Cognition 242 (2024) 105624

Contents lists available at ScienceDirect

Cognition

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

A computational modeling approach to investigating mind wandering-related adjustments to gaze behavior during scene viewing

Kristina Krasich[a,b,*,1], Kevin O'Neill[a,b,c,1], Samuel Murray[d], James R. Brockmole[e], Felipe De Brigard[a,b,c,f], Antje Nuthmann[g]

[a] Center for Cognitive Neuroscience, Duke University, Durham, NC, USA
[b] Duke Institute for Brain Sciences, Duke University, Durham, NC, USA
[c] Department of Psychology & Neuroscience, Duke University, Durham, NC, USA
[d] Philosophy Department, Providence College, Providence, RI, USA
[e] Department of Psychology, University of Notre Dame, Notre Dame, IN, USA
[f] Department of Philosophy, Duke University, Durham, NC, USA
[g] Institute of Psychology, Kiel University, Kiel, Germany

ARTICLE INFO

Keywords:
Visual-cognitive processing
Mind wandering
Fixation duration
Computational modeling
Scene viewing

ABSTRACT

Research on gaze control has long shown that increased visual-cognitive processing demands in scene viewing are associated with longer fixation durations. More recently, though, longer durations have also been linked to mind wandering, a perceptually decoupled state of attention marked by decreased visual-cognitive processing. Toward better understanding the relationship between fixation durations and visual-cognitive processing, we ran simulations using an established random-walk model for saccade timing and programming and assessed which model parameters best predicted modulations in fixation durations associated with mind wandering compared to attentive viewing. Mind wandering-related fixation durations were best described as an increase in the variability of the fixation-generating process, leading to more variable—sometimes very long—durations. In contrast, past research showed that increased processing demands increased the mean duration of the fixation-generating process. The findings thus illustrate that mind wandering and processing demands modulate fixation durations through different mechanisms in scene viewing. This suggests that processing demands cannot be inferred from changes in fixation durations without understanding the underlying mechanism by which these changes were generated.

1. Introduction

Confined by the physical and cognitive limitations of the visual system, people cannot adequately perceive everything in the environment simultaneously. Instead, visual perception is piecemeal, with people shifting their eyes frequently to acquire situational information. The nonrandom nature of gaze control suggests that eye movements index the information-processing priorities of the visual system (Just & Carpenter, 1976). For example, fixation durations—how long the eyes remain relatively still at one location—are thought to vary according to the time needed for acquiring and evaluating visual inputs toward comprehension (Rayner, 1978, 1998, 2009).

During the exploration of naturalistic scenes, fixation durations are affected by changes in *global* image features. For example, fixation durations are prolonged when the luminance of the entire scene is reduced (Henderson, Nuthmann, & Luke, 2013; Loftus, 1985) or when color is removed (von Wartburg et al., 2005). In addition, a number of studies have employed gaze-contingent scene manipulations to investigate the degree to which fixation durations can be immediately adjusted to ongoing visual-cognitive processing demands (Glaholt, Rayner, & Reingold, 2013; Henderson & Pierce, 2008; Pannasch, Schulz, & Velichkovsky, 2011; Walshe & Nuthmann, 2014). The results from these studies suggest that the timing of fixations adapts to stimulus changes (e.g., a reduction in scene luminance) that occurred on a fixation-to-fixation basis (Walshe & Nuthmann, 2021, for a review).

Fixation durations during scene viewing are also modulated by *local*

* Corresponding author at: Center for Cognitive Neuroscience, Duke University, Durham, NC, USA.
E-mail address: kristina.krasich@duke.edu (K. Krasich).
[1] Denotes equal contribution.

https://doi.org/10.1016/j.cognition.2023.105624
Received 8 August 2022; Received in revised form 8 July 2023; Accepted 14 September 2023
Available online 7 November 2023

scene processing difficulty. Nuthmann (2017) and Tatler, Brockmole, and Carpenter (2017) showed that low-, intermediate-, and higher-level information at the fovea is systematically related to fixation durations when inspecting scenes under different task instructions. For example, these studies found that fixation durations increase as the number of edges in foveal vision increases. When fixation durations are analyzed with regard to objects in scenes, gaze duration represents the summed duration of all fixations landing on the object before moving away from it (Henderson, Weeks Jr, & Hollingworth, 1999). Gaze durations tend to be longer for larger objects (Wang, Hwang, & Pomplun, 2010) and, independently, for higher-salience objects (Nuthmann, Schütz, & Einhäuser, 2020). Moreover, gaze durations are longer for objects that are out-of-place than for objects that cohere with the scene (Coco, Nuthmann, & Dimigen, 2020; Henderson et al., 1999; Loftus & Mackworth, 1978; Võ & Henderson, 2009).

Neuroscientific findings using co-registered eye tracking and magnetic resonance scanning showed that individual fixation durations under normal scene viewing conditions (i.e., without any experimental manipulations) were positively correlated with activation in brain regions that support visual-cognitive processing as well as the executive control of ocular motor behavior during scene viewing (Henderson & Choi, 2015). These findings thus suggested that fixation durations reflected naturally occurring modulations in real-time scene processing, with longer fixations indicating greater processing.

Considering past research on gaze control during scene viewing collectively, a longstanding conventional view has emerged according to which longer fixation durations reflect increasingly complex and more difficult visual-cognitive processing. One challenge with the conventional view, though, is that longer fixation durations have also been associated with mind wandering during scene viewing (Krasich et al., 2018; Zhang, Anderson, & Miller, 2021). Mind wandering is considered a state of attenuated visual-cognitive processing, as indicated by converging evidence from behavioral (Mason et al., 2007; Stawarczyk, Majerus, Maj, Van der Linden, & D'Argembeau, 2011), neuroimaging (Christoff, Irving, Fox, Spreng, & Andrews-Hanna, 2016; Fox, Spreng, Ellamil, Andrews-Hanna, & Christoff, 2015; Turnbull et al., 2019), and electroencephalogram (EEG) (Baird, Smallwood, Lutz, & Schooler, 2014; Barron, Riby, Greer, & Smallwood, 2011; Kam et al., 2021; Smallwood, Beach, Schooler, & Handy, 2008) research.

Krasich et al. (2018) asked participants to study pictures of naturalistic scenes for 45 to 75 s in anticipation of a later memory test. While viewing these scenes, participants also responded to pseudo-randomly distributed thought probes that asked participants to report whether they were mind wandering or paying attention to viewing the image at a given moment in time. Specifically, the prompts asked, "In the moments right before this message, were you paying attention to the picture or zoning out." Fixations made prior to reports of mind wandering were then compared to fixations made prior to reports of attentive viewing. The findings showed that at least 15 s prior to the self-report, fixations associated with self-reported mind wandering were on average significantly longer than the fixations made prior to reports of attentive viewing. Moreover, rates of reported mind wandering during initial scene memorization were negatively correlated with performance on the later memory test, suggesting that the longer durations did not correspond to better processing of the fixated content. Zhang et al. (2021) showed a similar link between probe-caught mind wandering, longer fixation durations, and worse scene memory in a scene memorization task that included a substantially larger stimulus set, shorter viewing times (i.e., 10 s), and more thought probes than in Krasich et al. (2018). Thus, across two separate studies, mind wandering—and its presumed attenuated processing—has been linked to longer fixation durations in scene viewing.

Research on mind wandering thus presents a difficulty for inferring visual-cognitive processing characteristics from fixation durations in scene viewing. Specifically, the same behavioral phenomenon (i.e., increased fixation durations) has been empirically associated with both bouts of increased (e.g., under degraded, complex, semantically interesting conditions) and decreased (i.e., during mind wandering) processing demands. Therefore, if an increase in fixation durations is observed, it is not clear which is driving this change. It is thus critical to disambiguate what underlying mechanisms generate increased fixation durations during mind wandering for a more accurate, comprehensive view of gaze control. Toward this end, we used an established computational model of saccade timing and programming to investigate the mechanisms underlying the increase in fixation durations associated with self-reported mind wandering that was observed in Krasich et al. (2018) and Zhang et al. (2021). We then interpreted our findings in conjunction with past work that had used the same computational model to characterize the mechanism underlying increased fixation durations associated with higher processing demands in scene viewing (Henderson et al., 2013; Henderson & Pierce, 2008; Nuthmann, Smith, Engbert, & Henderson, 2010; Walshe & Nuthmann, 2014, 2021). Thus, our current work ultimately seeks to determine whether increased processing demands and mind wandering impact fixation durations through similar or disparate mechanisms, with the further aim of understanding how visual-cognitive processing demands can be inferred from fixation durations during scene viewing.

1.1. The current work

We investigated the fixation duration distributions of the data reported in Krasich et al. (2018) and in Zhang et al. (2021). Typically, fixation duration analyses in scene viewing investigate mean durations, but this analytic approach implicitly assumes that the underlying distributions are symmetric and that mean measures can thus provide a good estimate of the central tendency of these distributions (Balota & Yap, 2011). However, it is well-known that fixation duration distributions in scene viewing are positively skewed and sometimes heavy tailed (Castelhano, Mack and Henderson, 2009; Nuthmann et al., 2010; Walshe & Nuthmann, 2014). Therefore, the distributions of two conditions can not only differ in the mean but also in the spread and the tail.

In the current work, we thus shifted our focus from analyzing central tendency alone to one that investigated the difference in the overall distributions of fixation durations between reports of mind wandering and reports of attentive viewing. We used a random-walk model for saccade timing and programming informally known as the Unnamed Computational Model (UCM) (Walshe & Nuthmann, 2021). Using a computational model allowed us to make specific predictions about which model parameters would best account for modulations in fixation durations associated with self-reported mind wandering compared to self-reported attentive viewing.

The UCM incorporates and develops principles that have been proposed to explain patterns of eye movements in high-level tasks including reading (SWIFT: Engbert, Nuthmann, Richter, & Kliegl, 2005; Schad & Engbert, 2012), visual search (ICAT: Trukenbrod & Engbert, 2014) and scene perception (CRISP: Nuthmann et al., 2010; Saez de Urabain, Nuthmann, Johnson, & Smith, 2017). At the heart of the model is an autonomous random saccade timer that keeps the eyes moving at a certain mean rate. Random-timer models acknowledge that eye-movement behavior is inherently rhythmic by nature and that the eyes never rest—even in the absence of cognitive processing demands (Lange, Pieczykolan, Trukenbrod, & Huestegge, 2018). In these models, the initiation of the saccade programming cascade is not directly coupled to aspects of visual-cognitive processing, which makes them particularly suitable for modeling fixation durations during mind wandering (see Nuthmann & Engbert, 2009).

Fig. 1 shows a simulation example of the UCM. The fixation-generating process involves several sub-processes. First, the random timer provides the signal to generate a new saccade program whose completion comprises multiple distinct stages (labile, non-labile, motor, and execution, see y-axis labels in Fig. 1). Both the timer and the different stages of saccade programming and execution are implemented

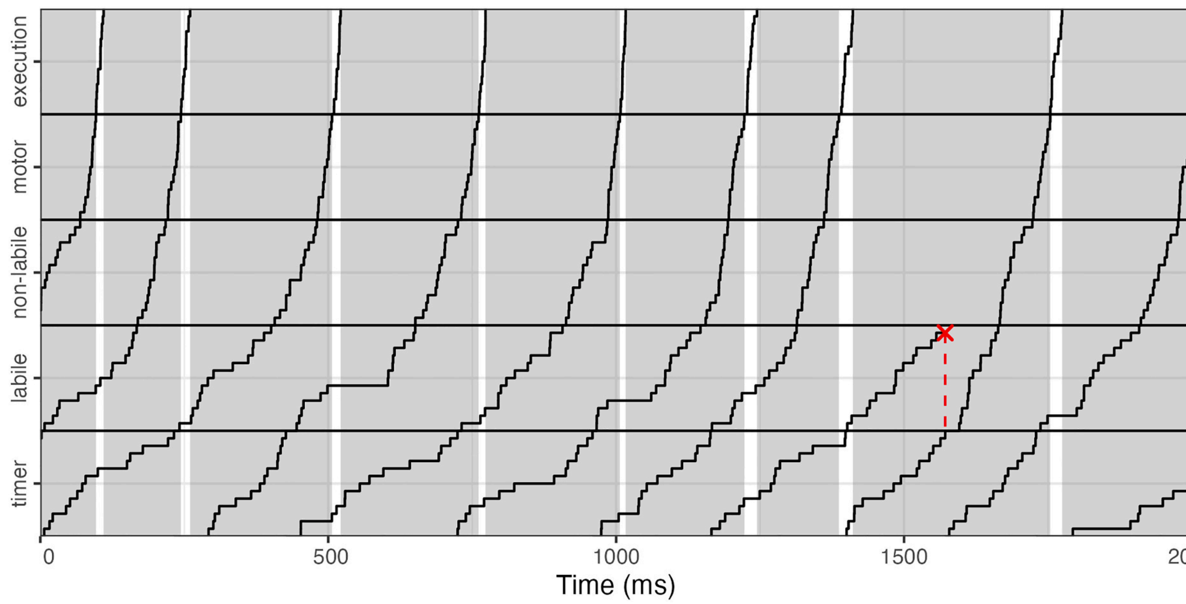

Fig. 1. Example trace generated by the UCM. The random timer cycles continuously over time, initiating a new labile saccade program upon each completion. The saccade program then continues into the nonlabile stage, the motor stage, and saccade execution. Fixations (represented by shaded areas) are the periods between subsequent saccades when the eyes remain relatively still. Occasionally, a saccade program in the labile stage is cancelled when the timer reaches threshold before that saccade program has reached the non-labile stage (highlighted in red). These cancellations are associated with longer fixation durations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

as independent discrete-state random walks that each rise toward a predefined threshold. As illustrated in the bottom part of Fig. 1, the saccade timer cycles continuously. Once the random walk of the timer reaches threshold, a new saccade program is initiated and the timer is reset to its initial state (e.g., Fig. 1 at ~25 ms). The new saccade program starts in a labile stage in which it can still be cancelled (Walshe & Nuthmann, 2015). Specifically, the saccade program will be cancelled if its random walk does not reach threshold before the random timer initiates a new saccade program (Fig. 1, highlighted in red). If a labile stage runs to completion, though, it is followed by a non-labile stage during which the saccade can no longer be cancelled (Becker & Jürgens, 1979; Ludwig, Mildinhall, & Gilchrist, 2007; Walshe & Nuthmann, 2015). The subsequent motor stage represents the time it takes for a neural signal to command the eyes to move, and the final stage represents the execution of the saccade.

For each internal sub-process S of the model, the time that is spent in S is determined by (1) the total number of discrete states N_s within the random walk and (2) the rate of state transitions r_s from state n to $n+1$. It can be shown that this duration has mean N_s/r_s and variance N_s/r_s^2. Notably, the mean and variance are necessarily related (see Supplementary Information for a derivation of this idea).

Past research has shown that modulations in fixations durations related to changes in processing difficulty could be modeled in terms of setting a new rate $r_s^{'} = \beta_{rate} r_s$ for the timer, the labile stage, and the nonlabile stage (i.e., rate modulation; Nuthmann et al., 2010; Walshe & Nuthmann, 2021). Fig. 2A provides a theoretical illustration of how rate modulation can impact the mean and variance of fixation durations (see Supplementary Information for a derivation). A lower rate will increase the amount of time spent in a given sub-process by a factor of $1/\beta_{rate}$ and will increase the variance of this duration by a factor of $1/\beta_{rate}^2$. In other words, random walks with a lower transition rate take longer to reach threshold on average, which in turn increases fixation durations. In principle, it is possible that, just like with increased processing difficulty, the longer fixation durations associated with mind wandering could be modeled via rate modulation, with mind wandering being associated with a slower rate than attentive viewing. Indeed, both Krasich et al. (2018) and Zhang et al. (2021) had speculated that bouts of mind

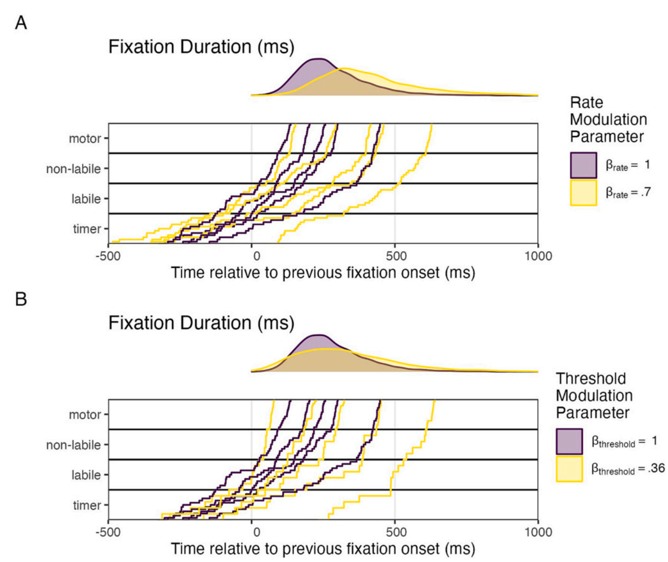

Fig. 2. Theoretical predictions of fixation duration distributions as a function of (A) rate modulation and (B) threshold modulation, which were applied to the random walks of the saccade timer and the labile and non-labile stages of saccade programming in the model. Five randomly sampled discrete-state random walks corresponding to five separate fixations are illustrated below the corresponding distribution. Each random walk is temporally aligned relative to the onset of the previous fixation, such that the fixation begins at 0 ms, and the endpoint of the random walk indicates the fixation duration. (A) Random walks with a lower state transition rate ($\beta_{rate} < 1$; in yellow) spend an increased duration within each sub-process on average, which will on average correspond with longer fixation durations. (B) Random walks with a lower number of discrete states ($\beta_{threshold} < 1$; in yellow) have a more variable duration within each sub-process, resulting in more saccade cancellations and longer fixation durations. All simulations were generated using the baseline parameters from Walshe and Nuthmann (2021) unless otherwise noted. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

wandering might be likened to a sort of 'processing difficulty.' That is, it might be more difficult to process relevant content during mind wandering, so people may need to fixate for longer periods of time for full comprehension. If mind wandering-related fixation durations are indeed best modeled as a decrease in rate modulation compared to attentive viewing, our findings would show that increased processing difficulty and mind wandering both modulate fixation durations through a similar mechanism.

In the UCM, fixation durations are also influenced by the total number of discrete states N_s within each random walk. Similar to rate modulation, the UCM allows for threshold modulation, which is a change in the number of steps required by the random walks for the timer, the labile stage, and the non-labile stage to reach threshold, $N_s^{'} = \beta_{threshold} N_s$. As illustrated in Fig. 2B, when a random walk requires fewer states to reach threshold, fixation durations become more variable because any exceptionally long or short state can deviate the corresponding fixation duration from the mean. Threshold modulation increases the variability of the time spent in a given sub-process by a factor of $1/\beta_{threshold}$, but it does not affect the mean duration of the sub-process. It can, however, impact mean fixation duration through modulating the rate of saccade cancellations (i.e., cancellation rate; see Supplementary Information for mathematical details). In random-timing models, saccade cancellations contribute to the variability in fixation durations and are particularly important for producing long tailed fixation duration distributions, which increases mean measures (Nuthmann et al., 2010; Saez de Urabain et al., 2017; Trukenbrod & Engbert, 2014; Walshe & Nuthmann, 2021). We considered the possibility that mind wandering-related fixations might be highly variable and sometimes very long (e.g., a blank stare). Thus, as part of our simulations, we assessed the degree to which mind wandering-related fixation durations could be modeled through threshold modulation. If mind wandering-related fixation durations are best modeled through threshold modulation, our findings would suggest that mind wandering increased mean measures of fixation through a different mechanism than increased processing difficulty. Moreover, our findings would show that the control mechanisms underlying fixation duration variability in scene viewing—which have received little attention to date—are in part linked to the attentional state of the observer.

To summarize our competing hypotheses, the increased mean fixation durations associated with mind wandering could be characterized in the UCM by a change in the state transition rate of the random walks (rate modulation) and/or a decrease in the number of states needed to reach threshold (threshold modulation) that in turn increases the number of saccade cancellations. To test these competing hypotheses, we fit the UCM to fixation durations measured in Krasich et al. (2018) and Zhang et al. (2021), testing which of the two mechanisms (i.e., rate or threshold modulation) would best account for fixation durations associated with self-reported mind wandering compared to self-reported attentive viewing. To preview our results, mind wandering-related fixation durations were best modeled as a decrease in threshold, which caused an increase in the number of cancellations compared to fixations associated with self-reported attentive viewing. This demonstrates a mechanism by which the attentive state of the observer impacts the variability of the fixation-generating process, leading to more variable—sometimes very long—durations. Moreover, when considered in conjunction with past research (Henderson et al., 2013; Henderson & Pierce, 2008; Nuthmann et al., 2010; Walshe & Nuthmann, 2014, 2021), our findings suggest that increased processing difficulty and mind wandering can both increase mean fixation durations albeit through different mechanisms, with the former increasing the mean duration of saccade programming and the latter increasing the variability of this programming.

2. Methods

2.1. Participants

Data were obtained from the main study of Krasich et al. (2018) and from Zhang et al. (2021). Participants from Krasich et al. (2018) were 51 volunteers from the University of Notre Dame, and those from Zhang et al. (2021) were 57 volunteers from the University of Michigan.

2.2. Stimuli and apparatuses

The stimuli from Krasich et al. (2018) were 12 color photographs of real-world urban scenes (800 × 600 pixels) presented at a viewing distance of 80 cm on a 20-in. monitor with a resolution of 1024 × 768 pixels. Eye movements were sampled at a rate of 1000 Hz using an EyeLink 2k tower-mounted eye tracking system (SR Research, Ltd., Kanata, Canada) with a chin and forehead rest. The stimuli from Zhang et al. (2021) were 180 color photographs (60 exteriors, 60 interiors, and 60 landscape) from the SUN (Xiao, Hays, Ehinger, Oliva, & Torralba, 2010) and the LabelMe (Russell, Torralba, Murphy, & Freeman, 2008) databases. The scenes were presented at a viewing distance of 70 cm on a 20.1-in. monitor in 1024 × 768 pixels resolution. Eye movements were sampled at a rate of 500 Hz using an Eyelink 1000 desk-mounted tracker without a chin or forehead rest.

2.3. Experimental procedures

Both Krasich et al. (2018) and Zhang et al. (2021) operationalized *mind wandering* as moments when participants were not avidly focused on the task and were instead engaging in task-unrelated thoughts (Smallwood & Schooler, 2006). This definition of mind wandering does not capture the origin or dynamic progression of mind wandering (Christoff et al., 2016; Christoff et al., 2018), but it is currently the most used operationalization (Mills, Raffaelli, Irving, Stan, & Christoff, 2018). *Attentive viewing*, on the other hand, refers to instances when participants self-reported having been paying attention to viewing the image rather than mind wandering.

Participants from Krasich et al. (2018) sequentially studied each image for 45 to 75 s ($M = 59.96$ s; $SD = 8.49$ s) in preparation for a later memory test. Thought probes were presented randomly after eight of the images, and asked, "In the moments right before this message, were you paying attention to the picture or were you zoning out?". Participants reported mind wandering on 27% of probes ($SD = 22\%$), for a total of 109 mind wandering observations. We focused on the fixations made within the 15-s time frame prior to the thought probe because this was the period in which fixations associated with self-reported mind wandering were significantly longer than fixations associated with self-reported attentive viewing. Just as in Krasich et al. (2018), fixations that were shorter than 50 ms (8 fixations; 0.05%) were excluded. Although Krasich et al. (2018) included fixations as long as 10,000 ms, here we excluded fixations longer than 2000 ms, which were rare (95 fixations, 0.64% of all fixations) and did not affect our main effect of interest. Overall, from Krasich et al. (2018), we included 3353 fixations preceding reports of mind wandering and 11,311 fixations preceding reports of attentive viewing.

Participants from Zhang et al. (2021) sequentially studied each image for 10 s also in preparation for a later memory test. Random thought probes appeared after 36 of the images, with each image-probe pairing being the same for each participant. The thought probes asked, "Where was your attention during the last picture?" and participants indicated, "I was focusing on the picture" or "I was thinking about something else." If participants indicated that they had been thinking of something else, they answered a second probe asking whether they had been doing so intentionally or unintentionally. On average, participants reported unintentional mind wandering on 22% ($SD = 19\%$) of probes and intentional mind wandering on 5% ($SD = 8\%$) of probes. Given the few observations of intentional mind wandering and that both types of mind wandering were associated with longer fixation durations, we excluded all trials in which participants reported intentional mind wandering in the current work. This resulted in a total of 428 mind wandering observations. We analyzed fixation durations across the entire 10-s viewing time. Just as in Zhang et al. (2021), fixations that were shorter than 80 ms (3539 fixations; 7.35%) and longer than 2000 ms (181 fixations; 0.38%) were excluded. Overall, from Zhang et al. (2021), we included 8516 fixations preceding reports of mind wandering and 35,992 fixations preceding reports of attentive viewing.

2.4. Computational modeling

To explore the mechanisms behind prolonged fixation durations during mind wandering, we fit the UCM (Walshe & Nuthmann, 2021) to each dataset. In this model, fixation durations are generated by simulating multiple random walks for saccade timing and different stages of saccade programming. The baseline model has a maximum of ten free parameters: a mean duration and a threshold for each of the five sub-processes of the model (i.e., saccade timing, labile and non-labile saccade programming, motor component, and saccade execution). To reduce the complexity of the model, we fixed the threshold to be equal for each sub-process (Walshe & Nuthmann, 2021). Also, the mean duration of the motor stage was fixed at 30 ms as a plausible value based on neurophysiological estimates (Becker, 1989; Becker, 1991). The mean duration of saccade execution was fixed at 40 ms, which is within a plausible range for saccade durations independent of task- or stimulus-features (Devillez, Guyader, Curran, & O'Reilly, 2020).

The UCM was previously used to model the degree to which the duration of individual fixations can be immediately adjusted to ongoing visual-cognitive processing demands (Walshe & Nuthmann, 2021). Here, we applied this general approach to model the control of fixation durations during mind wandering. Specifically, we contrasted two different influences on the timer, the labile stage, and the non-labile stage during empirically identified bouts of mind wandering: (a) a rate modulation parameter, which could decrease the rate of state transitions for these three sub-processes, and (b) a threshold modulation parameter, which could lower the number of state transitions required to proceed from one sub-process to the next (modulated thresholds were rounded to the nearest integer). In the Supplementary Information, we showed that while rate modulation impacts the mean and variance of the duration of each internal sub-process of the model, threshold modulation only impacts the variance, implying that the two modulation parameters are jointly identifiable. Notably, then, we applied both rate modulation and threshold modulation simultaneously during mind wandering, allowing us to distinguish between effects on each parameter. Our model therefore had a total of six free parameters: the mean durations for the random walks of the timer, the labile stage, and the non-labile stage, a fixed threshold for all random walks, the rate modulation parameter, and the threshold modulation parameter (Table 1).

We implemented the model using the R package *simmer* (Ucar, Smeets, & Azcorra, 2019). To find the maximum likelihood parameter values for the model, we used Bayesian optimization with the upper confidence bound utility function. Since the likelihood of UCM is not analytically tractable, we approximated the likelihood through simulation (Cranmer, Brehmer, & Louppe, 2020). That is, we binned the observed fixation durations and simulated fixation durations from 10,000 instances of the model into histograms and used the proportion of simulated data in each bin as the likelihood (Walshe & Nuthmann, 2021). Simulated fixations were excluded according to the same criteria as the participant-generated fixations. Specifically, for the data in Krasich et al. (2018), we placed fixation durations into 50-ms bins from 50 to 2000 ms. For the data in Zhang et al. (2021), we used 40-ms bins from 80 to 2000 ms.

3. Results

In Krasich et al. (2018), fixation durations were on average 342 ms ($SD = 251$ ms) for self-reported mind wandering and 318 ms ($SD = 205$ ms) for self-reported attentive viewing. In Zhang et al. (2021), fixation durations were on average 328 ms ($SD = 238$ ms) for self-reported mind wandering and 297 ms ($SD = 199$ ms) for self-reported attentive viewing.

The maximum-likelihood set of parameters for the model are presented in Table 1. Histograms of the actual and model-simulated fixation durations for self-reported mind wandering and self-reported attentive viewing are illustrated in Fig. 3. To highlight the differences between the fixation duration distributions in greater detail, Fig. 4 depicts the difference in the likelihood of fixation durations between self-reported mind wandering and attentive viewing for actual (red) and full-model-simulated (blue) fixation durations. Specifically, each line was generated by (1) creating histograms of fixation durations with a bin size of 0.1 ms, (2) subtracting the probability density within each bin between mind wandering and attentive viewing, and then (3) smoothing those differences using linear locally estimated scatterplot smoothing (LOESS) with a span of 0.033. In this plot, positive values indicate that there were more fixations of a given duration preceding reports of mind wandering compared to attentive viewing, and negative values indicate that there were more fixations of that duration preceding reports of attentive viewing compared to mind wandering.

Although the fit of the model was not exact (e.g., the model underestimated the likelihood of extremely long fixations >750 ms), the results showed that the model qualitatively captured the observed increase in mean fixation duration as well as differences in the shape of the distributions between mind wandering and attentive viewing. Specifically, in Krasich et al. (2018), mind wandering was associated with an

Table 1
Parameter values of the full and reduced models for Krasich et al. (2018) and Zhang et al. (2021).

Parameter	Symbol	Krasich et al. (2018)		Zhang et al. (2021)		Range
		Full	Reduced	Full	Reduced	
Threshold	N	15	13	18	18	[2,30]
Timer duration (ms)	t_{timer}	246	241	230	229	[150, 375]
Labile duration (ms)	t_{labile}	202	197	189	190	[100, 225]
Nonlabile duration (ms)	$t_{nonlabile}$	80	71	55	80	[25, 80]
Motor command (ms)	t_{motor}	30	30	30	30	fixed
Saccade execution (ms)	$t_{saccade}$	40	40	40	40	fixed
Rate modulation	β_{rate}	1 (0%)	0.98 (−2%)	1 (0%)	0.97 (−3%)	[0.25, 1]
Threshold modulation	$\beta_{threshold}$	0.65 (−35%)	1 (0%; fixed)	0.68 (−32%)	1 (0%; fixed)	[0.25, 1]

Note. The threshold denotes the number of states of a random walk. In a given simulation, the threshold was held constant for all five random walks involved in the fixation-generating process. The duration parameters represent the mean durations of the random walks. The rate modulation and threshold modulation parameters are multiplicative factors representing the reduction in the state transition rate and the threshold, respectively. The modulation parameters, which were applied to the timer and the labile and nonlabile stages of saccade programming, are reported as proportions with percentage changes in parentheses. During threshold modulation, the new threshold was rounded to the nearest integer. The Range column indicates the allowable parameter values explored during model optimization.

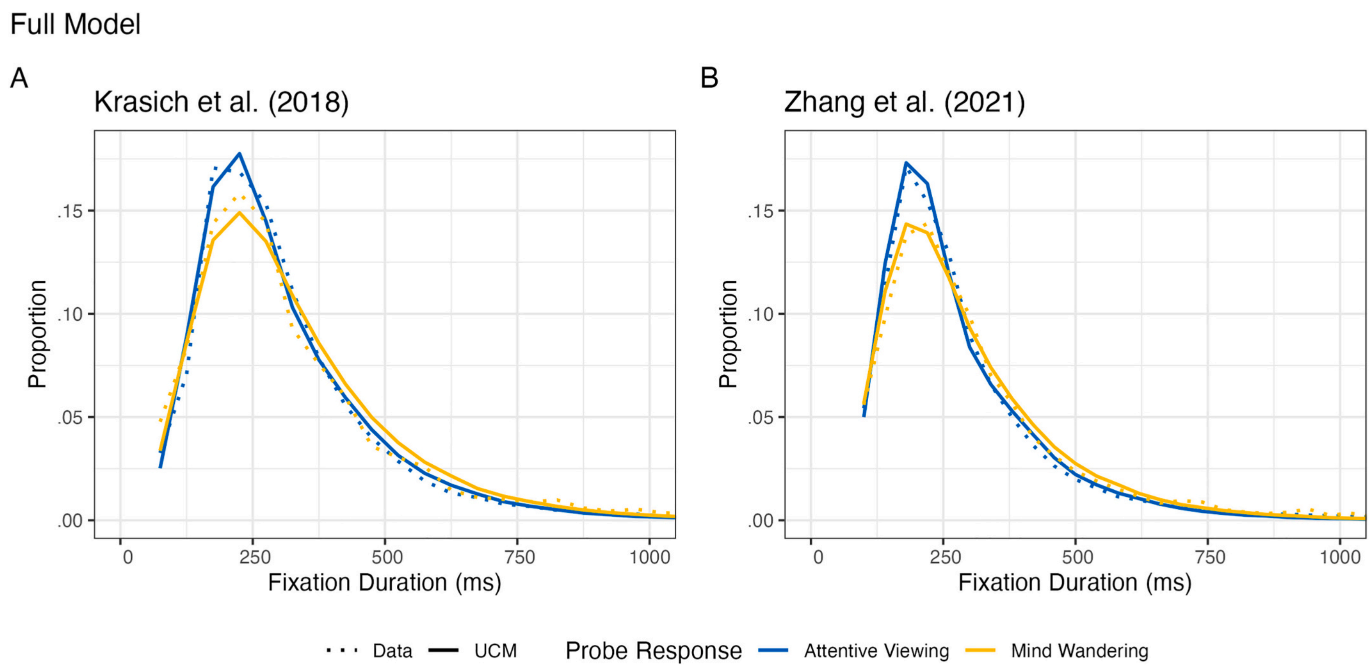

Fig. 3. Distributions of the observed (dotted lines) and the simulated (solid lines) fixation durations preceding reports of attentive viewing (blue) and mind wandering (yellow). The simulated data were generated by the full model that allowed for both rate and threshold modulation. (A) Data and model predictions for Krasich et al. (2018). (B) Data and model predictions for Zhang et al. (2021). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

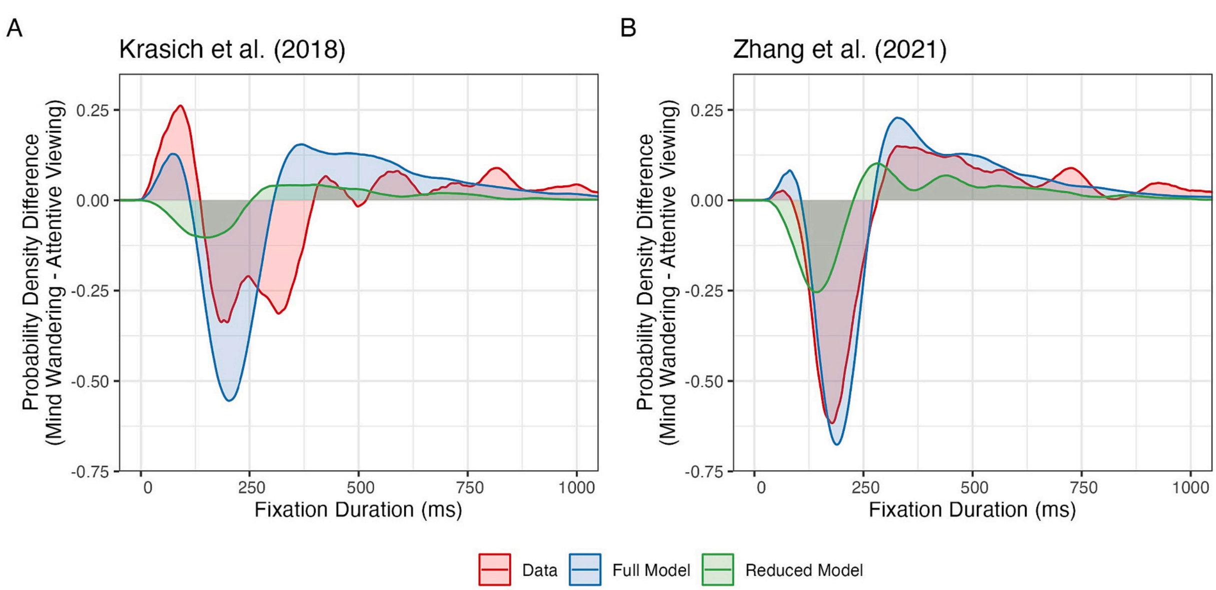

Fig. 4. Difference in the probability density of fixation durations preceding reports of mind wandering compared to attentive viewing for the actual data (red), simulated fixation durations from the full model (blue), and simulated fixation durations from the reduced model (green). Positive values indicate the presence of more fixations with a given duration preceding reports of mind wandering compared to attentive viewing, and negative values indicate the presence of more fixations with a given duration preceding reports of attentive viewing compared to mind wandering. (A) Data and model predictions for Krasich et al. (2018). (B) Data and model predictions for Zhang et al. (2021). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

increase in both shorter ($<$ 100 ms) and longer ($>$ 400 ms) durations but a decrease in intermediate-length durations (~250 ms) compared to attentive viewing. The full model largely reproduced this pattern, albeit it differed in the exact location and size of these effects. Similarly, in Zhang et al. (2021), mind wandering was associated with an increase in longer ($>$ 300 ms) fixations and a decrease in intermediate-length fixations (~200 ms) compared to attentive viewing. Here, the full model provided a close fit to the data. Overall, in both cases, the full model qualitatively reproduced the differences between distributions of fixation duration preceding reports of mind wandering and attentive viewing.

We next investigated whether the model accounted for the difference in the fixation duration distributions between self-reported mind wandering and self-reported attentive viewing as a difference in (a) rate modulation, as is seen during bouts of processing difficulties (Walshe & Nuthmann, 2021) or (b) threshold modulation and saccade cancellations. For both datasets, the full model estimated no (i.e., a 0%) reduction in the rate of state transitions (Table 1). In contrast, the threshold was reduced by 35% in Krasich et al. (2018) and by 32% in Zhang et al. (2021) preceding reports of mind wandering compared to attentive viewing. This collectively suggests that threshold modulation, but not rate modulation, was necessary to account for the increase in mean fixation duration associated with mind wandering.

To verify that threshold modulation was indeed necessary to account for mind wandering-related fixation durations, we fit a reduced version of the model that included rate modulation, but not threshold modulation, and compared it to the full model using a likelihood ratio test (LRT). The simulated distributions for this reduced model are plotted in Fig. 5, and the difference between simulated distributions preceding reports of mind wandering and attentive viewing are depicted in green in Fig. 4. For both studies, the reduced model performed worse than the full model. Specifically, for Krasich et al. (2018), the reduced model predicted a small difference (i.e., a 2% reduction in transition rate) in fixation durations preceding reports of mind wandering compared to attentive viewing. According to the LRT, it provided a significantly worse fit than the full model, $\chi^2(1) = 47.83$, $p < .001$. For Zhang et al. (2021), the reduced model also captured a small difference (i.e., a 3% reduction in transition rate) in fixation durations preceding reports of mind wandering compared to attentive viewing, though it was likewise outperformed by the full model, $\chi^2(1) = 41.50$, $p < .001$. Overall, the reduced model exhibited two main deficiencies. First, it underestimated the size of the differences between distributions of fixation durations during attentive viewing and mind wandering. Moreover, the reduced model also failed to reproduce the increase in very short fixations associated with mind wandering primarily present in Krasich et al. (2018).

Last, we explored whether saccade cancellations in the model contributed to the longer fixation durations associated with mind wandering. We found that this was indeed the case: for Krasich et al. (2018), the model predicted a more extensive involvement of saccade cancellations preceding reports of mind wandering ($M = 33.1\%$) than attentive viewing ($M = 29.5\%$). Similarly for Zhang et al. (2021), the model estimated that saccade programs were cancelled more frequently preceding reports of mind wandering ($M = 31.9\%$) than attentive viewing ($M = 28.1\%$). To verify this prediction of the model, we derived the probability of saccade cancellation in the model and demonstrated that threshold modulation, but not rate modulation, affected this probability (see Supplementary Information).

4. General discussion

It has been long thought that increased processing demands in scene viewing correspond with longer fixation durations (Rayner, 1978, 1998, 2009). Research on mind wandering, though, presents a difficulty for inferring processing from fixation durations, given that mind wandering is a state of attenuated visual-cognitive processing but is also associated with longer fixation durations in scene viewing (Krasich et al., 2018; Zhang et al., 2021). Toward overcoming this challenge, we further investigated how mind wandering-related changes in fixation durations may emerge. We used an established random-walk computational model of saccade timing and programming (UCM, Walshe & Nuthmann, 2021) to make specific predictions about which model parameters would best account for differences in the fixation duration distributions between reports of mind wandering and reports of attentive viewing. We then fit the UCM to fixation durations measured in Krasich et al. (2018) and Zhang et al. (2021), testing which mechanism (i.e., rate or threshold modulation) would best account for the longer fixation durations associated with self-reported mind wandering compared to self-reported attentive viewing.

Our findings showed that differences in fixation duration distributions were best modeled as a decrease in the number of steps required for each random walk to reach threshold. This threshold modulation increased the variability in when the random walks reached threshold and, consequently, increased the number of saccade cancellations. These effects collectively resulted in a more variable fixation-generating process, which sometimes led to very *long* fixations that accounted for the increased mean durations associated with mind wandering observed in Krasich et al. (2018) and Zhang et al. (2021). Thus, the current findings

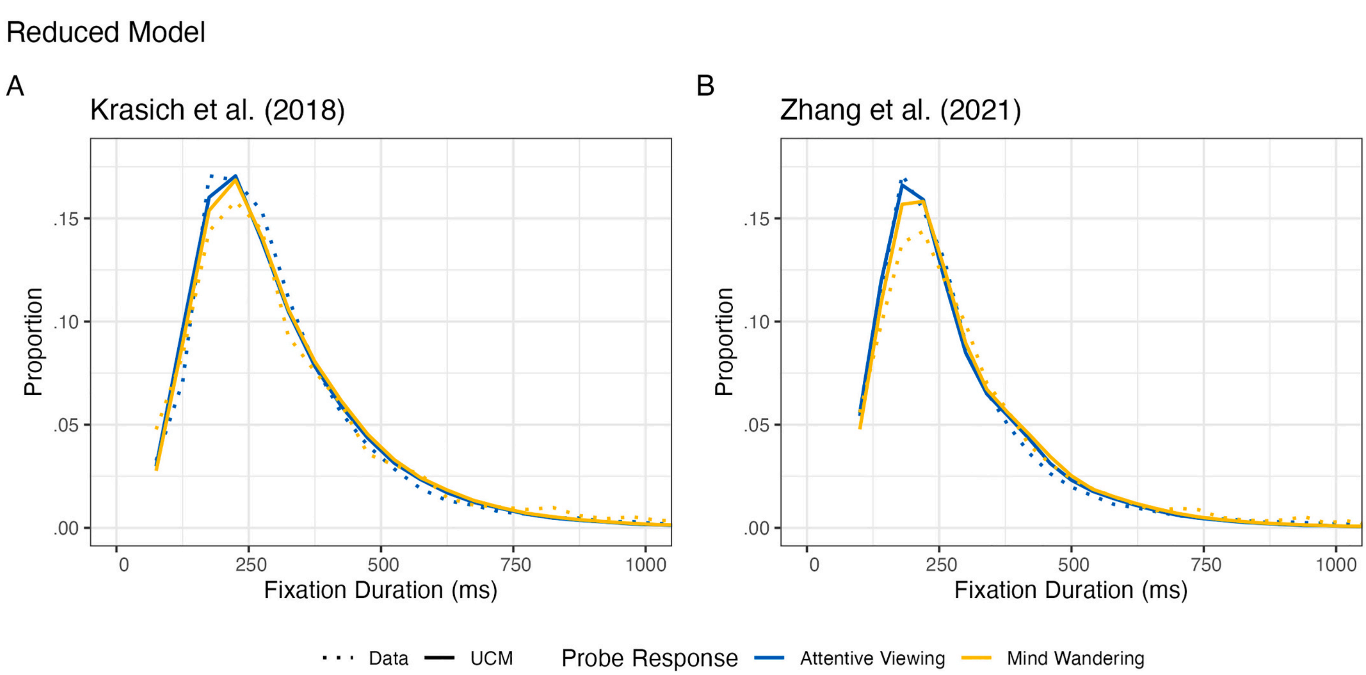

Fig. 5. Simulations with a reduced model in which rate modulation was applied, but not threshold modulation. Distributions of observed (dotted lines) and simulated (solid lines) fixation durations preceding reports of attentive viewing (blue) as opposed to mind wandering (yellow). (A) Data and model predictions for Krasich et al. (2018). (B) Data and model predictions for Zhang et al. (2021). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

demonstrated that threshold modulation was the mechanism by which fixation durations were on average longer for self-reported mind wandering than self-reported attentive viewing. Interestingly, though, the more variable fixation-generating process also sometimes resulted in an increase in very *short* fixations, especially in Krasich et al. (2018). This finding further indicated that mind wandering was not only associated with differences in the mean central tendency but also the spread and tails of the distribution. This highlights the importance of investigating differences in fixation distributions—not just measures of central tendency—for a more comprehensive view on how mind wandering impacts fixation durations and visual-cognitive processing more generally.

Importantly, mind wandering-related fixation durations were not well characterized by rate modulation even though this parameter was critical in past research for modeling increased fixation durations associated with increased visual-cognitive processing demands in scene viewing (Nuthmann et al., 2010; Walshe & Nuthmann, 2021). Thus, when considered concurrently with past research, our findings suggest that increased processing demands and mind wandering impact fixation durations through different mechanisms. That is, mind wandering does not seem to just engender a sort of 'visual-cognitive processing difficulty' as originally speculated in both Krasich et al. (2018) and Zhang et al. (2021). Instead, our results indicated that mind wandering was best interpreted as a state of increased variability in visual-cognitive processing (through threshold modulation), rather than as a global processing difficulty (through rate modulation).

This work suggests how the same behavioral outcome (increased fixation durations) can emerge from different states of visual-cognitive processing via disparate mechanisms. This poses a problem for the conventional view in scene viewing, which has long considered increased fixation durations to reflect increasingly complex and more difficult visual-cognitive processing rather than mind wandering. Specifically, visual-cognitive processing—and the underlying fixation-generating mechanisms—may not be easily inferred from mean increases in fixation durations alone. Instead, our work shows how opposing states of visual-cognitive processing may be distinguishable with the use of a computational model, such as the UCM, that considers differences in fixation duration distributions and the underlying mechanisms by which they emerge.

One remaining question is whether the observed changes in fixation duration can be used to determine the exact onset and dynamic progression of mind wandering over time (Christoff et al., 2016; Smallwood, 2013). We analyzed fixations that occurred 10 s (Zhang et al., 2021) or 15 s (Krasich et al., 2018) before the thought probes, which is where mind wandering-associated effects on fixation durations were observed in these studies. This is a standard approach for measuring correlates of mind wandering (Murray, Irving, & Krasich, 2022), although it is possible that there were dynamic fluctuations between attentive viewing and mind wandering preceding the thought probe (Christoff et al., 2016; Smallwood, 2013). If so, the analyzed time period in the current work could contain a mixture of mind wandering- and attentive viewing-related fixations, which, if true, would only reduce the observed differences between reported mind wandering and reported attentive viewing relative to their actual magnitude. Despite this possibility, though, we observed measurable differences in the fixations made prior to self-reported mind wandering and self-reported attentive viewing, which suggest that these self-reports reflected different mental states. A methodology that prevents cross-contamination of mind wandering- and attentive viewing-related fixations could potentially yield larger differences. Therefore, an important future question to ask is whether fixation durations can assist in tracking the temporal dynamics of mind wandering as well as the natural fluctuations between states of attentive viewing and mind wandering over time.

Further, the current work focused on global changes in fixation durations independently from the content being fixated. However, it is well-known that *local* scene features, such as visually salient and meaningful scene content, can also impact the location and duration of fixations, which are tightly linked (Einhäuser & Nuthmann, 2016; Nuthmann, 2017; Tatler et al., 2017). Moreover, mind wandering has been associated with a greater tendency to fixate on visually salient scene content (Krasich, Huffman, Faber, & Brockmole, 2020), although these findings varied with how image features were quantified (Krasich et al., 2020; Zhang, Anderson, & Miller, 2022). Thus, it is possible that what local image features and scene content were fixated during mind wandering could have in part impacted the more variable fixation-generating process. Whether the impact of mind wandering on fixation durations varied according to what was fixated has not yet been empirically tested and is thus an important avenue for future research. Moreover, future work could employ other computational models that rely on both the placement and timing of fixations to further investigate the mechanisms underlying the effects of mind wandering on gaze control (Kucharský, van Renswoude, Raijmakers, & Visser, 2021; Tatler et al., 2017).

In conclusion, the current work showed how mind wandering-related changes in fixation durations emerged in two separate scene-viewing tasks. As a result, the current work advances understanding of how the same behavioral phenomenon—an increase in mean fixation duration—can reflect opposing states of visual-cognitive processing. Considered collectively with past research, the current work has critical implications for research on gaze control in scene viewing, especially under normal viewing conditions when stimulus features are not directly manipulated. Specifically, if an increase in mean fixation duration is observed, visual-cognitive processing cannot be inferred—contra to the conventional view—without understanding the underlying mechanism by which this increase was generated. That is, did the longer fixation durations originate as a decrease in the rate of visual-cognitive processing (such as when viewing a degraded stimulus; Nuthmann et al., 2010; Walshe & Nuthmann, 2021) or as an increase in the variability of processing (such as during mind wandering)? The current work indicates that detailed computational modeling provides a way to distinguish between these two alternative possibilities.

Author statement

Kristina Krasich: conceptualization, investigation, data curation, writing-original draft, writing-review & editing, and project administration; **Kevin O'Neill:** methodology, software, formal analysis, visualization, investigation, writing-original draft, and writing-review & editing; **Samuel Murray:** conceptualization, writing-original draft, and writing-review & editing. **James R. Brockmole:** supervision, writing-original draft, and writing-review & editing; **Felipe De Brigard:** supervision, writing-original draft, writing-review & editing, and funding acquisition; **Antje Nuthmann:** methodology, software, formal analysis, writing-original draft, writing-review & editing, supervision.

Author note

The authors declare no conflicts of interest.

Data availability

All analysis scripts as well as data from Krasich et al. (2018) are available at https://osf.io/nuys7/. Data from Zhang et al. (2021) can be accessed at https://osf.io/6pj9m/ and downloaded by our main UCM.R script.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2023.105624.

References

Baird, B., Smallwood, J., Lutz, A., & Schooler, J. W. (2014). The decoupled mind: Mind wandering disrupts cortical phase-locking to perceptual events. *Journal of Cognitive Neuroscience, 26*(11), 2596–2607. https://doi.org/10.1162/jocn_a_00656

Balota, D. A., & Yap, M. J. (2011). Moving beyond the mean in studies of mental chronometry: The power of response time distributional analyses. *Current Directions in Psychological Science, 20*(3), 160–166. https://doi.org/10.1177/0963721411408885

Barron, E., Riby, L. M., Greer, J., & Smallwood, J. (2011). Absorbed in thought: The effect of mind wandering on the processing of relevant and irrelevant events. *Psychological Science, 22*(5), 596–601. https://doi.org/10.1177/0956797611404083

Becker, W. (1989). Metrics. In R. H. Wurtz, & M. E. Goldberg (Eds.), *The neurobiology of saccadic eye movements (pp. 13–67)*. Elsevier.

Becker, W. (1991). Saccades. In R. H. S. Carpenter (Ed.), *Eye movements (pp. 95–137)*. CRC Press.

Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision Research, 19*(9), 967–983. https://doi.org/10.1016/0042-6989(79)90222-0

Castelhano, M. S., Mack, M. L., & Henderson, J. M. (2009). Viewing task influences eye movement control during active scene perception. *Journal of Vision, 9*(3). https://doi.org/10.1167/9.3.6

Christoff, K., Irving, Z. C., Fox, K. C. R., Spreng, R. N., & Andrews-Hanna, J. R. (2016). Mind wandering as spontaneous thought: A dynamic framework. *Nature Reviews Neuroscience, 17*(11), 718–731. https://doi.org/10.1038/nrn.2016.113

Christoff, K., Mills, C., Andrews-Hanna, J. R., Irving, Z. C., Thompson, E., Fox, K. C. R., & Kam, J. W. Y. (2018). Mind wandering as a scientific concept: Cutting through the definitional haze. *Trends in Cognitive Sciences, 22*(11), 957–959. https://doi.org/10.1016/j.tics.2018.07.004

Coco, M. I., Nuthmann, A., & Dimigen, O. (2020). Fixation-related brain potentials during semantic integration of object–scene information. *Journal of Cognitive Neuroscience, 32*(4), 571–589. https://doi.org/10.1162/jocn_a_01504

Cranmer, K., Brehmer, J., & Louppe, G. (2020). The frontier of simulation-based inference. *Proceedings of the National Academy of Sciences, 117*(48), 30055–30062. https://doi.org/10.1073/pnas.1912789117

Devillez, H., Guyader, N., Curran, T., & O'Reilly, R. C. (2020). The bimodality of saccade duration during the exploration of visual scenes. *Visual Cognition, 28*(9), 484–512. https://doi.org/10.1080/13506285.2020.1830325

Einhäuser, W., & Nuthmann, A. (2016). Salient in space, salient in time: Fixation probability predicts fixation duration during natural scene viewing. *Journal of Vision, 16*(11). https://doi.org/10.1167/16.11.13

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112*(4), 777–813. https://doi.org/10.1037/0033-295X.112.4.777

Fox, K. C. R., Spreng, R. N., Ellamil, M., Andrews-Hanna, J. R., & Christoff, K. (2015). The wandering brain: Meta-analysis of functional neuroimaging studies of mind wandering and related spontaneous thought processes. *Neuroimage, 111*, 611–621. https://doi.org/10.1016/j.neuroimage.2015.02.039

Glaholt, M. G., Rayner, K., & Reingold, E. M. (2013). Spatial frequency filtering and the direct control of fixation durations during scene viewing. *Attention, Perception, & Psychophysics, 75*(8), 1761–1773. https://doi.org/10.3758/s13414-013-0522-1

Henderson, J. M., & Choi, W. (2015). Neural correlates of fixation duration during real-world scene viewing: Evidence from fixation-related (FIRE) fMRI. *Journal of Cognitive Neuroscience, 27*(6), 1137–1145. https://doi.org/10.1162/jocn_a_00769

Henderson, J. M., Nuthmann, A., & Luke, S. G. (2013). Eye movement control during scene viewing: Immediate effects of scene luminance on fixation durations. *Journal of Experimental Psychology: Human Perception and Performance, 39*(2), 318–322. https://doi.org/10.1037/a0031224

Henderson, J. M., & Pierce, G. L. (2008). Eye movements during scene viewing: Evidence for mixed control of fixation durations. *Psychonomic Bulletin & Review, 15*(3), 566–573. https://doi.org/10.3758/PBR.15.3.566

Henderson, J. M., Weeks, P. A., Jr., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance, 25*(1), 210–228. https://doi.org/10.1037//0096-1523.25.1.210

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology, 8*(4), 441–480. https://doi.org/10.1016/0010-0285(76)90015-3

Kam, J. W. Y., Helfrich, R. F., Solbakk, A.-K., Endestad, T., Larsson, P. G., Lin, J. J., & Knight, R. T. (2021). Top–down attentional modulation in human frontal cortex: Differential engagement during external and internal attention. *Cerebral Cortex, 31*(2), 873–883. https://doi.org/10.1093/cercor/bhaa262

Krasich, K., Huffman, G., Faber, M., & Brockmole, J. R. (2020). Where the eyes wander: The relationship between mind wandering and fixation allocation to visually salient and semantically informative static scene content. *Journal of Vision, 20*(9). https://doi.org/10.1167/jov.20.9.10

Krasich, K., McManus, R., Hutt, S., Faber, M., D'Mello, S. K., & Brockmole, J. R. (2018). Gaze-based signatures of mind wandering during real-world scene processing. *Journal of Experimental Psychology: General, 147*(8), 1111–1124. https://doi.org/10.1037/xge0000411

Kucharský, Š., van Renswoude, D., Raijmakers, M., & Visser, I. (2021). WALD-EM: Wald accumulation for locations and durations of eye movements. *Psychological Review, 128*(4), 667–689. https://doi.org/10.1037/rev0000292

Lange, E. B., Pieczykolan, A., Trukenbrod, H. A., & Huestegge, L. (2018). The rhythm of cognition–effects of an auditory beat on oculomotor control in reading and sequential scanning. *Journal of Eye Movement Research, 11*(2). https://doi.org/10.16910/jemr.11.2.9

Loftus, G. R. (1985). Picture perception: Effects of luminance on available information and information-extraction rate. *Journal of Experimental Psychology: General, 114*(3), 342–356. https://doi.org/10.1037/0096-3445.114.3.342

Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance, 4*(4), 565–572. https://doi.org/10.1037//0096-1523.4.4.565

Ludwig, C. J. H., Mildinhall, J. W., & Gilchrist, I. D. (2007). A population coding account for systematic variation in saccadic dead time. *Journal of Neurophysiology, 97*(1), 795–805. https://doi.org/10.1152/jn.00652.2006

Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., & Macrae, C. N. (2007). Wandering minds: The default network and stimulus-independent thought. *Science, 315*(5810), 393–395. https://doi.org/10.1126/science.1131295

Mills, C., Raffaelli, Q., Irving, Z. C., Stan, D., & Christoff, K. (2018). Is an off-task mind a freely-moving mind? Examining the relationship between different dimensions of thought. *Consciousness and Cognition, 58*, 20–33. https://doi.org/10.1016/j.concog.2017.10.003

Murray, S., Irving, Z. C., & Krasich, K. (2022). The scientific study of passive thinking: The methodology of mind wandering research. In F. De Brigard, & W. Sinnott-Armstrong (Eds.), *Neuroscience and philosophy (pp. 389–426)*. MIT Press.

Nuthmann, A. (2017). Fixation durations in scene viewing: Modeling the effects of local image features, oculomotor parameters, and task. *Psychonomic Bulletin & Review, 24*(2), 370–392. https://doi.org/10.3758/s13423-016-1124-4

Nuthmann, A., & Engbert, R. (2009). Mindless reading revisited: An analysis based on the SWIFT model of eye-movement control. *Vision Research, 49*(3), 322–336. https://doi.org/10.1016/j.visres.2008.10.022

Nuthmann, A., Schütz, I., & Einhäuser, W. (2020). Salience-based object prioritization during active viewing of naturalistic scenes in young and older adults. *Scientific Reports, 10*. https://doi.org/10.1038/s41598-020-78203-7

Nuthmann, A., Smith, T. J., Engbert, R., & Henderson, J. M. (2010). CRISP: A computational model of fixation durations in scene viewing. *Psychological Review, 117*(2), 382–405. https://doi.org/10.1037/a0018924

Pannasch, S., Schulz, J., & Velichkovsky, B. M. (2011). On the control of visual fixation durations in free viewing of complex images. *Attention, Perception, & Psychophysics, 73*(4), 1120–1132. https://doi.org/10.3758/s13414-011-0090-1

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin, 85*(3), 618–660. https://doi.org/10.1037/0033-2909.85.3.618

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422. https://doi.org/10.1037//0033-2909.124.3.372

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62*(8), 1457–1506. https://doi.org/10.1080/17470210902816461

Russell, B. C., Torralba, A., Murphy, K. P., & Freeman, W. T. (2008). LabelMe: A database and web-based tool for image annotation. *International Journal of Computer Vision, 77*(1), 157–173. https://doi.org/10.1007/s11263-007-0090-8

Saez de Urabain, I. R., Nuthmann, A., Johnson, M. H., & Smith, T. J. (2017). Disentangling the mechanisms underlying infant fixation durations in scene perception: A computational account. *Vision Research, 134*, 43–59. https://doi.org/10.1016/j.visres.2016.10.015

Schad, D. J., & Engbert, R. (2012). The zoom lens of attention: Simulating shuffled versus normal text reading using the SWIFT model. *Visual Cognition, 20*(4–5), 391–421. https://doi.org/10.1080/13506285.2012.670143

Smallwood, J. (2013). Distinguishing how from why the mind wanders: A process–occurrence framework for self-generated mental activity. *Psychological Bulletin, 139*(3), 519–535. https://doi.org/10.1037/a0030010

Smallwood, J., Beach, E., Schooler, J. W., & Handy, T. C. (2008). Going AWOL in the brain: Mind wandering reduces cortical analysis of external events. *Journal of Cognitive Neuroscience, 20*(3), 458–469. https://doi.org/10.1162/jocn.2008.20037

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin, 132*(6), 946–958. https://doi.org/10.1037/0033-2909.132.6.946

Stawarczyk, D., Majerus, S., Maj, M., Van der Linden, M., & D'Argembeau, A. (2011). Mind wandering: Phenomenology and function as assessed with a novel experience sampling method. *Acta Psychologica, 136*(3), 370–381. https://doi.org/10.1016/j.actpsy.2011.01.002

Tatler, B. W., Brockmole, J. R., & Carpenter, R. H. S. (2017). LATEST: A model of saccadic decisions in space and time. *Psychological Review, 124*(3), 267–300. https://doi.org/10.1037/rev0000054

Trukenbrod, H. A., & Engbert, R. (2014). ICAT: A computational model for the adaptive control of fixation durations. *Psychonomic Bulletin & Review, 21*(4), 907–934. https://doi.org/10.3758/s13423-013-0575-0

Turnbull, A., Wang, H. T., Murphy, C., Ho, N. S. P., Wang, X., Sormaz, M., , … Margulies, D. S., et al. (2019). Left dorsolateral prefrontal cortex supports context-dependent prioritisation of off-task thought. *Nature Communications, 10*(1). https://doi.org/10.1038/s41467-019-11764-y

Ucar, I., Smeets, B., & Azcorra, A. (2019). simmer: Discrete-event simulation for R. *Journal of Statistical Software, 90*(2), 1–30. https://doi.org/10.18637/jss.v090.i02

Võ, M. L.-H., & Henderson, J. M. (2009). Does gravity matter? Effects of semantic and syntactic inconsistencies on the allocation of attention during scene perception. *Journal of Vision, 9*(3). https://doi.org/10.1167/9.3.24

Walshe, R. C., & Nuthmann, A. (2014). Asymmetrical control of fixation durations in scene viewing. *Vision Research, 100*, 38–46. https://doi.org/10.1016/j.visres.2014.03.012

Walshe, R. C., & Nuthmann, A. (2015). Mechanisms of saccadic decision making while encoding naturalistic scenes. *Journal of Vision, 15*(5). https://doi.org/10.1167/15.5.21

Walshe, R. C., & Nuthmann, A. (2021). A computational dual-process model of fixation-duration control in natural scene viewing. *Computational Brain & Behavior, 4*(4), 463–484. https://doi.org/10.1007/s42113-021-00111-4

Wang, H.-C., Hwang, A. D., & Pomplun, M. (2010). Object frequency and predictability effects on eye fixation durations in real-world scene viewing. *Journal of Eye Movement Research, 3*(3). https://doi.org/10.16910/jemr.3.3.3

von Wartburg, R., Ouerhani, N., Pflugshaupt, T., Nyffeler, T., Wurtz, P., Hügli, H., & Müri, R. M. (2005). The influence of colour on oculomotor behaviour during image perception. *Neuroreport, 16*(14), 1557–1560. https://doi.org/10.1097/01.wnr.0000180146.84020.c4

Xiao, J., Hays, J., Ehinger, K. A., Oliva, A., & Torralba, A. (2010). SUN database: Large-scale scene recognition from abbey to zoo. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 3485–3492. https://doi.org/10.1109/CVPR.2010.5539970

Zhang, H., Anderson, N. C., & Miller, K. F. (2021). Refixation patterns of mind wandering during real-world scene perception. *Journal of Experimental Psychology: Human Perception and Performance, 47*(1), 36–52. https://doi.org/10.1037/xhp0000877

Zhang, H., Anderson, N. C., & Miller, K. F. (2022). Scene meaningfulness guides eye movements even during mind wandering. *Attention, Perception, & Psychophysics, 84*(4), 1130–1150. https://doi.org/10.3758/s13414-021-02370-6