Consciousness and Cognition 140 (2026) 104031

Contents lists available at ScienceDirect

Consciousness and Cognition

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

Eye movements reflect thought patterns while listening to literary narratives

Diane Mézière[a,*], Johanna K. Kaakinen[a,b], Jarkko Lehtola[c], Karin Kukkonen[d], Jonathan Smallwood[e], Jaana Simola[f]

[a] Department of Psychology and Speech-Language Pathology, University of Turku, Finland
[b] INVEST Research Flagship and Research Center, University of Turku, Finland
[c] Rehabilitation Service Area, Department of Psychology, Turku University Hospital, Turku, Finland
[d] Department of Literature, Area Studies and European Languages, University of Oslo, Norway
[e] Department of Psychology, Queen's University, Kingston, Ontario, Canada
[f] Department of Education, University of Helsinki, Finland

ARTICLE INFO

Keywords:
Eye movements
Audiobook
Immersion
Mind-wandering
Narratives

ABSTRACT

While listening to an audiobook, listeners' attention may shift between the text and self-generated thoughts occurring during mind-wandering. Despite the growing use of audiobooks, little is known about how listeners process text when listening to it. The present study investigated the types of thought patterns that listeners have while listening to an audiobook, and whether and how these are reflected in eye movements. Participants ($N_{Study1} = 63$, $N_{Study2} = 58$) listened to an audio recording of a novel and responded to a 13-item mind-wandering questionnaire and a simple memory question 30 times during the listening task. In study 2, participants' eye movements were also recorded while they looked at a fixation cross on a screen. A principal components analysis (PCA) of the mind-wandering questionnaire responses produced four components in Study 1, and three components in Study 2. Three components were replicated across the two studies: Immersion, Mind-wandering, and Sub-vocalization. We then examined how these thought patterns were reflected in eye movements in Study 2 including: fixation duration, fixation count, fixation dispersion, saccade amplitude, blink duration, and blink count. The results showed that higher levels of immersion was characterized by fewer and less dispersed fixations on the screen, shorter saccades, and longer blinks. Mind-wandering was related to more dispersed fixations. Sub-vocalization resulted in more fixations, higher dispersion across the screen, and more blinks. The results suggest that eye movements reflect shifts in attentional focus while listening to a literary text. The results provide important information about the processes underlying literary experience.

1. Introduction

While reading or listening to a literary text, people frequently experience feeling transported into the story world (Gerrig, 1993; Green & Brock, 2000; Kuijpers et al., 2014; Ryan, 2001). This phenomenon of transportation, also called immersion (e.g., Ryan, 2001) or story absorption (Kuijpers et al., 2014), is typically associated with listeners focusing their attention to the content of the text, such as wondering what will happen next, empathizing with the characters, or wondering how a particularly bad event could have been avoided. As such, this immersion into the story world often includes vivid imagery (e.g., feeling like you're *there*), as well as stronger emotional content (e.g., suspense when wondering how the story will end; Appel et al., 2015). Immersion has also been shown to be associated with more positive emotions (Mézière et al., 2025), and to lead to better text comprehension (Lange et al., 2022). The focus of attention inward and on task-relevant thoughts (i.e., related to the text), which is characteristic of immersive experiences, is accompanied with a diminished attention to the external world such as not being able to keep track of time or respond to story-irrelevant stimuli.

On the other hand, it is also common for listeners to experience internal distractions such as mind-wandering, that is instances where the attention is not focused on the text content but elsewhere such as wondering what you will have for dinner or thinking about a

* Corresponding author.
E-mail address: diane.meziere@utu.fi (D. Mézière).

https://doi.org/10.1016/j.concog.2026.104031
Received 17 December 2025; Received in revised form 24 February 2026; Accepted 24 February 2026

conversation that happened earlier. Such mind-wandering episodes are extremely common as research suggest they may occur as often as 68% of the time while listening to an audiobook (Faber et al., 2020). In these episodes, attention is decoupled from the stimuli (i.e., the text), which has been associated with poorer task performance such as poorer text comprehension during both silent reading and listening to text (Varao Sousa et al., 2013; 2018).

Most research carried out on immersion and mind-wandering during the processing of literary text has been done on reading (e.g., Eekhof et al., 2021; Lei et al., 2023; Mézière et al., 2025), and very little is yet known about the types of thoughts readers have while they listen to a literary text like an audiobook (however see, Kosch et al., 2024; Polychroni et al., 2022). The present study therefore aims to identify the kinds of thoughts readers have while listening to a novel (see Mulholland et al., 2023) and uses eye-tracking to examine how these thought patterns are reflected in eye movement behavior. This will provide information on how focusing of attention on internal, self-generated thoughts vs. text stimuli during a listening task are coupled with visual processing.

1.1. Attention while listening to narratives

In the present study, we will use the term *immersion* when referring to subjective experiences that some authors have called transportation (Gerrig, 1993; Green & Brock, 2000) or story absorption (Kuijpers et al., 2014). When a reader or listener is immersed into the story world, their attention is fully focused on the content of the text such that they may express a feeling "like being *there*" and they typically experience strong emotions and vivid imagery (Gerrig, 1993; Green & Brock, 2000; Kuijpers et al., 2014; Ryan, 2001). The listener may thus have vivid image representations of the characters or events, empathize with the characters, wonder how an unpleasant turn of event could have been avoided, or feel suspense about how the story is going to end. Critically, all these thoughts are triggered by and related to the events and the story world that are being described and are characterized by focused attention on the text content. Previous work suggests that immersion is indeed characterized by increased attention to the story and disengaged attention from the external environment. For example, Bezdek and Gerrig (2017) found that when watching narrative film clips, participants were slower to respond to an unrelated sound probe when they were immersed, thus indicating their lack of attention to stimuli outside of the narrative content. Using auditory narratives, Ballenghein et al. (2023) found that emotionally arousing story content, specifically erotica, induced immersion and captured listeners' attention more than neutral story content, as indicated by a steeper reduction in postural sway during the listening task. Such focused attention has also been linked to positive emotions such as enjoyment of the text and has been shown to be positively associated with text comprehension (Lange et al. 2022), possibly due to this higher attentional focus on the content of the text leading to a more detailed and accurate mental representation of the text content.

On the other hand, another common experience while listening to audiobooks is mind-wandering, in which the attention is focused on internal, text-unrelated thoughts. During mind-wandering, attention is decoupled from the text content, which typically leads to worse performance on the task at hand (Bonifacci et al., 2023; Varao Sousa et al., 2013; Varao-Sousa et al., 2018). According to the cascade model of inattention (Smallwood, 2011), this lack of attention on the text also leads to a poorer representation of the text content, which explains the poorer performance on text comprehension measures.

Hence, while little is yet understood about the types of thoughts that listeners experience while listening to an audiobook, previous research posits that they are likely to experience at least two types of thoughts, namely immersion and mind-wandering. Immersive episodes should be characterized by attention focused on the text content, as well as higher enjoyment, stronger emotions, and possibly better comprehension of the text. On the other hand, mind-wandering should be characterized by thoughts unrelated to the task at hand and lead to poorer comprehension outcomes. One way to investigate these differences in cognitive processes such as where the attention is allocated is through the use of eye-tracking.

1.2. Eye-movement behaviour while listening to narratives

Attention and thought patterns experienced while processing literary narratives have primarily been studied in the context of reading. As of today, very little work has been done looking at the relationship between eye-movement behaviour and literary text processing when listening to the text rather than reading it. Hence, little is known about what readers experience while listening to audiobooks (Kosch et al., 2024). Nevertheless, a handful of studies have investigated the relationship between eye movements as indicators of attention and common thought patterns experienced by readers such as immersion or mind-wandering.

In a study on listening of literary narratives, Kaakinen and Simola (2020) examined the relationship between immersion and eye-tracking measures reflecting engagement namely pupil size and blink rates. In their study, participants listened to extracts from Stephen King short stories that either contained horrific or neutral events. They found that participants experienced more immersion while listening to the horror extracts compared to the neutral ones, and that immersion was associated with lower blinking rates and stronger pupil size fluctuation reflecting higher engagement with the text. This suggests that immersion is characterized by higher attention on the text content, and that eye-movement measures do reflect these differences in cognitive engagement between immersive and non-immersive episodes while listening to text.

In a recent study, Lange and colleagues (2022) examined the relationship between immersion and and eye-movement behaviour while listening to short extracts from a range of literary works including fiction and poems. They found that higher levels of immersion into the narrative was associated with lower blinking rates, which reflected the idea that immersion is linked to higher attention to the narrative (see also Bezdek & Gerrig, 2017). The results suggests that eye-movement measures reflecting attentional focus such as blink rates may reflect participants' immersive experiences while they are listening to literary texts.

Another commonly-experienced phenomenon while listening to literature is mind-wandering, during which attention is decoupled from the external stimuli or in this case the text. Faber et al. (2020) explored the relationship between eye-movement behaviour and mind-wandering in a variety of contexts including listening to an audiobook. In this condition, participants listened to a passage from Walden by Henry Thoreau while fixating a fixation cross on a screen. They found that mind-wandering was extremely common during this task (68.19% of probes) and was associated with making shorter fixations compared to instances when the participants' thoughts were on the text. This result also suggests that thought patterns such as mind-wandering are reflected in eye-movement behaviour, which indicates how visual attention is allocated while listening to a literary text. This is particularly interesting because comprehending the text during audiobook listening does not require visual processing (i.e., sampling of the visual environment) in the same way as reading does (Faber et al., 2018).

Taken together, results from previous work suggest that readers experience varying types of thought patterns including immersion and mind-wandering while listening to literary narratives such as audiobooks. In addition, results suggest that eye movements can reflect differences in thought patterns while listening to literary texts, even though listening does not require visual processing. Yet little is known about how common thought patterns such as immersion and mind-wandering are reflected in eye-movement behaviour while processing literary texts. The primary aims of this study are thus: 1) to examine the types of thoughts listeners have while listening to literary texts; and 2) to examine the relationship between these thought patterns and eye-movement behaviour while listening to literary texts.

2. Aims and research questions

This article investigates the types of thoughts that listeners have while listening to literary texts such as an audiobook (Study 1), as well as how these thoughts are reflected in eye-movement behaviour during listening (Study 2). These two aims are linked to four research questions:

1) What kind of thought patterns do listeners experience while listening to a literary narrative (i.e., audiobook)?

2) Does text content, and specifically, valence (positive, neutral, negative) impact the type of thoughts that listeners experience?

3) Do thought patterns influence listeners' comprehension of the text?

4) How are these thought patterns reflected in eye movement behaviour while listening to literary narratives?

Given the limited number of previous studies on this topic, we cannot make strong predictions about the outcome of the study. However, it is probable that at least some of the thought patterns found in previous studies on reading such as immersion (Mézière et al., 2025) and mind-wandering (Varao-Sousa et al., 2018) will also be experienced by participants while listening to the text (Aim 1, RQ1. Similarly, we predict that some thought patterns will be associated with text valence, as previous studies have shown that immersion tends to be associated more with positive texts (RQ2; Ballenghein et al., 2023; Mézière et al., 2025). We also expect that some thought patterns will be associated with listening comprehension, as mind-wandering has been shown to negatively impact text comprehension in listening studies (RQ3; Varao-Sousa et al., 2018). Lastly, we expect that eye-movement behaviour will reflect differences in attention during listening associated with different thought types (Aim 2; RQ4). As only a few studies have been conducted on this topic, our analyses remain exploratory in nature.

3. Study 1: Online experiment

3.1. Method

3.1.1. Participants

Participants were recruited via student email lists, social media, and word of mouth. In total 105 participants started the experiment; of these, 63 adult (9 males, 2 non-binary, age range 19–47 years, mean age 24.5 years) native speakers of Finnish completed the experiment. The sample size was based on previous work using similar materials in a reading experiment (Mézière et al., 2025). Participants received either a course credit or a complimentary one-month subscription to an audiobook service. All participants gave informed consent prior to starting the experiment.

3.1.2. Text materials and comprehension questions

The text materials consisted of parts of the novel "Memories of the future" written by Siri Hustvedt (2019) and translated to Finnish by Kristiina Rikman. This novel was chosen because it is recent and it contains passages that vary in emotional content (negative, positive, and neutral), which can be thought to induce different types of thoughts (e.g., immersion and mind-wandering). It has also been used in previous research on thought patterns and eye movements during reading (Mézière et al., 2025). A pretest was used to assess the emotional valence and arousal induced by text paragraphs during reading. Pretest participants (N = 13) read the text paragraphs one at a time and rated valence and arousal with self-assessment manikins (SAM; Bradley & Lang, 1994). Based on the pretest, 30 paragraphs (10 neutral, 10 positive, and 10 negative) were selected as target paragraphs (text versions of the target paragraphs are available in OSF, https://osf.io/u3qhj/overview). The descriptive statistics of the valence, arousal, and paragraph length are presented in Table 1. Simple comprehension questions tapping into the memory for the exact contents of the target paragraphs were created (e.g. "Was the protagonist happy about the move?"). Participants responded to the questions by selecting either a "yes" or a "no" response. An audio recording of the text was created by recording a professional voice actor (female) read the text aloud. The total duration of the recording was 1 h 44 min.

3.1.3. Multi-dimensional experience sampling questionnaire

Participants' thoughts during the listening task were probed with the Multi-Dimensional Experience Sampling Questionnaire (Mézière et al., 2025; Turnbull et al., 2019). The questionnaire contained 13 items, and the first item presented to participants always queried whether the participant was focused on the task. The rest of the 12 items were presented in random order. Other items asked about the content of the thoughts including when (i.e., past, future), and who they were about (i.e., the participant, other people) as well as the form of thoughts (i.e., images or words). Items also included asking how vivid and detailed the thoughts were, whether they were recurrent in their theme(s), evolving (i.e., in steps), and spontaneous, as well as the emotional valence of the thoughts (i.e., positive or negative). Responses were given on a four-point Likert scale where 1 = not at all and 4 = completely, except for items tapping into emotions (1 = negative emotion, 4 = positive emotion) and spontaneity of the thoughts (1 = spontaneous, 4 = voluntary). Instructions for replying to the prompts and the prompts themselves are provided on OSF (https://osf.io/u3qhj/overview).

3.1.4. The transportation scale – Short form

Transportation to the story was measured with the Transportation Scale – Short Form (TS-SF; Appel et al., 2015). The scale contained five items tapping into the imagery, emotional experience, and engagement with the narrative, and responses were given on a 7-point Likert scale (1 = not at all, 7 = completely). The scale is available in OSF (https://osf.io/u3qhj/overview). An average score was calculated across the items, Cronbach α= 0.84.

3.1.5. Pleasure of listening

Participants rated the pleasantness of the audiobook listening experience with the SAM valence scale (Bradley & Lang, 1994). Participants were presented with nine pictures depicting a range of emotional experiences from extremely negative to extremely positive, and they were asked to choose the one that best described their listening experience. The responses were coded on a scale 1–9, where 1 = extremely negative and 9 = extremely positive.

Table 1
Descriptive statistics for valence, arousal, and length of the target paragraphs.

Paragraph	Valence[a] *M*	*SD*	Arousal[b] *M*	*SD*	Words *M*	*SD*	Sentences *M*	*SD*	Duration (s) *Mean*	*SD*
Positive	6.21	0.21	2.89	0.67	84.9	40.24	5.2	3.68	50.4	18.4
Neutral	5.00	0.18	2.85	0.18	89.7	33.77	5.5	2.51	45.1	16.2
Negative	3.74	0.54	3.5	0.37	98.5	24.74	8.2	5.96	42.3	12.1
All	4.98	1.08	3.08	0.53	91.03	32.85	6.3	4.36	45.9	16.1

[a] scale 1–9, where 1 = extremely negative and 9 = extremely positive. [b] scale 1–9, where 1 = extremely calm and 9 = extremely aroused.

3.1.6. Procedure

The study protocol was reviewed by the Ethics Committee for Human Sciences at the University of Turku. The experiment was programmed with the Gorilla Experiment Builder, and the data were collected on the Gorilla online platform (www.gorilla.sc; Anwyl-Irvine et al., 2020). Participants could complete the experiment on their own device (computer or a mobile device) at a time that best suited them.

After giving informed consent, participants were familiarized with the mDES questions, followed by the listening task. Participants were instructed that they would be listening to parts of a novel, and that they would occasionally be asked to respond to the mDES questionnaire and a simple comprehension question regarding the text. During the listening task, participants were probed with the mDES and the comprehension question regarding the previous paragraph after each of the 30 target paragraphs. A short break was offered in the middle of the experiment. In the end of the experiment, participants filled in the TS-SF and responded to the Pleasure of Listening question, and finally responded to background questions concerning their age, gender, first language, prior familiarity of the novel, and the number of fiction and non-fiction books they had read or listened within the previous year. All data, study materials and analysis code are available at OSF (https://osf.io/u3qhj/overview).

3.1.7. Data pre-processing and analysis

All analyses were carried out in R statistical software (R Core Team, 2023). To identify participants' thought patterns during the experiment, we conducted a *principal component analysis* (PCA) using the '*psych*' package with varimax rotation (Revelle, 2024). The number of components was determined through parallel analysis using the '*fa.parallel*' function and visualizing scree plots. The PCA was conducted on participants' responses to the 13 items of the mDES scale (1890 observations). We examined the reliability of the components by randomly splitting the dataset in half and running the PCA on each split-half again. These split-half scores were then compared to the original component scores by running Pearson correlations, with higher correlation coefficients indicating higher component reliability (see Mézière et al., 2025; Mulholland et al., 2023 for similar methods of estimating component reliability). The component scores extracted from the PCA were used in the subsequent analyses.

To examine the relationship between the components and transportation, we ran correlations using the '*stats*' package (R Core Team, 2023). As we only had one transportation and pleasure of listening score per participant, we first average participants' scores for each component identified in the PCA to obtain one score per component per participant to run the correlations. We ran Pearsons's correlations between the averaged components scores and participants' responses to the transportation scale, as well as their rating of the pleasure of listening.

To investigate the influence of text valence on participants' thought patterns, we ran linear regression models within the Bayesian framework using the '*brms*' package (Bürkner, 2017; Bürkner 2018; Bürkner,2021) and making inferences based on the 95% credible interval. We used the Bayesian framework as it puts the focus on the size and degree of certainty of an effect rather than whether or it not is significant. We ran a separate model for each PCA component as our outcome variable. In these models, we included text valence as a fixed effect using neutral texts as our baseline, and random intercepts for participants and items. The final model was Component ~ Valence + (1|Participant) + (1|Item). Post-hoc contrasts between positive, neutral, and negative texts were examined using the '*emmeans*' package (Lenth, 2024).

Lastly, we examined the relationship between participants' thought patterns and their text comprehension accuracy using linear regression within the Bayesian framework and making inferences based on the 95% credible interval. We included comprehension accuracy as our outcome variable, which was treated as a binary variable (correct = 1, incorrect = 0), and the component scores from all PCA components as fixed effects. In addition, we included random intercepts for participants and items. The final model was Accuracy ~ Component1 + Component2 + Component3 + Component4 + (1|Participant) + (1|Item).

3.2. Results

3.2.1. Thought patterns during listening

Results of the PCA identified four components based on the parallel analysis, explaining 55% of the variance in participants' responses to the mDES. The loadings for these components are shown in Table 2. The first component ("immersion") was associated with on-task thoughts, and characterized by specific thoughts, vivid imagery, and generally positive emotions. The second component ("mind-wandering") was characterized by off-task thoughts about the future and the participant themselves and tended to be evolving and habitual. Component 3 ("sub-vocalization") was characterized by deliberate verbal thoughts. Finally, component 4 ("social episodic thoughts"), was characterized by thoughts about the past and other people. These components are illustrated in word clouds in Fig. 1.

Results from the correlations between components scores calculated from the full dataset and scores calculated from the split-half datasets indicated very good component reliability with an average correlation coefficient of 0.96. Reliability was particularly high for the immersion, mind-wandering, and social episodic thoughts components with coefficients of 0.96, 0.96, and 0.97 respectively. Reliability was only a little lower for the sub-vocalizing component with a coefficient of 0.89.

Correlations between participants' responses to the transportation scale and their averaged component scores indicated that immersion was positively correlated with transportation scores ($rs = 0.50$, $p < 0.001$), suggesting that this component is indicative of immersive episodes. On the other hand, transportation was not significantly correlated with any of the other components (mind-wandering: $rs = -0.16$, $p = 0.21$; sub-vocalization $rs = 0.03$, $p = 0.83$; social episodic thoughts: $rs = 0.14$, $p = 0.26$).

Similarly, correlations between participants' mean component scores and their pleasantness ratings showed that pleasantness was positively correlated with immersion ($rs = 0.39$, $p < 0.002$), but was not correlated with any of the other components (mind-wandering: $rs = -0.11$, $p = 0.39$; sub-vocalization: $rs = 0.10$, $p = 0.45$; social episodic thoughts: $rs = 0.13$, $p = 0.31$).

3.2.2. Text valence and thought patterns

The means and standard deviations of component scores for each text valence condition are shown in Table 3. Results showed that immersion

Table 2
Study 1: Loadings of PCA Components.

	Component 1	Component 2	Component 3	Component 4
Q1: Task	**0.47**	**−0.46**	0.19	0.00
Q2: Future	−0.01	**0.73**	−0.03	−0.12
Q3: Past	0.04	0.15	0.02	**0.76**
Q4: Self	0.01	**0.79**	0.05	−0.02
Q5: Other people	0.20	−0.05	−0.07	**0.71**
Q6: Emotion	**0.53**	−0.00	−0.24	−0.27
Q7: Images	**0.78**	−0.05	−0.11	0.13
Q8: Words	−0.20	0.37	**0.56**	0.06
Q9: Vividness	**0.77**	−0.05	0.16	0.22
Q10: Specific	**0.64**	−0.01	0.38	0.23
Q11: Habit	−0.01	**0.65**	0.05	0.26
Q12: Evolving	−0.21	**0.53**	−0.39	0.21
Q13: Deliberate	0.15	−0.19	**0.70**	−0.08
SS Loadings	2.25	2.27	1.24	1.42
Proportion of Variance	17.3%	17.4%	9.5%	10.9%
Cumulative Variance	17.3%	34.7%	44.2%	55.2%

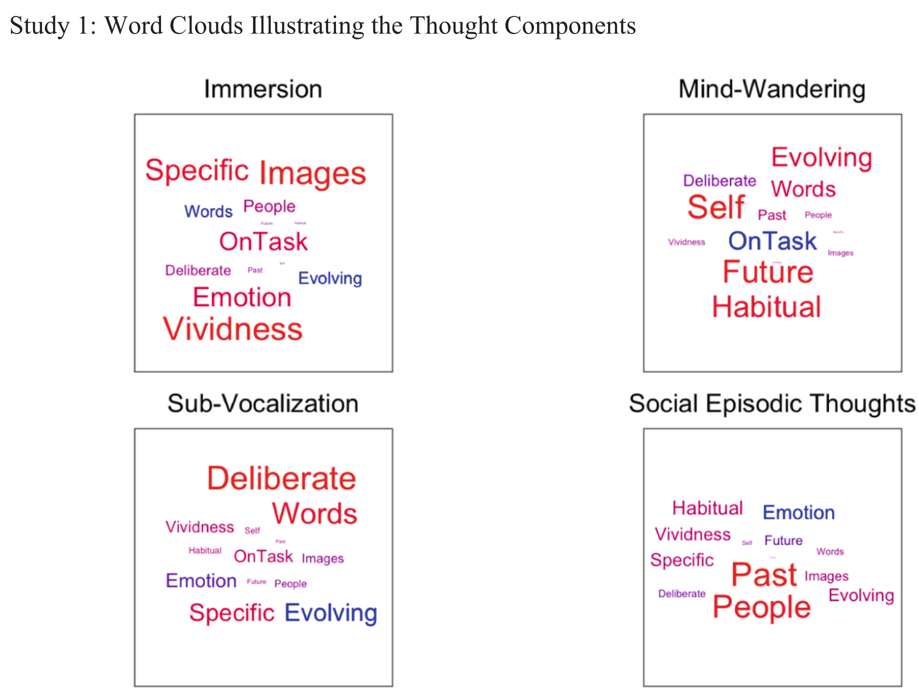

Fig. 1. Study 1: Word Clouds Illustrating the Thought Components. Note. Fig. 1 shows word clouds for the four PCA components found in Study 1. Words in red indicate a positive loading onto the component, and words in blue indicate negative loadings. Word in a larger font indicate higher loadings compared to words in a smaller font. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Study 1: Mean and Standard Deviation of Component Scores per Text Valence.

Valence	Immersion		Mind-Wandering		Sub-Vocalization		Social Episodic Thoughts	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Neutral	−0.07	0.98	0.07	0.98	−0.03	1.01	−0.03	0.97
Negative	−0.12	1.04	−0.06	1.01	0.04	1.01	0.13	0.98
Positive	0.19	0.95	−0.00	1.01	−0.02	0.98	−0.10	1.04

scores were lower for both neutral (b = 0.25, CI = 0.49 – 0.02) and negative (b = 0.30, CI = 0.06 – 0.54) texts compared to positive texts. There was no significant difference between neutral and negative texts. There were no effects of text valence on the mind-wandering or the sub-vocalization component. Lastly, scores for the 'social episodic thoughts' component were higher for negative texts compared to positive (b = 0.23, CI = 0.39 – 0.10) and neutral texts (b = 0.16, CI = 0.31—0.02). There was no difference in the occurrence of social episodic thoughts between positive and neutral texts.

3.2.3. Thought patterns and comprehension

Participants generally did well on the comprehension questions with a 75% accuracy on average (range 50% – 100%). The output of the model exploring the relationship between component scores and their comprehension accuracy is shown in Table 4. The results show that higher immersion was associated with higher probability of answering the comprehension question accurately (b = 0.22, CI = 0.07 – 0 36). On the other hand, mind-wandering was associated with lower probabilities of answering the comprehension question correctly (b = -0.23, CI = −0.37 – 0.09). Scores on the sub-vocalizing and social episodic thoughts components were not associated with comprehension accuracy.

Table 4
Study 1: Effect of Component Scores on Comprehension Accuracy.

Predictors	Estimate	CI (95%)
Intercept	1.40	0.96 – 1.85
Immersion	**0.22**	**0.07 – 0.36**
Mind-Wandering	**−0.23**	**−0.37 – −0.09**
Sub-Vocalization	0.11	−0.01 – 0.24
Social Episodic Thoughts	0.09	−0.05 – 0.23

Note. Table 4 shows the estimates and credible intervals (CI) for the relationship between participants' comprehension accuracy and scores on the four components.

4. Study 2: Laboratory experiment

4.1. Method

4.1.1. Participants

Participants were recruited via student email lists and word of mouth. In total 58 participants participated in the experiment (8 males, age range 19–61 years, mean age 25.5 years). The sample size was based on another study with similar materials focusing on reading (Mézière et al., 2025). All participants were native speakers of Finnish. Participants received either a course credit or a 10€ gift card for participating. All participants gave informed consent prior to starting the experiment.

4.1.2. Text materials and comprehension questions

The text materials and comprehension questions were identical to the ones used in Study 1.

4.1.3. Multi-dimensional experience sampling questionnaire

Participants' thoughts during the listening task were probed with the Multi-Dimensional Experience Sampling Questionnaire (Turnbull et al., 2019), similarly to Study 1.

4.1.4. The transportation scale – Short form

The Transportation Scale questionnaire was used to measure participants' transportation to the story. The scale is identical as the one used in Study 1. An average score was calculated across the items, Cronbach α= 0.81.

4.1.5. Equipment

The eye-tracking data were collected monocularly using a EyeLink 1000+ (SR Research), with a sampling rate of 500 Hz. Eye-tracking features including fixations and saccades were computed online with the EyeLink algorithm for fixation detection using the cognitive research settings (see EyeLink 1000 + manual for details). Participants were also fitted with a target sticker on their forehead which is used by the eye-tracker to find their head and pupil. EEG data were also collected simultaneously with a NeurOne system (Bittium Bisignals Oy) with a 64-channel cap, but these results will be reported elsewhere. During the listening task, a fixation cross was presented on a grey background (RGB: 254, 254, 254) at the center of the screen on a 21″ computer screen with 1920 x 1080 resolution and 144 Hz. The screen was placed at a viewing distance of about 70 cm from the chair.

4.1.6. Procedure

The experiment took place in a dimly-lit and sound-proof room. Participants were first sat in a comfortable chair and fitted with a 64-channel EEG cap. Once the EEG cap was set up, participants were given headphones, and the sound was tested to ensure a comfortable listening volume. Then, to collect the eye-tracking data, participants were fitted with a target on their forehead prior to calibration and data collection as no headrest or chin-rest were used. The experiment was divided into two sessions to provide the participants with a short break half-way through. The EEG and eye-tracking data were collected simultaneously from all participants in both sessions. Each session started with a 9-point calibration, and stimuli presentation started immediately after. Participants were instructed to listen to the audio-book, and to look at a fixation cross placed in the middle of the presentation screen. As in Study 1, participants were prompted 30 times in total (15 times per session) with the mDES questionnaire about the content of their thoughts and answered a comprehension question about the content of the previous paragraph with a yes/no answer. The prompts occurred at semi-random intervals every 3.5 min on average with listening times ranging from around 2 to 6 min between prompts. At the end of experimental sessions, participants completed the transportation scale, a questionnaire about their reading experience and habits, and their familiarity with the novel. In total, the testing session took about three hours. The procedure for the experiment is illustrated in Fig. 2. All data, study materials, and analysis code are available at: https://osf.io/u3qhj/overview.

4.1.7. Data pre-processing and analysis

All analyses were carried out in R statistical software (R Core Team, 2023). To identify participants' thought patterns during the experiment, we conducted a PCA using the '*psych*'; Revelle, 2024 with varimax rotation (1691 observations). The number of components was determined using parallel analysis, and component reliability was checked using split-half correlations. This analysis was conducted in the same manner as for Study 1 as described above. The component scores were then used in the other analyses.

Like in Study 1, we then examined the relationship between the components and transportation by running Pearson's correlations between the averaged component scores and participants' responses to the transportation scale, as well as their rating of the pleasure of listening. In addition, we examined the relationship between component scores and participants' accuracy in responding to the comprehension questions.

We then investigated the relationship between component scores and eye-movement behavior by running linear mixed models within the Bayesian framework using the '*brms*' package (Bürkner, 2017, 2018, 2021) and making inferences based on the 95% credible interval. We used the Bayesian framework as it puts the focus on the size and degree of certainty of an effect rather than whether or it not is significant. Prior to analysis, eye-tracking data were cleaned based on visualization of the eye movements during the trials, such that trials with too much data loss (e.g., no pupil detected) were excluded from further analysis (2.4% of eye-tracking data). From this data, we computed six eye-movement variables: (1) average fixation duration (in ms), (2) number of fixations, (3) fixation dispersion (mean Euclidean distance between the fixation-cross and the average fixation position in that trial/analysis window); (4) saccade amplitude (degrees), (5) blink duration (mean duration of blinks in ms); and (6) number of blinks. These measures were extracted from the analysis reports from DataViewer (SR Research). For the analysis, eye-movement measures were aggregated for the 10 s interval immediately prior to participants answering the prompts. We then ran a linear model for each eye-movement measure individually. Fixation duration, fixation dispersion, saccade amplitude and blink duration were log-transformed prior to modelling to ensure normality. Number of fixations and blinks were modeled with a Poisson distribution. Each model included the component scores as fixed effects, as well as random intercepts and random slopes for participants and items for components. Each model thus took the following form: eye-tracking measure ~ Component1 + Component2 + Component3 + Component4 + (Component1 + Component2 + Component3 + Component4 |Participant) + (Component1 + Component2 + Component3 + Component4 | Item). As these analyses were mostly exploratory, we used weak priors in the models.

4.2. Results

4.2.1. Thought patterns during listening

The parallel analysis indicated running the PCA for three components. The PCA identified three components which explained 48% of the variance in participants' responses to the mDES questionnaire. The loadings of these components are shown in Table 5. Component 1 ("immersion") was associated with on-task thoughts, as well as positive

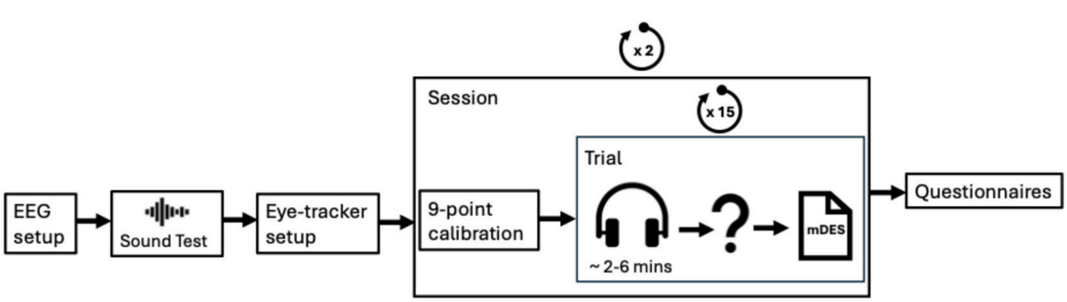

Fig. 2. Experimental Procedure for Study 2.

Table 5
Study 2: Loadings of the PCA Components.

	Component 1	Component 2	Component 3
Q1: Task	**0.57**	**−0.47**	0.08
Q2: Future	0.05	**0.61**	0.14
Q3: Past	0.19	**0.56**	−0.19
Q4: Self	−0.12	**0.64**	0.09
Q5: Other people	0.34	**0.52**	−0.23
Q6: Emotion	**0.35**	−0.28	−0.10
Q7: Images	**0.78**	−0.00	−0.20
Q8: Words	−0.22	0.21	**0.69**
Q9: Vividness	**0.81**	0.05	0.05
Q10: Specific	**0.73**	0.09	0.22
Q11: Habit	−0.00	**0.68**	0.26
Q12: Evolving	−0.23	**0.54**	−0.15
Q13: Deliberate	0.13	−0.11	**0.70**
SS Loadings	2.51	2.49	1.28
Proportion of Variance	19.3%	19.1%	9.9%
Cumulative Variance	19.3%	38.5%	48.3%

emotions and specific and vivid visual imagery. Component 2 (“mind-wandering”) was associated with off-task thoughts about the future, the past, other people, the participant themselves, and tended to consist of habitual and evolving thoughts. Component 3 (“sub-vocalization”) consisted of deliberate verbal thoughts. The split-half reliability correlations indicated very good reliability of the components with correlation coefficients of 0.99 for all three components. These components are illustrated in word-clouds in Fig. 3.

Correlations between these components and participants’ responses to the transportation scale showed that immersion was positively correlated with transportation (rs = 0.41, p < 0.001). Transportation scores were not significantly correlated to mind-wandering scores (rs = 0.05, p = 0.07), and were only weakly correlated to sub-vocalization scores (rs = -0.08, p = 0.003).

Correlations between participants’ ratings of the pleasure of listening and the component scores showed that pleasantness was positively correlated with immersion (rs = 0.31, p < 0.001), and weakly correlated with mind-wandering (rs = 0.08, p = 0.003). Pleasantness was not correlated with sub-vocalization (rs = -0.04, p = 0.17).

4.2.2. *Text valence and thought patterns*

The means and standard deviation of component scores for the three text valence conditions are shown in Table 6. Results showed that participants tended to have higher immersion while listening to positive text paragraphs compared to neutral ones (b = 0.27, CI = 0.51 – 0.01). There were no other significant effects of text valence on immersion. There were no significant effects of text valence on the mind-wandering or the sub-vocalization components.

4.2.3. *Thought patterns and comprehension*

Participants performed similarly on the comprehension questions as in Study 1 with an average accuracy of 75% correct answers (range 56 – 93%). The output of the model exploring the relationship between participants’ text comprehension accuracy and their thought patterns is shown in Table 7. The results showed no significant relationship between comprehension accuracy and any of the components (all CIs include 0).

Table 6
Study 2: Means and Standard Deviation of Component Scores per Text Valence.

Valence	Immersion		Mind-Wandering		Sub-Vocalization	
	Mean	SD	Mean	SD	Mean	SD
Neutral	−0.13	0.97	0.11	1.02	−0.09	0.98
Negative	0.02	1.01	0.11	0.95	−0.04	1.00
Positive	0.16	0.99	−0.07	1.01	−0.06	0.96

4.2.4. *Relationship between thought patterns and eye movements*

Outputs of the linear models on fixation duration, number, and dispersion are shown in Table 8 Results showed no significant effect of the three components on fixation durations. However, results showed that participants’ fixations tended to be less dispersed across the screen when they experienced higher immersion levels (b = -0.11), but more dispersed when they were mind-wandering (b = 0.04) or during sub-vocalization (b = 0.05). Similarly, participants made fewer fixations with higher levels of immersion (b = -0.03). The posterior distributions of the model estimates are shown in Figs. 4-6.

Outputs of the linear models for saccade amplitude, blink duration, and blink number are shown in Table 9. Results suggest that participants made shorter saccades while experiencing higher levels of immersion (b = -0.04), but that saccade amplitude was not affected by mind-wandering or sub-vocalizing. Results also show that participants made longer blinks during immersed episodes (b = -0.03), and blinked more during sub-vocalizing (b = 0.05). The posterior distributions for the model estimates are shown in Figs. 7-9.

Table 7
Study 2: Effect of Component Scores on Comprehension Accuracy.

Predictors	Estimate	CI (95%)
Intercept	1.42	0.95 – 1.91
Immersion	0.04	−0.10 – 0.19
Mind-Wandering	−0.06	−0.21 – 0.08
Sub-Vocalizing	−0.05	−0.20 – 0.10

Note. Table 4 shows the estimates and credible intervals (CI) for the relationship between participants’ comprehension accuracy and scores on the three components.

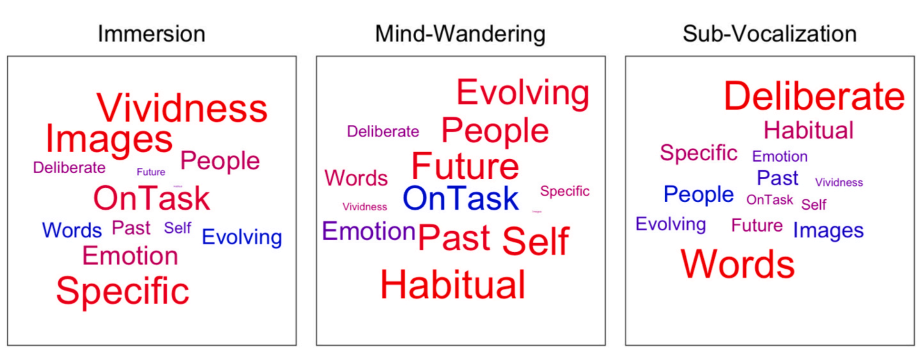

Fig. 3. Study 2: Word-Clouds of the PCA Components. Note. Fig. 3 shows word clouds for the three PCA components found in Study 2. Words in red indicate a positive loading onto the component, and words in blue indicate negative loadings. Word in a larger font indicate higher importance compared to words in a smaller font.

Table 8
Study 2: Relationship between Thought Patterns and Fixation Measures.

Predictors	Fixation Duration		Fixation Dispersion		Fixation Number	
	Estimate	*CI (95%)*	*Estimate*	*CI (95%)*	*Estimate*	*CI (95%)*
Intercept	5.60	5.43 – 5.77	8.45	8.25 – 8.63	2.52	2.39 – 2.66
Immersion	0.02	−0.00 – 0.04	**−0.11**	**−0.13 – -0.09**	**−0.03**	**−0.05 – −0.01**
Mind-Wandering	−0.02	−0.04 – 0.00	**0.04**	**0.02 – 0.06**	0.02	−0.00 – 0.04
Sub-Vocalization	−0.00	−0.02 – 0.02	**0.05**	**0.03 – 0.07**	−0.01	−0.02 – 0.01

Note. Table 1 shows the model estimates and 95% credible intervals (CI) for the fixation measures.

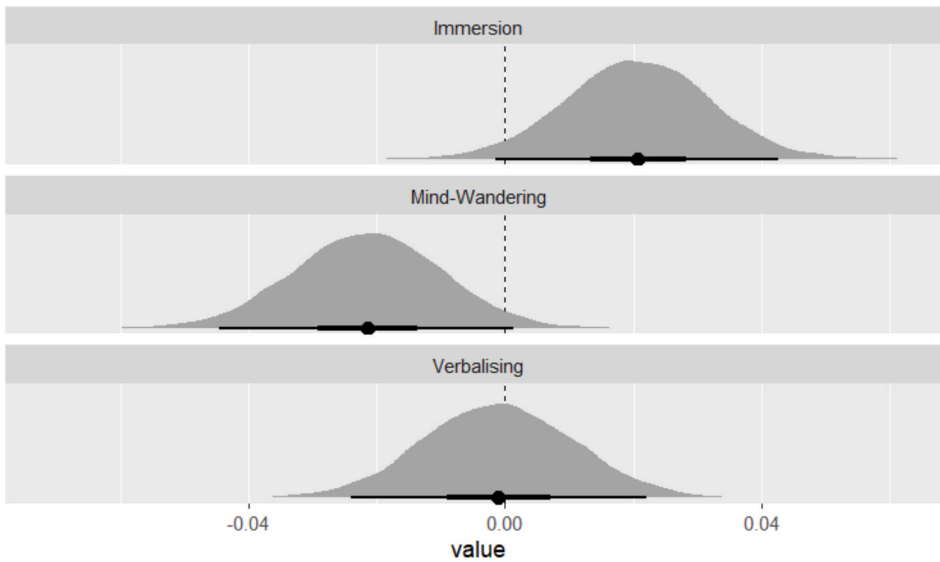

Fig. 4. Posterior Distribution of Model Estimates for Fixation Duration. Note. Fig. 4 shows the posterior distribution of model estimates for fixation durations for each though pattern. The black dot indicates the model estimates, and the black lines correspond to the 95% credible interval.

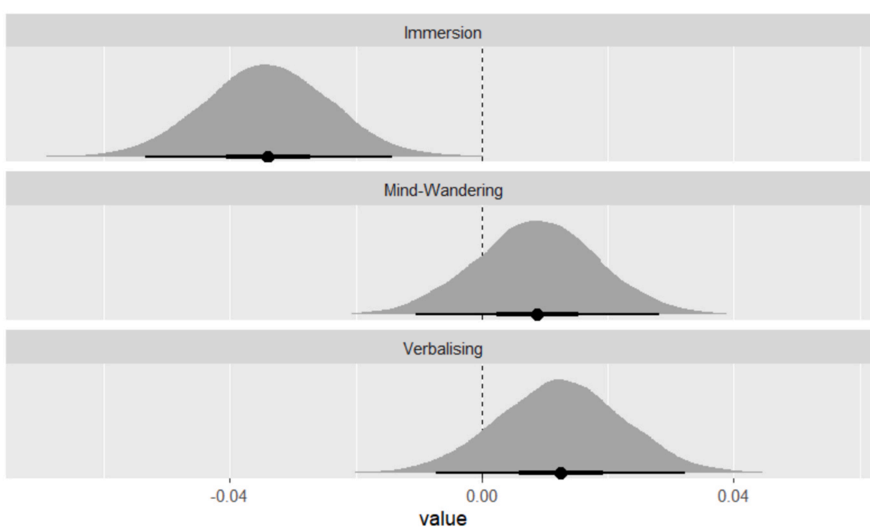

Fig. 5. Posterior Distribution of Model Estimates for Fixation Number. Note. Fig. 5 shows the posterior distribution of model estimates for the number of fixations for each though pattern. The black dot indicates the model estimates, and the black lines correspond to the 95% credible interval.

5. Discussion

5.1. Thought patterns experienced during listening to literary narratives

The first aim of this study was to examine the types of thought patterns participants have while listening to a literary narrative. Based on previous research, we expected to find evidence of at least two types of thoughts, namely, immersion and mind-wandering. The PCA results from both studies 1 and 2 are in line with these expectations and with previous research. In both studies, we found evidence for a component resembling immersion, which was associated with task-related thoughts, as well as vivid imagery and positive emotions (Gerrig, 1993; Green & Brock, 2000; Kuijpers et al., 2014; Lange et al., 2022; Mézière et al., 2025; Ryan, 2001). In both studies, immersion component scores were

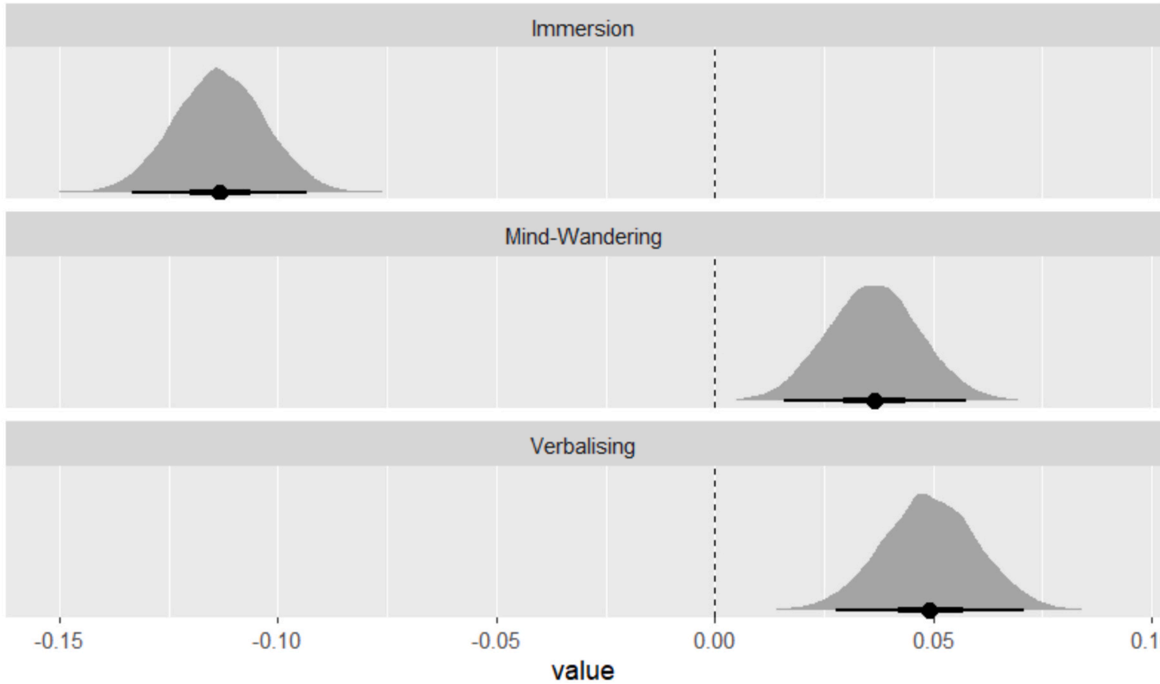

Fig. 6. Posterior Distribution of Model Estimates for Fixation Dispersion. Note. Fig. 6 shows the posterior distribution of model estimates for fixation dispersion for each though pattern. The black dot indicates the model estimates, and the black lines correspond to the 95% credible interval.

Table 9
Study 2: Relationship between Thought Patterns, Saccades, and Blink Measures.

Predictors	Saccade Amplitude		Blink Duration		Blink Number	
	Estimate	*CI (95%)*	*Estimate*	*CI (95%)*	*Estimate*	*CI (95%)*
Intercept	4.97	4.88 – 5.06	4.98	4.86 – 5.09	1.25	1.13 – 1.37
Immersion	**−0.04**	**−0.06 – -0.03**	**−0.03**	**−0.06 – -0.01**	−0.01	−0.05 – 0.03
Mind-Wandering	0.01	−0.01 – 0.03	−0.01	−0.04 – 0.02	−0.01	−0.04 – 0.03
Sub-Vocalization	0.01	−0.01 – 0.03	−0.03	−0.06 – 0.00	**0.05**	**0.01 – 0.08**

Note. Table 9 shows model estimates and 95% credible intervals (CI) for the saccade and blink measures.

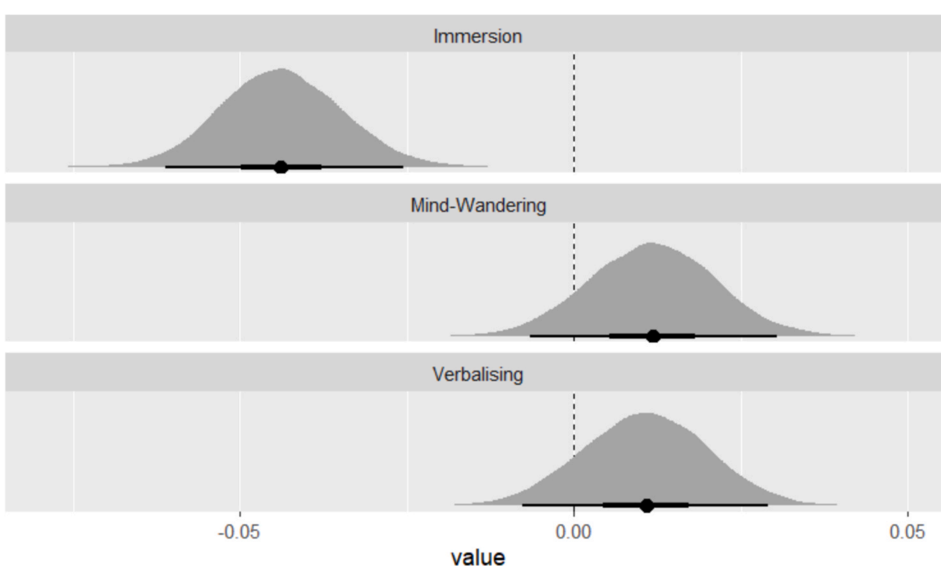

Fig. 7. Posterior Distribution of Model Estimates for Saccade Amplitude. Note. Fig. 7 shows the posterior distribution of model estimates for saccade amplitude for each though pattern. The black dot indicates the model estimates, and the black lines correspond to the 95% credible interval.

significantly correlated with the transportation and pleasantness scores, which is in line with the interpretation that this component represents

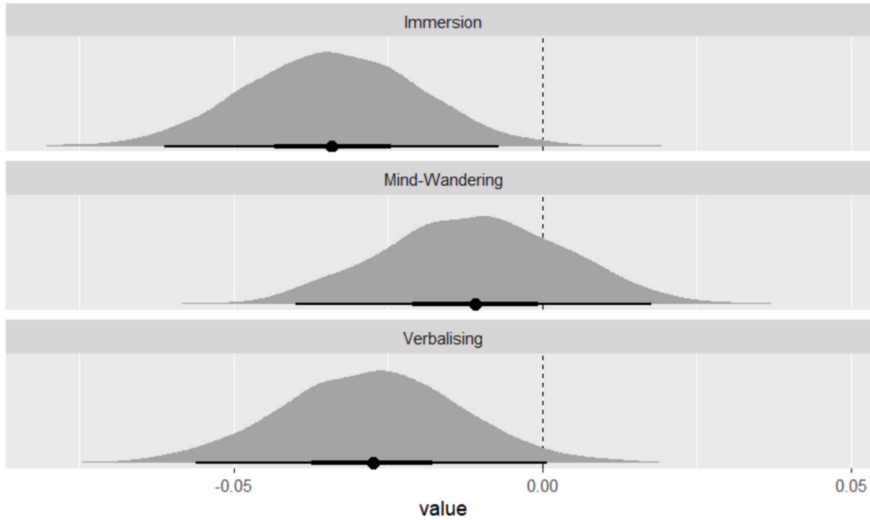

Fig. 8. Posterior Distribution of Model Estimates for Blink Duration. Note. Fig. 8 shows the posterior distribution of model estimates for blink duration for each though pattern. The black dot indicates the model estimates, and the black lines correspond to the 95% credible interval.

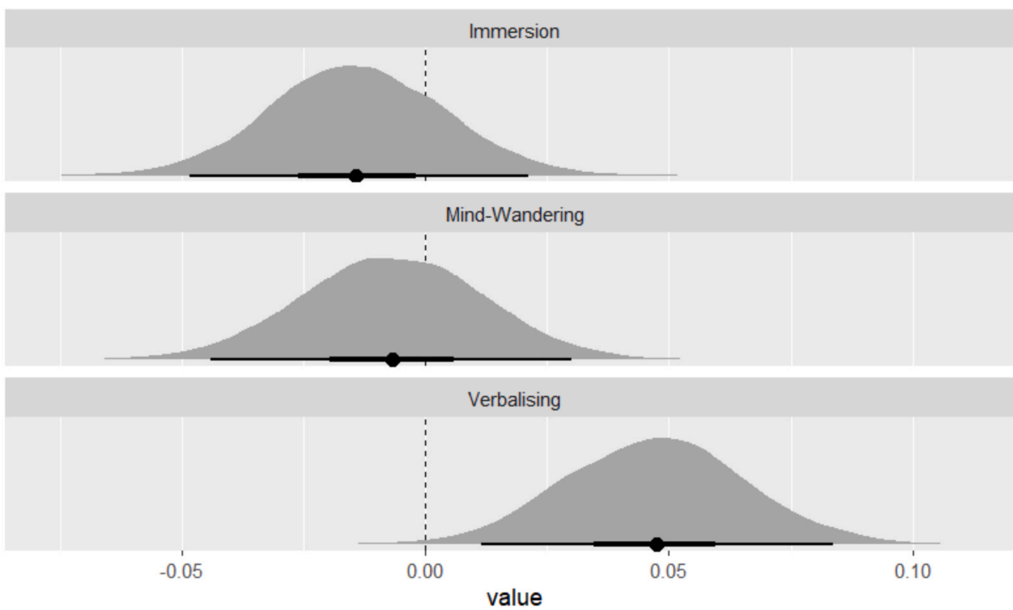

Fig. 9. Posterior Distribution of Model Estimates for Number of Blinks. Note. Fig. 9 shows the posterior distribution of model estimates for number of blinks for each though pattern. The black dot indicates the model estimates, and the black lines correspond to the 95% credible interval.

immersive episodes (Lange et al., 2022; see also Mézière et al., 2025 for similar findings with reading).

We also found evidence for a mind-wandering component, associated with off-task thoughts (Faber et al., 2020) as well as habitual and evolving thoughts about the future (Mézière et al., 2025). Of the types of thoughts found in this study, this component most closely resembles typical mind-wandering, as it loads negatively onto the off-task dimension of the PCA, in line with previous work looking at mind-wandering while listening to an audiobook (Faber et al., 2020; Varao-Sousa, 2013; Varao-Sousa, 2018). This component is highly similar to the mind-wandering component identified by Mézière and colleagues (2025) in an experiment using the same methodology (i.e., mDES) during reading of the same stimuli, such that the mind-wandering component went beyond the simple dichotomy of on– or off-task thoughts, and specified that these types of thoughts typically involved habitual and evolving thoughts about oneself and the future. Importantly, this component likely only represents one type of mind-wandering and may not be representative of other types of mind-wandering episodes such as task-related or voluntary mind-wandering (Seli et al., 2018).

We also found evidence for a third component, namely sub-vocalization which has also been found in previous work looking at reading of literary texts (Mézière et al., 2025). Sub-vocalization was associated with deliberate verbal thoughts and is in line with an “inner-voice” phenomenon previously reported during reading (Kosch et al., 2024; Mézière et al. 2025), during which the reader is “sounding out” the text to themselves while reading. During listening, this pattern may represent focus on the writing itself (e.g., appreciation for the form), or perhaps on the voice of the narrator, as the sound is already available via the audiobook. However, more research is necessary in order to better understand this phenomenon, particularly with regards to listening to literature, as this component may also be representative of other types of thoughts such as voluntary or task-related mind-wandering (Seli et al., 2018), or be related to specific forms of literary appreciation such as appreciation for the writing itself (i.e., beauty of the sentences used) for example.

Lastly, we identified a fourth component in Study 1, namely social episodic thought, which was associated with thoughts about the past and other people. Such a pattern is in line with previous work using a similar methodology looking at reading (Mézière et al., 2025) and may relate to thoughts about past events that are similar to the events present in the narrative (Fabry & Kukkonen, 2019). Contrary to previous findings, this component did not load negatively onto the off-task dimension of the PCA (Mézière et al., 2025) although it did have the second lowest loading out of the four components on the "on/off-task dimension", hence it is unclear whether this component is identical to the one found in reading studies or whether it represents a different type of thoughts specific to the listening modality. However, as this thought pattern is not well-understood yet, more research is necessary in order to understand what this though pattern is, how it differs from similar types of thoughts experienced during reading, and how it may relate to mind-wandering, text comprehension, emotions, and attention.

Overall, the thought patterns identified in the two studies are highly similar to the components identified in previous work using the same stimuli and methodology but while participants were reading the texts (Mézière et al., 2025), which suggests that, in general, the types of thoughts readers have while processing literary text may be similar whether they are reading or listening to the text, particularly with regards to immersion and mind-wandering.

5.2. Impact of text valence on thought patterns

The second aim of this study was to investigate the relationship between the emotional valence of a text and the types of thoughts listeners experience while listening to the text. In both studies, we found that listeners experienced higher levels of immersion during positive text paragraphs, which is in line with previous findings for reading the same text (Mézière et al., 2025), partially in line with previous work suggesting that immersion is associated with both positive and negative text content (Ballenghein et al., 2023; Kaakinen & Simola, 2020). The difference with previous work may be due to differences in the types of passages that were used, as previous work used horror stories which may be more likely to induce feelings of suspense and hence immersion for "negative" texts compared to the text used in the current study. This is further illustrated by the fact that our findings are identical to findings on reading of the same stimuli whereby only positive texts were found to elicit higher immersion scores (Mézière et al., 2025). Overall, this finding is also in line with the idea that immersion is typically associated with stronger emotional content.

Contrary to previous findings (Mézière et al., 2025), we found no association between emotional valence and sub-vocalization, although previous work with the same stimuli found that readers had higher tendency to sub-vocalize for both positive and negative texts compared to neutral ones. One plausible reason for this difference is that while the components are highly similar to each other, they may not represent the exact same type of thoughts. For example, during reading it may be that the readers are "sounding out" the text, which is less likely to happen during listening as the sound of the text is directly available. Hence, it may be that during listening, this component is representative of task-related mind-wandering, which may not be influenced by the emotional valence of the text. However, more research into this component is necessary to understand whether and how it differs from the sub-vocalization found in reading research, and if so what such component(s) represent in terms of reader's thought content, attention to the stimuli/task, and emotional processing of the text.

We also found that social episodic thoughts were associated with negative passages, such that listeners had higher levels of social episodic thoughts during negative passages compared to both neutral and positive ones. This relationship was not found in the reading study (Mézière et al., 2025), hence it may again be that this component represents something slightly different during listening to the text compared to reading it. Another possible explanation has to do with the fact that this component was only found in Study 1, in which participants listened to the book outside of a controlled laboratory setting which was not the case in study 2 or in Mézière et al.'s (2025) study. Hence it may be that the occurrence of this component, and the influence of text valence on its occurrence depend at least partially on both the modality of presentation (listening versus reading), as well as the environment (inside versus outside the lab). Prior research also suggests that listening to an audiobook outside of the lab can influence the occurrence of different thought patterns such as mind-wandering or being distracted (Varao-Sousa, 2018). Nevertheless, more research is necessary to understand this component, and whether and how it is related to text presentation modality and environment. Overall, the results of the current study suggests that the types of thoughts that listeners experience while listening to an audiobook is in part influenced by the emotional content of the text.

5.3. Impact of thought patterns on listening comprehension

The third research question of our study was about whether the type of thoughts listeners experience would impact comprehension of the text itself. In Study 1, we found that higher levels of immersion led to better comprehension scores. These results are in line with previous work (Lange et al., 2022) and with the idea that as attention is focused on the content of the text, readers would have a more detailed and accurate mental representation of the text content thus leading to better performance on the comprehension task. Importantly, we did not find such an effect in the laboratory experiment, Study 2. This may be in part due to differences between listening to an audiobook outside the laboratory such as at home or during a walk compared to listening to it in a controlled laboratory experiment while fixating a fixation cross. Hence, it may be that being "in the wild" while listening to an audiobook is more conducive to becoming immersed in the story world, thus impacting comprehension of the text, compared to being in the laboratory, as research suggests that laboratory versus in the wild settings can impact the types of thought readers have and their comprehension (Varao-Sousa et al., 2018).

In study 1, we also found that mind-wandering was negatively associated with comprehension, as higher levels of mind-wandering led to poorer comprehension outcomes. These results also align with previous work showing that mind-wandering negatively impacts comprehension while listening to text (e.g., Varao-Sousa et al., 2013). This outcome is also in line with the predictions of the cascade model of inattention (Smallwood, 2011) whereby off-task thoughts involve focused attention on stimuli/thoughts other than the text hence leading to a less detailed processing of the text content and ultimately a less detailed mental model of the text, thus leading to poorer comprehension. Interestingly, we did not find these effects in Study 2, which may be due to differences in the experiment setting, as Study 1 participants listened to the text in their own time in whatever environment they chose, while participants in Study 2 were in a controlled laboratory setting. Indeed, research suggests that mind-wandering and distraction in particular tends to be higher outside of the laboratory thus leading to stronger effects of distraction and mind-wandering on comprehension outcomes (Varao-Sousa et al., 2018). Hence, it may be that participants from Study 1 were more distracted than participants in Study 2 thus leading to significant effects of overall inattention (i.e., mind-wandering or distraction) on their comprehension outcomes. Overall, results from Study 1 suggest that, at least outside of controlled experiment settings, thought patterns such as immersion and mind-wandering do impact text comprehension while listening to a literary narrative.

5.4. Eye movements reflect immersion and mind-wandering during narrative listening

The final aim of the current study was to investigate the relationship between eye-movement behavior and thought patterns while listening to an audiobook. These analyses were exploratory in nature, hence we did not have clear hypotheses as to the presence or direction of this relationship. Firstly, we found that higher levels of immersion were associated with fewer and less dispersed fixations, shorter saccades, and shorter blinks. These findings are in line with related work about immersion while experiencing other forms of art such as music, which showed that immersion was related to making fewer saccades (Lange et al., 2017), suggesting that listeners also made fewer fixations. This is in line with the interpretation that immersion is characterized by attention being focused on the text content, reflected in eye movement behavior as less visual exploration (i.e., fewer and more concentrated fixations). Unlike previous work however we did not find evidence of lower blink rates as evidence for more focused attention compared to Kaakinen & Simola (2020) and Lange et al (2017; 2022). Indeed, previous work suggests that blinking rates are reflective of attention (de) coupling to the stimuli and has been shown to decrease during immersive episodes, indicating more focused attention to the stimuli such as literary text (audiobook listening: Kaakinen & Simola, 2020; Lange et al., 2022) and music (Lange et al., 2017). The difference of our results and previous studies may come from the difference in methodology between the studies, as we used PCA on mDES responses to measure immersion while Kaakinen & Simola used the Transportation Scale short form (Appel et al., 2015) and Lange and colleagues used the Absorption Scale short form (Lange et al., 2017; 2022). Indeed, while the immersion scores from the mDES are correlated with participants' scores on the transportation scale, these scores are not identical, which may have influenced the relationship between blink rates and immersion scores. However, we did find that immersion was associated with making shorter blinks, suggesting higher engagement and attention to the text stimuli. Indeed, while few studies have investigated blink duration as an indicator of attention de(coupling), Hollander & Huette (2022) recently showed that longer blink duration is associated with attention decoupling, engagement with the task, and mind-wandering in multiple tasks including reading and listening to text. The inverse relationship found in our study (i.e., shorter blinks associated with higher immersion) may therefore reflect readers' being more engaged and focused on the content of the text when they are immersed. Taken together, the results are in line with results from Kaakinen & Simola (2020) and Lange et al. (2022) as indicating that immersion is linked to more focused attention on the content of the text.

We also found that higher levels of mind-wandering were associated with more dispersed fixations, which is in line with the interpretation that mind-wandering is characterized by less focused attention. These results are partially in line with Faber et al. (2020) who found a trend for higher fixation dispersion and shorter fixations for mind-wandering in their 25-second interval immediately before the prompt. Unlike immersion, mind-wandering was not related to the number of fixations or saccade length, suggesting that while participants made fixations that were not focused on the target visual stimuli (i.e., the target is "task-relevant" in the sense that they are instructed to fixate it while also listening to the book), the fixations themselves were not necessarily far apart from each other or more numerous. This is partially in line with recent work using micro-saccades to investigate mind-wandering, which found no relationship between micro-saccade rates and mind-wandering as measured by prompts (Valsecchi et al., 2025), suggesting that the number or dispersion of fixations across the screen was similar between on-task and mind-wandering trials. Taken together, these findings suggest that a critical difference between immersion and mind-wandering may be how closely fixations are to the target visual stimuli as opposed to how closely fixations are to each other across the visual field (e.g., the screen). Overall, our results are in line with the idea that mind-wandering leads to less focused visual attention while listening to literary narratives (Smallwood, 2011). However, more research is necessary in order to investigate the relationship between these thought patterns and attentional focus on task-relevant stimuli including both visual and non-visual stimuli.

Lastly, we found that sub-vocalization was associated with more dispersed fixations and blinking more. This suggests that this component may be linked to attention that is less focused on the text, more similar to the mind-wandering component. A possible interpretation is that this component characterizes another form of mind-wandering such as voluntary or task-related mind-wandering (e.g., focus on the audio stimuli but not on the visual stimuli). However, as we cannot be sure what this component represents exactly, more research is necessary to understand what this component is, and how it relates to attention and eye-movement behaviour. Overall, the results of this analysis do suggest that eye-movement measures that reflect attentional focus are related to different thought patterns that listeners experience while listening to a literary narrative such as an audiobook.

6. Conclusions

The results from both experiments show that listeners experience a range of thought patterns while listening to a literary text, including mind-wandering and immersion. These findings further support the idea that the typically-used dichotomy of readers/listeners being on/off task may be too simplistic and does not allow for the full range of thoughts experienced while processing literary texts. Interestingly, these thought patterns seem to be relatively consistent both across environment (i.e., laboratory versus in the wild), and modality (listening and reading), although more research is necessary to better understand what these thought patterns are and how their occurrence/content is impacted by the environment and reading modality. Lastly, we show that different types of thought patterns are reflected in eye-movement behaviour, indicating differences in how visual attention is focused, even when the task itself does not require visual processing.

CRediT authorship contribution statement

Diane Mézière: . **Johanna K. Kaakinen:** . **Jarkko Lehtola:** Writing – review & editing, Formal analysis, Data curation, Conceptualization. **Karin Kukkonen:** Writing – review & editing, Conceptualization. **Jonathan Smallwood:** Writing – review & editing, Conceptualization. **Jaana Simola:** Writing – review & editing, Conceptualization.

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.

Data availability

All data, stimuli, and analysis code are available at a link shared in the manuscript.

References

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N. Z., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioural experiment builder. *Behavior Research Methods, 52*, 388–407. https://doi.org/10.3758/s13428-019-01237-x

Appel, M., Gnambs, T., Richter, T., & Green, M. C. (2015). The transportation scale–short form (TS–SF). *Media Psychology, 18*(2), 243–266. https://doi.org/10.1080/15213269.2014.987400

Ballenghein, U., Kaakinen, J. K., Tissier, G., & Baccino, T. (2023). Fluctuation in cognitive engagement during listening and reading of erotica and horror stories. *Cognition and Emotion, 37*(5), 874–890. https://doi.org/10.1080/02699931.2023.2215974

Bezdek, M. A., & Gerrig, R. J. (2017). When narrative transportation narrows attention: Changes in attentional focus during suspenseful film viewing. *Media Psychology, 20*(1), 60–89. https://doi.org/10.1080/15213269.2015.1121830

Bonifacci, P., Viroli, C., Vassura, C., Colombini, E., & Desideri, L. (2023). The relationship between mind wandering and reading comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 30*(1), 40–59. https://doi.org/10.3758/s13423-022-02141-w

Bradley, M., & Lang, P. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25*(1), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9

Bürkner, P., & C.. (2017). brms: An R Package for Bayesian Multilevel Models using Stan. *Journal of Statistical Software, 80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Bürkner, P., & C.. (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal, 10*(1), 395–411. https://doi.org/10.32614/RJ-2018-017

Bürkner, P., & C.. (2021). Bayesian Item Response Modeling in R with brms and Stan. *Journal of Statistical Software, 100*(5), 1–54. https://doi.org/10.18637/jss.v100.i05

Eekhof, L. S., Kuijpers, M. M., Faber, M., Gao, X., Mak, M., van den Hoven, E., & Willems, R. M. (2021). Lost in a story, detached from the Words. *Discourse Processes, 58*(7), 595–616. https://doi.org/10.1080/0163853X.2020.1857619

Faber, M., Bixler, R., & D'Mello, S. (2018). An automated behavioral measure of mind wandering during computerized learning. *Behavior Research Methods, 50*(1), 134–150. https://doi.org/10.3758/s13428-017-0857-y

Faber, M., Krasich, K., Bixler, R. E., Brockmole, J. R., & D'Mello, S. K. (2020). The eye–mind wandering link: Identifying gaze indices of mind wandering across tasks. *Journal of Experimental Psychology: Human Perception and Performance, 46*(10), 1201–1221. https://doi.org/10.1037/xhp0000743

Fabry, R. E., & Kukkonen, K. (2019). Reconsidering the mind-wandering reader: Predictive processing, probability designs, and enculturation. *Frontiers in Psychology, 2648*. https://doi.org/10.3389/fpsyg.2018.02648

Gerrig, R. J. (1993). *Experiencing narrative worlds: On the psychological activities of reading. Routledge*. https://www.jstor.org/stable/j.ctt1dr37cb.

Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology, 79*(5), 701–721. https://doi.org/10.1037/0022-3514.79.5.701

Hollander, J., & Huette, S. (2022). Extracting blinks from continuous eye-tracking data in a mind wandering paradigm. *Consciousness and Cognition, 100*, Article 103303. https://doi-org.ezproxy.utu.fi:2443/10.1016/j.concog.2022.103303.

Hustvedt, S. (2019). *Muistoja tulevaisuudesta* (K. Rikman, Suom.). Otava. (Alkuperäisteos julkaistu 2019).

Kaakinen, J., & Simola, J. (2020). Fluctuation in pupil size and spontaneous blinks reflect story transportation. *Journal of Eye Movement Research, 13*(3), 1–14. https://doi.org/10.16910/jemr.13.3.6.

Kosch, L., Schwabe, A., Boomgaarden, H., & Stocker, G. (2024). Experiencing Literary Audiobooks: A Framework for Theoretical and Empirical Investigations of the Auditory Reception of Literature. *Journal of Literary Theory, 18*(1), 67–88. https://doi.org/10.1515/jlt-2024-2005

Kuijpers, M. M., Hakemulder, F., Tan, E. S., & Doicaru, M. M. (2014). Exploring absorbing reading experiences: Developing and validating a self-report scale to measure story world absorption. *Scientific Study of Literature, 4*(1), 89–122. https://doi.org/10.1075/ssol.4.1.05kui

Lange, E. B., Thiele, D., & Kuijpers, M. M. (2022). Narrative aesthetic absorption in audiobooks is predicted by blink rate and acoustic features. *Psychology of Aesthetics, Creativity, and the Arts, 16*(1), 110–124. https://doi.org/10.1037/aca0000321

Lange, E. B., Zweck, F., & Sinn, P. (2017). Microsaccade-rate indicates absorption by music listening. *Consciousness and cognition, 55*, 59–78. https://doi.org/10.1016/j.concog.2017.07.009

Lei, A., Willems, R. M., & Eekhof, L. S. (2023). Emotions, fast and slow: Processing of emotion words is affected by individual differences in need for affect and narrative absorption. *Cognition and Emotion, 37*(5), 997–1005. https://doi.org/10.1080/02699931.2023.2216445

Lenth, R. (2024). emmeans: Estimated Marginal Means, aka Least-Squares Means. *R package version, 1*(10), 5. https://CRAN.R-project.org/package=emmeans.

Mézière, D. C., Kaakinen, J. K., Ranta, E., Kukkonen, K., Smallwood, J., & Simola, J. (2025). Do eye movements reflect readers' thoughts during reading? evidence from multidimensional experience sampling and eye movements. *Consciousness and Cognition, 134*, Article 103918. https://doi.org/10.1016/j.concog.2025.103918

Mulholland, B., Goodall-Halliwell, I., Wallace, R., Chitiz, L., Mckeown, B., Rastan, A., & Smallwood, J. (2023). Patterns of ongoing thought in the real world. *Consciousness and cognition, 114*, Article 103530. https://doi.org/10.1016/j.concog.2023.103530

Polychroni, N., Herrojo Ruiz, M., & Terhune, D. B. (2022). Introspection confidence predicts EEG decoding of self-generated thoughts and meta-awareness. *Human Brain Mapping, 43*(7), 2311–2327. https://doi.org/10.1002/hbm.25789

R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org.

Revelle, W. (2024). *psych: Procedures for Psychological, Psychometric, and Personality Research*. Evanston, Illinois: Northwestern University. R package version 2.4.12.

Ryan, M.-L. (2001). *Narrative as virtual reality: Immersion and interactivity in literature and electronic media*. Baltimore: Johns Hopkins University Press.

Seli, P., Kane, M. J., Smallwood, J., Schacter, D. L., Maillet, D., Schooler, J. W., & Smilek, D. (2018). Mind-wandering as a natural kind: A family-resemblances view. *Trends in Cognitive Sciences, 22*(6), 479–490. https://doi.org/10.1016/j.tics.2018.03.010

Smallwood, J. (2011). Mind-wandering while reading: Attentional decoupling, mindless reading and the cascade model of inattention. *Language and Linguistic Compass, 5*(2), 63–77. https://doi.org/10.1111/j.1749-818X.2010.00263.x

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

Varao Sousa, T. L., Carriere, J. S. A., & Smilek, D. (2013). The way we encounter reading material influences how frequently we mind wander. *Frontiers in Psychology, 4*.

Varao-Sousa, T. L., Smilek, D., & Kingstone, A. (2018). In the lab and in the wild: How distraction and mind wandering affect attention and memory. *Cognitive Research: Principles and Implications, 3*(1), 42. https://doi.org/10.1186/s41235-018-0137-0

Valsecchi, M., Dalmaso, M., Castelli, L., Baldini, E., & Galfano, G. (2025). Is mind wandering reflected in microsaccade dynamics? *Biological Psychology, 109109*. https://doi.org/10.1016/j.biopsycho.2025.109109