



Eye movements during visual imagery and perception show spatial correspondence but have unique temporal signatures

Lilla M. Gurtner^{a,*}, Matthias Hartmann^{a,b}, Fred W. Mast^a

^a Department of Psychology, University of Bern, Fabrikstrasse 8, 3012 Bern, Switzerland

^b Faculty of Psychology, UniDistance Suisse, Überlandstrasse 12, 3900 Brig, Switzerland

ARTICLE INFO

Keywords:

Visual imagery
Eye movement
Looking at nothing
Recurrence quantification analysis
Bayesian hierarchical regression
Spatial correspondence

ABSTRACT

Eye fixation patterns during mental imagery are similar to those during perception of the same picture, suggesting that oculomotor mechanisms play a role in mental imagery (i.e., the “looking at nothing” effect). Previous research has focused on the spatial similarities of eye movements during perception and mental imagery. The primary aim of this study was to assess whether the spatial similarity translates to the temporal domain. We used recurrence quantification analysis (RQA) to assess the temporal structure of eye fixations in visual perception and mental imagery and we compared the temporal as well as the spatial characteristics in mental imagery with perception by means of Bayesian hierarchical regression models. We further investigated how person and picture-specific characteristics contribute to eye movement behavior in mental imagery. Working memory capacity and mental imagery abilities were assessed to either predict gaze dynamics in visual imagery or to moderate a possible correspondence between spatial or temporal gaze dynamics in perception and mental imagery. We were able to show the spatial similarity of fixations between visual perception and imagery and we provide first evidence for its moderation by working memory capacity. Interestingly, the temporal gaze dynamics in mental imagery were unrelated to those in perception and their variance between participants was not explained by variance in visuo-spatial working memory capacity or vividness of mental images. The semantic content of the imagined pictures was the only meaningful predictor of temporal gaze dynamics. The spatial correspondence reflects shared spatial structure of mental images and perceived pictures, while the unique temporal gaze behavior could be driven by generation, maintenance and protection processes specific to visual imagery. The unique temporal gaze dynamics offer a window to new insights into the genuine process of mental imagery independent of its similarity to perception

1. Introduction

When people imagine a visual scene, their eyes move even though there is nothing to look at. Intriguingly, there is a spatial correspondence between fixation positions during perception and visual imagery of the same content (Brandt & Stark, 1997; Johansson, Holsanova, & Holmqvist, 2006; Laeng & Teodorescu, 2002; Spivey & Geng, 2001; Wynn, Olsen, Binns, Buchsbaum, & Ryan, 2018), suggesting that visual imagery has a pictorial format of representation (Albers, Kok, Toni, Dijkerman, & de Lange, 2013; Brogaard & Gatzia, 2017; Kosslyn, 1994; Pearson & Kosslyn, 2015; Senden, Emmerling, van Hoof, Frost, & Goebel, 2019). The spatial correspondence between fixation positions during perception and visual imagery (“looking at nothing” effect) is empirically well established, but its origin and role are still controversial

(Johansson, Holsanova, Dewhurst, & Holmqvist, 2012; Johansson & Johansson, 2014; Richardson & Spivey, 2000; Wynn et al., 2018). For example, the scan path theory assumes that an inspected picture is internally represented as a sequence of sensory and motor activities, and that eye movements during encoding are integrated into the underlying memory representation (Neisser, 1967; Noton & Stark, 1971a). Consequently, recalling a visual scene would reenact the eye movement sequence of encoding, leading to spatially corresponding fixation locations (Laeng & Teodorescu, 2002; Spivey & Geng, 2001; but see Foulsham & Kingstone, 2012; Johansson et al., 2012). Furthermore, there is consensus that eye movements on a blank screen can originate from attention shifts associated with the generation, maintenance or inspection of mental images (Johansson et al., 2012; Kosslyn, Thompson, & Ganis, 2006; Mast & Kosslyn, 2002). Interestingly, most studies that

* Corresponding author.

E-mail address: lilla.gurtner@psy.unibe.ch (L.M. Gurtner).

<https://doi.org/10.1016/j.cognition.2021.104597>

Received 3 March 2020; Received in revised form 7 January 2021; Accepted 8 January 2021

0010-0277/© 2021 The Author(s).

Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

investigated the role of eye fixations in visual imagery have focused on the spatial correspondence, for example by comparing the distribution of fixation locations across different areas of interest (AOIs) on the screen during encoding and visual imagery (or recall) (Johansson & Johansson, 2014; Laeng, Bloem, D'Ascenzo, & Tommasi, 2014; Laeng & Teodorescu, 2002; Martarelli & Mast, 2013; Richardson & Spivey, 2000; Spivey & Geng, 2001). While this type of analysis emphasizes the spatial similarity of eye movements during mental imagery and visual perception, it ignores the intrinsic temporal structure of these eye movements.

The temporal structure of eye fixations carries important information about the mental imagery process and its relationship to perception (Bone et al., 2019; Lanata et al., 2020). Firstly, if the temporal characteristics of eye fixations in mental imagery resemble those in perception, this would support the idea that oculomotor traces are reactivated during visual imagery, as implied by a strict interpretation of the scan path theory (Neisser, 1967; Noton & Stark, 1971b). Secondly, the temporal structure of eye fixations indicates how the pictorial representation is generated and maintained. For example, returning to a previously inspected area during mental imagery could indicate maintenance of its content (Ferreira, Apel, & Henderson, 2008; Foerster, 2018; Laeng & Teodorescu, 2002; Mast & Kosslyn, 2002; Scholz, Klichowicz, & Krems, 2018; Wynn et al., 2016a). The temporal characteristics of refixations could provide information about how the single parts of a mental image are (re)activated in order to generate and maintain a holistic representation of the imagined content (e.g., how often are parts of the image reactivated? How are characteristics of the temporal dynamics of eye fixations related to later memory accuracy of the imagined content?). Thirdly, the analysis of temporal dynamics can help to specify the relationship between perception and mental imagery. While the spatial overlap in eye fixations ("looking at nothing" effect) suggests a spatial resemblance of mental imagery and visual perception, the temporal structure of eye movements can shed light on processes that are distinct for mental imagery (Gurtner, Bischof, & Mast, 2019). For example, if eye fixations during mental imagery were directed to the same locations visited during perception but scan path sequences would be repeated more often or in a different order, we would observe a spatial correspondence without a temporal correspondence. A difference in the temporal characteristics of eye fixations between perception and imagery may be related to the processes that are unique to mental imagery. For example, in mental imagery there is no external information in the visual periphery to which one can shift one's gaze for closer inspection. What is more, maintenance of a mental image might be particularly susceptible to interference from concurrent visual input (Pearson, Clifford, & Tong, 2008; Teng & Kravitz, 2019), and even from unrelated, intermittent fixations (Bochynska & Laeng, 2015). Avoiding such interference could result in fixations being less spread out in mental imagery (Brandt & Stark, 1997; Johansson et al., 2006). Taken together, studying the temporal gaze dynamics of eye fixations in mental imagery can enrich our understanding of the mechanisms that underlie mental imagery by extending the current focus on rather static spatial similarities towards a more multifaceted view.

Previous studies that considered temporal aspects of fixations during mental imagery remain inconclusive. On the one hand, not only over-all fixation distributions are spatially similar in perception and mental imagery. Correspondence between scan paths in perception and visual imagery has been found for simple checker-board stimuli (Brandt & Stark, 1997) and to some extent also for naturalistic pictures (Humphrey & Underwood, 2008; Johansson et al., 2006). On the other hand, scan paths during visual imagery can differ from those during perception (Gurtner et al., 2019; Johansson et al., 2012), especially in terms of their temporal structure (Gurtner et al., 2019). Particularly, gaze returns more often and sooner to previously inspected areas and the scan path sequences are more repetitive during visual imagery when compared to visual perception. Finally, previous studies that considered the order of fixations relied on rigid borders between the AOIs on the screen (Brandt & Stark, 1997; Humphrey & Underwood, 2008), potentially leading to

misclassifications because fixations on the same objects fall into different AOIs (Anderson, Anderson, Kingstone, & Bischof, 2015). To sum up, there are inconsistencies in the literature as well as methodological limitations of the analysis of AOIs. Therefore, it remains unclear whether and to what extent the temporal characteristics of eye fixations during visual imagery are influenced by eye fixations during perception, or conversely, to what extent they are triggered by mechanisms unique to visual imagery.

The primary aim of this study was to investigate the origin of the temporal characteristics of eye fixations during visual imagery. The temporal structure of eye fixations can be assessed by means of recurrence quantification analysis (RQA). RQA has originally been used to describe complex, nonlinear systems that evolve over time (Webber Jr & Zbilut, 2005) and was successfully applied to the analysis of eye movements (Anderson, Bischof, Laidlaw, Risko, & Kingstone, 2013; Farnand, Vaidyanathan, & Pelz, 2016; Vaidyanathan, Pelz, Alm, Shi, & Haake, 2014). RQA computes recurrent fixations, or refixations of previously inspected areas, and further characterizes the timing and systematicity of these refixations: 1) The center-of-recurrence-mass (CORM) value describes whether in general refixations are temporally close to the initial fixation of an area or whether there are many intermediate fixations between refixations (see also Anderson et al., 2015). This timing of refixations has been associated with viewing strategies in perception, where temporally close refixations of an area can indicate poor attention or encoding of the content displayed in this area (Meghanathan, Nikolaev, & van Leeuwen, 2019; Wu, Anderson, Bischof, & Kingstone, 2014). Unlike perceptual input, mental images decay over time (De Beni, Pazzaglia, & Gardini, 2007; Farah, 1989; Kosslyn, 1994). Temporally close refixations (i.e. lower CORM values) could serve to counteract the decay of mental images or they might indicate avoidance of interference by reducing intermittent fixations. In fact, we found evidence in our previous study (Gurtner et al., 2019) that CORM is lower in mental imagery and we expect similar results with a new sample of participants in this study. 2) On a more local scale, the determinism value describes the percentage of recurrent fixations that are part of the re-enactment of a previous fixation sequence. It reflects how stereotypical a sequence of refixations is (Anderson et al., 2013) or how ordered the refixations in a sequence are. It has been related to mnemonic rehearsal strategies in free viewing (Meghanathan et al., 2019). Please note that determinism indicates the repetitiveness of fixation sequences within each trial and does not reflect the similarity in the order of fixations between trials.

Recurrence, determinism and CORM are based on fixation locations. Nevertheless, RQA analyzes temporal patterns because the percentage of refixations does not contain any information about where the refixations were allocated. Given this independence of absolute fixation locations, RQA describes the temporal organization of eye fixations in a different way than measures expressing the similarity between scan paths (Johansson et al., 2006; Johansson et al., 2012; Laeng & Teodorescu, 2002). Scan path similarity measures compare two scan paths and they can only be meaningfully applied when comparing gaze behavior in identical or very similar stimuli. Moreover, the similarity value resulting from the comparison is difficult to relate to the different sources of variance such as the stimulus, the task, or the observer (Anderson et al., 2013). Conversely, RQA parameters are not restricted to comparisons between similar stimuli but are a descriptive measures of the overall temporal exploration style, which is driven by the stimulus (Wu et al., 2014) as well as by characteristics of the observers (Chapman, Heath, Westwood, & Roy, 2001; Gandomkar, Tay, Brennan, & Mello-Thoms, 2017). We have recently applied RQA to eye movement data during visual imagery and have shown that recurrence values differ between categories of mental images (natural scenes, artwork, faces). This shows that the RQA parameters are related to the content of mental images (Gurtner et al., 2019). In the present study, we want to test whether temporal gaze dynamics in mental imagery are determined by those in perception, as predicted by scan path theory, or whether the temporal

gaze dynamics reflect processes unique to mental imagery such as maintaining a mental image or avoiding interference. Furthermore, we want to thoroughly characterize the mechanisms underlying the temporal gaze dynamics in mental imagery.

A further aim of this study was to investigate the role of individual differences in eye movements during visual imagery. We observed remarkable individual differences in RQA-parameters of eye movements during visual imagery in our previous study (Gurtner et al., 2019). In the present study, we tested whether these differences are linked to individual differences in working memory capacity, vividness of mental imagery and whether they are related to memory for the imagined content. Working memory is conceptually related to mental imagery (Kosslyn, Thompson, Kim, Rauch, & Alpert, 1996; Tong, 2013) and it has been shown to correlate with sensory strength of mental images (Keogh & Pearson, 2011, 2014). Moreover, eye fixations on a blank screen reflect the temporal information of items in working memory (Rinaldi, Brugger, Bockisch, Bertolini, & Girelli, 2015). Interestingly, working memory could exert its influence on the looking at nothing effect in two ways: people with higher working memory capacity could have a stronger spatial structure of their mental image, leading to more similar fixation locations (Kosslyn et al., 2006). Conversely, if (corresponding) eye fixations during visual imagery play a functional role in memory retrieval (Johansson et al., 2012; Johansson & Johansson, 2014), people with poorer visuospatial working memory capacities could use eye fixations in mental imagery as spatial indexes (Ballard, Hayhoe, Pook, & Rao, 1997; Wynn et al., 2018), and visual exploration behavior during mental imagery could resemble the exploration behavior during perception. Thus, it remains open how the looking at nothing effect could be influenced by working memory. Working memory could furthermore influence temporal gaze dynamics. If eye movement sequences are part of the memory trace, as proposed by scan path theory (Noton & Stark, 1971b), then these sequences are expected to be longer for participants with higher working memory capacity. Longer eye movement sequences lead to later refixations as indicated by higher CORM values. Additionally, working memory capacity can influence determinism. It is possible that participants with a higher working memory capacity use mnemonic strategies in order to maintain more items in visual working memory. This would be reflected in increased determinism values.

In addition to working memory, the experienced vividness could be reflected in gaze behavior. Vividness is modulated by the overlap of brain activation between mental imagery and perception (Dijkstra et al., 2017; Lee, Kravitz, & Baker, 2012; St-Laurent, Abdi, & Buchsbaum, 2015) and the strength of top-down influences in the occipital lobe (Dijkstra, Zeidman, et al., 2017). Furthermore, similar brain activation patterns in imagery and perception have been correlated with fixation reinstatements (Bone et al., 2019), suggesting a possible link between gaze behavior and experienced vividness (Bone et al., 2019). Such a link could go beyond the spatial structure of eye movements. The temporal orderliness of fixations, measured by determinism, could indicate that the different parts of a mental image are reactivated in a systematic manner, making mental images more robust. In turn, more robust mental images could be experienced more vividly. In addition to this, robust mental images could retain different parts of the mental image more reliably which might counteract forgetting and lead to improved picture memory. Memory for imagined content has already been linked to the spatial dimension of eye movements in mental imagery: the degree of spatial correspondence improved detection of location changes (Olsen, Chiew, Buchsbaum, & Ryan, 2014) and was associated with better memory for the the presence (Kinjo, Fookan, & Spering, 2020) and the size of a stimulus (Laeng et al., 2014), but no relationship was found between fixation reinstatement and the detection of small modifications in pictures (Bone et al., 2019). The exact relationship between eye movements during visual imagery and visuo-spatial working memory capacities, experienced vividness and memory for imagined content needs to be investigated in further studies. In these studies, it is essential

to focus on both the spatial and temporal characteristics of eye movements.

To test whether there is a temporal correspondence between gaze dynamics in mental imagery and perception and to explore the role of interindividual differences in temporal gaze dynamics, we measured participants' eye movements on a blank screen while they imagined previously presented pictures. We first analyzed the spatial correspondence of fixations in mental imagery and perception, and more importantly, whether a temporal correspondence follows from the spatial correspondence. To this end, we modeled the spatial and temporal signature of eye fixations during visual imagery by the respective signatures (spatial and temporal) in perception using Bayesian generalized hierarchical regression models. The spatial similarity of eye movements in perception and visual imagery was quantified by comparing the percentage of fixations in the same AOI in the two conditions, along the lines of the analysis by Johansson and Johansson (2014). Temporal gaze dynamics were quantified by means of three parameters from RQA that describe specific aspects of the temporal course of fixations (as described above): percentage of refixations, the CORM value to capture the global timing of refixations and the determinism value to capture the repetitiveness of participants' scan paths (Anderson et al., 2013; Gurtner et al., 2019). After investigating the spatial and temporal correspondence, we investigated the origin of individual differences in eye movements during visual imagery. Individual differences in visual working memory capacities were assessed using an n-back task, and the vividness of visual imagery was assessed by means of VVIQ-questionnaire (VVIQ, Marks, 1973) and by asking participants how well they had been able to imagine the stimulus. Moreover, we also assessed the memory performance for specific contents of the images by means of an old-new recognition task and by asking questions about details in the pictures. These two additional tasks capture the ability to remember the content of the pictures, and thus serve to measure the amount of detailed information contained in mental images (Bochynska & Laeng, 2015). Individual differences in working memory capacity and visual imagery abilities could be promising candidates to either predict the temporal gaze dynamics in visual imagery or to moderate a possible correspondence between spatial or temporal gaze dynamics in perception and mental imagery. Therefore, we included them as predictors or, respectively, as moderator variables into the models.

2. Materials and methods

2.1. Participants

Thirty-two participants (mean age = 23.8, SD = 6.6, 8 male) took part in the study (the data of 8 participants was lost due to technical problems and five of these participants were replaced. Thus, data of 29 participants were analyzed). The study was approved by the ethical committee of the Human Sciences Faculty of the University of Bern. All participants gave their informed consent prior to participation.

2.2. Apparatus

We used an EyeLink 1000 (SR Research, Canada) with a sampling rate of 500 Hz. Eye fixations were parsed using the default parameters of the manufacturer, which operationalize fixations as follows: For each sample, velocity and acceleration are computed and compared against a threshold ($30^\circ/\text{sec}$ and $8000^\circ/\text{sec}^2$ respectively). Samples are labeled as part of a fixation if neither the velocity nor the acceleration are above the threshold and if the pupil size is not very small, distorted or missing in the camera image. Thus, conceptually, the software defines fixations as the absence of a saccade or blink (Hessels, Niehorster, Nyström, Andersson, & Hooge, 2018). Fixation data were exported using SR Research Data Viewer for further processing in R (R Core Team, 2015) with RStudio (RStudio: Integrated Development Environment for R, 2016) and in Matlab (MATLAB, 2015). Stimuli were presented on a

1902 × 1080 pixel screen (26"). Participants rested their head on a chinrest at a distance of 75 cm from the screen. Participants' picture memory was measured on a notebook using PsychoPy software (Peirce et al., 2019).

2.3. Materials

We used a set of 30 pictures (10 face, 10 landscape, and 10 art pictures). For the face pictures, 5 male and 5 female portraits (smiling version) were chosen from the data set of the European Conference on Visual Perception ([dataset] 2D face sets, 2008). Landscape and art pictures were retrieved from the internet. In landscape pictures, the upper half of the picture was a uniform blue sky with no clouds or other objects. The art pictures depicted scenes containing several human beings located in different parts of the picture, see Fig. 1 for example stimuli. To assess picture memory, we used an old-new recognition task and then asked questions about details in the pictures. For the old-new recognition task, the 30 pictures presented during the experiment were mixed with 30 new pictures (10 for each category with the same characteristics) such that the chance probability of a correct answer was 50%. Detailed questions served to assess picture memory for each of the 30 pictures seen in the imagery task. In each picture, we covered the part the respective question was referring to (e.g., "Was there a chair in the covered part?"). For half of the questions, the correct answer was "true". We used the VVIQ to assess the vividness of visual imagery (Marks, 1973), and a visuo-spatial and auditory n-back task from the BrainTwister2 collection of cognitive training tasks (Studer-Luethi, Kodzhabashev, Hogrefe, & Perrig, 2015) to assess participants' working memory capacity. The auditory n-back task was used as an additional measurement for the specificity of working memory for gaze dynamics: If both working memory tasks were equally related to gaze dynamics, then gaze dynamics might be determined by domain-general working memory capacity rather than by specific visuo-spatial working memory capacities.

2.4. Procedure

The experiment consisted of an eye tracking part and an assessment part. At the beginning of the experiment, participants read a cover story about the relationship between pupil dilation and cognitive load, telling them that the purpose of eye tracking was to measure the pupil size. The aim of this cover story was to divert participants' attention away from eye movements (Hartmann, Mast, & Fischer, 2016; Martarelli, Mast, & Hartmann, 2017). In the subsequent eye tracking part, participants were instructed to explore a series of pictures, and to vividly imagine after each picture what they just saw while keeping their eyes open and directed at the blank screen. After successful 9-point-calibration procedure of the eye tracker, the task started. At the beginning of each trial, a drift correction point was presented at the center of the screen. Participants then explored the picture for 15 s, followed by a 1 s fixation

cross and a visual imagery phase in which a blank gray screen was presented for another 15 s. Such a long perception and imagery time is needed in order to obtain enough fixations from participants for the RQA (Anderson et al., 2015). At the end of each trial, participants indicated in an imaginability rating how well they had been able to imagine the stimulus on a 7-point Likert scale. The 30 pictures were presented in random order. The eye-tracking part lasted for about 30 min. In the subsequent assessment part, we measured participants' memory for the pictures in the eye tracking part, their working memory capacity and the vividness of their visual imagery. In order to assess the picture memory, participants first completed an unannounced old-new recognition task and were then also asked about details of the images. In the old-new recognition task, 30 old and 30 new pictures (see section 2.3) were presented after a fixation cross. A pre-test revealed that participants were faster and better for face pictures when compared to art and landscape pictures. To ensure comparable task difficulty across picture categories, landscape and art pictures were presented for 300 ms while face pictures were presented for 180 ms. The performance of the participants in the experiment proper are reported in the Supplementary Material (Table 1) and the results indicate comparable task difficulty. Participants had to indicate via button press whether they had seen the picture before (old) or not (new). Next, detailed questions were presented on the screen (for example: "Was there a chair in the covered part?"). Participants pressed the space bar to continue and one of the pictures of the eye tracking part appeared on the screen (with the relevant area covered). Participants responded by pushing one of two keys (yes or no) on the laptop keyboard. Working memory capacity was then assessed by means of a visuo-spatial and auditory n-back task (see section 2.3). We balanced the order in which the visual and auditory version were assessed. Both tasks started with the 2-back condition where participants were instructed to press a key if the location of a square in a 3 × 3 grid (visuo-spatial task) or the identity of an auditorily presented letter (auditory task) matched the location (or identity) of two presentations (n) before. One block consisted of 20 + n trials and lasted about one minute. If the participant was correct in 90% of the trials in a block, the n in the next block was increased by one, if the performance was below 70%, the n was reduced by one, otherwise the n remained constant. Participants completed five such blocks. We chose the n of the last block weighted by the participant's performance in that block (e.g., 90% correctness) as measure for visuo-spatial and auditory working memory performance. We divided this weighted performance by 100 to ensure that the beta coefficient of the n-back performance (range 150–475) was not too small when predicting percentages of fixations in an AOI or recurrent fixations (range: 0–1). We also ran our models with several possible implementations such as the weighted n of the last correctly performed block, the performance-weighted average n of the 4th and 5th block, and lastly, the n of a potential 6th block. Although participants did not perform the 6th block, their n for this block can be derived from the n and performance in the 5th block. All implementations lead to the same pattern of results. Finally, participants completed



Fig. 1. Example stimuli used in the experiment. Due to personality rights, the face in the original portrait was replaced by a similar, publicly available portrait.

the VVIQ (Marks, 1973). Before debriefing, we asked participants about the purpose of the experiment. None of the participants guessed the true purpose of the experiment.

2.5. Data analysis

The results section is comprised of three parts: 1) basic eye movement analysis, 2) the correspondence of the eye movements between perception and imagery, both in terms of their spatial fixation locations and their temporal gaze dynamics, 3) the influence of person-level predictors on gaze dynamics during visual imagery. Data are made available in a OSF repository (see Supplementary Material).

2.5.1. Basic eye movement analysis

2.5.1.1. Correspondence of eye movements between visual imagery and perception (temporal and spatial). We examined whether the spatial correspondence of fixation locations between visual imagery and perception generalized to the temporal domain, that is, whether temporal gaze dynamics in perception predicted those in visual imagery. The spatial allocation of fixations are influenced by the picture category (for example, in all landscape pictures, fixations will mostly be directed at the lower half of the picture) and so are the temporal gaze dynamics (Anderson et al., 2013; Gurtner et al., 2019; Wu et al., 2014). Thus, in order to show how far the specific gaze behavior in perception in each trial determined gaze behavior in visual imagery, we controlled for the picture category by including it as a predictor in the correspondence analyses described below.

The analysis of spatial correspondence followed the procedure employed by Johansson and Johansson (2014). We split the screen into four equally sized areas of interest (AOIs). Then, we predicted the percentage of fixations spent in each AOI during visual imagery by the percentage of fixations spent in the same AOI during perception by a zero-one-inflated beta regression. We included the interaction between picture category and AOI-number into the model in order to control for the picture category. The interaction predicted the percentage of fixations in a specific AOI, given the picture category of that particular trial. If the percentage of fixations in each AOI during perception was still predictive for the percentage in the AOI during visual imagery, this would indicate a picture-specific correspondence in fixation locations between perception and imagery in addition to the overall effect of the picture category on fixation locations.

The analysis of the temporal gaze dynamics included the percentage of recurrent fixations, CORM and determinism. These parameters in visual imagery were predicted by their counterpart in perception in zero-one-inflated beta regressions. Here, we not only included the picture category as a predictor, but we also controlled for the spread of fixations since the closer fixations are arranged, the higher their recurrence values become.

2.5.1.2. Influence of person-level predictors. Person-level characteristics such as imagery vividness, working memory and picture memory can potentially predict gaze dynamics and spatial correspondence in visual imagery. The spatial location of fixations in visual imagery can be predicted by those in perception (i.e. the “looking at nothing” effect, Laeng & Teodorescu, 2002), and we analyzed whether the spatial correspondence was moderated by the person-level predictors. Two-way interactions between each of the person-level predictors with the percentage in the respective AOI in perception were computed. A significant interaction coefficient would indicate that the degree of spatial correspondence of fixations depends, for example, on the expression of the imagery vividness of a person.

In the analysis of the temporal gaze dynamics, we focused on the random intercepts for participants and stimuli, which both make an individual contribution to the prediction of the temporal gaze dynamics

in each trial. The standard deviation of the random intercepts shows the amount of variance between the contributions of all participants or all pictures respectively. The standard deviation of random intercepts is small when all participants exhibit similar gaze dynamics or, respectively, when all pictures lead to similar gaze dynamics. Conversely, the standard deviation of the random intercepts is large when gaze dynamics vary between individuals or when different pictures lead to different gaze dynamics. Thus, the standard deviations of the random intercepts for participants and for pictures make it possible to distinguish between variance between participants and stimuli. Similar to the logic of model comparisons (where predictors are iteratively included to see whether they increment the predictive power of the model), we compare an intercept-only model to models that contain a meaningful predictor. Any predictor that captures a common structure in the variance between participants or between pictures is expected to reduce the “random” variance between participants or pictures by explaining some of it in a meaningful way. For example, we expect the semantic picture category (landscapes, art pictures and faces) to account for the variance between gaze dynamics of the individual pictures. But such a model will most likely not reduce the standard deviation of the random intercepts between participants. The same logic can be applied to person-level predictors: any variable that captures a common structure in the variance between participants is expected to reduce the variance between the random intercepts of individual participants but not the variance between the random intercepts of pictures. We tested whether the standard deviation of random intercepts for participants and pictures from an intercept-only model changes when the semantic picture category is included as a fixed effect (semantic model, Fig. 5A) and when person-level predictors (i.e., visual working memory, VVIQ, individual imaginability ratings and picture memory, see also section 2.6) are further included as fixed effects (individual model, Fig. 5B).

2.6. Implementation details

Data analysis was performed in R (R Core Team, 2015) with RStudio (RStudio: Integrated Development Environment for R, 2016). Recurrence values were calculated with the functions provided by (Anderson et al., 2013) in Matlab (MATLAB, 2015). The threshold distance for two fixations to be considered recurrent was set to approximately 2.5 degrees of visual angle (64 pixels). In all models for recurrence parameters, we controlled for the spread of fixations since near fixations are associated with higher recurrence values. Bayesian hierarchical generalized linear regression models were used (implemented with the brms package, Bürkner (2018)). All model specifications can be found in the section 2 in the Supplementary Material. The output of a Bayesian model does not provide a point estimate of a regression weight (like a frequentist model would). Rather, the model provides posterior distributions of the estimated regression weights. If zero is part of the 95% density interval of such a posterior distribution, the effect is typically absent, whereas an effect is typically present if the 95% density interval of the posterior distribution does not include zero. Thus, for the estimated parameters, we provide the upper and lower bound for the 95% posterior density interval (credibility interval, 95% CI). In all our models, we used 4 Markov-Chains, each using 6000 iterations to sample from the posterior distributions. Where necessary, tree depth was set to 11 and delta was set to 0.99 to ensure that the chains converged and that sampling was unbiased. The link function for all models was a zero-one-inflated-beta function to handle trials with 100% or 0% recurrence or correspondence instead of excluding such trials from data analysis. Random intercepts for participants and picture were included in all models and we used dummy coding such that the picture category “art” was the reference category. For all models, goodness of fit was assessed by posterior-predictive checks: Based on the statistical model with the posterior parameter estimates (that is, the parameter estimates that best fit the collected data), artificial data was generated. The artificial data was compared to the real data to assess the model fit. We visually

verified that all models fit the data. Furthermore, we tested the model fit by approximating a leave-one-out-cross-validation (LOO function in brms). The results are reported in Table 2 of the Supplementary Material. We also evaluated the pareto-k diagnostics. All pareto-k estimates were below 0.5 indicating that no model had severe misspecification issues or overly influential data points (outliers) that distorted the fit. Finally, we ensured that the models of the person-level effects did not suffer from multicollinearity, that is, we assured that no effect was masked by the presence of other predictors. Multicollinearity arises when two or more predictors are correlated. Correlations between the person-level predictors can be found in Fig. 2. The two picture memory tasks were correlated, suggesting that the two tasks both measure memory for the pictures' content. Furthermore, the imaginability rating was correlated with performance in the old-new-recognition task. In order to control whether these correlations lead to multicollinearity, we also fit the models in a separate analyses, entering one predictor at a time instead of all predictors at once. We entered performance in the old-new-recognition task first in one analysis and, likewise, we entered the imaginability rating first in another. The estimated beta weights were similar in both cases, suggesting that multicollinearity is not a problem in the data.

3. Results

3.1. Basic eye movement analysis

All fixations outside the screen and all fixations longer than 5 s were excluded from analysis (0.04% of fixations). To ensure data quality and compliance of participants, we first confirmed previous findings on eye fixations in mental imagery.

Participants made longer fixations in visual imagery than in perception (M = 595 ms, SD = 406 and M = 277 ms, SD = 152 respectively) (estimated increase of duration in imagery by a lognormal hierarchical regression model was 273 ms, 95% CI [183, 422], which is in accordance with Brandt and Stark (1997) and Recarte and Nunes (2000). In imagery trials, participants made fewer fixations (M = 21.9, SD = 9.27) than in perception (M = 42.5, SD = 9.98) (estimated increase

in number of fixations in the perception condition in a Poisson hierarchical model was 20.70 (95% CI [11.40, 29.80]). The perception and imagery phases were both 15 s long and if fixations are longer, fewer fixations can be made in total.

Fixations in visual imagery were also less spread, as previously reported in Gurtner et al. (2019), Brandt and Stark (1997) and Johansson et al. (2006). The median distance of all fixations from the center of fixations was smaller in visual imagery than in perception (Median = 128 px, SD = 99.40; Median = 242 px, SD = 107, respectively). The estimated increase in fixation spread for perception in a lognormal hierarchical regression was 95.70 px (95% CI [32.90, 193]).

Finally, we confirmed the results of our previous study (Gurtner et al., 2019). Picture category and experimental phase (perception vs. imagery) influenced gaze dynamics. Recurrence and determinism were higher in visual imagery. This means that in mental imagery, participants made more refixations and the refixations were part of more repetitive patterns. CORM values were lower in visual imagery, showing that refixations occur earlier in time in mental imagery compared to perception. Recurrence, CORM and determinism were highest for face pictures and similar for art and landscape pictures (see Table 3 in the Supplementary Material). Our results show that gaze dynamics differ between visual imagery and perception and they are influenced by the semantic content of the mental image. This demonstrates that participants complied with task instructions, and it rules out that gaze dynamics during visual imagery are a product of random eye movements.

3.2. Correspondence of eye movements between visual imagery and perception (temporal and spatial)

First, we tested whether the spatial correspondence is present in our data. We found a positive relationship between the percentage of fixations spent in an AOI in perception and visual imagery (beta = 1.68, 95% CI [1.46, 1.92], zero-one-inflated beta regression), demonstrating the spatial correspondence (see Fig. 3A). This correspondence could be a general effect of the picture category: in perception and visual imagery, the picture category could influence fixation positions in a similar way (for example, most fixations in a landscape picture will be directed at the lower half) without a specific effect of each individual picture. However, we included the picture category as a predictor, and the fact that the fixation locations in perception still predicted those in imagery shows that the spatial correspondence goes beyond what could be expected due to the effect of the picture category alone. Rather, the spatial correspondence reflects the specific content of the individual pictures. This extends previous studies of the “looking at nothing” effect which usually do not control for the confounding effect of picture category (Johansson et al., 2006).

Second, temporal gaze dynamics (recurrence, CORM and determinism) in visual imagery were modeled as a function of those in perception in zero-one-inflated beta regressions, each controlled for the spread of fixations as well as for the picture category. The gaze dynamics in perception did not predict those in visual imagery for recurrence, CORM and determinism (exact beta-weights for the regressions for recurrence, CORM and determinism can be found in Table 4 in the Supplementary Material). Thus, while the spatial location of fixations in visual imagery is influenced by fixation locations in perception, no such relationship can be found in the temporal structure of the eye movements (see Fig. 3B, C and D). This raises the question whether RQA measures in perception can predict RQA measures in future perception. With a separate sample of participants, we ensured that RQA measures

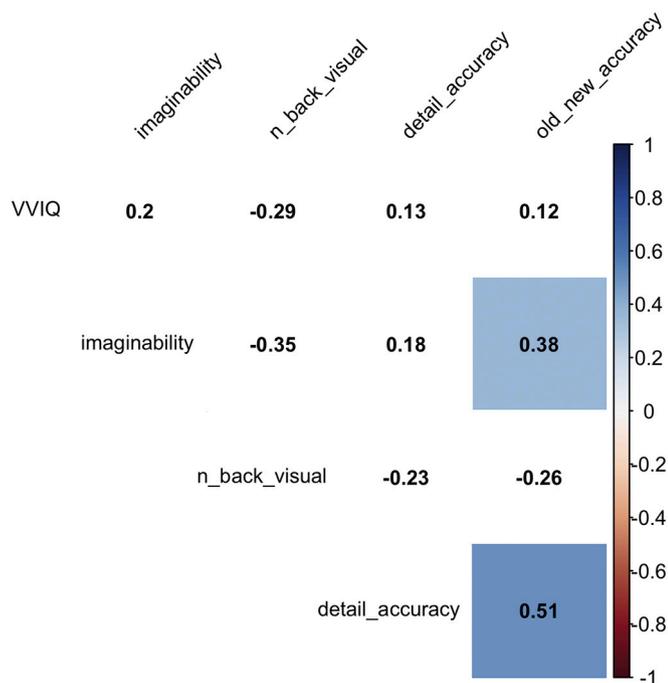


Fig. 2. Correlation matrix of person-level predictors of the individual model. Correlations in colored squares are significant with $p < 0.05$ (not adjusted for multiple comparisons).

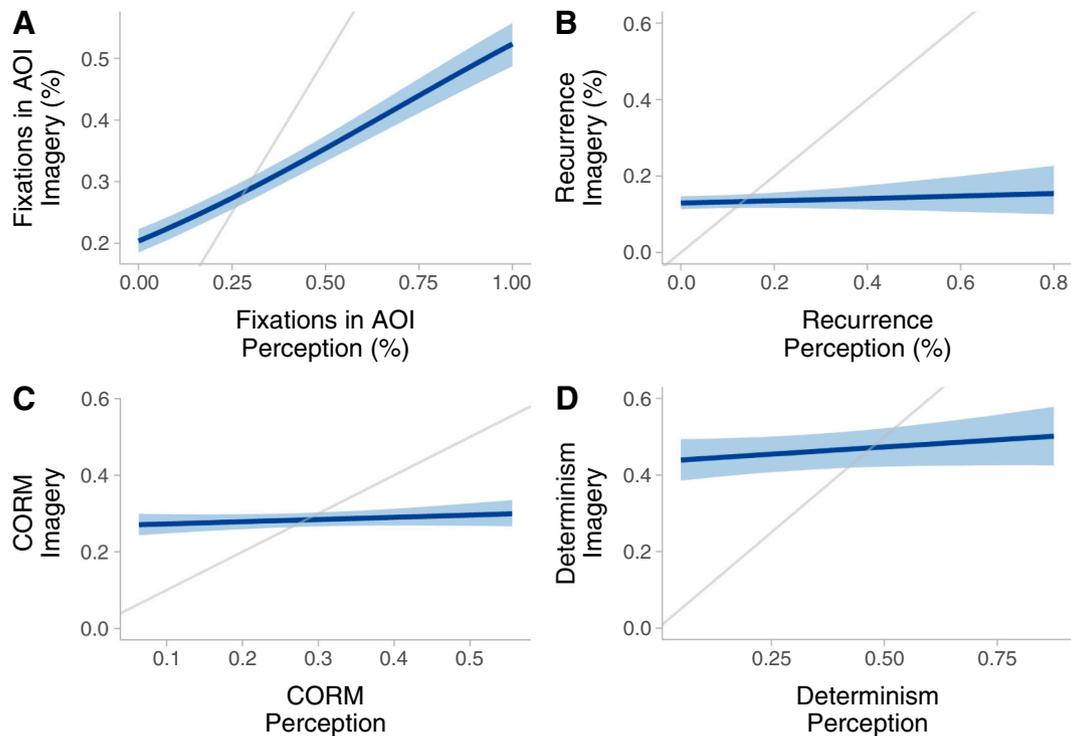


Fig. 3. Effects of eye fixations in perception on eye fixations in visual imagery, controlled for the effect of the picture category. A: The “looking at nothing” effect, showing a strong dependence of fixation locations in visual imagery on the fixation locations in perception. B, C and D represent the RQA gaze dynamics. B: Percent of recurrent fixations in perception does not predict the percent of recurrent fixations in visual imagery. C: Center-of-recurrence-mass (CORM) values and D: Determinism values no clear effects either. Lines represent the median of the predicted values with the 95% uncertainty interval of the prediction around them. The grey line is a unity line showing hypothetical perfect correspondence.

in perception are predicted by preceding RQA measures.¹ We therefore conclude that temporal gaze dynamics in visual imagery are different from those made in perception, and yet, they are specific to the content of the picture that is imagined.

3.3. Influence of person-level predictors on gaze behavior during visual imagery

We tested whether the person-level predictors (working memory capacity, imaginability of the stimuli and the picture memory) moderated the spatial correspondence and whether they predicted the temporal gaze dynamics in mental imagery. Descriptive statistics of these predictors can be found in the Supplementary Material (Table 5). We assessed differences in vividness of mental imagery in two ways: by means of the VVIQ questionnaire and by asking how well participants were able to imagine the picture after each trial, where participants answered on a 7-point Likert-Scale. Both measures yielded sufficient variance between participants.

For the moderation of the spatial similarity, we predicted the percentage of fixation in each AOI in mental imagery by 2-way interactions between the person-level predictors and the respective percentage in perception. These two-way interactions indicate that the predictive power of fixation locations in perception depends on the manifestation of the person-level variable. Only the performance in the visual n-back task interacted with the percentage of fixations spent in the respective AOI. Thus, the “looking at nothing” effect is moderated by the visual

working memory capacity in that the better a participant’s visual working memory, the stronger the relationship between fixation locations in perception and visual imagery (beta-coefficient of the interaction = 0.3, 95% CI [0.044, 0.56], see Fig. 4). The evidence ratio of the moderation effect being larger than zero was 82.92. The same moderation pattern was also found by using three different ways to assess the performance in the n-back task (see section 2.4). Model comparisons (reported in Table 2 in the Supplementary Material) confirmed that the individual model (with the person-level predictors) was the most adequate model to describe the percentage of fixations in an AOI in mental imagery.

For the analysis of the temporal gaze dynamics, the “semantic model” was compared with an intercept-only model. We show that including the picture category as predictor substantially reduces the variance between individual pictures in all temporal gaze dynamics. This result was to be expected based on the results in the basic eye movement analysis part and serves as a proof of concept for this novel analysis approach. Fig. 5A shows the change in the standard deviation of random intercepts for participants and pictures when the picture category is included as predictor. The standard deviation of the random intercepts between pictures (blue bars) is substantially reduced for all three temporal gaze dynamic measures (no blue bar includes zero in its 95% density interval). As expected, the standard deviation between participants does not change (all green bars include zero) because the picture category does not explain systematic differences between individual participants. In a next step, we included the person-level predictors into the models of the temporal gaze dynamics. We extended the semantic model by VVIQ, visual working memory performance and the average accuracy in the picture memory tasks and average imaginability ratings. None of the person-level predictors was able to predict any of the temporal gaze dynamics. Thus, neither differences in VVIQ and imaginability nor differences in picture memory and visual working

¹ A separate sample of 10 participants looked at the pictures two times for 15 s each. Recurrence in the first perception phase predicted recurrence in the second phase. Thus, the patterns in temporal correspondence are reliable when looking at the same picture for a second time. See section 6 in the Supplementary Material for details.

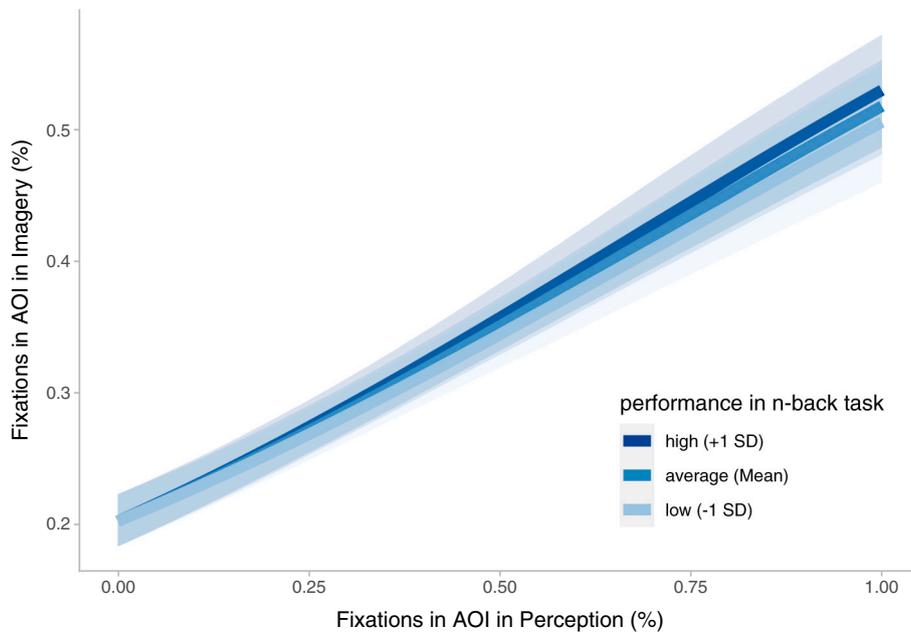


Fig. 4. Moderation of the “looking at nothing” effect by working memory capacity: The higher working memory capacity as measured by the performance in the n-back task, the steeper the slope of the fitted regression line, that is, the stronger the correspondence between fixation locations in visual imagery and perception.

Change in the Estimated Standard Deviation of Random Intercepts

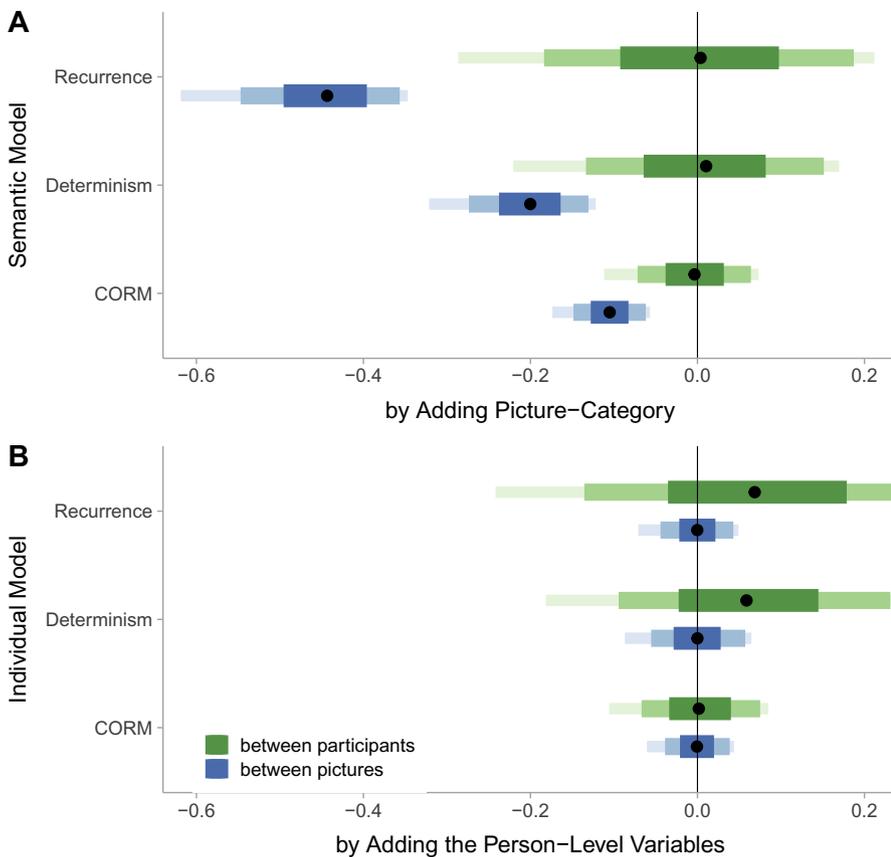


Fig. 5. A: Change in posterior distributions of standard deviation of random intercepts for the semantic model, where the picture category is included into the model, compared to an intercept-only model. The inclusion of the picture category reduces the standard deviation of random intercepts between pictures in all RQA values (blue bars do not include zero). B: Change in estimated standard deviation of random intercepts when visual working memory, VVIQ, imaginability of the stimulus and picture memory are further added as predictors to the semantic model. The standard deviation of random intercepts between participants is not reduced by the person-level predictors (green bars all include zero). This means none of the person-level predictors is able to reduce unexplained variance between persons in temporal gaze dynamics. The segments of the bars represent 90%, 80% and 50% of the posterior distributions and the black point represents the median. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

memory capacity translated to differences in the temporal gaze dynamics between individuals, see Fig. 5B (all blue posterior intervals include zero). Model comparisons (Table 2 in the Supplementary Material) confirm that adding the person-level predictors does not increase

model fit, which further supports the absence of effects. For all recurrence parameters, the semantic model is most adequate to explain the data.

4. Discussion

The primary aim of this study was to assess to what extent the spatial and temporal properties of eye movements during mental imagery were determined by the visual exploration behavior in perception. Furthermore, we wanted to explore how person-level variables contribute to eye movement behavior in visual imagery. We demonstrate that the spatial distribution of fixations during visual imagery relates to the fixation distribution in perception (Johansson et al., 2012; Johansson & Johansson, 2014; Laeng & Teodorescu, 2002; Richardson & Spivey, 2000). Participants tended to fixate areas during visual imagery that they also visited during perception. This “looking at nothing” effect was moderated by visual working memory performance. Individuals with larger visual working memory capacity had a more pronounced “looking at nothing” effect. To our knowledge, this is the first demonstration of this kind of moderation effect. Higher spatial working memory capacity could help to maintain more information about the parts of the mental image, such as the spatial configuration of the perceived picture and the absolute location of each part on the screen (Hebb, 1968; Laeng et al., 2014). This, in turn, could provide the mental image with a fine-grained spatial structure similar to that of the original picture, leading to a higher spatial correspondence of fixation locations (Olsen et al., 2014). This interpretation supports the notion that mental imagery and working memory are conceptually related processes (Kosslyn et al., 2006; Kosslyn & Thompson, 2003; Pearson, 2019; Tong, 2013) and it could explain the results of Keogh and Pearson (2011) and Keogh and Pearson (2014), who found that working memory correlates with the sensory strength of mental imagery. Improved sensory strength of the mental image in their study could originate from the more fine-grained spatial structure of the mental image in participants with large working memory capacity. Contrary to the findings of Bochynska and Laeng (2015), Johansson et al. (2012) or Johansson and Johansson (2014), who found that reenacting the original scan path is beneficial to picture memory compared to following a predefined new scan path, in our experiment, spatially similar fixation distribution did not lead to a better picture memory. The fixation distributions of our participants were similar but not identical (as in some previous studies, where participants were forced to follow their exact encoding scan path, e.g. Bochynska and Laeng (2015)) and it is possible that the degree of spatial similarities shown spontaneously by our participants was not high enough to effect picture memory.

In addition to the spatial correspondence, we investigated whether the temporal structure of eye movements in visual imagery was related to their temporal structure in perception. Unlike the spatial layout of fixations, the temporal gaze dynamics in visual imagery were not determined by the gaze dynamics in perception (see Fig. 3). This novel finding suggests that the temporal gaze dynamics reflect aspects of the ongoing processes of visual imagery that do not have a counterpart in gaze behavior in perception. We show that refixations are made more often in visual imagery, they are temporally closer and more stereotypical (Table 3, Supplementary Material). The RQA parameters further depend on the picture category which shows that the temporal structure of eye movements is not random. Specifically, the lower CORM in art and landscape pictures could result from the fact that such pictures have more features. In general, mental images have fewer features than actual visual pictures (Xie, Kaiser, & Cichy, 2020). The more features a picture contains, the more challenging it could be to maintain the corresponding mental image. This challenge might be met by avoiding intermittent fixations, that could interfere with maintaining a fragile mental image (leading to lower CORM values) (Bochynska & Laeng, 2015). These systematic effects confirm that the analysis of eye movements on a blank screen during a visual imagery task is informative about the underlying mental imagery process (Bone et al., 2019; Hebb, 1968). The temporal gaze dynamics thus reflect unique, picture-specific aspects of the visual imagery process. The person-level predictors, however, had no influence on temporal gaze dynamics: Gaze dynamics of participants were not

explained by their visual working memory capacity or VVIQ score and temporal gaze dynamics did not systematically influence experienced vividness of a mental image or picture memory (see Fig. 5). Apparently, vividness of a mental image, a meta-cognitive evaluation, can result from different processes. Making no eye movements in imagery (i.e. maximal avoidance of intermittent fixations, as indicated by lower CORM values) can lead to a vivid mental image just as much as making many eye movements (i.e. reactivating different parts repeatedly). Such a distinction between mental imagery processes and their meta-cognitive evaluation has also been found on the level of the brain, where vividness judgments have been associated with the volume of the prefrontal cortex whereas the sensory strength of mental images was associated with the volume of the primary visual cortex (Bergmann, Genç, Kohler, Singer, & Pearson, 2016). Similarly, different strategies, associated with different gaze behavior, could lead to good picture memory. We measured general picture memory (with an old-new recognition test) and the memory for details (by asking whether something had been present or not in a certain area). However, it is possible that there is a link between memory for *spatial details* and temporal gaze dynamics in mental imagery. The spatial structure of a picture is reflected in the eye movements made during mental imagery (“looking at nothing” effect) and the more pronounced this effect, the better the memory for stimulus location (Olsen et al., 2014) and size (Laeng et al., 2014). It is possible that participants following the strategy of avoiding intermittent fixations (showing low CORM values) have reduced spatial picture memory while participants who use their gaze to actively regenerate and maintain the location of a particular item (showing higher CORM values) improve memory for spatial details. While both might have good general picture memory, the different strategies might lead to differences in specific memory tasks. Consistent with the spatial correspondence we found, refixations in mental imagery might help to sustain the general spatial structure of the perceived picture rather than the recreation of details and features. This proposed function of refixations in mental imagery is different from their function in perception where refixations lead to increased processing of visual details. Taken together, temporal gaze dynamics in visual mental imagery are unique, picture-specific and independent of the person-level predictors we assessed in this study.

How can we interpret the temporal gaze dynamics during visual imagery? In perception, recurrence parameters are related to spatial attention allocation and the quality of the encoding of visual information (Meghanathan et al., 2019; Wu et al., 2014). They are thus determined by the way external visual information is acquired. Evidently, this cannot be their function in mental imagery, since no relevant sensory information can be acquired from a blank screen. The recurrence values in mental imagery are unlike those in perception and we propose an interpretation of recurrence parameters (and gaze behavior in general) in visual imagery that reflects the differences between mental imagery and perception.

One aspect that sets imagery apart from perception is that in mental imagery, internal pictorial representations need to be actively reconstructed from memory (Hassabis & Maguire, 2009). Indeed, eye movements during visual imagery have been interpreted as reflecting this process of reconstruction (Laeng & Teodorescu, 2002; Mast & Kosslyn, 2002; Wynn et al., 2016b). Reconstruction of a mental image in the visual buffer (Kosslyn et al., 2006) would thus be accompanied by repeated refixations to its different parts (Gurtner et al., 2019) and repeating entire fixation sequences could be a mnemonic strategy to optimally handle different parts of a mental image at the same time. Scan path theory (Neisser, 1967; Noton & Stark, 1971b) suggests that, as we reconstruct a mental image, the eye movements made during perception are reactivated and replayed. The differences in the temporal structure of eye movements in visual imagery that we found in this study, along with the results from Johansson et al. (2012), are not in line with scan path theory, which would predict spatial *and* temporal correspondence (at least in its strict interpretation). Rather, the findings

from this study are more consistent with the idea that the same central processing mechanisms (e.g. attention distribution) that are responsible for eye movements in perception can be applied to mental imagery as well (Ballard et al., 1997; Brandt & Stark, 1997; Hebb, 1968). Specifically, our findings suggest that the mechanisms determining visual exploration during perception are also employed in visual mental imagery, leading to spatial correspondence, but with a different temporal signature: Higher recurrence in visual imagery suggests that the focus is shifted towards re-inspecting already recreated parts rather than creating new parts (as in perceptual exploration), leading to a spatial but not to a temporal correspondence. Thus, the unique pattern of gaze dynamics observed in mental imagery could reflect the ongoing process of generating and maintaining a mental image.

The fact that mental images need to be (re)created is not the only aspect that sets mental imagery apart from perception. Mental representations are fragile (De Beni et al., 2007; Farah, 1989; Kosslyn, 1994) and easily disturbed by concurrent perceptual input. For example, mental image strength diminishes when participants are exposed to high luminance during mental imagery (Pearson et al., 2008) and mental representations are biased towards distracting or surprising visual inputs (Teng & Kravitz, 2019). One way of minimizing interference from visual input is to minimize disruptive or unexpected changes in visual input, for example by revisiting previously inspected areas. This would lead to fewer, and less spread fixations as well as to higher recurrence values, just as we could show it in this study. At the same time, some eye movements are necessary as they may aid the generation and maintenance of mental images (Johansson et al., 2006; Johansson, Holsanova, & Holmqvist, 2011; Johansson & Johansson, 2014; Laeng & Teodorescu, 2002; Wynn et al., 2016b). Thus, eye movements in visual mental imagery could balance two constraints: minimizing interference from perceptual input and using eye movements as a strategy to generate and maintain the mental image. This could explain why, unlike the spatial fixation locations, the temporal gaze dynamics in visual imagery are unique and have no correspondence with those in perception. In perception, the temporal gaze dynamics reflect exploration behavior and attention distribution (Heaton & Freeth, 2016; Meghanathan et al., 2019; Wu et al., 2014), while in visual imagery they could reflect the generation and maintenance as well as the protection of the mental representations. The uniqueness of the temporal characteristics of eye fixations during mental imagery does not contradict the similarity in the spatial domain because the temporal domain is informative about *how* a mental image is generated, maintained and how much concurrent visual input is avoided, while the purely spatial domain is informative about the spatial layout of a mental image. In the light of these two constraints, participants might have individual strategies how they weight the need of reactivation of parts of the image versus avoidance of interference. Such an individual “tradeoff between the benefit and cost of making saccades” has recently been proposed (Kinjo et al., 2020, p.87) and it could lead to the large interindividual differences in gaze dynamics. The focus on temporal dynamics is new and may be able to highlight differences in the underlying processes of mental imagery and perception that are otherwise overseen when solely relying on the spatial summary analysis of fixations over time.

To our surprise, the temporal gaze dynamics were independent of the person-level predictors that we assessed. It is possible that our sample size was too small to detect an effect on the person level. At the same time, small samples can lead to spurious correlations. Therefore, both the absence of the effects of person-level predictors on temporal gaze dynamics as well as the moderation effect of the looking at nothing effect by working memory should be replicated with a larger sample. We are confident that the null-findings reported here can be replicated. The model diagnostics showed no signs of overly influential data points (all pareto-k diagnostics were above 0.5) which could mask a possible effect. Furthermore, there were no signs of any trends of person level effects for the temporal dynamics of recurrence and determinism and only small trends of in inconsistent directions in the case of CORM where lower

values were associated with worse performance in the memory task for picture details and with better performance in the old-new-recognition task, see Table 6 in the Supplementary Material.

We cannot exclude that the lack of a relationship is due to the specific choice of measures for the person-level predictors. To date, the evidence for the relationship between VVIQ and the temporal dynamics of mental imagery processes is mixed. In two recent papers, Dijkstra, Mostert, Lange, Bosch, and Gerven (2018) and Shatek and Carlson (2019) found that responses in the VVIQ were not able to disambiguate temporal brain activation dynamics during mental imagery. Our findings are in line with these two recent findings supporting the absence of a relationship between VVIQ and the temporal dynamics of visual imagery. Furthermore, our results suggest no relationship between VVIQ and the degree of spatial correspondence between eye movements in visual imagery and perception. Borst and Kosslyn (2010) report no evidence for a correlation between VVIQ and spatial mental imagery ability either. Thus, a mental image need not be spatially accurate to be experienced as vivid. The Object-Spatial Imagery and Verbal Questionnaire (OSIVQ, Blazhenkova and Kozhevnikov (2009)) could be a promising addition to the VVIQ. The OSIVQ focuses on the object-based vs. spatial processes of mental imagery rather than on the experienced vividness of the mental image (Roldan, 2017) and it has been shown to correlate with the spread of fixations in visual imagery (Johansson et al., 2011).

Furthermore, the absence of an effect of the person-level predictors on recurrence parameters could be due to the fact that the relationship between visual imagery, eye movements and experienced vividness is more complex than initially assumed. As discussed above, vividness could be high because one person reduces disruptive perceptual input by minimizing the occurrence of eye movements, while another person could achieve a vivid mental image by executing eye movements to create a more spatially accurate mental image (Johansson et al., 2012; Johansson & Johansson, 2014; Wynn et al., 2018). Different imagery strategies could lead to different eye movements (Pearson & Keogh, 2019), yet both could lead to high vividness ratings. The absence of an effect of vividness of visual imagery on spatial and temporal gaze dynamics could further be due to our participants who are part of the “healthy” population. It would be interesting to study the gaze dynamics of persons with post-traumatic stress disorder who often suffer from intrusive and highly vivid images of traumatic events. On the opposite end of the range of imagery vividness are participants with aphantasia (Jacobs, Schwarzkopf, & Silvanto, 2018). They lack the ability to voluntarily generate mental images (Keogh & Pearson, 2017; Zeman, 2001) and their gaze dynamics in visual imagery could deviate considerably from those of normal participants.

A promising avenue for future research is to explore the origin of interindividual variance of gaze dynamics during visual imagery. It would be interesting to see whether differences in temporal gaze dynamics translate to individual differences in temporal patterns of brain activations as found by Dijkstra et al. (2018) and Shatek and Carlson (2019). Finding such a relationship could further clarify the nature of the underlying processes reflected in gaze dynamics. The nature of imagery processes could be also tackled by simulation of eye movements in mental imagery. Here, the framework of active inference in predictive coding would be promising. The framework proposes an internal representation, or a generative model of visual input, that is tested against real incoming sensory input by means of eye movements. Implementations of eye movement control in perception on the basis of internal representations have already been successful (Friston, Adams, Perrinet, & Breakspear, 2012; Parr, Mirza, Cagnan, & Friston, 2019). In the case of mental imagery however, no inference can be made because there is no relevant external stimulus. Incorporating mental imagery would thus require an extension of the framework. For example, this could entail a dynamic stimulus such as a rolling marble for which a generative model can be formulated based on physical laws. On the one hand, including mental imagery could be an important contribution to the validation of the framework and, on the other hand, it offers an interesting

opportunity to better understand eye movements during mental imagery. Of particular interest is the role of the prediction error. In perception, the prediction error is crucial for eye movement control since it confirms or rejects prior hypothesis about the real causes of sensory input. In mental imagery, we would argue that the prediction error would have to be strongly attenuated or even be ignored entirely (Mast & Ellis, 2015), so as to not overrule the mental image and switch perception to the actual incoming sensory data. The role of prediction errors in mental imagery could clarify how the brain generates a non-factual representation without losing touch with reality.

Additionally, the cognitive representation style of each individual might relate to gaze dynamics (Roldan, 2017). Eye movements differentiate between learning styles (Cao & Nishihara, 2012). Piccardi et al. (2016) showed that participants with different navigation styles show differences in how many consecutive fixations are spent in an AOI during recall of a route. Similarly, the influence of the cognitive representation style could cause the known differences in eye movements between experts and novices: Vaidyanathan et al. (2014) showed that expert dermatologists exhibit higher recurrence patterns when inspecting melanoma. Experts could represent information differently than novices which could translate to differences in recurrence parameters during mental imagery too. Humans are arguably “face experts” and accordingly, our participants showed higher recurrence values for faces, both in perception and in visual imagery. Lastly, other measures of individual mental imagery abilities need to be included such as mental chronometry, where participants imagine an ongoing process and indicate its start and end time. Indeed, (Lanata et al., 2020) showed that higher imagery ability, as measured by mental chronometry, is associated with lower recurrence and determinism in eye movements during mental imagery.

To conclude, spatial and temporal gaze behavior during visual imagery depend on the semantic content of the mental image. The spatial distribution of eye fixations in visual imagery corresponds to that in perception, and this correspondence is moderated by visuo-spatial working memory capacity. This supports the notion of a spatial structure of mental images and that higher working memory capacity allows for incorporation of this spatial structure. Interestingly, the temporal dynamics of eye fixations during visual imagery cannot be predicted by the temporal dynamics during perception. Based on these findings, we proposed that the temporal dynamics of eye movements in mental imagery are determined by two constraining functions. On the one hand, refixations in mental imagery could support the generation and maintenance of an internal pictorial representation. On the other hand, refixations in mental imagery could be a means to minimize interference from visual input. This dual constraint might explain why the temporal pattern of eye movements during visual imagery is distinct from and unrelated to the (temporal) exploration behavior during perception. The balance between these dual constraints could be at the heart of the individual differences in eye movement behavior in mental imagery. Our results highlight that visual imagery needs to be understood in terms of its own temporal signature in gaze dynamics and not only in terms of its similarity to perception. While similarities to perception are informative about the spatial layout of a mental image, the unique temporal gaze dynamics allow for new insights into the genuine process of mental imagery independent of its similarity to perception.

Declaration of interests

None.

Acknowledgments

We would like to thank Teodora Moser, Katja Näf and Therese Albert for their help in data collection and Boris Mayer for his helpful suggestions on the analysis of the data. Furthermore, we thank two anonymous reviewers and Bruno Laeng for their expert advice on previous

versions of the paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cognition.2021.104597>.

References

- [dataset] 2D face sets. Retrieved from http://pics.stir.ac.uk/2D_face_sets.htm%20%5BData%20set%5D, (2008).
- Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & de Lange, F. P. (2013). Shared representations for working memory and mental imagery in early visual cortex. *Current Biology*, 23(15), 1427–1431. <https://doi.org/10.1016/j.cub.2013.05.065>.
- Anderson, N. C., Anderson, F., Kingstone, A., & Bischof, W. F. (2015). A comparison of scanpath comparison methods. *Behavior Research Methods*, 47(4), 1377–1392. <https://doi.org/10.3758/s13428-014-0550-3>.
- Anderson, N. C., Bischof, W. F., Laidlaw, E. W. K., Risko, F. E., & Kingstone, A. (2013). Recurrence quantification analysis of eye movements. *Behavior Research Methods*, 45, 842–856. <https://doi.org/10.3758/s13428-012-0299-5>.
- 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(4), 723–767. <https://doi.org/10.1017/S0140525X97001611>.
- Bergmann, J., Genç, E., Kohler, A., Singer, W., & Pearson, J. (2016). Smaller primary visual cortex is associated with stronger, but less precise mental imagery. *Cerebral Cortex*, 26(9), 3838–3850. <https://doi.org/10.1093/cercor/bhv186>.
- Blazhenkova, O., & Kozhevnikov, M. (2009). The new object-spatial-verbal cognitive style model: Theory and measurement. *Applied Cognitive Psychology*, 23, 638–663. <https://doi.org/10.1002/acp>.
- Bochynska, A., & Laeng, B. (2015). Tracking down the path of memory: Eye scanpaths facilitate retrieval of visuospatial information. *Cognitive Processing*, 16(1), 159–163. <https://doi.org/10.1007/s10339-015-0690-0>.
- Bone, M. B., St-Laurent, M., Dang, C., McQuiggin, D. A., Ryan, J. D., & Buchsbaum, B. R. (2019). Eye movement reinstatement and neural reactivation during mental imagery. *Cerebral Cortex*, 29(3), 1075–1089. <https://doi.org/10.1093/cercor/bhy014>.
- Borst, G., & Kosslyn, S. M. (2010). Individual differences in spatial mental imagery. *Quarterly Journal of Experimental Psychology*, 63(10), 2031–2050. <https://doi.org/10.1080/17470211003802459>.
- 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. <https://doi.org/10.1162/jocn.1997.9.1.27>.
- Brogard, B., & Gatzia, D. E. (2017). Unconscious imagination and the mental imagery debate. *Frontiers in Psychology*, 8(MAY), 1–14. <https://doi.org/10.3389/fpsyg.2017.00799>.
- Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, 10(1), 395–411.
- Cao, J., & Nishihara, A. (2012). Understand learning style by eye tracking in slide video learning. *Journal of Educational Multimedia and Hypermedia*, 21(4), 335–358. Retrieved from <https://www.learnlib.org/p/39182>.
- Chapman, C. D., Heath, M. D., Westwood, D. A., & Roy, E. A. (2001). Memory for kinesthetically defined target location: Evidence for manual asymmetries. In 46. *Brain and cognition* (pp. 62–66). Academic Press Inc. [https://doi.org/10.1016/S0278-2626\(01\)80035-X](https://doi.org/10.1016/S0278-2626(01)80035-X).
- De Beni, R., Pazzaglia, F., & Gardini, S. (2007). The generation and maintenance of visual mental images: Evidence from image type and aging. *Brain and Cognition*, 63(3), 271–278. <https://doi.org/10.1016/j.bandc.2006.09.004>.
- Dijkstra, N., Bosch, S. E., & van Gerven, M. A. J. (2017). Vividness of visual imagery depends on the neural overlap with perception in visual areas. *Journal of Neuroscience*, 37(5), 1367–1373. <https://doi.org/10.1523/JNEUROSCI.3022-16.2016>.
- Dijkstra, N., Mostert, P., Lange, F. P. D., Bosch, S., & Gerven, M. A. J. V. (2018). Differential temporal dynamics during visual imagery and perception. *eLife*, 7, 1–16.
- Dijkstra, N., Zeidman, P., Ondobaka, S., van Gerven, M. A. J., Gerven, M. A. J. V., & Friston, K. J. (2017). Distinct top-down and bottom-up brain connectivity during visual perception and imagery. *Scientific Reports*, 7(1), 5677. <https://doi.org/10.1038/s41598-017-05888-8>.
- Farah, M. J. (1989). The neural basis of mental imagery. *Trends in Neurosciences*, 12(10), 395–399. [https://doi.org/10.1016/0166-2236\(89\)90079-9](https://doi.org/10.1016/0166-2236(89)90079-9).
- Farnand, S. P., Vaidyanathan, P., & Pelz, J. B. (2016). Recurrence metrics for eye movements in perceptual experiments. *Journal of Eye Movement Research*, 9(4), 1–12. <https://doi.org/10.16910/jemr.9.4.1>.
- Ferreira, F., Apel, J., & Henderson, J. M. (2008). Taking a new look at looking at nothing. *Trends in Cognitive Sciences*, 12(11), 405–410. <https://doi.org/10.1016/j.tics.2008.07.007>.
- Foerster, R. M. (2018). “Looking-at-nothing” during sequential sensorimotor actions: Long-term memory-based eye scanning of remembered target locations. *Vision Research*, 144(June 2017), 29–37. <https://doi.org/10.1016/j.visres.2018.01.005>.
- Foulsham, T., & Kingstone, A. (2012). Modelling the influence of central and peripheral information on saccade biases in gaze-contingent scene viewing. *Visual Cognition*, 20(4–5), 546–579. <https://doi.org/10.1080/13506285.2012.680934>.

- Friston, K. J., Adams, R. A., Perrinet, L., & Breakspear, M. (2012). Perceptions as hypotheses: Saccades as experiments. *Frontiers in Psychology*, 3, 151. <https://doi.org/10.3389/fpsyg.2012.00151>.
- Gandomkar, Z., Tay, K., Brennan, P. C., & Mello-Thoms, C. (2017). A model based on temporal dynamics of fixations for distinguishing expert radiologists' scanpaths. In *10136. Medical Imaging 2017: Image perception, Observer Performance, and Technology Assessment* (p. 1013606). <https://doi.org/10.1117/12.2254527>.
- Gurtner, L. M., Bischof, W. F., & Mast, F. W. (2019). Recurrence quantification analysis of eye movements during mental imagery. *Journal of Vision*, 19, 1–17.
- Hartmann, M., Mast, F. W., & Fischer, M. H. (2016). Counting is a spatial process: Evidence from eye movements. *Psychological Research*, 80(3), 399–409. <https://doi.org/10.1007/s00426-015-0722-5>.
- Hassabis, D., & Maguire, E. A. (2009). The construction system of the brain. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 364(1521), 1263–1271. <https://doi.org/10.1098/rstb.2008.0296>.
- Heaton, T. J., & Freeth, M. (2016). Reduced visual exploration when viewing photographic scenes in individuals with autism spectrum disorder. *Journal of Abnormal Psychology*, 125(3), 1–13. <https://doi.org/10.1037/abn0000145>.
- Hebb, D. O. (1968). Concerning imagery. *Psychological Review*, 75(6), 466–477. <https://doi.org/10.1037/h0026771>.
- Hessels, R. S., Niehorster, D. C., Nyström, M., Andersson, R., & Hooge, I. T. C. (2018). Is the eye-movement field confused about fixations and saccades? A survey among 124 researchers. *Royal Society Open Science*, 5(8). <https://doi.org/10.1098/rsos.180502>.
- Humphrey, K., & Underwood, G. (2008). Fixation sequences in imagery and in recognition during the processing of pictures of real-world scenes. *Journal of Eye Movement Research*, 2(2), 1–15. <https://doi.org/10.16910/jemr.2.2.3>.
- Jacobs, C., Schwarzkopf, D. S., & Silvanto, J. (2018). Visual working memory performance in aphantasia. *Cortex*, 105, 61–73. <https://doi.org/10.1016/j.cortex.2017.10.014>.
- 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. <https://doi.org/10.1037/a0026585>.
- 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. <https://doi.org/10.1207/s15516709cog000086>.
- 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 *Proceedings of the 33rd annual meeting of the cognitive science society* (pp. 1200–1205). Austin, TX: Cognitive Science Society. Retrieved from <https://mindmodeling.org/cogsci2011/papers/0284/paper0284.pdf>.
- Johansson, R., & Johansson, M. (2014). Look here, eye movements play a functional role in memory retrieval. *Psychological Science*, 25(1), 236–242. <https://doi.org/10.1177/0956797613498260>.
- Keogh, R., & Pearson, J. (2011). Mental imagery and visual working memory. *PLoS One*, 6(12). <https://doi.org/10.1371/journal.pone.0029221>.
- Keogh, R., & Pearson, J. (2014). The sensory strength of voluntary visual imagery predicts visual working memory capacity. *Journal of Vision*, 14(12). <https://doi.org/10.1167/14.12.7>.
- Keogh, R., & Pearson, J. (2017). The blind mind : No sensory visual imagery in aphantasia. *Cortex*, 105(2015), 53–60. <https://doi.org/10.1016/j.cortex.2017.10.012>.
- Kinjo, H., Fookan, J., & Spering, M. (2020). Do eye movements enhance visual memory retrieval? *Vision Research*, 176(July), 80–90. <https://doi.org/10.1016/j.visres.2020.07.013>.
- Kosslyn, S. M. (1994). *Image and brain: The resolution of the imagery debate*. Cambridge, MA: MIT Press.
- Kosslyn, S. M., & Thompson, W. L. (2003). When is early visual cortex activated during visual mental imagery? *Psychological Bulletin*, 129(5), 723–746. <https://doi.org/10.1037/0033-2909.129.5.723>.
- Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). *The case for mental imagery*. Oxford University Press.
- 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(1), 78–82. <https://doi.org/10.1162/jocn.1996.8.1.78>.
- 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. <https://doi.org/10.1016/j.cognition.2014.01.003>.
- Laeng, B., & Teodorescu, D.-S. (2002). Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cognitive Science*, 26, 207–231. <https://doi.org/10.1207/s15516709cog2602.3>.
- Lanata, A., Sebastiani, L., Di Gruttola, F., Modica, S. D., Scilingo, E. P., & Greco, A. (2020). Nonlinear analysis of eye-tracking information for motor imagery assessments. *Frontiers in Neuroscience*, 13(January), 1–11. <https://doi.org/10.3389/fnins.2019.01431>.
- Lee, S.-H., Kravitz, D. J., & Baker, C. I. (2012). Disentangling visual imagery and perception of real-word objects. *NeuroImage*, 59, 4064–4073. <https://doi.org/10.1016/j.neuroimage.2011.10.055>.
- Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology*, 64(1), 17–24. <https://doi.org/10.1111/j.2044-8295.1973.tb01322.x>.
- Martarelli, C. S., & Mast, F. W. (2013). Eye movements during long-term pictorial recall. *Psychological Research*, 77(3), 303–309. <https://doi.org/10.1007/s00426-012-0439-7>.
- Martarelli, C. S., Mast, F. W., & Hartmann, M. (2017). Time in the eye of the beholder: Gaze position reveals spatial-temporal associations during encoding and memory retrieval of future and past. *Memory and Cognition*, 45(1), 40–48. <https://doi.org/10.3758/s13421-016-0639-2>.
- Mast, F. W., & Ellis, A. W. (2015). Internal models, vestibular cognition, and mental imagery: Conceptual considerations. *Multisensory Research*, 28(5–6), 443–460. <https://doi.org/10.1163/22134808-00002503>.
- Mast, F. W., & Kosslyn, S. M. (2002). Visual mental images can be ambiguous: Insights from individual differences in spatial transformation abilities. *Cognition*, 86(1), 57–70. [https://doi.org/10.1016/S0010-0277\(02\)00137-3](https://doi.org/10.1016/S0010-0277(02)00137-3).
- MATLAB. (2015). *Version 8.5.0.197613 (R2015a)*. Natick, Massachusetts: The Mathworks, Inc.
- Meghanathan, R. N., Nikolaev, A. R., & van Leeuwen, C. (2019). Refixation patterns reveal memory-encoding strategies in free viewing. *Attention, Perception, & Psychophysics*, 2011. <https://doi.org/10.3758/s13414-019-01735-2>.
- Neisser, U. (1967). *Cognitive psychology* (Appleton-Century-Crofts).
- Noton, D., & Stark, L. W. (1971a). Scanpaths in eye movements during pattern perception. *Science*, 171(3968), 308–311.
- Noton, D., & Stark, L. W. (1971b). Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vision Research*, 11(9). [https://doi.org/10.1016/0042-6989\(71\)90213-6](https://doi.org/10.1016/0042-6989(71)90213-6).
- Olsen, R. K., Chiew, M., Buchsbaum, B. R., & Ryan, J. D. (2014). The relationship between delay period eye movements and visuospatial memory. *Journal of Vision*, 14(1). <https://doi.org/10.1167/14.1.8>.
- Parr, T., Mirza, M. B., Cagnan, H., & Friston, K. J. (2019). Dynamic causal modelling of active vision. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 39(32), 6265–6275. <https://doi.org/10.1523/JNEUROSCI.2459-18.2019>.
- Pearson, J. (2019, October). The human imagination: The cognitive neuroscience of visual mental imagery. *Nature Publishing Group*. <https://doi.org/10.1038/s41583-019-0202-9>.
- Pearson, J., Clifford, C. W., & Tong, F. (2008). The functional impact of mental imagery on conscious perception. *Current Biology*, 18(13), 982–986. <https://doi.org/10.1016/j.cub.2008.05.048>.
- Pearson, J., & Keogh, R. (2019). Redefining visual working memory: A cognitive-strategy, brain-region approach. *Current Directions in Psychological Science*, 28(3), 266–273. <https://doi.org/10.1177/0963721419835210>.
- Pearson, J., & Kosslyn, S. M. (2015). The heterogeneity of mental representation: Ending the imagery debate. In *112. Proceedings of the national academy of sciences* (pp. 10089–10092). <https://doi.org/10.1073/pnas.1504933112>.
- Pearce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51(1), 195–203. <https://doi.org/10.3758/s13428-018-01193-y>.
- Piccardi, L., De Luca, M., Nori, R., Palermo, L., Iachini, F., & Guariglia, C. (2016). Navigational style influences eye movement pattern during exploration and learning of an environmental map. *Frontiers in Behavioral Neuroscience*, 10(Jun), 1–15. <https://doi.org/10.3389/fnbeh.2016.00140>.
- R Core Team. (2015). *R: A language and Environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.r-project.org/>.
- Recarte, M. A., & Nunes, L. M. (2000). Effects of verbal and spatial-imagery tasks on eye fixations while driving. *Journal of Experimental Psychology: Applied*, 6(1), 31–43. <https://doi.org/10.1037/1076-898X.6.1.31>.
- 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. [https://doi.org/10.1016/S0010-0277\(00\)0084-6](https://doi.org/10.1016/S0010-0277(00)0084-6).
- Rinaldi, L., Brugger, P., Bockisch, C. J., Bertolini, G., & Girelli, L. (2015). Keeping an eye on serial order: Ocular movements bind space and time. *Cognition*, 142, 291–298. <https://doi.org/10.1016/j.cognition.2015.05.022>.
- Roldan, S. M. (2017). Object recognition in mental representations: Directions for exploring diagnostic features through visual mental imagery. *Frontiers in Psychology*, 8(May), 1–15. <https://doi.org/10.3389/fpsyg.2017.00833>.
- RStudio: Integrated Development Environment for R. (2016). Boston, MA: RStudio, Inc. Retrieved from <http://www.rstudio.com/>.
- Scholz, A., Klichowicz, A., & Krems, J. F. (2018). Covert shifts of attention can account for the functional role of “eye movements to nothing”. *Memory and Cognition*, 46(2), 230–243. <https://doi.org/10.3758/s13421-017-0760-x>.
- Senden, M., Emmerling, T. C., van Hoof, R., Frost, M. A., & Goebel, R. (2019). Reconstructing imagined letters from early visual cortex reveals tight topographic correspondence between visual mental imagery and perception. *Brain Structure and Function*, 224(3), 1167–1183. <https://doi.org/10.1007/s00429-019-01828-6>.
- Shatek, S. M., & Carlson, T. A. (2019). Decoding images in the mind's eye: The temporal dynamics of visual imagery. *Vision*, 3(53), 1–16. <https://doi.org/10.1101/637603>.
- 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. <https://doi.org/10.1007/s004260100059>.
- St-Laurent, M., Abdi, H., & Buchsbaum, B. R. (2015). Distributed patterns of reactivation predict vividness of recollection. *Journal of Cognitive Neuroscience*, 27(10), 2000–2018. <https://doi.org/10.1162/jocn.2014.00839>.
- Studer-Luetthi, B., Kodzhabashev, S., Hogrefe, A., & Perrig, W. J. (2015). *BrainTwister2 – An extended collection of cognitive training tasks*. Bern: Fachstelle für Lernen und Gedächtnis, Universität Bern.
- Teng, C., & Kravitz, D. J. (2019). Visual working memory directly alters perception. *Nature Human Behaviour*, 3(8), 827–836. <https://doi.org/10.1038/s41562-019-0640-4>.

- Tong, F. (2013). Imagery and visual working memory: One and the same? *Trends in Cognitive Sciences*, 17(10), 489–490. <https://doi.org/10.1016/j.tics.2013.08.005>.
- Vaidyanathan, P., Pelz, J., Alm, C., Shi, P., & Haake, A. (2014). Recurrence quantification analysis reveals eye-movement behavior differences between experts and novices. In *Proceedings of the symposium on eye tracking research and applications* (pp. 303–306). Safety Harbor, Florida: ACM. <https://doi.org/10.1145/2578153.2578207>.
- Webber, C. L., Jr., & Zbilut, J. P. (2005). Recurrence quantification analysis of nonlinear dynamical systems. *Tutorials in Contemporary Nonlinear Methods for the Behavioural Sciences*, 26–94.
- Wu, D. W., Anderson, N. C., Bischof, W. F., & Kingstone, A. (2014). Temporal dynamics of eye movements are related to differences in scene complexity and clutter. *Journal of Vision*, 14(9). <https://doi.org/10.1167/14.9.8>.
- Wynn, J. S., Bone, M. B., Dragan, M. C., Hoffman, K. L., Buchsbaum, B. R., & Ryan, J. D. (2016a). Selective scanpath repetition during memory-guided visual search. *Visual Cognition*, 24(1), 15–37. <https://doi.org/10.1080/13506285.2016.1175531>.
- Wynn, J. S., Bone, M. B., Dragan, M. C., Hoffman, K. L., Buchsbaum, B. R., & Ryan, J. D. (2016b). Selective scanpath repetition during memory-guided visual search. *Visual Cognition*, 24(1), 15–37. <https://doi.org/10.1080/13506285.2016.1175531>.
- Wynn, J. S., Olsen, R. K., Binns, M. A., Buchsbaum, B. R., & Ryan, J. D. (2018). Fixation reinstatement supports visuospatial memory in older adults. *Journal of Experimental Psychology: Human Perception and Performance*, 44(7), 1119–1127. <https://doi.org/10.1037/xhp0000522>.
- Xie, S., Kaiser, D., & Cichy, R. M. (2020). Visual imagery and perception share neural representations in the alpha frequency band. *Current Biology*, 30(13), 2621–2627. e5 <https://doi.org/10.1016/j.cub.2020.04.074>.
- Zeman, A. (2001). Consciousness. *Brain: A Journal of Neurology*, 124(Pt 7), 1263–1289. <https://doi.org/10.1093/Brain/124.7.1263>.