (RTs > M + 3 x SD for each participant for each task – in total, 2% of all trials), we performed a repeated measures ANOVA on RTs for correct trials, with task (image generation, image inspection, old/new recognition) and recognition (old, new) as independent variables. The analysis revealed a significant interaction between task and recognition, F(1.701, 39.124) = 8.434, p = .002, partial η^2 = .268. Bonferroni-corrected post hoc tests showed significant differences between old (M = 3637, SD = 830) and new items (M = 4019, SD = 1323) in the image generation task (p = .013), and between old (M = 5667, SD = 872) and new items (M = 5457, SD = 794) in the image inspection task (p = .045), whereas the difference between old (M = 1931, SD = 786) and new items (M = 1983, SD = 871) in the old/new recognition task was non-significant (p = .462). Interestingly, participants were slower in generating new images, but faster in inspecting new images, as compared to old images. The analysis also revealed a significant main effect of task, F(2, 46) = 97.000, p < .001, partial η^2 = .808. Bonferroni-corrected post hoc tests showed significant differences between all tasks, with the RTs in the image inspection task being the longest and those in the old/new recognition task being the shortest (p < .001). The main effect of recognition (old, new) was non-significant, F(1, 23) = 1.187, p = .287, partial η^2 = .049.

The corresponding area effect

Perceptual encoding. As a manipulation check, we compared the time spent in the area in which the stimuli were presented (5251 ms, SD = 374) with the mean time spent in the other three areas (64 ms, SD = 45), t(23) = 63.862, p < .001, d = 13.036. Participants

spent significantly more time in the areas in which the stimuli were presented, thus confirming proper encoding of the stimuli.

Recall phase. Analyses of gaze position during the recall phase were based on correct trials of the old/new recognition task for both new and old items and on correct trials of the inspection task for old items. We conducted a repeated measures ANOVA with gaze position (corresponding area, non-corresponding area), recognition (old, new), and task (image generation, image inspection, and old/new recognition) as within-subject factors. The dependent variable was dwell time in the AOIs (ms). Means are reported in Figure 2. The analysis revealed a significant two-way interaction between task (image generation, image inspection, old/new recognition) and gaze position (corresponding area, non-corresponding area), $F(2, 46) = 6.372$, $p = .004$, partial $\eta^2 = .217$. Post hoc tests with Bonferroni correction indicated that, in all tasks, participants spent more time in the corresponding area than in the non-corresponding area ($p < .002$). The effect was the largest in the image inspection task (1630 ms, $SD = 706$, in the corresponding area vs. 928 ms, $SD = 398$, in the non-corresponding area), followed by the image generation task (1053 ms, $SD = 542$, in the corresponding area vs. 697 ms, $SD = 333$, in the non-corresponding area), and the old/new recognition task (549 ms, $SD = 324$, in the corresponding area vs. 351 ms, $SD = 180$, in the non-corresponding area).

The two-way interaction between task (image generation, image inspection, old/new recognition) and recognition (old, new) was also significant, $F(1.64, 37.63) = 6.005$, $p = .008$, partial $\eta^2 = .207$. Post hoc tests with Bonferroni correction showed that only the difference between old items and new items in the image generation task was significant ($p = .005$). As already illustrated in the RT analyses, participants were slower in the image generation task with new items as compared to old items. The two-way interaction between recognition

(old, new) and gaze position (corresponding area, non-corresponding area) was not significant (p = .107), suggesting that the pattern of results (more time spent in the corresponding area than in the other areas) was similar for old and new items. The three-way interaction was not significant (p = .600). The main effect of task (image generation, image inspection, old/new recognition) yielded a significant result, $F(2, 46)$ = 65.664, $p <$ .001, partial η^2 = .741. Post hoc tests with Bonferroni correction revealed that all differences were significant ($p <$ .001), with the slowest responses being given in the image inspection task, followed by the image generation task, and the fastest responses being given in the old/new recognition task (see the RT analyses above). The main effect of gaze position (corresponding area, non-corresponding area) was also significant, $F(1, 23)$ = 18.61, $p <$ .001, partial η^2 = .447 (with more time spent in the corresponding areas than in the non-corresponding areas). However, the main effect of recognition (old, new) was not significant (p = .721), showing that there was no difference in time between old and new items overall (see RT analyses). The same repeated measures ANOVA with number of fixations as the dependent variable yielded the same results (no differences in rejecting or accepting the null hypothesis).

Insert Figure 2 about here

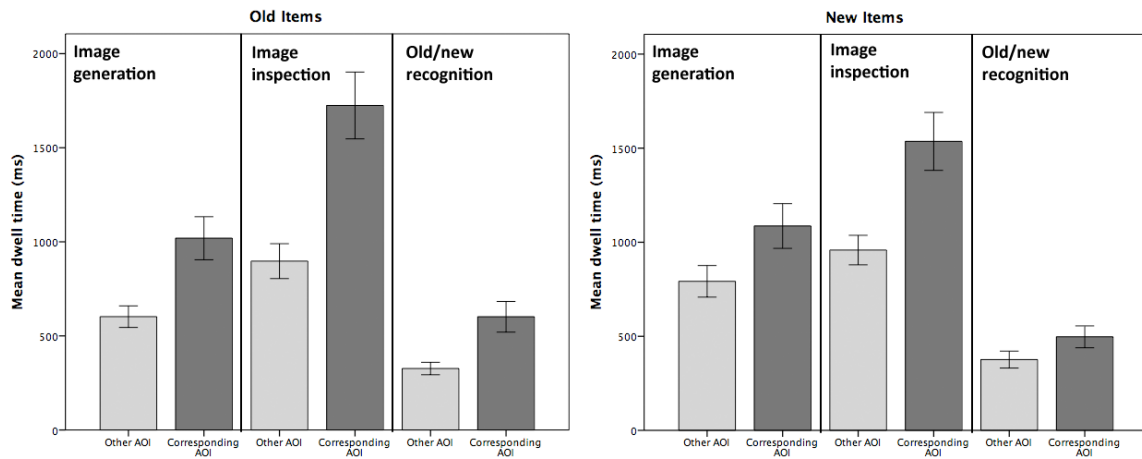---

Figure 2. Mean dwell time (in ms) in the corresponding vs. other AOIs for old and new items in the image generation, image inspection, and old/new recognition tasks. Error bars indicate 1 *SEM*.

Given that percentages allow for comparisons, in Appendix B, we also report the percentage of time spent in the corresponding AOIs (chance level 25%). In addition, we performed a repeated measures ANOVA on the percentage of time spent in the corresponding AOIs, with recognition (old, new) and task (image generation, image inspection, and old/new recognition) as within-subject factors. Only the main effect of recognition (old, new) was significant, $F(1, 23) = 5.55$, $p = .027$, partial $\eta^2 = .195$. Participants spent significantly more time in the corresponding area with old items (38%, *SEM* = .035) than with new items (34%, *SEM* = .026). The main effect of task and the interaction were non-significant ($p > .257$), suggesting that the pattern of results (time spent in the corresponding AOIs) was roughly equal across tasks.

In the analyses presented here, we pooled the dwell time in the non-corresponding areas. In order to ensure that pooling did not introduce bias, we computed separate analyses with unpooled data. The results are reported in Appendix C.

Error trials. We also conducted a repeated measures ANOVA on fixations during error trials with gaze position (corresponding area, non-corresponding area) and task (image

generation, image inspection, and old/new recognition) as within-subject factors.

Participants spent 902 ms (SD = 767) in the corresponding area vs. 807 ms (SD = 718) in the

non-corresponding area in the image generation task; they spent 1405 ms (SD = 1056) in the

corresponding area vs. 1000 ms (SD = 588) in the non-corresponding area in the image

inspection task; and 512 ms (SD = 442) in the corresponding area vs. 510 ms (SD = 371) in the

non-corresponding area in the old/new recognition task. The analyses revealed a main effect

of task, $F(2, 46) = 30.647$, $p < .001$, partial η^2 = .571. However, no significant interaction

between task and gaze position, $F(2, 46) = 2.130$, $p = .130$, partial η^2 = .085, and no main

effect of gaze position, $F(1, 23) = .761$, $p = .392$, partial η^2 = .032, was found.

Discussion

In this experiment, participants were presented with objects from four categories.

Objects from a particular category always appeared in the quadrant assigned to that

category. During visualization of the objects, participants spent more time in the

corresponding area, for both old and new items. Participants were neither asked to encode

the location of the objects, nor were they informed about the four categories and their

spatial information. Nonetheless, the eye gaze position indicated that they not only encoded

the specific spatial information of the objects along with other visual and semantic

properties, but that the encoded spatial information generalized to novel objects from the

four categories. The present research extends findings on eye gaze position during visual

memory by providing information about the representation of objects and categories.

Specifically, participants spent more time in the same area with remembered items

(e.g., a Vespa scooter) and, interestingly, also when imagining novel items from the same

category (e.g., a bicycle). Hence, location memory transferred to other objects from the

same category. To explain this transfer, we assume that each object's category is automatically activated when the object is memorized (e.g., Hintzman, 1986; Jamieson, Crump, & Hannah, 2012). Thus, memorizing an object in a given location strengthened the connections between the object and its position as well as with the object's category. Since different objects that belong to the same category were consistently presented at the same position, the present findings show that participants imagined new objects being in the same location as previous objects that belonged to the same category. We also suggest that the new object activates a similar category that can be found in episodic memory, which in turn activates a position that is congruent with past viewing experience.

Given that the new items were not previously associated to oculomotor or spatial information during encoding, we conclude that fixations to locations during encoding are not required to cause systematic eye movements with the blank screen paradigm. This finding lends support to the theory that eye movements during retrieval can be launched by spatial representations associated with a semantic category (e.g., Richardson & Spivey, 2000). One possibility is that not only object location but also category location is encoded along with visual and semantic information and thus will trigger eye movements to the relevant areas during both visual memory and mental imagery of objects belonging to the same categories. Eye movements represent the spatial extent that real and possible images embody. Spatial information as revealed by eye movements can play an active role in the representation of categories. Alternatively, category is automatically accessed after an item (whether previously encoded or novel) has been presented and the locations associated with the same or neighboring category are prioritized when a position is assigned to the visual image.

Interestingly, participants were slower in generating new images, but faster in inspecting new images, as compared to old images. This pattern illustrates that it is more

demanding for participants to create a new mental image than to retrieve the mental image of a previously inspected item. However, when it comes to image inspection, it is easier to inspect a mental image that has been created by the individual than an image that has been learned from an external template. Alternatively, it is possible that generating the mental image of a new item is more related to a prototypical or abstract version of that item (e.g., a four-legged animal), which is less concrete or detailed than an encoded exemplar from a specific category. The possible absence of specific details in the new images could be an alternative explanation for faster responses with this class of items.

Furthermore, we were unable to find the corresponding area effect for error trials, thus replicating the previous findings of Martarelli and Mast (2011) with visual material and Scholz et al. (2016) with verbal material. This result supports the suggestion by Ferreira, Apel, and Henderson (2008) and Richardson, Altmann, Spivey, and Hoover (2009) that many aspects of an event, including spatial information, are activated while retrieving semantic information related to that event. The absence of the corresponding area effect for error trials supports research showing the functional role of eye movements during mental imagery (e.g., Johansson & Johansson, 2014; Laeng et al., 2014). However, the best way to understand the sort of location that is being encoded remains the manipulation of eye position. One caveat is that the present results are of correlational nature.

Interestingly, similar eye movements for objects belonging to the same category suggest that there are tight links between spatial and conceptual representations. This result is consistent with perceptual-motor theories of cognitive representation (e.g., Barsalou, 2008). Lakoff and Johnson (1980) proposed that there is a metaphorical mapping for the concept of "category," which is represented by the image of a container. Boot and Pecher (2011) found that the understanding of the concept "category" is indeed grounded in the

concrete representation of the image of a container. They presented pictures of animals and vehicles outside or inside a frame and the participants were to decide whether two images presented on the screen were either animals (or vehicles, respectively) or not. The authors found faster responses when items that belonged to the same category were both presented in a frame. In our paradigm, there was no frame surrounding the items (except for the computer screen), but we think that our task activated, at least to some extent, the container image schema. Indeed, the position of the items was highly predictable (same position for each category). Thus, the gaze to a specific location structured the relationship between item and category.

Future research will need to be carried out in order to better investigate the role of prediction in memory performance. For example, we kept the typicality of the items constant, but it would be interesting to consider the degree of typicality (distance from the prototype, Rosch & Mervis, 1975). Another point that needs further investigation is perceptual similarity in order to disentangle the potential influence of perceptual similarity and the category the stimuli belong to. Future research should also vary the position of the objects of the same category (in the encoding phase), so that participants are unable to predict the location of an object belonging to a given category and thus not use this information to organize their knowledge (in the recall phase).

In conclusion, the results of this study show that eye gaze can be used strategically to organize knowledge. The eye gaze effects observed with the blank screen paradigm strongly suggest that conceptual knowledge is grounded in sensorimotor experience.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

Altmann, G. T. (2004). Language-mediated eye movements in the absence of a visual world: The

'blank screen paradigm'. *Cognition, 93*(2), B79-87.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617-645.

Boot, I., & Pecher, D. (2011). Representation of categories. *Experimental Psychology, 58*(2), 162-170.

Bochynska, A. & Laeng, B. (2015). Tracking down the path of memory: Eye scanpaths facilitate

retrieval of visuospatial information. *Cognitive Processing, 16*(1), 159-163.

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*(1), 27-38.

Ferreira, F., Apel, J., & Henderson, J. M. (2008). Taking a new look at looking at nothing. *Trends in

Cognitive Sciences, 12*(11), 405-410.

Foerster, R. M., Carbone, E., Koesling, H., & Schneider, W. X. (2012). Saccadic eye movements in the

dark while performing an automatized sequential high-speed sensorimotor task. *Journal of

Vision, 12*(2), 1-15.

Foulsham, T., & Kingstone, A. (2013). Fixation-dependent memory for natural scenes: An

experimental test of scanpath theory. *Journal of Experimental Psychology: General, 142*(1),

41-56.

Fourtassi, M., Hajjioui, A., Urquizar, C., Rossetti, Y., Rode, G., & Pisella, L. (2013). Iterative

fragmentation of cognitive maps in a visual imagery task. *PLoS ONE, 8*(7), e68560.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological

Review, 93*(4), 411–428.

Hoover, M. A., & Richardson, D. C. (2008). When facts go down the rabbit hole: Contrasting features

and objecthood as indexes to memory. *Cognition, 108*(2), 533-542.

Jamieson, R. K., Crump, M. J. C., & Hannah, S. D. (2012). An instance theory of associative learning.

Learning and Behavior, 40(1), 61–82.

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*(5), 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*(6), 1053-1079.

Johansson, R., & Johansson, M. (2014). Look here, eye movements play a functional role in memory retrieval. *Psychological Science, 25*(1), 236-242.

Laeng, B., Bloem, I. M., D'Ascenzo, S., & Tommasi, L. (2014). Scrutinizing visual images: The role of gaze in mental imagery and memory. *Cognition, 131*(2), 263-283.

Laeng, B., & Teodorescu, D.-S. (2002). Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cognitive Science, 26*(2), 207-231.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: Chicago University Press.

Martarelli, C. S., & Mast, F. W. (2011). Preschool children's eye-movements during pictorial recall. *British Journal of Developmental Psychology, 29*, 425-436.

Martarelli, C. S., & Mast, F. W. (2013). Eye movements during long-term pictorial recall. *Psychological Research, 77*(3), 303-309.

Maxfield, J. T., Stalder, W. D., & Zelinsky, G. J. (2014). Effects of target typicality on categorical search. *Journal of Vision, 14*(12), 1-11.

Noton, D., & Stark, L. W. (1971). Scanpaths in saccadic eye movements while viewing and recognizing pattern. *Vision Research, 11*(9), 929-942.

Richardson, D. C., Altmann, G. T. M., Spivey, M. J., & Hoover, M. A. (2009). Much ado about eye

movements to nothing: A response to Ferreira et al.: Taking a new look at looking at nothing. *Trends in Cognitive Sciences, 13*(6), 235-236.

Richardson, D. C., & Spivey, M. J. (2000). Representation, space and Hollywood Squares: Looking at things that aren't there anymore. *Cognition, 76*(3), 269-295.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in internal structure of categories. *Cognitive Psychology, 7*(4), 573-605.

Scholz, A., Mehlhorn, K., & Krems, J. F. (2016). Listen up, eye movements play a role in verbal memory retrieval. *Psychological Research, 80*(1), 149-58.

Spivey, M. J., & Geng, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research, 65*(4), 235-241.

Thompson, W. L., Kosslyn, S. M., Sukel, K. E., & Alpert, N. M. (2001). Mental imagery of high- and low-resolution gratings activates area 17. *NeuroImage, 14*(2), 454-464.

Wantz, A. L., Martarelli, C. S., Cazzoli, D., Kalla, R., Muri, R., & Mast, F. W. (2016). Disrupting frontal eye-field activity impairs memory recall. *Neuroreport, 27*(6), 374-378.

Wantz, A. L., Martarelli, C. S., & Mast, F. W. (2015). When looking back to nothing goes back to nothing. *Cognitive Processing, 17*(1), 105-114.

Zelinsky, G. J., Peng, Y., & Samaras, D. (2013). Eye can read your mind: Decoding gaze fixations to reveal categorical search targets. *Journal of Vision, 13*(14), 1-13.

Appendix A: Category typicality rating

We conducted a mixed ANOVA with recognition (old, new) and category (mammals, birds, machines, and home furniture) as within-subject factors, experiment version (A, B) as a between-subjects factor, and typicality ratings as the dependent variable. The means are reported in Table 1. The results revealed a significant three-way interaction, $F(3, 66) = 3.38$, $p = .023$, partial $\eta^2 = .133$. Bonferroni-corrected post hoc tests showed significant differences in Version A of the experiment in the birds category (Old: 4.21, $SD = .79$, New: 3.83, $SD = .62$), $p = .021$, and in the home furniture category (Old: 3.82, $SD = .56$, New: 4.18, $SD = .42$), $p = .008$. The two-way interaction between category and recognition also yielded a significant result, $F(3, 66) = 3.79$, $p = .014$, partial $\eta^2 = .147$. Bonferroni-corrected post hoc tests illustrated a significant difference in the mammals category (Old: 4.32, $SD = .55$, New: 4.21, $SD = .63$), $p = .040$. The results also revealed a significant effect of category, $F(3, 66) = 13.54$, $p < .001$, partial $\eta^2 = .381$. Bonferroni-corrected post hoc tests illustrated significant differences between the birds (4.04, $SD = .56$) and the machines (3.61, $SD = .62$), $p = .003$, the home furniture (4.12, $SD = .44$) and the machines (3.61, $SD = .62$), $p = .002$, and the mammals (4.26, $SD = .58$) and the machines (3.61, $SD = .62$), $p < .001$. Machines were judged as less category-typical than mammals, birds, and home furniture. The analysis further revealed no main effect of recognition (old vs. new), no main effect of experiment version (A, B), and no significant interaction between category and condition ($p > .328$).

Table 1

Participants' mean typicality ratings (and standard deviations) for old vs. new items in the four categories (mammals, birds, home furniture, and machines) by experiment version (A, B).

Category	Recognition	Version A of experiment	Version B of experiment
Mammals	Old	4.25 (*SD* = .56)	4.40 (*SD* = .54)
	New	4.12 (*SD* = .67)	4.30 (*SD* = .61)
Birds	Old	**4.21 (*SD* = .79)**	4.08 (*SD* = .54)
	New	**3.83 (*SD* = .62)**	4.05 (*SD* = .54)
Home furniture	Old	**3.82 (*SD* = .56)**	4.25 (*SD* = .43)
	New	**4.18 (*SD* = .42)**	4.25 (*SD* = .47)
Machines	Old	3.42 (*SD* = .68)	3.80 (*SD* = .59)
	New	3.53 (*SD* = .73)	3.69 (*SD* = .53)

Note. The areas in bold indicate significant differences within categories.

Appendix B: The corresponding area effect in percentages

Percentage of time spent in the corresponding AOIs (where the stimuli where displayed previously) during the image generation, image inspection, and old/new recognition tasks separated for old items (correct trials according to the specific question and to the old/new recognition task) and new items (correct trials according to the old/new recognition task). One-sample t tests were computed to compare the percentages of time with a chance level of 25% of the total time (four areas).

Dwell time in the corresponding AOIs (%)	Old items	New items
	36% (*SD* = 14%)	33% (*SD* = 10%)
Image generation task	*t*(23) = 3.934, *p* = .001	*t*(23) = 3.515, *p* = .002
	d = 1.641	*d* = 1.466
	40% (*SD* = 23%)	36% (*SD* = 16%)
Image inspection task	*t*(23) = 3.264, *p* = .003	*t*(23) = 3.240, *p* = .004
	d = 1.361	*d* = 1.351
	38% (*SD* = 19%)	33% (*SD* = 15%)
Old/new recognition task	*t*(23) = 3.358, *p* = .003	*t*(23) = 2.545, *p* = .018
	d = 1.400	*d* = 1.061

Appendix C: The corresponding area effect compared to the three non-corresponding AOIs

In order to exclude a possible bias caused by pooling the non-corresponding AOIs, we also computed analyses with unpooled data. As did Richardson and Spivey (2000), we "clock coded" the data. The corresponding AOI was labeled 0. The other three areas, moving clockwise, were labeled 1 to 3 (Tables 2, 3).

Table 2

Mean dwell time (in ms) during mental imagery of old items

Mean dwell time in the areas	Image generation task	Image inspection task	Old-new task
	M (SD)	M (SD)	M (SD)
Corresponding AOI	1005 (573)	1708 (900)	568 (359)
AOI 01	544 (267)	912 (686)	311 (233)
AOI 02	594 (308)	832 (458)	333 (213)
AOI 03	669 (415)	946 (619)	353 (198)

Note. The analysis of eye gaze position during mental imagery of old items showed that, in all three tasks (image generation, image inspection, old/new recognition), participants spent more time in the corresponding area (where they had previously seen the object) than in the non-corresponding AOIs (AOI 01, AOI 02, AOI 03). A repeated measures ANOVA revealed a significant interaction between gaze position (corresponding area, non-corresponding area) and task (image generation, image inspection, old/new recognition), $F(3.44, 79.08) = 4.567$, $p = .004$, partial $\eta^2 = .166$, a significant main effect of gaze position, $F(1.96, 45.05) = 10.689$, $p < .001$, partial $\eta^2 = .317$, and a significant main effect of task, $F(2, 46) = 89.126$, $p < .001$, partial $\eta^2 = .795$. Only correct trials (according to both the specific question and the old/new task) were considered.

Table 3

Mean dwell time (in ms) during mental imagery of new items

Mean dwell time in the areas	Image generation	Image inspection	Old-New task
	M (SD)	M (SD)	M (SD)
Corresponding AOI	1084 (496)	1519 (632)	495 (269)
AOI 01	785 (506)	920 (469)	408 (265)
AOI 02	781 (409)	1009 (476)	363 (265)
AOI 03	758 (427)	903 (377)	336 (226)

Note. The analysis of eye gaze position during mental imagery of new items showed that, in all three tasks (image generation, image inspection, old/new recognition), participants spent more time in the corresponding area (where objects from the same category had appeared previously) than in the non-corresponding AOIs (AOI 01, AOI 02, AOI 03). A repeated measures ANOVA revealed a significant interaction between gaze position (corresponding area, non-corresponding area) and task (image generation, image inspection, old/new recognition), $F(3.12, 71.87) = 4.773$, $p = .004$, partial $\eta^2 = .172$, a significant main effect of gaze position, $F(1.98, 45.45) = 10.010$, $p < .001$, partial $\eta^2 = .303$, and a significant main effect of task, 53.541, $p < .001$, partial $\eta^2 = .700$. Only correct trials (according to the old/new recognition task, i.e., new items correctly identified as new) were considered.