Consciousness and Cognition 100 (2022) 103302

Contents lists available at ScienceDirect

Consciousness and Cognition

journal homepage: www.elsevier.com/locate/concog

Schema-related eye movements support episodic simulation

Jordana S. Wynn[a,*], Ruben D.I. Van Genugten[a], Signy Sheldon[b], Daniel L. Schacter[a]

[a] *Department of Psychology, Harvard University, Cambridge, USA*
[b] *Department of Psychology, McGill University, Montreal, Canada*

ARTICLE INFO

Keywords:
Simulation
Eye movements
Schemas
Memory
Imagination

ABSTRACT

Recent work indicates that eye movements support the retrieval of episodic memories by reactivating the spatiotemporal context in which they were encoded. Although similar mechanisms have been thought to support simulation of future episodes, there is currently no evidence favoring this proposal. In the present study, we investigated the role of eye movements in episodic simulation by comparing the gaze patterns of individual participants imagining future scene and event scenarios to across-participant gaze templates for those same scenarios, reflecting their shared features (i.e., schemas). Our results provide novel evidence that eye movements during episodic simulation in the face of distracting visual noise are (1) schema-specific and (2) predictive of simulation success. Together, these findings suggest that eye movements support episodic simulation via reinstatement of scene and event schemas, and more broadly, that interactions between the memory and oculomotor effector systems may underlie critical cognitive processes including constructive episodic simulation.

1. Introduction

Recall the last time you went to the beach. You might find yourself looking around as you mentally reconstruct the sights, smells, and sounds of that event. In fact, research using eye movement monitoring indicates that successful memory retrieval is accompanied by spontaneous gaze shifts reflecting spatial and temporal mnemonic content (e.g., Bochynska & Laeng, 2015; Foulsham & Kingstone, 2013; Johansson & Johansson, 2013; Olsen, Chiew, Buchsbaum, & Ryan, 2014; for review, see Wynn, Shen, & Ryan, 2019). Even when looking at a blank computer screen in which there is no visual input present to attract gaze (for examples of classic models of physical saliency-based visual attention see Treisman & Gelade, 1980; Itti & Koch, 2000), eye movements during retrieval tend towards regions of the screen associated with previously present and salient encoded information. This pattern of "looking at nothing" has been proposed to support memory retrieval by projecting internal representations of spatial and/or temporal relations onto the external world (Altmann, 2004; Richardson & Spivey, 2000; for review, see Ferreira, Apel, & Henderson, 2008). Although looking at nothing studies have largely focused on internal episodic representations (e.g., of learned stimuli), both memory retrieval and construction (e.g., of future events) rely heavily on schematic representations, knowledge networks extracted from multiple repeated experiences (for review, see Gilboa & Marlatte, 2017). Thus, extending previous work, in the current study we present the first examination of how eye movements, and specifically looking at nothing, reflect and potentially support the reinstatement of mentally stored scene and event schemas during the construction of simulated future episodes.

Research over the past several decades has established a critical link between remembering the past and imagining the future. Both

* Corresponding author at: Psychology Department, Harvard University, 33 Kirkland St., Cambridge, MA 02138, USA.
E-mail address: jordanawynn@g.harvard.edu (J.S. Wynn).

https://doi.org/10.1016/j.concog.2022.103302
Received 26 November 2021; Received in revised form 2 February 2022; Accepted 16 February 2022
Available online 28 February 2022
1053-8100/

processes rely on *constructive episodic simulation*, the ability to reactivate and recombine spatial and temporal elements from one's personal past (Schacter & Addis, 2007, 2020; for review, see Addis, Wong, & Schacter, 2007). In addition to this shared cognitive basis, both episodic remembering and future imagining recruit a similar *core network* of brain regions, including the prefrontal cortex, prarahippocampal cortex, and hippocampus (Benoit & Schacter, 2015). Importantly, recent work suggests that this network, and specifically the hippocampus, is both anatomically and functionally connected to regions involved in visual imagery (Bird, Capponi, King, Doeller, & Burgess, 2010; Hassabis, Kumaran, & Maguire, 2007; for review, see Pearson, 2019) and oculomotor control (Ryan et al., 2019; Shen, Bezgin, Selvam, McIntosh, & Ryan, 2016; for review, see Ryan, Shen, & Liu, 2020). Activity in these (memory- and imagery-related) regions has also been linked to eye-movement-based mnemonic reinstatement (Ryals, Wang, Polnaszek, & Voss, 2015; Wynn, Liu, & Ryan, 2021; see also, Bone et al., 2019), further suggesting that the neural mechanisms that support constructive episodic simulation and gaze reinstatement are tightly linked.

Recent work suggests that the ability to freely and voluntarily execute eye movements (as opposed to maintaining gaze on a static or moving stimulus) is important for successful retrieval of autobiographical memories (Lenoble, Janssen, & El Haj, 2019) and simulation of possible future events (de Vito, Buonocore, Bonnefon, & Della Sala, 2015). However, whereas gaze fixations have been positively correlated with autobiographical memory retrieval (Armson, Diamond, Levesque, Ryan, & Levine, 2019), other work indicates that higher fixation rates are negatively associated with future imagining (Sheldon, Cool, & El-Asmar, 2019). Given these discrepant findings, it remains unclear what role eye movements play in constructive episodic simulation (i.e., what accounts for the above-noted free-viewing advantage). A promising hypothesis is that eye movements facilitate the broad reactivation of spatial and/or temporal contextual representations from memory, supporting the retrieval or generation of further episodic details (see Wynn et al., 2019; see also, Conti & Irish, 2021). Importantly however, such reactivation would not be evident in fixations rates, but rather in reinstated gaze patterns.

In the present study, we utilized gaze similarity analyses (see Wynn, Ryan, & Buchsbaum, 2020) to probe the role of eye movements in constructive episodic simulation and specifically, *how* eye movements during simulation reflect, and potentially support, the retrieval and reactivation of contextual details from memory. To this end, we leveraged existing data from Sheldon et al. (2019), in which participants imagined and described (out loud) future scene (e.g., beach) and event (e.g., wedding) scenarios while looking at a blank screen or Dynamic Visual Noise[1]. This latter condition was intended to reduce access to visual imagery processes thought to be critical for simulation (Sheldon et al., 2019; see also, Anderson, Dewhurst, & Dean, 2017) and was indeed found to reduce the number of internal details produced during simulation (Sheldon et al., 2019). To further investigate the role of eye movements in simulation, we generated gaze templates for cued schemas by aggregating the eye movements of all participants imagining future scene and event scenarios. Because schemas are based on gist-like details extracted across multiple repeated experiences, they show striking consistency across individuals (for review, see Ghosh & Gilboa, 2014; Gilboa & Marlatte, 2017; see also, Baldassano, Hasson, & Norman, 2018). This shared feature space allows us to measure the degree to which a schema is reactivated as a function of the degree to which those shared features, indexed here via aggregated (across-participant) eye movements, are expressed at the individual level, indexed here via individual (within-participant) eye movements. Accordingly, in the current study we present a novel multivariate eye-movement-based analysis for quantifying the schematicity of simulated events, which we validate with a novel linguistic measure (using natural language processing) for quantifying schematic content in narrative data.

Based on prior evidence of functional gaze reinstatement (i.e., gaze reinstatement that supports memory retrieval: for review, see Wynn et al., 2019), we predicted that: 1) if eye movements express and/or support the reactivation of schema-related contextual details, they should be more similar to the gaze template for the corresponding (matching) cue than to the gaze templates for the other (mismatching) cues, and 2) if this reactivation contributes to successful simulation, schema-related eye movements should be positively correlated with objective (i.e., number of episodic details) or subjective (i.e., familiarity and pre-experiencing ratings, see Table S1) measures of simulation success. Moreover, given that eye movements have been proposed to support memory retrieval particularly when task difficulty is high (see Wynn et al., 2019), we expected the predicted effects to be larger in the visual noise condition, in which internal details were significantly reduced as a result of limited access to visual imagery processes (Sheldon et al., 2019; see also, Anderson, Dewhurst, & Dean, 2017), relative to the control condition. Likewise, based on prior work linking eye movements to reinstatement of visuospatial details (see Wynn et al., 2019), we anticipated that the predicted effects would be larger for simulated scenes, which rely more heavily on perceptual details (Sheldon et al, 2019), than simulated events. In line with our predictions, we provide novel evidence that eye movements during simulation (in the face of distracting visual noise) are both schema-specific and predictive of simulation success, suggesting a critical role for overt gaze shifts in constructive episodic simulation.

2. Methods

2.1. Participants

Participants were 40 young adults (32 female) aged 18–35 ($M = 21$, $SD = 1.4$) with normal or corrected-to-normal vision, English fluency, and no history or psychiatric or neurological disorders. Participants were recruited through the McGill University participant pool and online classified ads. All participants provided written informed consent in accordance with the McGill code of ethics. Two participants were excluded from analysis due to incomplete data collection and failure to follow instructions.

[1] These manipulations were originally included to 1) examine differences in the content used to construct simulations of scenes vs. events, and 2) investigate the role of visual imagery processes in episodic simulation.

Based on significant effects in the data from Wynn et al., 2020 (match similarity > mismatch similarity, $d = 0.47$), power analysis (using the pwr package in R) indicated that the current sample size was sufficient to detect a reliable main effect of gaze similarity using a paired-samples t-test ($\alpha = 0.05$) with 80% power.

2.2. Stimuli

Stimuli consisted of 4 scene cues and 4 event cues (see Table 1) selected from a larger set of 30 cues based on ratings of scene and event representativeness, respectively (for selection criteria, see Sheldon et al., 2019). During simulation, the screen was occupied by either a grey square (control condition) or one of 6 possible Dynamic Visual Noise stimuli[2] (visual noise condition).

2.3. Apparatus

Stimuli were presented using E-Prime experimental software on a 24-inch monitor (1274 × 962 pixel resolution) positioned 30 cm away from participants. Monocular eye movements were recorded using a head mounted EyeLink 1000 Plus eyetracking system at 1000 Hz sampling rate (SR Research Ltd., Mississauga, Canada) while participants were positioned in a chin rest to reduce head movements. Saccades and blinks were defined by EyeLink as saccades greater than 0.5° of visual angle and the period in which saccade signal was missing for three or more consecutive samples, respectively. All remaining samples were classified as fixations.

2.4. Procedure

The current study used the same stimuli and procedure as was previously reported in Sheldon et al. (2019). Participants in the present study completed 8 trials of a simulation task, in which they were instructed to imagine and describe, in as much detail as possible, cued scene and event scenarios (Table 1, Fig. 1), while maintaining their gaze on the screen. Participants were specifically instructed to imagine a plausible future scenario and not to recall a past memory. Scene and event cues were counterbalanced across participants, with each cue being presented evenly in both the control condition (grey screen) and the visual noise condition (Dynamic Visual Noise screen). Within participants, 2 of each of the scene and event cues were presented in each condition. The temporal order of cue presentation was randomized within participants. On each trial, participants heard one of 8 scene or event cues presented auditorily over headphones. Participants were instructed to construct a scenario based on the presented cue and to press a button to indicate that they had done so (access phase). Following the button press, participants saw a grey screen or Dynamic Visual Noise and were instructed to maintain their gaze on the screen while they described the scenario in as much detail as possible for up to 3 min or until they were finished, as indicated by a button press (elaboration phase). If participants pressed the button prior to the 3-minute mark, they were presented with the prompt “can you think of any other details?”. After the button press or after 3 min elapsed, participants rated their simulation for sense of familiarity and pre-experiencing on 7 scales (see Table S1). Finally, following all 8 trials, participants completed 3 neuropsychological tests; for the purposes of the present study, data from these tests was not analyzed. During the elaboration phase, simulations were audio recorded and transcribed and eye movements were recorded.

2.5. Data analysis

2.5.1. Description scoring

Simulations were scored using an adapted version of the Autobiographical Interview scoring procedure (Levine, Svoboda, Hay, Winocur, & Moscovitch, 2002) for imagined scenarios (for examples, see Madore, Gaesser, & Schacter, 2014; Madore, Szpunar, Addis, & Schacter, 2016; Sheldon et al., 2015). Briefly, simulations were segmented into details, which were then classified as either internal or external, reflecting episodic and non-episodic processes, respectively. Whereas internal details are directly related to the described scenario, external details are tangential to the described scenario (e.g., commentary, semantic knowledge, etc.). For further details on description scoring, see the Supplemental Materials.

2.5.2. Subjective ratings

Subjective ratings were reduced to 2 principal components using a principal component factor analysis with varimax rotation (see Table S1). The first component, defined here as *Pre-Experiencing*, included ratings for Vividness (“How vividly can you picture this event in your mind?” Factor loading = 0.83), Sense of Presence (“How much did you feel as if you were experiencing this imagined scenario?” Factor loading = 0.76), and Spatial Arrangement (“How much did this scenario involve a sense of how things were arranged in space?” Factor loading = 0.73) and is thought to reflect the recruitment of imaginative processes during simulation. The second component, defined here as *Familiarity*, included ratings for Scenario Familiarity (“How familiar are you with the imagined scenario?” Factor loading = 0.94) and Reminders of Past Events (“Did this scenario remind you of a past personal event?” Factor loading = 0.94). Scores for each component were obtained by averaging the scores from the constituent ratings. For further details regarding factors and comparisons of scores, see the Supplemental Materials.

[2] The matching of visual noise stimuli to cues was fully randomized

Table 1
Simulation Cues.

Cue	Scenario
Event	A future and plausible award you will receive
Event	A future and plausible exam you will write
Event	A future and plausible wedding reception that you will attend
Event	A future and plausible holiday meal
Scene	You are standing in the aisles at a public library
Scene	You are lying on a white sand beach in a beautiful tropical bay
Scene	You are standing by a tombstone in a cemetery
Scene	You are standing by a small stream somewhere deep in a forest

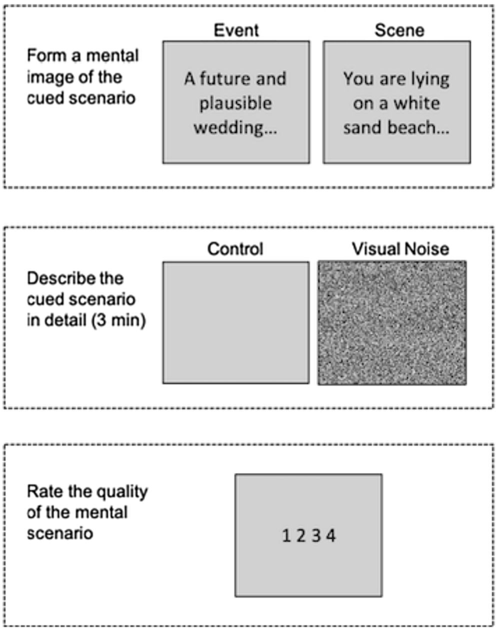

Fig. 1. Visualization of the experimental procedure (adapted from Sheldon et al., 2019). Participants are cued with a scene or event schema and must then describe the cued scenario in detail while looking at a blank (control) screen or visual noise stimulus. After 3 min or a button press, participants must rate the subjective quality of their imagined scenario.

2.6. Eye movement similarity analysis

Analysis of eye movement data was based on eye movements made during the elaboration phase[3] (when participants were describing the scenario), excluding the first fixation and all fixations off screen. To quantify cue-specific viewing, we used a leave-one-subject-out (LOSO) cross validation procedure. Specifically, to investigate how well individual gaze patterns could be predicted by the group-average gaze patterns for the same cue, we correlated the smoothed and duration-weighted fixation density maps for each cue (generated using the x-y coordinate fixations of all participants) with the corresponding left out participant's density map for the same cue (match similarity), using the eyesim package (https://github.com/bbuchsbaum/eyesim, Buchsbaum, 2021; see Wynn et al., 2020; for visualization of method, see Ryan, Wynn, Shen, & Liu, 2021). To rule out the possibility that similarity between participant-specific and group-average gaze patterns was driven by cue-invariant viewing biases (e.g., center bias), we additionally computed the similarity between participant-specific density maps and group-average density maps for all other cues. The resulting scores were averaged to obtain a single mismatch similarity score. Mismatch similarity scores were subtracted from match similarity scores to obtain controlled gaze similarity scores, reflecting cue-specific viewing that is not driven by generic viewing tendencies. For further details regarding the similarity analysis, see the Supplemental Materials.

[3] For mean elaboration times for each cue, see Table S2.

3. Results

3.1. Eye movements during visual noise simulation are schema-specific

Statistical analyses were conducted with R (R Core Team, 2017). To investigate whether eye movements during simulation aligned with the gaze template for the corresponding schema, we ran an ANOVA using the *afex* package (Singmann et al., 2016) with similarity value as the dependent variable, and cue (event, scene), condition (control, Dynamic Visual Noise), and similarity template (match, mismatch) as independent variables. If individual gaze patterns indeed reflect the cued schema, they should be more similar to the gaze template (i.e., group-average density map) for the same schema than to the gaze templates for all other schemas (i.e., match similarity should be significantly greater than mismatch similarity).

Results of the ANOVA revealed a significant interaction of similarity template by condition [$F(1,37) = 4.76, p = .036, \eta_p^2 = 0.11$], as well as significant main effects of condition [$M(SE)_{Control} = 0.402(0.01)$, $M(SE)_{Visual\ Noise} = 0.415(0.01)$; $F(1,37) = 4.22, p = .047, \eta_p^2 = 0.10$] and cue [$M(SE)_{Event} = 0.405(0.01)$, $M(SE)_{Scene} = 0.412(0.01)$; $F(1,37) = 4.51, p = .040, \eta_p^2 = 0.11$]. Follow up paired-samples *t*-tests indicated that controlled gaze similarity, indexed as the difference in match and mismatch similarity scores, was significantly greater in the visual noise condition relative to the control condition [$t(37) = 2.10, p = .043, d = 0.34$] but did not differ for scenes and events [$t(37) = 0.22, p = .83, d = 0.04$]. Further paired-samples *t*-tests of the difference in mean match and mismatch similarity scores in each condition (control, visual noise) indicated that match similarity was significantly greater than mismatch similarity in the visual noise condition only [$t(37) = 2.17, p = .036$[4], $d = 0.35$, see Fig. 2A; control condition: $t(37) = -2.01, p = .051, d = 0.33$]. Critically, these findings indicate that participant-specific eye movements during simulation in the face of visual noise reflect the specific cued schema. For all following analyses, only data from the visual noise condition were included, as these produced a significant difference between match and mismatch similarity scores in the expected direction (match > mismatch).

To further investigate whether the described eye movement measure captured the schematicity of a simulated event, we tested for a relationship between eye movements and a novel linguistic measure of schematicity based solely on participants' transcribed simulations[5]. We derived this linguistic measure by counting the number of schema-relevant words in each simulation, then dividing by the total number of words in the simulation, to obtain the proportion of content that is schema-relevant. To identify schema-relevant words, we used methods from natural language processing (specifically, Global Vectors for Word Representation (GloVe), an unsupervised learning algorithm for obtaining vector representations for words; Pennington, Socher, & Manning, 2014). For each cue word, we extracted a list of 10,000 related words based on the frequency of word co-occurrences in a corpus of webpages. This approach yielded word lists that sufficiently captured schematic (cue-related) words, while excluding non-schematic words. We scored each word in the transcribed simulations based on its match to the cue-related word list. Any word appearing in the cue-related word list received a score of 1, while all other words received a score of 0. For each simulation, we then derived the proportion of total words that were related to the cue word (% schema words). This measure was validated against two external measures of schematicity. For further details on this scoring approach, including validation procedures, see the Supplemental Materials.

Prior to analysis, one participant was removed on the basis of mean % schema words > 2.5 sd from the mean. A bootstrapped correlation ($n = 1000$) of mean controlled gaze similarity scores (indexed as the difference in match and mismatch similarity scores, for visual noise trials only), and mean % schema words (using the *wBoot* package, Weiss, 2016) was significant ($r = 0.28, p = .027$, *95 %CI* [0.038, 0.544], see Fig. 2B), indicating that participants who on average exhibited more schema-specific gaze patterns also produced a greater number of schema-related details during simulation.

3.2. Schema-specific eye movements support simulation success

To investigate whether schema-related eye movements were associated with simulation success, we ran a linear mixed effects model (LMEM; using the *lme4* package, Bates, Mächler, Bolker, & Walker, 2015) on controlled gaze similarity (match similarity-mismatch similarity for visual noise trials) with all interactions of cue, number of internal details, number of external details, pre-experiencing ratings, and familiarity ratings as predictors[6] and a random intercept for participant. We used backwards model comparison ($\alpha = 0.1$) to determine the best fit model (see Table 2).

Results of the best fit model of controlled gaze similarity (Table 2) revealed a significant effect of internal details for scenes, and this effect was significantly attenuated for events. Consistent with our predictions, this finding indicates that successful simulations, indexed by a high number of internal details, are characterized by a high degree of consistency with the gaze template for the corresponding scene schema. The model also revealed a marginally significant effect of familiarity, with more familiar cues producing greater eye movement similarity. Notably, both external details and pre-experiencing ratings were omitted from the model, indicating that they were not significantly predictive of gaze similarity. Although we did not observe significant schema-related eye movements in the control condition, a LMEM on gaze similarity for control trials (see Table S4) revealed a significant positive effect of external details and a significant negative effect of familiarity on gaze similarity.

[4] Mean match similarity (across all cues) was greater than all 7 mean mismatch similarity scores; $p = .03 \sim 0.17$.

[5] This measure is blind to episodicity. That is, words identified as schematic may also be scored (either independently or as part of a larger clause) as internal or external using the Autobiographical Interview scoring technique. These two methods are meant to be separate but complementary, providing distinct information about the nature of the described scenarios.

[6] Cue was dummy coded as: Scene = 0, Event = 1. All other variables were z-scored.

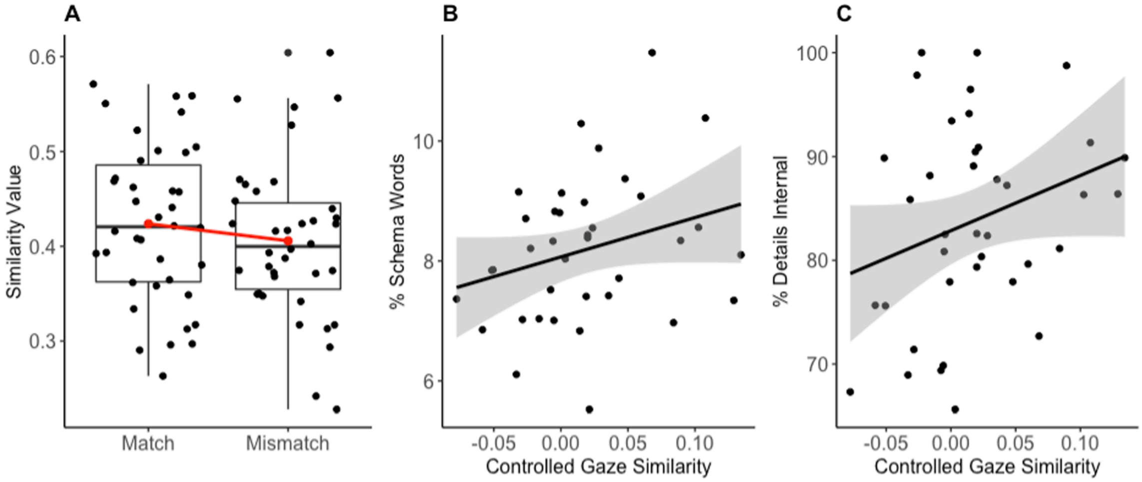

Fig. 2. (A) Similarity between schema-specific gaze patterns and gaze templates for matching and mismatching schema cues. (B) Correlation between controlled gaze similarity (match – mismatch) and the proportion of total words that are classified as schematic using the natural language processing scoring approach. (C) Correlation between controlled gaze similarity (match – mismatch) and the proportion of total words that are classified as internal using the autobiographical interview scoring approach. All plots contain data from the visual noise condition only.

Table 2
Controlled Gaze Similarity.

Fixed Effects					
	β	*95% CI*	*SE*	*t*	*p*
(Intercept)	0.021	0.001, 0.041	0.010	2.081	0.041 *
Internal Details	0.022	0.003, 0.042	0.009	2.277	0.024 *
Cue	−0.001	−0.024, 0.021	0.011	−0.094	0.925
Familiarity	0.011	−0.001, 0.023	0.006	1.845	0.067
Internal Details X Cue	−0.028	−0.051, −0.004	0.012	−2.350	0.020 *
Total Observations = 149					
		Variance	*SD*		
Random Effect for Participant (Intercept)		0.002	0.039		
Model equation: Controlled Gaze Similarity ~ Internal Details X Cue + Familiarity + (1 | Participant)					

Given that Sheldon et al. (2019) previously observed a significant negative effect of fixation rate (number of fixations per minute) on internal detail generation, we ran an additional model on the number of internal details (for external details, see Table S5) with both fixation rate, controlled gaze similarity, and their interaction as predictors (Table 3). Results of the best fit model revealed significant and opposing effects of both gaze similarity and fixation rate[7], neither of which was significantly predictive of external details (see Table S5). Together, these findings indicate that whereas a greater rate of fixations is related to a reduced number of internal details generated, similarity of those fixations to the cued template is related to a greater number of internal details generated. Additional gaze metrics are reported in Table S3.

Finally, to investigate whether the effect of gaze similarity on simulation success extended across participants, we correlated mean controlled gaze similarity scores with the mean proportion of total details that were scored as internal (as opposed to external). A bootstrapped correlation ($n = 1000$) of mean controlled gaze similarity scores and mean % internal details was significant ($r = 0.28$, *95 %CI* [0.018, 0.526], $p = .041$, see Fig. 2C), indicating that participants who on average exhibited more schema-specific gaze patterns also produced a greater number of internal details and a lower number of external details during simulation.

4. Discussion

Research using eye movement monitoring has provided critical evidence that eye movements not only reflect, but also actively support memory retrieval (for review, see Wynn et al., 2019). However, whereas the relationship between eye movements and episodic memory has received significant attention and study, the role of overt gaze shifts in episodic simulation remains unclear. Recent work suggests that eye movements play an important role in imagining future events, such that restricting them (e.g., to a static or moving fixation cross) impairs simulation (de Vito et al., 2015; see also, de Vito, Buonocore, Bonnefon, & Della Sala, 2014). However,

[7] The interaction of gaze similarity and fixation rate was omitted from the model

Table 3
Internal Details.

Fixed Effects					
	β	*95% CI*	*SE*	*t*	*p*
(Intercept)	−0.089	−0.335, 0.156	0.122	−0.732	0.468
Controlled Gaze Similarity	0.178	0.054, 0.302	0.063	2.839	0.005 **
Fixation Rate	−0.315	−0.500, −0.126	0.092	−3.415	<0.001 ***
Total Observations = 149					
		Variance		*SD*	
Random Effect for Participant (Intercept)		0.488		0.698	
Model equation: Internal Details ~ Controlled Gaze Similarity + Fixation Rate + (1 | Participant)					

executing more gaze fixations disrupts future simulation rather than supports it (Sheldon et al., 2019). Given that future simulation relies on many of the same cognitive and neural mechanisms that support episodic memory (Benoit & Schacter, 2015; Schacter & Addis, 2007, 2020), we proposed that eye movements during simulation would serve a similar role as they do during episodic memory – that is expressing, and likely facilitating, the retrieval of stored spatiotemporal contextual (in this case, schematic) details. Accordingly, in the present study, we applied gaze similarity and linguistic analyses to eye movement and narrative data from Sheldon et al. (2019), in which participants imagined and described plausible future scene and event scenarios, to test the hypothesis that overt gaze shifts during episodic simulation reflect and support the reactivation of schematic details from memory.

Lending support to our prediction, our results indicated that when imagining future scenes and events (while looking at a blank screen), participants spontaneously moved their eyes to screen regions associated with the cued schema (i.e., looking at nothing). Specifically, we found that in the visual noise condition, eye movements during simulation were more similar to the gaze template for the corresponding schema cue (generated from the eye movements of all other participants simulating the same scenario) than to the gaze templates for the other schema cues. This finding is consistent with previous work showing modulation of eye movements by prior knowledge (e.g., schemas: Võ & Wolfe, 2013; Wynn, Ryan, & Moscovitch, 2020), and with research indicating that gaze similarity supports memory retrieval, particularly when cognitive load exceeds cognitive resources (e.g., by virtue of age, Wynn et al., 2018; or task difficulty, Wynn et al., 2020; for review, see Wynn et al., 2019). Thus, when visual noise rendered simulation more difficult (as indicated by a significant reduction in internal details relative to the control condition; see Sheldon et al., 2019), presumably by interfering with visual imagery (see Anderson, Dewhurst, & Dean, 2017), participants may have executed overt gaze shifts reflecting the cued scenario in an effort to retrieve and visualize schema-specific contextual details.

Given the relationship between eye movements and visuospatial details evidenced by previous work (for review, see Wynn et al., 2019), we predicted that gaze similarity would be greater for simulated scenes than events. Notably, although match and mismatch similarity were significantly increased for simulated scenes relative to events, controlled gaze similarity (i.e., schema-specific viewing) did not differ significantly by cue. This finding suggests that eye movements may express both scene- and event-related schematic details. Indeed, prior work indicates that fixations correspond to both remembered or described spatial relations (e.g., Spivey & Geng, 2001; Johansson, Holsanova, & Holmqvist, 2006) as well as implied temporal order. Specifically, when retrieving past events, fixations trend more leftward than when imagining future events (Martarelli, Mast, & Hartmann, 2017), suggesting that like spatial context, temporal context (a key feature of schematic events) may be mapped using a spatial frame of reference. Accordingly, when simulating future scene and event schemas, eye movements may invoke visuospatial, temporal, or other contextual features that constitute a schema (i.e., that are shared across individuals).

To further validate gaze similarity as an index of the schematicity of a simulated scenario, we developed a novel linguistic analysis to quantify the number of schema-specific details present in narrative data. Leveraging tools from natural language processing (specifically, GloVe), we generated lexicons for each of the cue words and used these to identify schema-related words in described simulations. Results of the linguistic analysis revealed that the proportion of total words that were characterized as schematic was significantly correlated with gaze similarity, suggesting that eye movements capture the schematic content of simulations similarly to linguistic details. While further work will be required to fully probe the manner in which schematic and other mnemonic details are expressed via eye movements and natural language, these findings provide initial evidence that these modalities may be more related than previously thought (see also, Ferreira et al., 2008).

The described results support our hypothesis that eye movements during future simulation express reinstatement of schema-specific visuospatial contextual features. Importantly, however, a central feature of *functional* gaze reinstatement is that it predicts mnemonic performance (see Wynn et al., 2019). Accordingly, we proposed that processes central to constructive episodic simulation (e.g., retrieval) would not only be evident in the gaze patterns of participants during elaboration, but that such patterns would also be predictive of simulation success. In line with this prediction, we found that gaze similarity (to the corresponding schema gaze template) during visual noise simulation was positively correlated with the proportion of internal details generated, an objective measure of simulation success. This finding is consistent with recent work indicating that during autobiographical memory retrieval, activity in brain regions associated with stored knowledge networks or *schemas* precedes activity in the hippocampus, and further suggests that schemas may provide a foundation upon which episodic memories and simulations are (re)constructed (McCormick, Barry, Jafarian, Barnes, & Maguire, 2020). Moreover, when modeled together, gaze similarity was positively predictive of internal details, while gaze fixation rate was negatively predictive. Thus, while a greater rate of fixations might be indicative of a failed mental search (see Sheldon et al., 2019), schema-specific fixations appear to reflect and support active retrieval and construction processes (i.e., the generation of a schema). In other words, simulation success is determined not by how much we look, but rather by where we look.

In sum, the present results indicate that eye movements during episodic simulation in the face of distracting visual noise are both schema-specific and predictive of successful simulation. Considered together with prior evidence of functional gaze reinstatement, these findings suggest that eye movements may support the retrieval and reactivation of relevant (i.e., schematic) contextual details from memory by shifting attention to regions of the screen corresponding to imagined content, and in doing so, may provide a foundation on which to generate and construct an imagined scenario (see also, Hassabis & Maguire, 2007, 2009; McCormick et al., 2020). This interpretation is consistent with the proposed role of functional gaze reinstatement during episodic memory retrieval (for review, see Wynn et al., 2019) and simulation (see Conti & Irish, 2021), and provides a potential explanation for the detrimental effects of restricted viewing on autobiographical memory retrieval (Lenoble et al., 2019) and future imagining. (de Vito et al., 2015). Extending this work, the present results suggest that constructive memory is likely supported by complex interactions between processes involved in gaze control, stored knowledge, visual imagery, and memory retrieval (see also, Conti and Irish, 2021). The present findings also raise important questions for future work. For example, how does construction-related gaze reinstatement respond to changes in task difficulty (manipulated parametrically)? And how does gaze reinstatement similarly or differentially reflect, and potentially support, the imagination of spatial and temporal contextual features? Future work should continue to build on the present study by exploring the ways in which eye movements (and other effector systems) might support constructive episodic simulation across various tasks by reactivating stored contextual information.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors wish to thank Kelly Cool and Nadim El-Asmar for assistance with data collection and Tarek Amer for helpful comments on the manuscript. This work was supported by an NSERC Discovery Grant (Grant Number: RGPIN-04241) awarded to Signy Sheldon and a National Institute on Aging Grant (Grant Number: R01 AG008441) awarded to Daniel Schacter.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.concog.2022.103302.

References

Addis, D. R., Wong, A. T., & Schacter, D. L. (2007). Remembering the past and imagining the future: Common and distinct neural substrates during event construction and elaboration. *Neuropsychologia, 45*(7), 1363–1377. https://doi.org/10.1016/j.neuropsychologia.2006.10.016

Altmann, G. T. M. (2004). Language-mediated eye movements in the absence of a visual world: The "blank screen paradigm". *Cognition, 93*(2), 79–87. https://doi.org/10.1016/j.cognition.2004.02.005

Anderson, R. J., Dewhurst, S. A., & Dean, G. M. (2017). Direct and generative retrieval of autobiographical memories: The roles of visual imagery and executive processes. *Consciousness and Cognition, 49*, 163–171.

Armson, M. J., Diamond, N. B., Levesque, L., Ryan, J. D., & Levine, B. (2019). Vividness of recollection is supported by eye movements in individuals with high, but not low trait autobiographical memory. *Memory*, *206*(March 2020), 1–37. https://doi.org/10.1016/j.cognition.2020.104487.

Baldassano, C., Hasson, U., & Norman, K. A. (2018). Representation of real-world event schemas during narrative perception. *Journal of Neuroscience, 38*(45), 9689–9699. https://doi.org/10.1523/JNEUROSCI.0251-18.2018

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1). https://doi.org/10.18637/jss.v067.i01

Benoit, R. G., & Schacter, D. L. (2015). Specifying the core network supporting episodic simulation and episodic memory by activation likelihood estimation. *Neuropsychologia, 75*(3), 450–457. https://doi.org/10.1016/j.neuropsychologia.2015.06.034

Bird, C. M., Capponi, C., King, J. A., Doeller, C. F., & Burgess, N. (2010). Establishing the boundaries: The hippocampal contribution to imagining scenes. *Journal of Neuroscience, 30*, 11688–11695. https://doi-org.ezp-prod1.hul.harvard.edu/10.1523/JNEUROSCI.0723-10.2010.

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., McQuiggan, 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

Bradley Buchsbaum, eyesim, (2021), GitHub repository.

Conti, F., & Irish, M. (2021). Harnessing visual imagery and oculomotor behaviour to understand prospection. *Trends in Cognitive Sciences, 25*(4), 272–283. https://doi.org/10.1016/j.tics.2021.01.009

de Vito, S., Buonocore, A., Bonnefon, J. F., & Della Sala, S. (2014). Eye movements disrupt spatial but not visual mental imagery. *Cognitive Processing, 15*(4), 543–549. https://doi.org/10.1007/s10339-014-0617-1

de Vito, S., Buonocore, A., Bonnefon, J. F., & Della Sala, S. (2015). Eye movements disrupt episodic future thinking. *Memory, 23*(6), 796–805. https://doi.org/10.1080/09658211.2014.927888

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

Foulsham, T., & Kingstone, A. (2013). Fixation-dependent memory for natural scenes: An experimental test of scanpath theory. *Journal of Experimental Psychology. General, 142*(1), 41–56. https://doi.org/10.1037/a0028227

Ghosh, V. E., & Gilboa, A. (2014). What is a memory schema? A historical perspective on current neuroscience literature. *Neuropsychologia, 53*(1), 104–114. https://doi.org/10.1016/j.neuropsychologia.2013.11.010

Gilboa, A., & Marlatte, H. (2017). Neurobiology of schemas and schema-mediated memory. *Trends in Cognitive Sciences, 21*(8), 618–631. https://doi.org/10.1016/j.tics.2017.04.013

Hassabis, D., Kumaran, D., & Maguire, E. A. (2007). Using imagination to understand the neural basis of episodic memory. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience, 27*(52), 14365–14374. https://doi.org/10.1523/JNEUROSCI.4549-07.2007

Hassabis, D., & Maguire, E. A. (2007). Deconstructing episodic memory with construction. *Trends in Cognitive Sciences, 11*(7), 299–306. https://doi.org/10.1016/j.tics.2007.05.001

Hassabis, D., & Maguire, E. A. (2009). The construction system of the brain. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences, 364*(1521), 1263–1271. https://doi.org/10.1098/rstb.2008.0296

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. Retrieved from *Vision Research, 40*(10–12), 1489–1506 http://www.ncbi.nlm.nih.gov/pubmed/10788654.

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/s15516709cog0000_86

Johansson, R., & Johansson, M. (2013). Look here, eye movements play a functional role in memory retrieval. *Psychological Science, 25*(1), 236–242. https://doi.org/10.1177/0956797613498260

Lenoble, Q., Janssen, S. M. J., & El Haj, M. (2019). Don't stare, unless you don't want to remember: Maintaining fixation compromises autobiographical memory retrieval. *Memory, 27*(2), 231–238. https://doi.org/10.1080/09658211.2018.1501068

Levine, B., Svoboda, E., Hay, J. F., Winocur, G., & Moscovitch, M. (2002). Aging and autobiographical memory: Dissociating episodic from semantic retrieval. *Psychology and Aging, 17*(4), 677–689. https://doi.org/10.1037//0882-7974.17.4.677

Madore, K. P., Gaesser, B., & Schacter, D. L. (2014). Constructive episodic simulation: Dissociable effects of a specificity induction on remembering, imagining, and describing in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(3), 609–622. https://doi.org/10.1037/a0034885

Madore, K. P., Szpunar, K. K., Addis, D. R., & Schacter, D. L. (2016). Episodic specificity induction impacts activity in a core brain network during construction of imagined future experiences. *Proceedings of the National Academy of Sciences, 113*(38), 10696–10701. https://doi.org/10.1073/pnas. 1612278113

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

McCormick, C., Barry, D. N., Jafarian, A., Barnes, G. R., & Maguire, E. A. (2020). VmPFC drives hippocampal processing during autobiographical memory recall regardless of remoteness. *Cerebral Cortex, 30*(11), 5972–5987. https://doi.org/10.1093/cercor/bhaa172

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), 8. https://doi.org/10.1167/14.1.8

Pearson, J. (2019). The human imagination: The cognitive neuroscience of visual mental imagery. *Nature Reviews Neuroscience, 20*(10), 624–634. https://doi.org/10.1038/s41583-019-0202-9

Pennington, J., Socher, R., & Manning, C. D. (2014). Glove: Global vectors for word representation. In *Proceedings of the Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1532–1543).

R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation.

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)00084-6

Ryals, A. J., Wang, J. X., Polnaszek, K. L., & Voss, J. L. (2015). Hippocampal contribution to implicit configuration memory expressed via eye movements during scene exploration. *Hippocampus, 25*(9), 1028–1041. https://doi.org/10.1002/hipo.22425

Ryan, J. D., Shen, K., Kacollja, A., Tian, H., Griffiths, J., Bezgin, G., & McIntosh, A. R. (2019). Modeling the influence of the hippocampal memory system on the oculomotor system. *Network Neuroscience, 1–17*. https://doi.org/10.1162/netn_a_00120

Ryan, J. D., Shen, K., & Liu, Z. (2020). The intersection between the oculomotor and hippocampal memory systems: Empirical developments and clinical implications. *Annals of the New York Academy of Sciences, 1464*(1), 115–141. https://doi.org/10.1111/nyas.14256

Ryan, J. D., Wynn, J. S., Shen, K., & Liu, Z.-X. (2021). Aging changes the interactions between the oculomotor and memory systems. *Aging, Neuropsychology, and Cognition, 00*(00), 1–25. https://doi.org/10.1080/13825585.2021.2007841

Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: Remembering the past and imagining the future. *Philosophical Transactions of the Royal Society B: Biological Sciences, 362*(1481), 773–786. https://doi.org/10.1098/rstb.2007.2087

Schacter, D. L., & Addis, D. R. (2020). Memory and imagination: Perspectives on constructive episodic simulation. In A. Abraham (Ed.), *The Cambridge Handbook of the Neuroscience of Imagination*. Cambridge University Press.

Sheldon, S., Cool, K., & El-Asmar, N. (2019). The processes involved in mentally constructing event- and scene-based autobiographical representations. *Journal of Cognitive Psychology, 31*(3), 261–275. https://doi.org/10.1080/20445911.2019.1614004

Sheldon, S., Vandermorris, S., Al-Haj, M., Cohen, S., Winocur, G., & Moscovitch, M. (2015). Ill-defined problem solving in amnestic mild cognitive impairment: Linking episodic memory to effective solution generation. *Neuropsychologia, 68*, 168–175. https://doi.org/10.1016/j. neuropsychologia.2015.01.005

Shen, K., Bezgin, G., Selvam, R., McIntosh, A. R., & Ryan, J. D. (2016). An anatomical interface between memory and oculomotor systems. *Journal of Cognitive Neuroscience, 28*(11), 1772–1783. https://doi.org/10.1162/jocn_a_01007

Singmann, H., Bolker, B., Westfall, J., Aust, F. (2016). afex: analysis of factorial experiments. Vienna, Austria: R Foundation.

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

Treisman, a M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7351125.

Võ, M.-L.-H., & Wolfe, J. M. (2013). The interplay of episodic and semantic memory in guiding repeated search in scenes. *Cognition, 126*(2), 198–212. https://doi.org/10.1016/j.cognition.2012.09.017

Weiss, N. A. (2016). *WBoot: Bootstrap Methods*. Vienna, Austria: R Foundation.

Wynn, J. S., Liu, Z.-X., & Ryan, J. D. (2021). Neural correlates of subsequent memory-related gaze reinstatement. *Journal of cognitive neuroscience, 1–15*. https://doi.org/10.1162/jocn_a_01761

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

Wynn, J. S., Ryan, J. D., & Buchsbaum, B. R. (2020). Eye movements support behavioral pattern completion. *Proceedings of the National Academy of Sciences of the United States of America, 117*(11), 6246–6254. https://doi.org/10.1073/pnas.1917586117

Wynn, J. S., Ryan, J. D., & Moscovitch, M. (2020). Effects of prior knowledge on active vision and memory in younger and older adults. *Journal of Experimental Psychology: General.*. https://doi.org/10.1037/xge0000657

Wynn, J. S., Shen, K., & Ryan, J. D. (2019). Eye movements actively reinstate spatiotemporal mnemonic content. *Vision, 3*(2), 21. https://doi.org/10.3390/vision3020021