Journal of Experimental Psychology: Learning, Memory, and Cognition

The Role of Working Memory in the Occurrence of Involuntary Thoughts About the Past and Future: An Experimental Investigation

Krystian Barzykowski, Ewa Ilczuk, and Lia Kvavilashvili

Online First Publication, August 10, 2026. https://dx.doi.org/10.1037/xlm0001624

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: Learning, Memory, and Cognition

ISSN: 0278-7393

https://doi.org/10.1037/xlm0001624

The Role of Working Memory in the Occurrence of Involuntary Thoughts About the Past and Future: An Experimental Investigation

Krystian Barzykowski[1], Ewa Ilczuk[1], and Lia Kvavilashvili[2]
[1] Applied Memory Research Laboratory, Institute of Psychology, Faculty of Philosophy, Jagiellonian University
[2] Department of Psychology and Sport Sciences, University of Hertfordshire

Previous research has demonstrated that although involuntary autobiographical memories (IAMs) and involuntary future thoughts (IFTs) occur spontaneously with minimal effort, their frequency is significantly reduced during cognitively demanding activities, which leaves their precise underlying mechanism(s) yet to be revealed. Across two experiments, we tested the hypotheses that the frequency of IAMs and IFTs may be reduced due to (a) the limited working memory capacity that enables the processing of only a finite number of thoughts and/or (b) the reduced likelihood of noticing incidental, task-irrelevant cues that may potentially trigger IAMs and IFTs. To this end, participants completed a laboratory vigilance task with irrelevant cue words to trigger IAMs and IFTs and random thought probes to measure their frequency. Study 1 manipulated the level of working memory load (none, low, high) while participants performed a letter version of the *N*-back task during the vigilance task. The presentation of working memory task items and incidental cue words in the vigilance task was also manipulated (simultaneous vs. sequential). Study 2 further manipulated the type of working memory load by using either verbal or visuospatial *N*-back task while using a sequential presentation of working memory task items and cue words. In both studies, to further examine levels of incidental cue noticing, participants also completed an unexpected cue-recognition task after completing the vigilance task. Cognitive load reduced cue recognition, which positively predicted IAMs frequency. While working memory type had no robust effect, IFTs tended to be more frequent under visuospatial load. Overall, working memory load, cue noticing, and participants' strategic behaviors jointly shaped IAMs and IFTs retrieval. The results provide new insights about the underlying mechanisms of IAMs and IFTs in terms of their dependency on cognitive load and may open interesting avenues for future research.

Keywords: involuntary autobiographical memories, working memory, mental time travel, cognitive load, cue noticing

The present article focuses on the past- and future-oriented spontaneous thoughts that occur quite frequently in our daily life (e.g., Finnbogadóttir & Berntsen, 2013; Rasmussen et al., 2015) and are often triggered by incidental cues in one's environment or thoughts (e.g., Berntsen, 1996; Berntsen & Jacobsen, 2008; Cole et al., 2016; Mace, 2004; Mazzoni et al., 2014; Schlagman et al., 2007). While the former have been termed involuntary autobiographical memories (IAMs) and relate to our personal past, the latter

Nash Unsworth served as action editor.
Krystian Barzykowski https://orcid.org/0000-0003-4016-3966
The data that support the findings of this study are available at https://uj.rodbuk.pl/dataset.xhtml?persistentId=doi:10.57903/UJ/UXMKX6 and are described and can be accessed in Barzykowski et al. (2024c). The used laboratory procedures are available upon reasonable request. The procedure is also described in further details in Barzykowski et al. (2024b).

The authors have no conflicts of interest to declare. Because this research was funded in whole, or in part, by National Science Centre, Poland (Grant 2019/35/B/HS6/00528), for the purpose of open access, the author has applied a CC BY public copyright license to any author accepted manuscript version arising from this submission.

The authors thank Marcin Draszczuk for IT help in preparing the Involuntary Past and Future Thoughts Programme. The authors thank (in alphabetical order) Marcin Błażejowski, Julia Mądro, and Zuzanna Kapka for their invaluable assistance with data coding. Finally, Article Processing Charges were funded by a grant from the Strategic Program Excellence Initiative at the Jagiellonian University to Krystian Barzykowski.

During the preparation of this work, Krystian Barzykowski used ChatGPT 3.5 and 5 to proofread some of the sentences and enhance the language and readability. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

Krystian Barzykowski played a lead role in conceptualization, data curation, formal analysis, funding acquisition, methodology, project administration, resources, software, supervision, validation, visualization, writing–original draft, and writing–review and editing. Ewa Ilczuk played a lead role in investigation and visualisation and a supporting role in conceptualization, methodology, resources, software, writing–original draft, and writing–review and editing. Lia Kvavilashvili played a supporting role in conceptualization, methodology, and writing–review and editing.

Correspondence concerning this article should be addressed to Krystian Barzykowski, Applied Memory Research Laboratory, Institute of Psychology, Faculty of Philosophy, Jagiellonian University, ul. Ingardena 6, 30-060 Kraków, Poland. Email: krystian.barzykowski@uj.edu.pl

have been called involuntary future thoughts (IFTs) and relate to future possible events. A key feature shared by both IAMs and IFTs is that they tend to pop into mind unexpectedly without any preceding attempts to deliberately think about them (e.g., Barzykowski & Kvavilashvili, 2024; Berntsen, 1996; Berntsen & Jacobsen, 2008). For example, during a lecture, when a speaker mentions the phrase "summer travel," one might inadvertently recall a trip taken a few years ago or spontaneously think about an upcoming holiday that has already been planned.

Although IAMs and IFTs have been studied separately from each other, recently, there has been a significant increase not only in research on IFTs (Cole & Kvavilashvili, 2019) but also in research studying IAMs and IFTs within a unified framework of spontaneous cognitions (e.g., Barzykowski et al., 2021, 2022; Barzykowski, Radel, et al., 2019; Berntsen & Jacobsen, 2008; Cole et al., 2016; Finnbogadóttir & Berntsen, 2013; Mazzoni, 2019; Plimpton et al., 2015; Vannucci et al., 2017; Warden et al., 2019; Zareen et al., 2026). A clear advantage of adopting this unified approach is the possibility of investigating their underlying cognitive mechanisms.

A consistent result of previous diary studies of IAMs and IFTs is that they are both most likely to occur in response to incidental cues and importantly when people are not engaged in attentionally demanding activities (e.g., Berntsen, 1998; Berntsen & Jacobsen, 2008; Finnbogadóttir & Berntsen, 2013; Schlagman et al., 2007). Thus, it is more likely that spontaneous thoughts about our past or future will pop into our awareness when engaged in washing up (or driving on a motorway) than when answering an important phone call. However, although these spontaneous phenomena appear to depend on available attentional resources and cognitive load, their exact mechanism remains unknown.

So far, only a handful laboratory studies have investigated the cognitive load dependency of IAMs and IFTs under well-controlled experimental conditions (e.g., Ball, 2007, Experiment 2; Barzykowski et al., 2021; Barzykowski & Niedźwieńska, 2018a; Mazzoni, 2019; Vannucci et al., 2015, 2019). Most of these studies (e.g., Barzykowski & Niedźwieńska, 2018a; Mazzoni, 2019; Vannucci et al., 2015, 2019) have used modified versions of the experimental method originally developed by Schlagman and Kvavilashvili (2008) to elicit involuntary memories in the laboratory. In this paradigm, participants perform an undemanding vigilance task (i.e., responding to infrequent target slides with a pattern of vertical lines in a stream of slides with horizontal lines) while being exposed to short verbal phrases in the center of the screen, some of which may incidentally trigger involuntary memories (see also the Method section). To measure the frequency of IAMs and/or IFTs, participants also report involuntary thoughts that come to mind during the vigilance task. The effects of cognitive load were examined in these studies by varying parameters of the ongoing vigilance task itself (e.g., Mazzoni, 2019; Vannucci et al., 2015, 2019) or using a dual-task paradigm in which participants completed an additional cognitive task while performing the vigilance task (e.g., responding to a particular digit in a stream of visually presented digits; Ball, 2007; Barzykowski & Niedźwieńska, 2018a). Importantly, all these studies (i.e., Barzykowski & Niedźwieńska, 2018a; Mazzoni, 2019; Vannucci et al., 2015, 2019) consistently showed a significant decrease in the frequency of IAMs and IFTs as the cognitive demands increased.

Although initial findings appear to support the cognitive load dependency assumption, the nature of the cognitive load was not clearly examined in these studies, and therefore the precise mechanism(s) underlying this dependency remains still to be revealed (for a similar argument, see Vannucci et al., 2016, p. 1083). Moreover, cognitive load may affect the frequency of IAMs and IFTs via multiple routes, and therefore, the explanations proposed below are not mutually exclusive. For example, according to Mandler (1994), the process of retrieving spontaneous thoughts may not be entirely automatic and may depend on the same executive resources that are needed for controlling and carrying out cognitively demanding activities (for a similar argument, see Berntsen, 2009, p. 97). However, this view has not been supported by more recent neuroscientific evidence (e.g., N. M. Hall et al., 2008; S. A. Hall et al., 2014) that has demonstrated a marked absence of prefrontal cortex involvement in recalling involuntary memories.[1]

By contrast, Kvavilashvili and Mandler (2004) suggested that the cognitive dependency could result from a limited working memory capacity that enables the processing of only a finite number of thoughts in one's conscious mind. Therefore, even if a particular IAM or IFT is ready to pop into one's mind in response to a given cue if, at that moment, one's mind is already attending to the environment or is engaged in other activities or thoughts, this may be sufficient to prevent this thought coming to one's mind. This explanation highlights the idea that although retrieval of spontaneous thoughts may be mostly automatic and effortless, keeping the end product of this processing in one's conscious mind may still depend on available resources in one's working memory.

Broadly speaking, working memory is a complex system allowing us to maintain, store, and process information (that is no longer perceptually present) in the face of possible interference and distraction (e.g., Baddeley, 1986, 2007; Baddeley & Hitch, 1994; A. R. A. Conway et al., 2005; Diamond, 2013; Smith & Jonides, 1999). The distinction between verbal and visuospatial working memory (Diamond, 2013), which both have limited capacity and minimally interfere with each other, may be particularly useful in assessing the view that increased cognitive load interferes with keeping involuntary thoughts in mind postretrieval. In particular, there are important phenomenological differences between IAMs and IFTs that may differentially interfere with verbal and visuospatial working memory load.

First, existing theoretical models of autobiographical memory emphasize the importance of visual imagery in the subjective experience of autobiographical memories (e.g., Brewer, 1996; M. A. Conway & Pleydell-Pearce, 2000; Rubin, 2006), and there is ample evidence to demonstrate that thoughts about the past are more vivid and contain more sensory-perceptual details than thoughts about the future (e.g., Berntsen & Bohn, 2010; D'Argembeau & Van der Linden, 2004, 2006; Gamboz et al., 2010). Vannucci et al. (2016)

[1] The review of functional magnetic resonance imaging studies on spontaneous cognition and its neural mechanisms demonstrates the significant involvement of the key hubs of the default mode network such as the medial temporal lobe and posterior cingulate cortex (Andrews-Hanna et al., 2010; Philippi et al., 2021; Raichle, 2015; Van Calster et al., 2017). Thus, in contrast to voluntary retrieval, spontaneous retrieval may not need any additional engagement of the prefrontal cortex and, therefore, appears mostly to be a bottom–up process. Importantly, the key feature of the default mode network is that it increases its activity during passive states (when attention is diffused) which are conducive to internal thought processes (especially self-referential thoughts).

have also found that individuals with high visual imagery ability were better at remembering their personal past compared to individuals with lower visual imagery ability. Therefore, given "the centrality of imagery to autobiographical memory" (M. A. Conway & Pleydell-Pearce, 2000, p. 263), it may be expected that high visuospatial working memory load should more negatively affect the retrieval of IAMs than IFTs. On the other hand, given the fact that higher frequency of IFTs was observed when no verbal stimuli was presented (see Vannucci et al., 2017), it may be expected that verbal working memory load should more negatively influence the retrieval of IFTs than IAMs. As a result, we can expect different effects of the level of working memory load on the occurrence of IAMs and IFTs depending on the type of working memory load.

Finally, one other explanation suggested in the literature is that the diffuse state of attention, accompanying the low cognitive load activities, simply boosts the likelihood of noticing incidental or task-irrelevant cues that may act as potential triggers for IAMs and IFTs (Kvavilashvili & Mandler, 2004). Thus, fewer involuntary thoughts experienced during attentionally demanding activities may be a result of poorer cue noticing rather than of the cognitive resource dependency of involuntary retrieval per se. So far, only one previous study has partially addressed this possibility in the context of studying IAMs and IFTs (see Barzykowski et al., 2022). In this study, after finishing a cognitively demanding vigilance task, participants had to complete an unexpected recognition task for incidental word cues presented during the vigilance task. Results showed that participants in the high cognitive load condition were poorer at recognizing the incidental cues encountered in the vigilance task. While this result provides some initial support for the idea that poor cue noticing may be a driving force behind the powerful effect of cognitive load on IAMs and IFTs, Barzykowski et al. (2022) did not directly manipulate the presentation of cues during the vigilance task, and therefore, it is still unknown to what extent the cognitive load actually impaired the noticing of the cues during the vigilance.

Moreover, no previous study has directly investigated the possible role of working memory load and noticing the cues on the frequency of IFTs and IAMs by systematic experimental manipulation of the level of cognitive load (none, low, high) and the type of working memory (e.g., spatial, verbal) in participants while keeping the noticing of cues under strict control. Consequently, more research is needed to reveal the precise mechanisms of the cognitive load dependency of IAMs and IFTs.

The Present Research (Studies 1 and 2)

Across two planned studies, we wanted to directly test and further examine two possible mechanisms, as suggested by Kvavilashvili and Mandler (2004), underlying the cognitive load dependency of IAMs and IFTs. Specifically, we directly tested the possibility that the cognitive load dependency results from the decreased likelihood of noticing incidental or task-irrelevant cues that may act as potential triggers for IAMs and IFTs (Study 1) and/or the limited working memory capacity that is able to process only a finite number of thoughts at any given time (Studies 1 and 2). In addition, Study 2 specifically examined the role of verbal and visuospatial working memory load on the frequency of reported IAMs and IFTs.

To examine these possibilities, in Studies 1 and 2, we used a modified version of the standard laboratory paradigm for studying involuntary thoughts that was originally developed by Schlagman and Kvavilashvili (2008). Briefly, as described by Barzykowski et al. (2021, 2022), participants had to detect 15 infrequent target slides with vertical lines and ignore the nontarget slides with patterns of horizontal lines (785 slides). In addition, they were exposed to 270 short verbal phrases, some of which could incidentally trigger task-unrelated thoughts, including IFTs and IAMs. Throughout the vigilance task, participants were probed 23 times at random intervals to record their thoughts at the moment they were stopped (cf. Plimpton et al., 2015) and indicate whether the recorded thought had occurred spontaneously or deliberately. On completion of the vigilance task, participants were given their thought descriptions and asked to indicate whether their thoughts referred to past memories or future events. After completing the study, each thought description was coded by independent judges as related to the experimental task that participants were completing or something else. The number of recorded spontaneous task-unrelated thoughts (both IAMs and IFTs) was the focus of analyses.

Importantly, in both studies, the working memory load was manipulated by using one of the most popular and commonly used measures of working memory (especially in the neuroimaging literature), a verbal (Studies 1 and 2) and visuospatial (Study 2) *N*-back task (A. R. A. Conway et al., 2005; Kane & Engle, 2002). In this task, participants are presented with a stream of letters (Studies 1 and 2) or locations (Study 2), displayed one-by-one on the screen, and have to respond each time when the current letter (Studies 1 and 2) or location (Study 2) matches the one displayed *N*-trials back.

We decided to use this task for two reasons. First, it can be easily incorporated into the ongoing vigilance task without affecting its flow, and second, it directly taps into the main functions of working memory. Indeed, not only does the *N*-back task require encoding and storing relevant items, but it also strongly engages participants in the processing and manipulating these items such as continuous updating, counting, and matching as well as inhibiting irrelevant stimuli (see Gajewski et al., 2018, p. 2; Rac-Lubashevsky, & Kessler, 2016). In addition, as suggested by Jaeggi et al. (2010, p. 409), "it is a very useful tool for the experimental investigation of working memory processes because it allows load to be manipulated in a very simple, straightforward way."

Therefore, in both studies participants were engaged in a laboratory vigilance task under different working memory load conditions: none (with no working memory task), low (with 1-back), and high (with 3-back) working memory load task. In addition, to examine the possibility that the focused state of attention, induced by high cognitive load, reduces the likelihood of noticing incidental cues, in Study 1, we manipulated the presentation of the *N*-back working memory task stimuli (i.e., letters) and cue words in the vigilance task (simultaneous vs. sequential). In the simultaneous condition, incidental cue words in the vigilance task and letter stimuli in the *N*-back task were presented simultaneously on each slide. This should reduce the possibility of noticing incidental cue words because participants were processing the *N*-back stimulus on each slide in addition to monitoring for target vertical lines in the ongoing vigilance task. By contrast, in the sequential condition, *N*-back stimuli were not presented on those slides which contained the incidental word cues, which could enhance the likelihood of noticing these cues and hence increase the frequency of reporting IAMs and IFTs in the vigilance task (see Table 1).[2]

Given that the impaired noticing of cues (which is important for triggering spontaneous thoughts) and working memory load may have independent effects on the occurrence of IAMs and IFTs, in Study 1, we expected to observe significant main effects of (a) cue presentation (simultaneous vs. sequential) and (b) the level of working memory load (no load vs. low vs. high) on the number of reported IAMs and IFTs. In addition, we also expected these two factors to have a cumulative effect in further reducing the frequency of IAMs and ITFs, and thus we expected to observe a significant interaction with the strongest decrease in the number of IAMs and IFTs between low and high working memory load conditions in the simultaneous cue than in the sequential cue condition (for a summary of the research hypotheses, see Table 2).

Finally, in Study 2, we assessed more directly the possibility, suggested by Kvavilashvili and Mandler (2004), that increased cognitive load may impact primarily one's ability to keep the end products of involuntary retrieval in the mind (i.e., IAMs and IFTs) rather than noticing the cues per se. Therefore, in Study 2, we used only the sequential condition in which the *N*-back letter stimuli and vigilance task cue words were presented separately on each trial (which should maximize the likelihood of noticing the cues), and we manipulated the type of working memory load (verbal vs. visuospatial). Because verbal working memory load interferes with the processing of verbal material, high verbal working memory load should mainly reduce the occurrence of IFTs which have been shown to have verbally based contents (e.g., mind wandering about future involves more inner verbal speech, less visual imagery and sensory/visual details; Berntsen & Bohn, 2010; D'Argembeau & Van der Linden, 2004, 2006; Stawarczyk et al., 2013). By contrast, because visuospatial working memory interferes with the processing of visual information, high visuospatial working memory load should mainly reduce the number of IAMs because of their strong visual component. Therefore, if the retrieval of IAMs and IFTs is primarily limited by different types of working memory capacity, we should observe an interaction between the level of working memory load, the type of working memory manipulation, and the type of spontaneous thought. In particular, it was predicted that compared to the low cognitive load and control (no load) conditions (a) in the high visuospatial working memory load condition, participants would report IAMs less frequently than IFTs, and (b) in the high verbal working memory load condition, they would report IFTs less frequently than IAMs (for a summary of the research hypotheses, see also Table 2).

In summary, across two studies, we used different versions of the vigilance task where we manipulated (a) the level of working memory load (none, low, high; Studies 1 and 2), (b) the presentation of working memory task items and cue words in the vigilance task (simultaneous vs. sequential; Study 1), and (c) the type of working memory load (verbal, visuospatial; Study 2).

Study 1

Studies 1 and 2 were approved by the Jagiellonian University Research Ethics Committee (No. KE/39_2021). Written consent for participation was obtained prior to data collection. Participants were informed that they were free to withdraw from the study at any point. Both studies were conducted in accordance with the University guidelines for safe working during the COVID-19 pandemic.

Method

Design

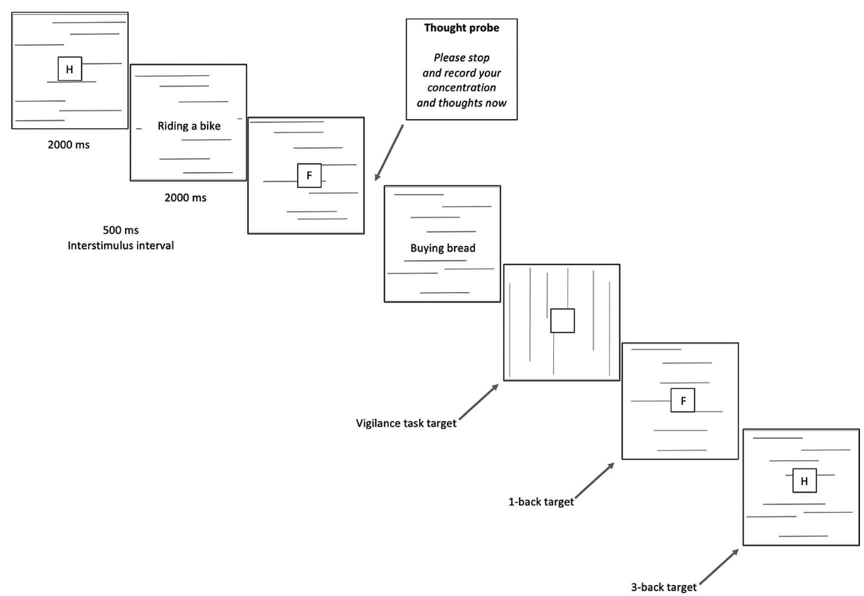
Table 3 shows an overview of all experimental groups in Studies 1 and 2. We employed a 3 (working memory load: none-control, low, high) × 2 (presentation of cues and working memory task: simultaneous vs. sequential) between-subjects experimental design. This allowed us to examine the role of working memory load and its possible interference with cue noticing in the retrieval of IAMs and IFTs. Specifically, in all six groups, participants were instructed to write down a content of their thoughts at the moment the program was stopped, regardless of what it was or how interesting they thought it to be. Working memory load was operationalized by engaging participants (or not engaging as in the control condition) in either 1-back (low working memory load) or 3-back verbal task (high working memory load). In addition, the incidental cue words in the vigilance task were presented either with the working memory task items (i.e., letter stimuli) on the same slide or on separate slides (see Table 1).

Statistical Power and Sample Size Analyses

Since no previous studies have used the manipulation of working memory load as described in the present project, there was no information to determine an expected effect size. However, to estimate the minimum sample size for Studies 1 and 2, the G*Power 3.1.9.6 (Faul et al., 2007) was used. Since it was planned to run two-way analyses of variance (ANOVAs) to analyze the effects of working memory load (none, low, high) and the effect of cue presentation (Study 1: simultaneous vs. sequential) or the effect of the type of working memory load (Study 2: verbal, spatial) on the frequency of IAMs/IFTs, the required sample sizes for 90% of power were 1,269, 206, and 83 for small (0.10), medium (0.25), and large (0.40) effect sizes, respectively (G*Power software; Faul et al., 2007). Therefore, in the present Studies 1 and 2, a minimum total sample size of 240 (40 participants per group) was planned and then applied. This sample size may be considered as adequate to detect significant effect of working memory load with a large effect size of .40 and a medium fixed effect size of 0.25, with the significance level of .05 and statistical power of .90. Importantly, we decided to increase the sample size from 206 to 240 to account for possible dropouts from the study and/or getting invalid or missing data (e.g., due to technical problems). Therefore, a sample size of 240, used in Studies 1 and 2, allowed a good power to detect medium- to large-sized effects. If the working memory hypothesis is valid, then one would expect medium-to-large effects between the conditions especially when considering effect sizes of cognitive load observed in previous studies that ranged from $d = 0.61$ to 0.67 (Studies 1 and 2, respectively, in Barzykowski & Niedźwieńska, 2018a). Conversely, if the effect of working memory load is weak or nonexistent, then we would not expect any noticeable differences between the conditions.

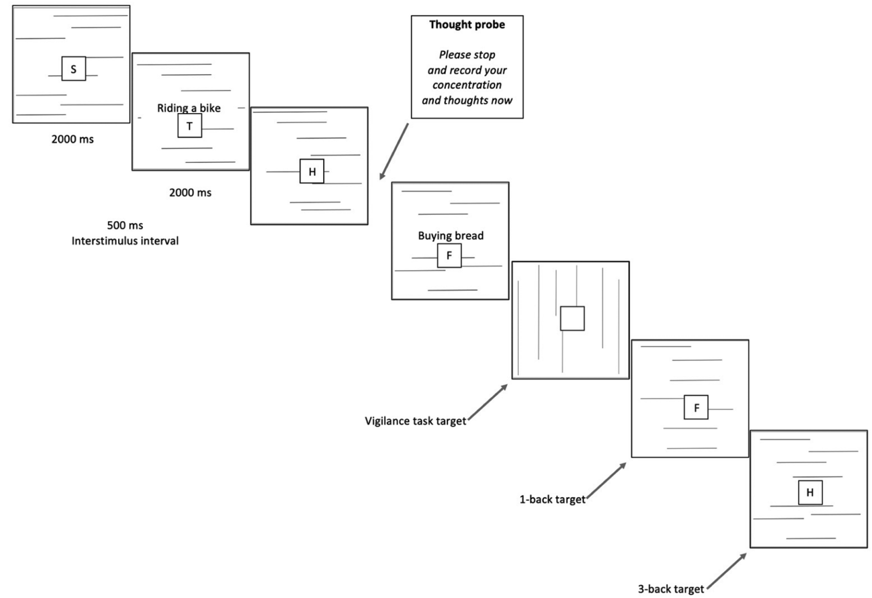
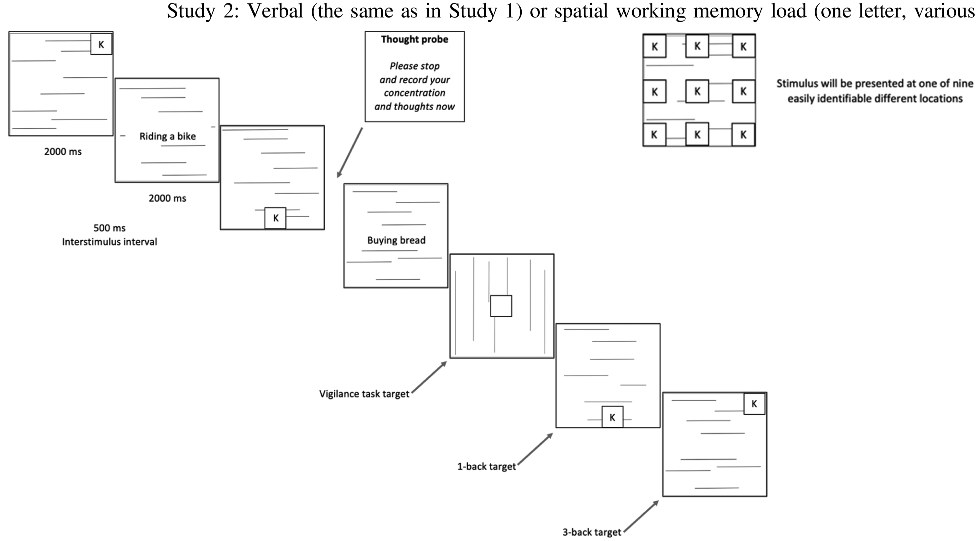
[2] To obtain further evidence for the role of cognitive load in cue noticing, in both studies, after finishing the vigilance task, participants also completed an unexpected recognition memory task for cues displayed in the vigilance task (cf. Barzykowski et al., 2022).

Table 1
An Overview of the Experimental Procedures in Studies 1 and 2

Procedural detail of Study 1 and Study 2	Working memory loading task		
	Control: No working memory task	Low working memory load	High working memory load
General overview	• The vigilance task: – 800 trials (2 s each). – 23 thought probes. – 270 trials with verbal cues. – 15 target trials with vertical lines. – Additional stimuli on the screen used for the *N*-back task (but without instruction to respond to them): ■ Study 1 (only verbal stimuli): Various letters inside in the square presented in the center of the screen (various letters, one location). ■ Study 2 (both verbal and spatial stimuli): ■ Verbal version: As in Study 1. ■ Spatial version: One letter in the square changing its location (one letter, various locations). • Unexpected cue-word recognition task (84 cues): – 42 presented during the task. – 42 not presented during the task.	• The vigilance task: – 800 trials (2 s each). – 23 thought probes. – 270 trials with verbal cues. – 15 target trials with vertical lines. – 1-back task: ■ Study 1: Verbal 1-back task: *Respond only when the consonant presented on the screen is the same as the one that had been presented 1-trial before*. ■ Study 2: Verbal or spatial 1-back task. Spatial: *Respond only when the letter-square is presented at the same position as it had been presented 1-trial before*. • Unexpected cue-word recognition task (84 cues).	• The vigilance task: – 800 trials (2 s each). – 23 thought probes. – 270 trails with verbal cues. – 15 target trials with vertical lines. – 3-back task: ■ Study 1: Verbal task: *Respond only when the consonant presented on the screen is the same as the one that had been presented 3-trials before*. ■ Study 2: Verbal and spatial 3-back task. Spatial: *Respond only when the letter-square is presented at the same position as it had been presented 3-trials before*. • Unexpected cue-word recognition task (84 cues).
Cues and working memory stimuli are separated from each other and are presented on different slides (sequential conditions). In summary there are • *N* of verbal cues = 270. • *N* of letters = 515: – Presented together with cues = 0. – Presented in-between cues = 515.	Study 1: Verbal working memory load—various letters (in uppercase: B, C, D, F, G, H, K, M, N, P, R, S, T, W, Z) but one location 		

(*table continues*)

Table 1 (*continued*)

Procedural detail of Study 1 and Study 2	Working memory loading task: Control: No working memory task	Low working memory load	High working memory load
Cues and working memory stimuli are combined with each other on the screen (simultaneous conditions). In summary there are • *N* of verbal cues = 270. • *N* of letters = 785: – Presented together with cues = 270. – Presented in-between cues = 515.			
Study 2: Verbal (the same as in Study 1) or spatial working memory load (one letter, various locations)			
Spatial working memory load: Cues and working memory stimuli are separated from each other and are presented on different slides (sequential conditions). In summary there are • *N* of cues = 270. • *N* of letters/ locations = 515: – Presented together with cues = 0. – Presented in-between cues = 515.			

Note. Text in italics highlights the key response rules that differed across experimental conditions.

Table 2
The Summary of Research Hypotheses in Studies 1 and 2

Hypothesis	Expectation
Study 1	
1. Working memory load plays a role in the retrieval of IAMs and IFTs.	Significant main effect working memory load (none, low, high): • The frequency of IAMs/IFTs will decrease in direct proportion to the increase in the working memory load with highest frequency of IAMs/IFTs in the no working memory load and the lowest frequency in the high working memory load conditions.
2. People are less likely to notice cues that may accidently trigger spontaneous thoughts when the cognitive load of the ongoing task increases.	Significant main effect of cue presentation (simultaneous, sequential): • The frequency of IAMs/IFTs will be the highest in sequential than simultaneous conditions.
3. The effect of working memory load and cue presentation have a cumulative effect on the retrieval of IAMs and IFTs.	Significant main effect of the Working Memory Load × Cue Presentation interaction: • The strongest decrease in the number of IAMs/IFTs between low and high working memory load conditions will be observed in the simultaneous cue condition than in the sequential cue condition.
Study 2	
1. The limited working memory capacity enables the processing of only a finite number of thoughts at any given time, and this relationship may be restricted to a certain type of working memory component used.	Significant main effect working memory load (none, low, high): • The frequency of IAMs/IFTs will decrease in direct proportion to increase in the working memory load with highest frequency of IAMs/IFTs in the no working memory load and the lowest in the high working memory load conditions.
1a: IAMs should be least frequently observed in the high visuospatial working memory load condition.	Significant main effect of type of working memory load (verbal, spatial): • The frequency of IAMs will be the lowest in spatial than verbal conditions.
1b: IFTs should be least frequently observed in the high verbal working memory load condition.	Significant main effect of type of working memory load (verbal, spatial): • The frequency of IFTs will be the lowest in verbal than spatial conditions.
2. The effect of type of working memory load and magnitude of working memory load have a cumulative effect on the retrieval of IAMs and IFTs.	
2a: As to IAMs	Significant main effect of the Type of Working Memory Load (verbal, visuospatial) × the Magnitude of Working Memory Load (none, low, high) interaction: • The strongest decrease in the number of IAMs between low and high working memory load conditions will be observed in the visuospatial condition than in the verbal cue condition.
2b: As to IFTs	Significant main effect of the Type of Working Memory Load (verbal, visuospatial) × the Magnitude of Working Memory Load (none, low, high) interaction: • The strongest decrease in the number of IFTs between low and high working memory load conditions will be observed in the verbal than in the visuospatial condition.

Note. IAMs = involuntary autobiographical memories; IFTs = involuntary future thoughts.

Participants

Two hundred forty young adults (age range 18–56 years; for details see Table 3) participated in Study 1 in return for modest financial reward (PLN 50, ca. $13). As indicated in Table 3, they were randomly assigned to one of six groups: (a) no working memory task, control conditions with cues and *N*-back task stimuli being presented on separate slides or on the same slides; (b) low working memory task conditions with sequential or simultaneous cue presentation and a 1-back task; and (c) high working memory task conditions with sequential or simultaneous cue presentation and a 3-back task.

Criteria for Data Inclusion and Exclusion

We excluded one participant due to technical problems during the experimental session preventing them from successfully completing the vigilance task. In addition, 11 participants having less than 50% of correct responses on the vigilance task and/or on the *N*-back task suggesting not being engaged in the task performance were excluded from the analysis. Next, three participants who had guessed the true purpose of the study (i.e., the main interest in past and future thoughts) were excluded from the analysis concerning the frequency of involuntary task-unrelated thoughts (ITUTs), including IFTs and IAMs. Finally, when analyzing reaction times (RTs), we excluded RTs under 100 ms, as the threshold of 100 ms is considered the minimum for valid and genuine responses (e.g., Whelan, 2008).

Materials

The Vigilance Task to Study Involuntary Past and Future Thoughts. We used a fully computerized version of the vigilance task with 800 slides (trials), which is described elsewhere in more detail (e.g., Barzykowski & Niedźwieńska, 2016, pp. 5–6; also Barzykowski & Staugaard, 2016, p. 524) and was very similar to the procedure used by Plimpton et al. (2015) and Barzykowski et al. (2022) to study IAMs and IFTs under various laboratory conditions (adapted from Schlagman & Kvavilashvili, 2008). See also study protocol as described by Barzykowski et al. (2024b).

Table 3
Overview of Experimental Groups in Studies 1 and 2 (Final Sample Sizes After Exclusions Reported in the Article)

Condition	Working memory load: None—control condition	Low working memory load condition (1-back task)	High working memory load condition (3-back task)
Study 1			
Cues and working memory task stimuli presentation			
Sequential	Group 1 (N = 39; 27 females, M_{age} = 23.90, SD = 6.67, range 18–56 years)	Group 3 (N = 39; 31 females, M_{age} = 22.86, SD = 4.14, range 19–40 years)	Group 5 (N = 35; 28 females, M_{age} = 22.81, SD = 4.02, range 19–42 years)
Simultaneous	Group 2 (N = 39; 29 females, M_{age} = 21.91, SD = 2.08, range 19–26 years)	Group 4 (N = 38; 25 females, M_{age} = 23.26, SD = 3.29, range 19–33 years)	Group 6 (N = 35; 27 females, M_{age} = 22.27, SD = 2.21, range 19–29 years)
Study 2			
Type of working memory load			
Verbal	Group 1 (N = 40; 30 females, M_{age} = 23.39, SD = 7.35, range 18–65 years)	Group 3 (N = 39; 36 females, M_{age} = 21.82, SD = 2.35, range 19–31 years)	Group 5 (N = 36; 30 females, M_{age} = 21.86, SD = 4.06, range 18–41 years)
Spatial	Group 2 (N = 40; 26 females, M_{age} = 22.89, SD = 6.12, range 19–51 years)	Group 4 (N = 38; 28 females, M_{age} = 23.17, SD = 5.20, range 19–47 years)	Group 6 (N = 32; 23 females, M_{age} = 22.41, SD = 3.68, range 19–33 years)

Briefly, as described in studies by Barzykowski, Radel, et al. (2019, p. 670; also, Barzykowski et al., 2021), the vigilance task involved detecting patterns of vertical lines (15 target slides) in a stream of 785 nontarget slides with horizontal lines. Slides were presented for 2 s, with an interstimulus interval of 0.5 s, and short verbal phrases (e.g., *riding a bike*) displayed in the center of some of the slides. There were equal numbers of neutral ($N = 90$; e.g., *buying a bread*), positive ($N = 90$; e.g., *a wonderful smile*), and negative ($N = 90$; e.g., *unpleasant conversation*) phrases, which constituted the final pool of 270 phrases used in previous studies (e.g., Barzykowski & Mazzoni, 2022; Barzykowski & Niedźwieńska, 2016; Barzykowski, Niedźwieńska, & Mazzoni, 2019; Barzykowski & Staugaard, 2016, 2018). One set of word phrases was used during the laboratory sessions. They were presented in two fixed pseudorandom orders and occurred at varying intervals consisting of a minimum of one (about 2.5 s) and a maximum of seven (about 17.5 s) slides. The mean interval between the cues was about three slides (i.e., 7.5 s). Similar to our previous studies (Barzykowski et al., 2022; Barzykowski & Niedźwieńska, 2018a), in conditions in which word cues and working memory stimuli were presented on separate slides, a square (approximately 1.5 cm × 1.5 cm) containing a random consonant (B, C, D, F, G, H, K, M, N, P, R, S, T, W, Z) was presented in the center of the screen on each trial with no word cue (for details of the *N*-back task, see the Verbal *N*-Back Task section). In the simultaneous conditions in which word cues and working memory letter stimuli were presented together on some of the slides, if there was a word cue on a given slide, the square was presented underneath it (see Table 1).

The program stopped automatically at 23 fixed points during the presentation, and the following message appeared on the screen: "Please stop and record your concentration and thoughts now." Participants typed into the computer program a brief description of the content of their thoughts; they indicated how much they were concentrating on the task when stopped (1 = *not at all*; 7 = *fully concentrating*; all points along the scale were clearly labeled during the task) and whether the thought occurred deliberately (they decided to think about it) or involuntarily (it simply popped into their mind). The probes, which were modeled on previous literature (Barzykowski et al., 2021, 2022; Barzykowski, Radel, et al., 2019; Plimpton et al., 2015), were presented in a fixed pseudorandom order and occurred at intervals of between 32 (about 80 s) and 40 (about 100 s) slides (the mean of 86.85 s). These intervals between the stops were comparable to similar previous studies (e.g., Barzykowski et al., 2022; Barzykowski, Radel, et al., 2019; Plimpton et al., 2015).

Finally, we used several attention control questions throughout the experiment where participants were asked to press a certain number key on their computer keyboard (e.g., when participants rated their concentration on the vigilance task in the thought probe, they could be presented with an additional question asking them to press "3" as an answer to the question). The detection of vertical lines can also be considered as a measure of attention to the task.

The Verbal *N*-Back Task. To further verify the role of working memory load on the reported experience of IAMs and IFTs, the *N*-back task was presented to participants during the vigilance task. An overview of the *N*-back task is presented in Table 1. Specifically, participants were instructed to respond by pressing a green button ("X" on the keyboard) each time the letter currently presented on the screen was the same as the one presented either one trial ago (low working memory load conditions) or three trials ago (high working memory load conditions). In general, participants were exposed to a series of consonants (in uppercase: B, C, D, F, G, H, K, M, N, P, R, S, T, W, Z), one-by-one, presented in a square appearing in the center of the screen for 2 s with an interstimulus interval of 0.5 s. Therefore, participants had 2.5 s to respond if needed. Maximum RT of 2.5 s and a minimum RT of 100 ms were allowed. Premature and late responses were not recorded and were categorized as misses.

In total, there were 515 slides with the working memory letter stimuli (presented separately from word cues and target vertical line slides) and 270 slides in which letters were presented together with word cues. No letter stimuli and cue words were presented on

15 slides with target vertical lines in the vigilance task. Importantly, across all working memory task conditions, there was the same number of 515 letters presented across all slides with cue words. This kept a target–foil ratio of 1:2 (i.e., 33% targets for the *N*-back task) constant across all working memory task conditions with 176 *N*-back task targets presented only on trials in which no cue words and no vigilance task targets (i.e., vertical lines) were presented. This way, responding to *N*-back targets did not interfere with cue noticing and performance on the vigilance task.

The Cue-Recognition Task. To further verify whether the working memory load did indeed influence the processing of word cues that could incidentally trigger IAMs and IFTs, the unexpected cue recognition task was presented to participants immediately at the end of the vigilance task. As in Barzykowski et al. (2022), participants were presented with a total of 84 cues, half of which (i.e., 42 cues) were randomly selected from the pool of cues presented during the vigilance task. The remaining 42 cues, which were the same for all participants, were presented for the first time and were randomly selected from the rest of the Polish pool of 800 cues used in previous studies on spontaneous past and future thoughts (e.g., Barzykowski & Niedźwieńska, 2016; Barzykowski, Niedźwieńska, & Mazzoni, 2019; Barzykowski & Staugaard, 2016, 2018). Importantly, there were equal numbers of neutral (n = 14), positive (n = 14), and negative (n = 14) phrases in each set of old and new cue phrases. Cues were presented in two fixed pseudorandom orders and were counterbalanced across groups and conditions. Participants responded to each cue, presented on the screen, in a self-paced manner without any time restrictions.

To investigate the comparability of cues used in the cue-recognition task, we calculated the mean ratings for each phrase as a function of cue-recognition type (presented during the vigilance task or not presented at all) and entered them into three separate independent t tests with concreteness, imagery, and typicality as dependent variables. There were no significant main effects (all ps > .53) of the cue-recognition type for any of the characteristics.

The Positive and Negative Affect Schedule (Brzozowski, 2010). To control for possible differences between groups, we used Positive and Negative Affect Schedule (PANAS; 30 items), which measures the strength of negative and positive current emotional states. Participants had to rate on a 5-point scale to what extent the given adjectives correspond with their current state. The reliability coefficients (internal consistency and stability) of the Polish version of PANAS range from 0.73 to 0.95 (Brzozowski, 2010).

Procedure

The experimental session lasted up to 2.5 hr. Participants were tested in groups of two to 11 in one laboratory-based session. They individually followed the procedure on individual computer stations. They were free to withdraw from the study at any time. In addition, they could refrain from reporting personally sensitive thoughts by typing "X" as an answer or by providing a generic description. Participants were briefly introduced to the procedure by the experimenter and then read detailed instructions on the computer screen. Figure 1 provides a general overview of the procedure. For a demonstration of the procedure, please refer to Barzykowski et al. (2024b), where a movie clip presenting a general overview of the procedure is also provided.

Control Condition With No Working Memory Loading Task (a Standard Version of the Vigilance Task). The procedure in this condition was very similar to the one used in the control condition by Barzykowski et al. (2022). First, participants were verbally informed that the experiment examined how people concentrate on monotonous and boring tasks. Next, they started the computerized vigilance task, which provided more detailed written instructions. In particular, participants were instructed to identify slides with vertical lines by pressing a red button ("m" on the keyboard). They were also informed that they would see word phrases and a square with a letter in the center of each slide, but they were asked not to respond to them during the experiment. Importantly, we did not ask them to specifically ignore these stimuli. Instead, participants were told they did not need to do anything in response to these stimuli because their task was to pay attention to the orientation of line patterns (horizontal or vertical) on each slide.

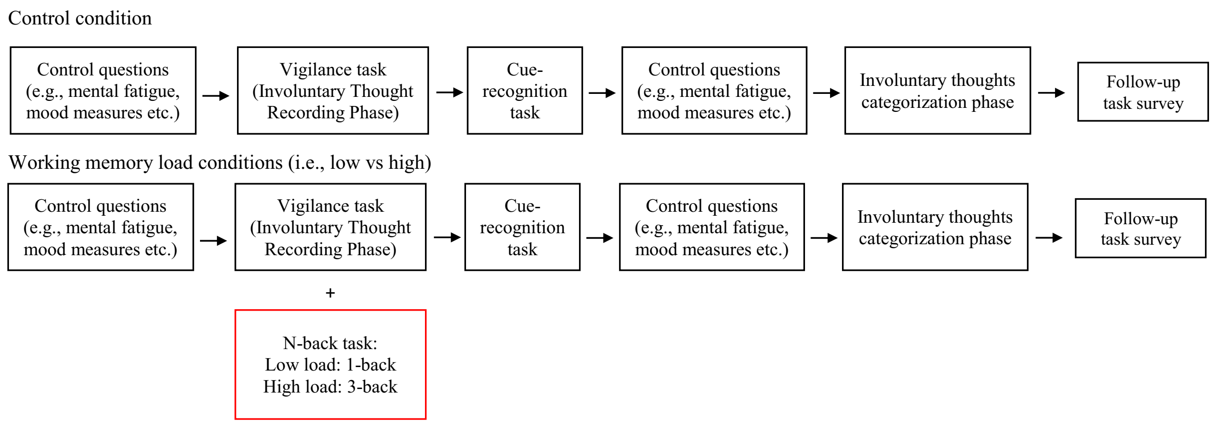
Next, participants completed the first practice session that consisted of 25 nontarget and two target stimuli, with no stop trials with thought probes. After this practice task, participants were informed that during the vigilance task they might experience different kinds of task-unrelated thoughts, and they were provided with examples of such thoughts. However, we did not put any particular emphasis on memories and future-oriented thoughts during the briefing. Participants were only informed that thoughts could be diverse (i.e., specific, general) and could pertain either to the present, past, or future. Importantly, participants were assured that these thoughts could be about anything and could simply pop into their mind spontaneously or they could deliberately choose to think about them. It was explained that since the study was about people's attention and concentration, the program would occasionally stop, and they would be asked to record their concentration level and the content of their thoughts at the exact moment they were stopped, regardless of what it was.

Participants then performed the second practice task that was the same as the first one with the addition of one stop trial. Once the practice session was completed and before the main vigilance task started, participants filled in the PANAS (Brzozowski, 2010) and rated their current level of physical and mental fatigue on 7-point scales (1 = *corresponding to not endorsing the item at all*; 4 = *corresponding to medium endorsement*; 7 = *corresponding to highly endorsing the item*; all points along the scales were clearly and explicitly labeled).[3] Each time the program stopped, participants provided a brief description of their thoughts (by typing it into the program), rated their concentration (on a 7-point scale), and indicated if the thought occurred deliberately (they decided to think about it) or involuntarily (it simply popped into their mind). After selecting either "involuntarily" or "deliberately," they also specified what triggered the content (1 = *something in the program*, 2 = *something in my mind*, 3 = *something in the surroundings*, 4 = *nothing*). Participants were also asked to provide a brief description of the trigger.

Just after completing the vigilance task, the cue-recognition task was introduced. It was explained to participants that they would be provided with verbal phrases and instructed to decide (*yes*/*no*)

[3] All points along the scale were clearly labeled during the task. For example, the scale for the physical fatigue was as follows: To what extent are you physically fatigued at this very moment? 1 = *I am not physically fatigued at all*, 2 = *I am not physically fatigued*, 3 = *I am rather not physically fatigued*, 4 = *I am somewhat physically fatigued*, 5 = *I am rather physically fatigued*, 6 = *I am physically fatigued*, 7 = *I am very physically fatigued*.

Figure 1
Overview of the Study Procedure



Note. See the online article for the color version of this figure.

whether each phrase had been presented to them during the vigilance task. Next, participants filled in PANAS and rated their current level of physical and mental fatigue for the second time. They responded also to additional manipulation-check questions. First, they answered open-ended questions concerning what they thought was the true goal of the study. They rated on a 7-point scale the extent to which the task was tiring, attention demanding, difficult to perform, and interesting.[4] Participants also rated the extent to which they had been concentrating on the task (in general), the verbal phrases, the vertical lines, and the square with letters, as well as the importance of performing the computer task as well as they could. Finally, participants also rated the extent to which verbal phrases and square letters were experienced as interfering and how much they suppressed (involuntary thoughts) or ignored (verbal phrases) them.

Participants were then provided with brief verbal instructions describing the nature of autobiographical memory (e.g., in Schlagman et al., 2009, p. 410) and future thoughts. During this involuntary thought categorization phase, participants reviewed all the thoughts they had recorded during the vigilance task. They did this one at a time and in the same order as they had been recorded; they were instructed to decide whether each thought was an autobiographical memory, future-oriented thought, thought relating to the current situation, or other type of thought by clicking the appropriately labeled button. In addition, they also answered several questions relating to each mental content (e.g., their phenomenological characteristics). Once all the mental contents were reviewed (and rated), the follow-up task questionnaire was provided assessing whether participants had adhered to the experiment instructions (e.g., whether the procedure was completed in an environment free from noise and distractors), but as the phenomenology ratings and follow-up task questionnaire were part of a separate study, neither the exact details nor the results of these separate questions will be reported here.

Low Working Memory Load Condition (Vigilance Task + 1-Back Verbal Task). The only difference between the low cognitive load and no cognitive load control conditions was that participants were performing an additional 1-back task while engaged in the vigilance task. Specifically, they were asked to respond each time a letter in the square, presented on the screen, was the same as the one presented on a previous slide showing the square with a letter in it. Importantly, participants familiarized themselves with the 1-back task during the first practice session. The threshold for proceeding to the next section of the study was 60% accuracy (used in both the first and second practice sessions). During practice sessions, participants' accuracy was displayed at the end of the practice session for feedback. The practice session was repeated until the required level of 60% accuracy was achieved by a given individual.

High Working Memory Load Condition (Vigilance Task + 3-Back Verbal Task). The only difference between the high and low cognitive load conditions was that participants were asked to perform a 3-back task instead of 1-back task. Specifically, they were asked to respond each time when a letter in a square was the same as the one presented three slides (with squares) earlier.

Results

For all statistical analyses reported below, the level of significance was set at $p < .05$, and the effect size was measured by partial eta-squared (η_p^2) with small, medium, and large effects defined as .01, .06, and .16, respectively (see Cohen, 1977). We closely followed the analyses outlined in the Stage 1 Registered Report. However, we also report nonpreregistered analyses in the Additional and Not-Pre-Registered Findings: Cue Recognition and Frequency of IAMs and IFTs section (Study 1) and Additional and Not-Pre-Registered Findings: Cue Recognition and Frequency of IAMs and IFTs section (Study 2). In addition, although all preregistered analyses were conducted as planned, some were moved to the Appendix to improve the clarity and flow of the Results section, as they were less central to the primary research questions. The data set from Studies 1 and 2 is published and described in detail in Barzykowski et al. (2024a, 2024c). For majority of analyses, the data were entered into a

[4] For example, the scale for difficulty was as follows: To what extent performing the vigilance task was difficult for you? 1 = *not difficult at all*, 2 = *not difficult*, 3 = *rather not difficult*, 4 = *somewhat difficult*, 5 = *rather difficult*, 6 = *it was difficult*, 7 = *it was very difficult*.

3 (working memory load: none, low, high) × 2 (cue presentation: simultaneous, sequential) between-subjects ANOVAs.

Equivalence of Experimental Groups Before, During, and After the Vigilance Task

Table 4 presents the mean ratings of various cognitive and noncognitive variables reported by participants before, during, and after the vigilance task. A detailed examination of the equivalence of experimental groups across these phases is provided in Appendix A. Briefly, the working memory load manipulation was effective: Participants in the high-load (3-back) condition consistently reported greater perceived task difficulty, lower concentration on task stimuli, and higher levels of thought suppression compared to the control condition. Across conditions, the vigilance task was associated with reductions in both positive and negative affect, as well as increases in mental and physical fatigue over time. Importantly, no systematic pretask differences between conditions were observed, indicating that the groups were comparable prior to the experimental manipulation.

Performance on the Vigilance Task

The results are presented in Table 4. We run a series of 3 (working memory load) × 2 (cue presentation) between-subjects ANOVAs on the mean proportions of correctly detected targets (i.e., vertical lines), response times to targets, and levels of self-rated concentration. Significant results were obtained only for response times to correctly detected targets. Thus, there was a main effect of working memory load with participants in the control condition responding significantly faster to vertical lines compared to the low (i.e., 1-back; p = .003) and high working memory load (i.e., 3-back; p = .011) conditions, which did not significantly differ from each other (p = .678). The Working Memory Load × Cue Presentation interaction was also significant. While there were no differences between sequential and simultaneous cue presentation in the 1-back and 3-back conditions, participants in the control condition were faster when targets were presented sequentially than simultaneously (p = .002). In summary, the findings indicate that the vigilance task was somewhat more challenging in the two working memory load conditions than in the control condition.

Performance on the *N*-Back Task

A 2 (working memory load: low, high) × 2 (cue presentation: simultaneous, sequential) between-subjects ANOVA was conducted on the proportion of correct *N*-back responses and *N*-back target response times (see Table 4). For the proportion of correct responses, there was a significant main effect of working memory load, with participants in the 1-back task outperforming those in the 3-back task (p = .001). The interaction between working memory load and cue presentation was also significant, indicating that accuracy in the 1-back condition did not differ significantly between simultaneous and sequential cue presentation (p = .053), whereas in the 3-back condition accuracy was significantly higher in the sequential than in the simultaneous presentation condition (p = .024). For *N*-back target response times, there was a significant main effect of working memory load, with slower responses in the 3-back than in the 1-back task (p = .001). The main effect of cue presentation was also significant, with sequential presentation resulting in longer response times than simultaneous presentation (p = .008). Overall, these results confirm that the 3-back task imposed a substantially higher cognitive demand than the 1-back task.

Frequency and Type of Recorded Thoughts

For data preparation and coding, we followed a comprehensive guide for investigating mind wandering and related phenomena such as involuntary past and future thinking by Barzykowski et al. (2024b). Thus, all recorded thoughts (n = 5,175) were coded independently by two judges as either task-unrelated (e.g., *swimming in a sea*), task-related (*don't forget to push the button*), or task-related interference (*this is so boring, I am tired*; cf. Plimpton et al., 2015; Smallwood, Obonsawin, & Reid, 2003; Smallwood, Obsonsawin, et al., 2003). Disagreements were solved by discussion. Thoughts coded as task-related or task-related interference (total n = 2,052), as well as thoughts classified as "nothing" (n = 137), were excluded from the main analyses presented below (but see Appendix B for analyses of spontaneous task-unrelated thoughts). Next, out of all task-unrelated thoughts (n = 2,986), 469 were classified by participants as occurring deliberately rather than coming to mind involuntarily. The exclusion of these task-unrelated deliberate thoughts resulted in 2,517 spontaneous task-unrelated thoughts which were the focus of analyses presented below.

Temporal Focus of Thoughts

After completing the vigilance task, participants were asked to decide whether each reported thought was a memory of a past event, a future-oriented thought, or something else. To ensure the thoughts had been categorized correctly by participants, all the entries were also screened by the judges (for a detailed guideline of categorization process, see Barzykowski et al., 2024b). However, if any of the thoughts coded by the judges as autobiographical memories or future thoughts were not originally identified as such by the participants, they were discussed by judges to arrive at a final decision. Such entries (e.g., *the first time I kissed a girl* coded by participant as "other thought" instead of "memory") were included in the analysis only if there were 100% agreement between judges. If judges disagreed, then the participant's original code was retained. The exact number of reevaluated entries was 500 out of the total 2,517 ITUTs.[5]

We then entered the mean number of IAMs and IFTs into two separate two-way ANOVAs with working memory load (none, low, high) and cue presentation (simultaneous, sequential) as between-subjects variables. The means are presented in Table 5 (see also Figure 2). This analysis showed that the working memory load significantly influenced the frequency of both IAMs and IFTs. For IAMs, post hoc comparisons showed that the number of reported involuntary thoughts about the past was significantly higher in the control condition (M = 4.89, SD = 3.38) than in the 1-back (M = 3.59, SD = 2.82; p = .010) and 3-back (M = 3.85, SD = 3.12;

[5] Please note that this procedure is identical to the one used in our previous studies (e.g., Barzykowski & Niedźwieńska, 2018a, p. 123; Barzykowski, Radel, et al., 2019, p. 676; Barzykowski et al., 2021, 2022). This additional coding is necessary because the decision whether a mental content is or is not a memory is irreversible, and this may result in some errors that the participants commit in the categorization task.

Table 4

Means and Standard Deviations of Background Task Performance Measures Across Conditions in Study 1

Outcome measure	Control: No working memory load — Sequential (N = 39)	Control: No working memory load — Simultaneous (N = 39)	Low working memory load (1-back task) — Sequential (N = 39)	Low working memory load (1-back task) — Simultaneous (N = 38)	High working memory load (3-back task) — Sequential (N = 35)	High working memory load (3-back task) — Simultaneous (N = 35)	Effect of working memory load (low, high, none)	Effect of cue presentation (simultaneous, sequential)	Interaction
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	Test		
Mood before vigilance task									
PANAS: Positive affect	42.13 (12.44)	36.77 (11.23)	41.87 (10.35)	39.97 (9.15)	36.40 (9.14)	40.63 (11.77)	Please refer to the Equivalence of Experimental Groups Before, During, and After the Vigilance Task section in the article		
PANAS: Negative affect	24.10 (8.41)	24.44 (10.16)	24.49 (11.19)	20.89 (6.16)	20.51 (8.23)	19.66 (7.54)			
Physical fatigue	3.46 (1.23)	3.56 (1.35)	3.23 (1.25)	3.13 (1.07)	3.34 (1.08)	3.14 (1.31)			
Mental fatigue	4.00 (1.36)	3.79 (1.59)	4.05 (1.47)	3.74 (1.45)	4.14 (1.38)	3.74 (1.29)			
Mood after vigilance task									
PANAS: Positive affect	34.28 (12.77)	32.44 (11.21)	31.41 (11.42)	33.03 (10.69)	28.83 (10.87)	30.40 (12.82)	Please refer to the Equivalence of Experimental Groups Before, During, and After the Vigilance Task section in the article		
PANAS: Negative affect	20.90 (6.50)	22.38 (10.03)	22.13 (9.55)	19.16 (6.50)	21.63 (7.40)	20.94 (8.22)			
Physical fatigue	3.92 (1.63)	4.05 (1.56)	4.59 (1.25)	3.71 (1.18)	4.23 (1.52)	4.20 (1.73)			
Mental fatigue	4.26 (1.31)	4.51 (1.27)	4.69 (1.28)	4.21 (1.14)	4.97 (1.34)	5.23 (1.17)			
Performance on vigilance task (self-reporting)									
Concentration on the task	5.05 (0.86)	4.90 (1.21)	5.05 (0.94)	5.00 (1.16)	5.11 (0.96)	5.37 (1.14)	$F(2, 219) = 1.33$ $\eta^2 = .01, p = .267$	$F(1, 219) = 0.02$ $\eta^2 = .01, p = .902$	$F(2, 219) = 0.75$ $\eta^2 = .01, p = .474$
Concentration on vertical lines	5.62 (1.29)	5.51 (1.45)	5.23 (1.80)	4.18 (1.86)	4.31 (1.71)	4.11 (2.19)	$F(2, 219) = 11.61$ $\eta^2 = .10, p = .001$	$F(1, 219) = 3.78$ $\eta^2 = .02, p = .053$	$F(2, 219) = 1.72$ $\eta^2 = .02, p = .182$
Concentration on phrases	4.44 (1.54)	4.13 (1.76)	3.87 (1.42)	3.50 (1.64)	3.11 (1.64)	2.94 (1.35)	$F(2, 219) = 11.81$ $\eta^2 = .10, p = .001$	$F(2, 219) = 1.84$ $\eta^2 = .01, p = .176$	$F(2, 219) = 0.08$ $\eta^2 = .01, p = .925$
Performance on vigilance task (self-reporting)									
How much were verbal phrases ignored	3.18 (1.52)	3.36 (1.51)	4.54 (1.52)	4.08 (1.76)	4.40 (1.97)	5.31 (1.69)	$F(2, 219) = 17.54$ $\eta^2 = .14, p = .001$	$F(1, 219) = 0.90$ $\eta^2 = .01, p = .343$	$F(2, 219) = 3.12$ $\eta^2 = .03, p = .046$
How much did verbal phrases interfere with the vigilance task	2.21 (1.32)	2.36 (1.29)	3.85 (1.51)	3.03 (1.67)	3.86 (1.68)	3.66 (1.61)	$F(2, 219) = 19.75$ $\eta^2 = .15, p = .001$	$F(1, 219) = 2.03$ $\eta^2 = .01, p = .155$	$F(2, 219) = 2.04$ $\eta^2 = .02, p = .132$
How much were involuntary thoughts being suppressed	3.13 (1.72)	3.15 (1.73)	3.41 (1.79)	2.97 (1.46)	3.86 (1.65)	4.46 (2.05)	$F(2, 219) = 7.86$ $\eta^2 = .07, p = .001$	$F(1, 219) = 0.07$ $\eta^2 = .01, p = .786$	$F(2, 219) = 1.63$ $\eta^2 = .01, p = .198$
Perceived task difficulty	3.03 (1.60)	2.92 (1.16)	3.54(1.39)	3.26 (1.33)	5.31 (1.18)	5.34 (1.30)	$F(2, 219) = 63.75$ $\eta^2 = .37, p = .001$	$F(1, 219) = 0.42$ $\eta^2 = .01, p = .515$	$F(2, 219) = 0.24$ $\eta^2 = .01, p = .788$
Importance of performing the task well	5.92 (0.90)	5.69 (1.06)	5.82 (0.94)	6.34 (0.81)	5.86 (1.03)	6.23 (0.77)	$F(2, 219) = 1.97$ $\eta^2 = .02, p = .142$	$F(1, 219) = 3.19$ $\eta^2 = .01, p = .076$	$F(2, 219) = 3.56$ $\eta^2 = .03, p = .030$
How interesting was the task	3.79 (1.45)	3.77 (1.61)	3.67 (1.38)	3.66 (1.49)	3.60 (1.50)	3.74 (1.56)	$F(2, 219) = 0.15$ $\eta^2 = .01, p = .860$	$F(1, 219) = 0.03$ $\eta^2 = .01, p = .857$	$F(2, 219) = 0.07$ $\eta^2 = .01, p = .933$
Performance on vigilance task (objective indicators)									
Proportion of vigilance targets detected	0.95 (0.09)	0.95 (0.10)	0.96 (0.08)	0.97 (0.07)	0.97 (0.05)	0.96 (0.07)	$F(2, 219) = 0.88$ $\eta^2 = .01, p = .417$	$F(1, 219) = 0.04$ $\eta^2 = .01, p = .838$	$F(2, 219) = 0.27$ $\eta^2 = .01, p = .764$
Correct response reaction time (in section) to vigilance targets	0.71 (0.12)	0.82 (0.19)	0.86 (0.17)	0.81 (0.14)	0.84 (0.14)	0.81 (0.11)	$F(2, 219) = 5.35$ $\eta^2 = .05, p = .005$	$F(1, 219) = 0.31$ $\eta^2 = .01, p = .579$	$F(2, 219) = 6.19$ $\eta^2 = .05, p = .002$

(*table continues*)

Table 4 (*continued*)

	Working memory load condition								
	Control: No working memory load		Low working memory load (1-back task)		High working memory load (3-back task)				
	The presentation of cues and working memory load stimuli								
	Sequential ($N = 39$)	Simultaneous ($N = 39$)	Sequential ($N = 39$)	Simultaneous ($N = 38$)	Sequential ($N = 35$)	Simultaneous ($N = 35$)	Effect of working memory load (low, high, none)	Effect of cue presentation (simultaneous, sequential)	Interaction
Outcome measure	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	Test		
Concentration rating (probe measured)	4.50 (1.18)	4.20 (1.27)	4.35 (0.96)	4.38 (1.31)	4.24 (0.99)	4.45 (0.99)	$F(2, 219) = 0.01$ $\eta^2 = .01, p = .995$	$F(1, 219) = 0.02$ $\eta^2 = .01, p = .898$	$F(2, 219) = 0.96$ $\eta^2 = .01, p = .384$
Correct N-back response			0.93 (0.08)	0.97 (0.03)	0.76 (0.12)	0.71 (0.12)	$F(1, 143) = 182.19$ $\eta^2 = .56, p = .001$	$F(1, 143) = 0.09$ $\eta^2 = .01, p = .763$	$F(1, 143) = 8.94$ $\eta^2 = .06, p = .003$
Correct N-back response time			0.78 (0.13)	0.69 (0.10)	1.02 (0.16)	0.98 (0.18)	$F(1, 143) = 124.81$ $\eta^2 = .47, p = .001$	$F(1, 143) = 7.25$ $\eta^2 = .05, p = .008$	$F(1, 143) = 0.95$ $\eta^2 = .01, p = .330$
Successful cue recognition	0.64 (0.08)	0.63 (0.08)	0.62 (0.09)	0.60 (0.07)	0.62 (0.07)	0.57 (0.06)	$F(2, 219) = 4.86$ $\eta^2 = .04, p = .009$	$F(1, 219) = 5.59$ $\eta^2 = .02, p = .019$	$F(2, 219) = 1.43$ $\eta^2 = .01, p = .242$
Attention checks	0.82 (0.14)	0.78 (0.18)	0.83 (0.11)	0.79 (0.24)	0.82 (0.12)	0.85 (0.11)	$F(2, 219) = 0.93$ $\eta^2 = .01, p = .398$	$F(1, 219) = 0.44$ $\eta^2 = .01, p = .509$	$F(2, 219) = 0.90$ $\eta^2 = .01, p = .410$
Number of IAMs	$r(37) = .01$ $p = .987$	$r(39) = .27$ $p = .093$	$r(39) = .32$ $p = .048$	$r(37) = .36$ $p = .030$	$r(34) = .31$ $p = .075$	$r(34) = -.02$ $p = .913$			
Number of IFTs	$r(37) = -.10$ $p = .552$	$r(39) = .09$ $p = .592$	$r(39) = -.06$ $p = .710$	$r(337) = .34$ $p = .042$	$r(34) = .15$ $p = .400$	$r(34) = .36$ $p = .036$			

Note. Study 1: Means and standard deviations for variables measuring mood, fatigue, concentration, motivation, and performance on the vigilance task, working memory task as a function of working memory load condition (low, high, none), and presentation of cues and working memory stimuli (simultaneous, sequential). Additional and not preregistered findings: Correlations (Pearson's correlation coefficient) between successful cue-recognition and the frequency of IAMs and IFTs across groups and conditions in Study 1.All self-reported questions except PANAS were rated on 7-point scales (1 = *low* to 7 = *high*). Significant results are highlighted in grey. PANAS = Positive and Negative Affect Schedule; IAMs = involuntary autobiographical memories; IFTs = involuntary future thoughts.

Table 5
Study 1: Frequency of Involuntary Thoughts as a Function of Working Memory Load and Cue Presentation

Involuntary thought category	Working memory load condition												Effect of working memory load (none, low, high)	Effect of cue presentation (simultaneous, sequential)	Interaction
	Control: No working memory load				Low working memory load (1-back task)				High working memory load (3-back task)						
	The type of working memory load														
	Sequential (N = 37)		Simultaneous (N = 39)		Sequential (N = 39)		Simultaneous (N = 37)		Sequential (N = 34)		Simultaneous (N = 34)			Test	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*			
All involuntary task-unrelated thoughts	13.65	4.66	12.13	4.95	12.03	5.17	11.22	4.44	9.47	4.69	9.79	5.03	$F(2, 214) = 8.22$, $\eta^2 = .07$, $p = .001$	$F(1, 214) = 1.05$, $\eta^2 = .01$, $p = .307$	$F(2, 214) = 0.66$, $\eta^2 = .01$, $p = .518$
Involuntary task-oriented and task-interference thoughts	6.67	4.64	6.19	3.32	6.58	3.82	7.37	4.39	8.67	4.95	8.97	5.69	$F(2, 210) = 5.36$, $\eta^2 = .05$, $p = .005$	$F(1, 210) = 0.11$, $\eta^2 = .01$, $p = .738$	$F(2, 210) = 0.37$, $\eta^2 = .01$, $p = .688$
Involuntary autobiographical memories	5.57	3.28	4.26	3.39	3.77	2.90	3.41	2.75	3.65	2.59	4.06	3.59	$F(2, 214) = 3.84$, $\eta^2 = .03$, $p = .023$	$F(1, 214) = 1.01$, $\eta^2 = .01$, $p = .316$	$F(2, 214) = 1.39$, $\eta^2 = .01$, $p = .251$
Involuntary future-oriented thoughts	4.00	3.10	4.31	3.44	4.05	3.14	4.32	3.25	2.44	2.20	2.56	1.78	$F(2, 214) = 7.73$, $\eta^2 = .07$, $p = .001$	$F(1, 214) = 0.35$, $\eta^2 = .01$, $p = .555$	$F(2, 214) = 0.02$, $\eta^2 = .01$, $p = .979$

Note. Study 1: Means and standard deviations for the number of involuntary task-unrelated and task-related thoughts, involuntary autobiographical memories and future-oriented thoughts as a function of working memory load (low, high, none), and presentation of cues and working memory task stimuli (simultaneous, sequential). Significant results are highlighted in grey.

$p = .046$) task conditions, which did not differ from each other ($p = .615$). For IFTs, their reported frequency was significantly higher in the control ($M = 4.16$, $SD = 3.26$) and the low working memory load conditions ($M = 4.18$, $SD = 3.17$) compared to the 3-back condition ($M = 2.50$, $SD = 1.99$; $ps < .001$), but no significant differences were observed between the control and low working memory load conditions ($p = .956$). Importantly, neither the main effect of cue presentation nor its interaction with the working memory load was significant for IAMs (Fs < 1.39) and IFTs (Fs < 1).

Since we observed a null result for the main effect of cue presentation (sequential vs. simultaneous), we supplemented this analysis by examining Bayes factors (BFs) to better assess the results as favoring the null or alternative hypothesis. The Bayesian ANOVA provided a BF_{10} of 0.237 and 0.010 for the main effect of cue presentation on the frequency of IAMs and IFTs, respectively. Following Wetzels et al.'s (2011) recommendations for the interpretation of BFs, this suggests substantial evidence in favor of the null hypothesis.

Cue-Recognition Ratio Across Participants and Conditions

The mean cue-recognition ratio, which refers to the percentage of cues correctly identified as either previously presented or not presented during the task, was entered into a two-way ANOVA with working memory load and cue presentation as a between-subjects variables (see Table 4; see also Figure 3). This analysis resulted in the significant main effect of working memory load with participants in the control condition showing superior performance in the cue-recognition task compared to the 3-back task ($p = .002$) but not compared to the 1-back task conditions ($p = .121$). The main effect of cue presentation was also significant. Specifically, sequential cue presentation resulted in higher cue recognition ratio ($M = 0.63$, $SD = 0.08$) compared to simultaneous cue presentation ($M = 0.60$, $SD = 0.08$), indicating that presenting verbal cues and N-back stimuli simultaneously on some of the vigilance task slides may have led to poorer cue processing. The Working Memory Load × Cue Presentation interaction effect was not significant.

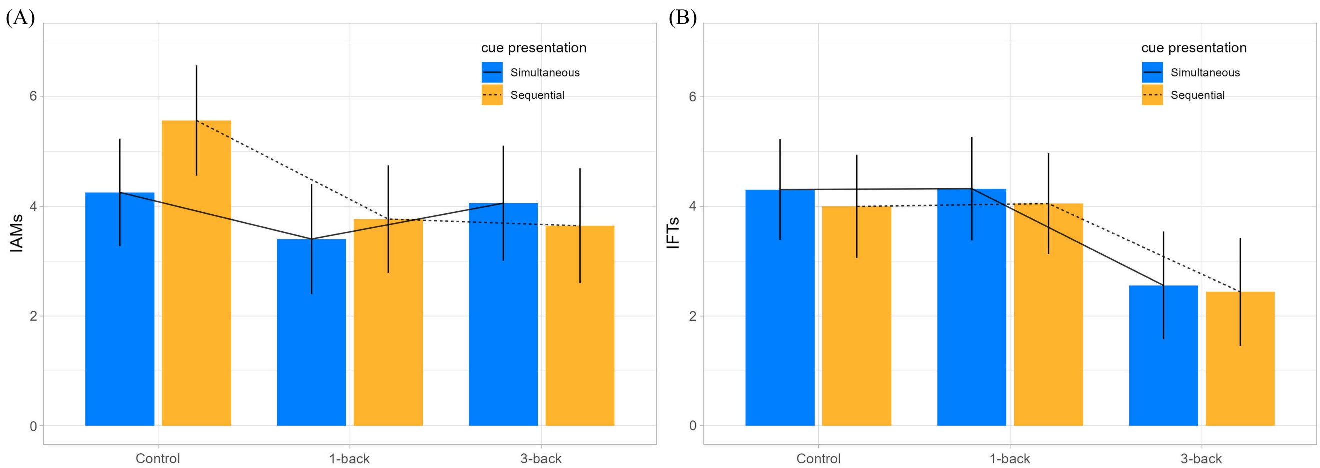
A more detailed analysis of cue-recognition performance as a function of cue valence (positive, neutral, negative) is provided in Appendix C, where neutral cues were found to be recognized less accurately than emotional cues.

Additional and Not-Pre-Registered Findings: Cue Recognition and Frequency of IAMs and IFTs

Similar to Barzykowski et al. (2022), we also examined the relationship between successful cue-recognition ratio and the frequency of IAMs and IFTs. First, across all groups and conditions, the recognition ratio was positively and significantly correlated with the number of IAMs, $r(220) = .21$, $p = .002$, but not IFTs, $r(220) = .12$, $p = .078$, indicating that participants who recognized more cues correctly (presumably because they were noticed during the vigilance task) were also reporting higher number of IAMs than participants who recognized fewer cues presented in the vigilance task (for correlations across groups and conditions, see Table 4). Second, we used the linear regression model to further examine the relationship between the frequency of IAMs and IFTs but with control of cognitive load condition. The relationship between IAMs and cue

Figure 2

Frequency of Involuntary Autobiographical Memories and Future Thoughts Across Conditions



Note. Study 1: The mean number of (Panel A) IAMs and (Panel B) IFTs as a function of working memory load (none, low, high) and presentation of cues and working memory task stimuli (simultaneous, sequential). Error bars represent 95% confidence intervals. IAMs = involuntary autobiographical memories; IFTs = involuntary future thoughts. See the online article for the color version of this figure.

recognition remains statistically significant and positive, $t(216) = 2.88$, $p = .004$, while between IFTs and cue recognition remains statistically insignificant, $t(216) = 1.20$, $p = .231$. Importantly, models including the interaction of cue recognition and cognitive load were not significantly better than simpler models explaining both IAMs, $F(2, 214) = 0.72$, $p = .488$, and IFTs, $F(2, 214) = 0.50$, $p = .609$, which suggests that the relationship between cue recognition and IAMs/IFTs frequency is similar regardless of the level of manipulation of cognitive load. Therefore, the visualizations present models without interaction effects (see Figure 3B–3D).

Discussion

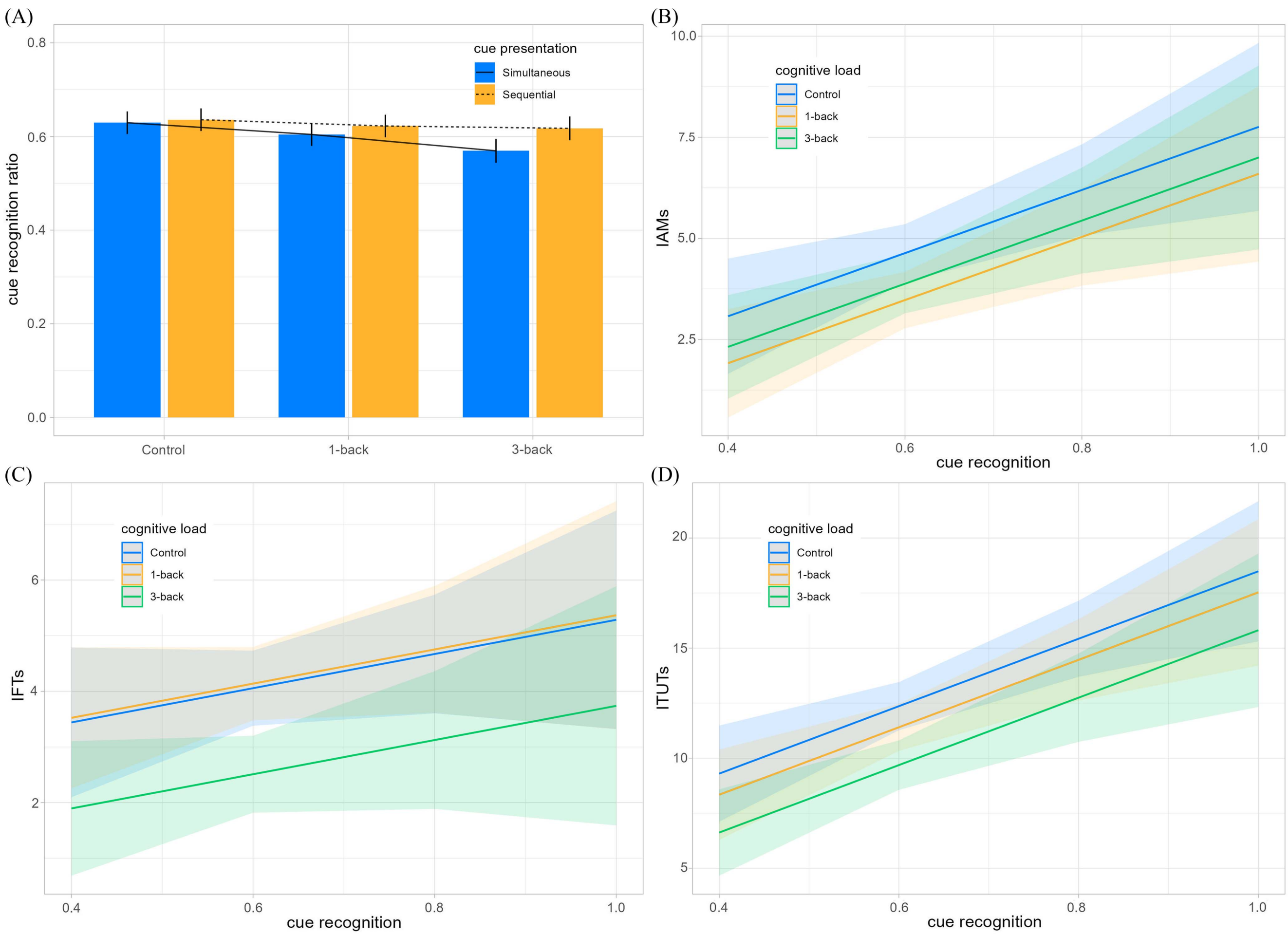
Study 1 investigated possible underlying mechanisms of the cognitive load effect, which typically results in a decreased number of involuntary thoughts in general and IAMs and IFTs in particular. More specifically, we aimed to examine whether the reduced frequency of these spontaneous thoughts is primarily due to working memory load per se or whether it also reflects a reduced likelihood of noticing incidental cues that could trigger such thoughts. To address this question, we manipulated working memory load using a well-established task (*N*-back) and varied cue presentation. In the simultaneous cue-presentation condition, verbal cue phrases were embedded within the *N*-back task slides, reducing the likelihood of being noticed. In the sequential cue-presentation condition, verbal cues were presented separately from the slides presenting the *N*-back task, maximizing the chances of being noticed by participants. To further assess cue awareness, we included an unexpected cue-recognition test. As summarized in Table 2, we expected to observe robust main effects of working memory load with fewer spontaneous thoughts under high load and the strongest reduction in involuntary thoughts in the simultaneous condition. Importantly, beyond manipulating key parameters for IAMs and IFTs retrieval (i.e., working memory load and cue presentation), we also examined background variables reflecting participants' own responses and strategies across conditions, providing insight into how individuals managed this cognitively demanding activity.

Before turning to the discussion of results, it is important to establish that we successfully manipulated both cognitive load and cue presentation. Specifically, subjective ratings of task difficulty and objective performance indicators (e.g., RTs for correct responses in the vigilance and *N*-back tasks) confirmed that working memory load conditions were more challenging than control conditions. As intended, high working memory load was consistently associated with greater perceived and objective task difficulty. Creating these distinct load conditions was therefore instrumental for assessing the cognitive mechanisms involved in IAMs and IFTs retrieval. Importantly, by collecting additional measures related to background variables (e.g., the extent to which verbal phrases were ignored or involuntary thoughts were deliberately suppressed), we were able to capture aspects of participants' cognitive strategies. These variables suggest that participants did not merely respond passively to the imposed cognitive constraints but also engaged in regulatory efforts that may have shaped the frequency of IAMs and IFTs. For instance, high cognitive load was consistently associated with lower reported concentration on verbal phrases, while in control conditions, verbal phrases were judged as least interfering with the vigilance task and least likely to be ignored. Finally, participants reported the strongest thought suppression under high working memory load. Taken together, these findings indicate that, beyond automatic processes outside of participants' control (e.g., working memory limitations, priming, automated cue noticing[6]), individuals may also strategically attempt to regulate the occurrence of spontaneous thoughts. This suggests that part of the cognitive load effect on reducing IAMs and IFTs may actually stem from the strategies individuals employ to cope with increased task demands.

[6] As people may also engage in deliberate strategies to (not) notice cues, here we specifically refer to cue noticing as occurring outside of one's voluntary control.

Figure 3

Cue Recognition Across Conditions and Its Relationship With IAMs and IFTs in Study 1



Note. Study 1: Successful cue recognition as (Panel A) a function of working memory load condition (none, low, high) and presentation of cues and working memory stimuli (simultaneous, sequential), and as a predictor of the number of (Panel B) IAMs, (Panel C) IFTs, and (Panel D) involuntary task-unrelated thoughts (ITUTs). Shaded areas and erros bars represents 95% confidence intervals. IAMs = involuntary autobiographical memories; IFTs = involuntary future thoughts. See the online article for the color version of this figure.

As to the *effects of working memory load*, our results first replicated well-established findings from previous studies (e.g., Barzykowski et al., 2022; Barzykowski & Niedźwieńska, 2018a; Mazzoni, 2019; Vannucci et al., 2015): Participants were more likely to report spontaneous task-unrelated thoughts in the low-load (1-back and control) than in the high-load (3-back) conditions. Additionally, when considering task-related and task-interference thoughts, these were reported more frequently in the 3-back condition. This suggests that a demanding working memory task not only taxes resources during performance but also induces more thoughts about the task itself, thereby limiting cognitive space for other types of thoughts, including IAMs and IFTs. Given the verbal nature of the working memory load task in our study, it may have been particularly prone to triggering thoughts related to task material (e.g., the letters in the *N*-back task). Future research could benefit from employing more procedural, nonverbal tasks to examine whether such tasks reduce this overlap and provide a clearer test of cognitive load effects.

Importantly, we also observed differences between IAMs and IFTs. Specifically, IAMs were reduced in both low and high working memory load conditions relative to control, suggesting that even a moderate increase in load disproportionately affects IAM retrieval. By contrast, IFTs were reduced only under high working memory load. This pattern is consistent with Barzykowski et al. (2022), who also found that IFTs were less strongly disrupted by concurrent tasks than IAMs. In summary, this pattern of findings suggests a differential sensitivity of IAMs and IFTs to cognitive load. It may be speculated that greater engagement in task-oriented and task-interference thoughts reduce the availability of cognitive resources for IAMs retrieval, which in turn may create more space for the emergence of IFTs. Alternatively, future-oriented thoughts may more often concern duties or plans, making them particularly relevant for monitoring current goals and thus less susceptible to the disruptive effects of cognitive load. This interpretation is consistent with evidence that prospective thoughts are closely tied to goal management and self-regulation (Cole & Kvavilashvili, 2021; Kvavilashvili & Rummel, 2020), which may grant them a higher degree of resilience compared to autobiographical memories. These possibilities clearly warrant further investigation, and we discuss them in more detail in the general discussion.

As to the *effects of cue presentation*, contrary to our expectations, we did not observe significant effects of cue presentation neither on involuntary thoughts in general nor on IAMs and IFTs in particular. As BFs provided only anecdotal evidence for the null hypothesis, this finding may be worth replicating in future studies with different types of tasks and cue-presentation manipulations. In our study, the absence of effects may suggest that beyond merely noticing a cue, it may be necessary to engage with it at a deeper cognitive level (Craik & Lockhart, 1972). Even when a cue is noticed, high cognitive load may leave insufficient working memory resources to elaborate its meaning, connect it to autobiographical memory, launch spreading activation, and eventually achieve successful retrieval. It is also possible that simply increasing the number of noticed cues—by manipulating cue frequency—does not genuinely enhance the likelihood of "inspiring" involuntary thoughts. Indeed, Vannucci et al. (2015) demonstrated that when cues become too numerous, they no longer increase the number of thoughts but instead lead to a decrease. Alternatively, our manipulation may not have been sufficiently strong to prevent participants from noticing the cue words even in the simultaneous condition. Although the cue-word recognition task indicated that words were better recognized in the sequential than in the simultaneous condition, this task primarily reflects long-term recognition processes, which may differ from the more immediate cue noticing and spreading activation that occur during task performance. Future studies could verify this explanation by systematically manipulating cognitive load and cue noticing while additionally assessing potential priming effects resulting from cue exposure.

Finally, in relation to *cue noticing measured by recognition*, performing a highly demanding task (3-back) significantly impaired cue recognition. Moreover, recognition performance was better under sequential than simultaneous cue presentation, supporting the idea that simultaneous presentation leads to poorer cue processing. Importantly, the findings suggest that it is necessary to control for the possibility that ongoing task demands may overlap with, or suppress, the processing of cues when studying the retrieval processes of IAMs and IFTs (cf. Barzykowski & Niedźwieńska, 2018a). Thus, to capture the "pure" cognitive mechanisms of spontaneous thought retrieval, future studies should minimize such confounds by separating stimuli across conditions and ensuring that task-related materials do not interfere with the cues of interest.

Taken together, the findings suggest that when individuals are engaged in cognitively demanding tasks, cues may be processed less efficiently, which in turn may reduce the likelihood of triggering IAMs (but not IFTs). This interpretation is further supported by additional, exploratory analyses showing that cue recognition correlated significantly with IAMs frequency, even when controlling for working memory load. However, cue recognition cannot fully explain the cognitive load effect: While recognition performance in the 3-back task was poorer than in the control condition, no such difference was observed between the 1-back and control conditions—yet IAMs were already reduced in the 1-back condition. In general, the pattern of finding observed in Study 1 suggests additive effects of both working memory load, cue noticing, and participants' strategies on IAMs occurrence.

Study 2

The aim of Study 2 was to replicate and extend the results of Study 1 by manipulating different types of working memory load. Thus, in addition to using the same verbal working memory task as used in Study 1, we also used a visuospatial working memory task. In this task, participants were provided with a location-changing letter-square, and they were asked to respond only when the current letter-square was presented in the same position as it had been presented *N*-trials before (see Table 1). By manipulating the type (verbal, visuospatial) and level of working memory load (none, low, high) while controlling the likelihood of noticing the cues across verbal and visuospatial working memory conditions by presenting cues and working memory task items sequentially, we were able to directly verify the assumption that the limited working memory capacity enables the processing of only a finite number of thoughts at any given time and that this relationship may be restricted to a certain type of working memory component used. More specifically, we tested the idea that the retrieval of IAMs and IFTs depends on different working memory components. Therefore, IAMs should be least frequently observed in the high visuospatial working memory load condition while IFTs should be least frequently observed in the high verbal working memory load condition.

These predictions were based on the assumption that performance in 1- and 3-back conditions would be comparable across verbal and visual working memory tasks and hence any differences obtained in the number of reported IAMs and IFTs in these two conditions would not be due to the fact that one type of working memory task (e.g., visuospatial) was more demanding than the other (e.g., verbal) as measured by performance accuracy on these tasks and by self-rated levels of concentration, motivation, fatigue, and so forth.

To ensure the comparability across the working memory task conditions, we conducted a pilot study with two conditions to be used in the main study: (a) the verbal 3-back task and (b) the visuospatial 3-back task, with 20 participants in each task condition and without using thought probes. Next, we ran a series of independent *t* tests on (a) the mean ratings of various cognitive and noncognitive variables (those presented in Table 4) provided by participants before, during, or after the vigilance task; (b) on the mean proportions of vigilance task targets (i.e., vertical lines) detected, response times to targets, and levels of concentration; and (c) the proportion of correct 3-back responses and the 3-back response times.

First, when comparing two conditions in terms of pre- and posttask mood (e.g., PANAS positive and negative, physical and psychological fatigue, general well-being) and variables related to vigilance task performance (e.g., self-reported: importance of the task, perceived task difficulty, concentration on lines and verbal phrases, verbal cue ignorance, concentration engagement), we did not observe any significant differences (all $ps > .072$) except for two variables: the perceived level of interest in the task, $t(38) = 2.07$, $p = .046$, Cohen's $d = 0.65$, and the degree of interference from verbal cues, $t(38) = 2.08$, $p = .044$, Cohen's $d = 0.66$. In both cases, the visuospatial 3-back task was perceived as more interesting and associated with a higher level of perceived verbal cue interference. Second, when comparing the two conditions in terms of objective parameters such as percentage of correct *N*-back response (verbal = .75, visuospatial = .69) and correct *N*-back response time (verbal = 1.00 s, visuospatial = 1.01 s), no significant differences were observed (all $ps > .305$). In summary, these findings clearly demonstrated that the verbal and visuospatial conditions were equally difficult and attention demanding, leading us to select them for use in Study 2.

Method

Design

In Study 2, we employed a 2 (type of working memory load: verbal, spatial) × 3 (working memory load: none, low, high) between-subjects factorial design. This allowed us to examine the effects of working memory type and working memory load in the retrieval of IAMs and IFTs while keeping exposure to incidental verbal cue phrases constant across the conditions. The working memory type was operationalized by engaging participants (or not engaging as in the control condition) in either a spatial or verbal version of the *N*-back task, and the working memory load was operationalized by engaging participants (or not as in the control condition) in verbal/visual versions of 1-back (low working memory load) or 3-back tasks (high working memory load). Table 3 shows an overview of all experimental groups in Study 2.

Participants

As in Study 1, 240 participants (age range 18–65 years; for details see Table 3) took part in Study 2 in return for modest financial reward (PLN 50, ca. $13). As indicated in Table 3, they were randomly assigned to one of the six conditions: (a) no working memory task, verbal or spatial *N*-back task stimuli displayed on the screen; (b) low working memory conditions with a verbal or spatial 1-back task; and (c) high working memory conditions with a verbal or spatial 3-back task.

Materials and Procedure

As shown in Table 1, Study 2 used two types of *N*-back tasks: the verbal *N*-back and spatial *N*-back task. The verbal *N*-back task was the same as the one used in Study 1 with incidental word cues and the*N*-back task stimuli presented on separate slides. While the spatial *N*-back task was based on the same parameters as the verbal task used in Study 1 (e.g., each slide was presented for 2 s with a 0.5 s interstimulus intervals), the stimuli (square with capital letter “K”) were presented at one of the nine different locations on the computer screen (see Table 1). Specifically, participants were instructed to respond whenever the location of the letter-square presented on the screen matched either with the last presented location (low working memory load) or with the third previously presented location (high working memory load) in the sequence. As the square was presented only on slides where no cue phrase was provided, there were a total of 515 locations from which 176 were targets and 345 served as nontargets (foils). Finally, as in Study 1, we also tested participants recognition memory for incidental verbal cues presented in the vigilance task.

Results

We repeated all the analyses as in Study 1. For instance, the number of IAMs and IFTs was entered into a two-way ANOVA with the type of working memory *N*-back task (verbal, spatial) and the working memory load condition (none, low, high) as between-subjects variables (see Tables 6 and 7). Finally, to test the assumption of different effects of working memory components (verbal, spatial) on IAMs and IFTs retrieval, the number of IAMs and IFTs was entered into separate 2 (type of working memory load: verbal, spatial) × 3 (working memory load condition: none, low, high) × 2 (type of thought: IAMs vs. IFTs) mixed ANOVA with repeated measures on the last factor.

Equivalence of Experimental Groups Before, During, and After the Vigilance Task

A detailed examination of the equivalence of experimental groups before, during, and after the vigilance task is provided in Appendix D. In brief, the vigilance task was associated with reductions in both positive and negative affect and increases in mental and physical fatigue over time, replicating the pattern observed in Study 1. Importantly, no systematic pretask differences between conditions were observed. The working memory load manipulation was effective. Participants in the high working memory load (3-back) condition consistently reported greater perceived task difficulty and reduced concentration on task-relevant stimuli. Notably, they reported being least concentrated on incidental verbal phrases compared to both the control and low working memory load conditions. Furthermore, participants in the high load condition reported higher levels of ignoring verbal phrases and greater suppression of involuntary thoughts relative to the other conditions. Together, these findings confirm that the high working memory load condition imposed greater cognitive demands and altered participants’ attentional engagement with task stimuli.

Performance on the Vigilance Task

The results are presented in Table 6. As in Study 1, the vigilance task was easy, with success rates ranging between 95% and 97% across all conditions. A series of 2 (type of working memory task: verbal, spatial) × 3 (working memory load: control, low, high) ANOVAs were conducted on (a) the mean proportions of correctly detected targets (i.e., vertical lines) and response times to these targets and (b) the levels of probe-based concentration ratings.

Mean Proportions of Correctly Detected Targets and Response Times. The main effects and interactions for the proportion of correctly detected targets were not significant, indicating that performance accuracy in detecting vertical lines was stable across all conditions. For response times to correctly detected targets, several significant effects emerged. A main effect of type of working memory task showed that participants in the spatial working memory condition responded significantly more slowly to vertical lines than those in the verbal condition. This suggests that spatial working memory conditions were objectively more demanding. One possible explanation is that, in the spatial version, participants needed to scan the entire screen to locate the target, which likely increased processing time. In contrast, in the verbal version—where stimuli were consistently presented at the center—participants could maintain a more stable gaze, leading to faster responses. Environmental factors, such as visual distractors preceding target presentation, may also have contributed to these differences in response times. A significant main effect of *working memory load* was also observed. Participants in the control condition responded significantly faster to vertical lines compared to those in the 1-back ($p = .001$) and 3-back ($p = .025$) conditions, while the latter two did not differ from each other ($p = .190$). Thus, higher working

Table 6
Means and Standard Deviations of Background Task Performance Measures Across Conditions in Study 2

	The type of working memory load								
	Verbal working memory load			Spatial working memory load					
	Working memory load condition								
	Control: No working memory load (N = 40)	Low working memory load (N = 39)	High working memory load (N = 36)	Control: No working memory load (N = 40)	Low working memory load (N = 38)	High working memory load (N = 32)	Effect of type of working memory load (verbal, spatial)	Effect of working memory load (none, low, high)	Interaction
Outcome measure	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	Test		
Mood before vigilance task									
PANAS: Positive affect	39.65 (11.31)	39.33 (10.23)	38.22 (10.85)	40.10 (9.73)	38.82 (8.94)	44.97 (12.86)	Please refer to the Equivalence of Experimental Groups Before, During, and After the Vigilance Task section in the article (Study 2)		
PANAS: Negative affect	22.48 (9.28)	21.72 (9.27)	22.81 (10.50)	21.20 (6.85)	20.39 (9.76)	19.34 (5.12)			
Physical fatigue	3.35 (1.12)	3.10 (1.14)	3.31 (1.09)	3.38 (1.37)	3.26 (1.39)	3.19 (1.42)			
Mental fatigue	3.65 (1.33)	3.92 (1.33)	3.61 (1.50)	4.08 (1.31)	3.53 (1.29)	3.25 (1.41)			
Mood after vigilance task									
PANAS: Positive affect	32.98 (10.77)	30.67 (9.24)	30.06 (9.60)	32.85 (10.75)	32.97 (12.73)	34.66 (13.97)	Please refer to the Equivalence of Experimental Groups Before, During, and After the Vigilance Task section in the article (Study 2)		
PANAS: Negative affect	7.76 (1.23)	8.69 (1.39)	8.96 (1.49)	5.49 (0.87)	4.68 (0.76)	4.88 (0.86)			
Physical fatigue	3.80 (1.20)	3.97 (1.61)	4.11 (1.37)	4.38 (1.50)	4.00 (1.56)	3.50 (1.48)			
Mental fatigue	4.05 (1.36)	4.69 (1.08)	4.75 (1.40)	4.60 (1.37)	4.05 (1.41)	4.28 (1.37)			
Performance on vigilance task (self-reporting)									
Concentration on the task	4.58 (1.13)	5.18 (1.12)	5.23 (0.84)	4.73 (1.40)	4.87 (1.09)	5.28 (1.08)	F(1, 218) = 0.06 η^2 = .01, p = .812	F(2, 218) = 5.41 η^2 = .05, p = .005	F(2, 218) = 0.89 η^2 = .01, p = .411
Concentration on vertical lines	5.63 (1.19)	5.08 (1.87)	4.20 (2.07)	5.60 (1.41)	4.13 (1.86)	5.00 (2.16)	F(1, 218) = 0.06 η^2 = .01, p = .811	F(2, 218) = 8.36 η^2 = .07, p = .001	F(2, 218) = 4.37 η^2 = .04, p = .014
Concentration on phrases	4.30 (1.32)	3.92 (1.40)	3.26 (1.12)	4.03 (1.67)	3.71 (1.41)	3.34 (1.43)	F(1, 218) = 0.50 η^2 = .01, p = .480	F(2, 218) = 6.86 η^2 = .06, p = .001	F(2, 218) = 0.33 η^2 = .01, p = .717
Performance on vigilance task (self-reporting)									
How much were verbal phrases ignored	2.95 (1.62)	3.85 (1.68)	4.51 (1.67)	3.50 (1.68)	4.18 (1.47)	4.28 (1.71)	F(1, 218) = 0.99 η^2 = .01, p = .321	F(2, 218) = 9.97 η^2 = .08, p = .001	F(2, 218) = 1.09 η^2 = .01, p = .338
How much did verbal phrases interfere with the vigilance task	2.30 (1.59)	3.36 (1.56)	4.37 (1.66)	2.43 (1.52)	3.61 (1.65)	3.88 (1.64)	F(1, 218) = 0.04 η^2 = .01, p = .846	F(2, 218) = 22.98 η^2 = .17, p = .001	F(2, 218) = 1.09 η^2 = .01, p = .339
How much were involuntary thoughts being suppressed	3.15 (1.56)	4.13 (1.69)	3.91 (1.80)	3.53 (1.75)	3.08 (1.76)	4.16 (1.72)	F(1, 218) = 0.39 η^2 = .01, p = .532	F(2, 218) = 3.04 η^2 = .03, p = .050	F(2, 218) = 4.02 η^2 = .04, p = .019
Perceived task difficulty	2.90 (1.46)	3.38 (1.41)	5.20 (1.21)	3.30 (1.54)	3.24 (1.53)	5.06 (1.24)	F(1, 218) = 0.04 η^2 = .01, p = .840	F(2, 218) = 44.05 η^2 = .29, p = .001	F(2, 218) = 0.95 η^2 = .01, p = .389
Importance of performing the task well	5.55 (1.24)	6.08 (0.77)	6.06 (0.73)	5.95 (1.11)	5.89 (0.83)	6.00 (0.92)	F(1, 218) = 0.17 η^2 = .01, p = .677	F(2, 218) = 1.87 η^2 = .02, p = .157	F(2, 218) = 2.00 η^2 = .02, p = .138
How interesting was the task	4.10 (1.58)	3.92 (1.55)	3.74 (1.24)	3.65 (1.69)	4.03 (1.48)	4.09 (1.63)	F(1, 218) = 0.01 η^2 = .01, p = .995	F(2, 218) = 0.08 η^2 = .01, p = .921	F(2, 218) = 1.33 η^2 = .01, p = .267
Performance on vigilance task (objective indicators)									
Proportion of vigilance targets detected	0.97 (0.07)	0.96 (0.08)	0.95 (0.08)	0.94 (0.09)	0.96 (0.09)	0.95 (0.06)	F(1, 218) = 0.80 η^2 = .01, p = .373	F(2, 218) = 0.48 η^2 = .01, p = .622	F(2, 218) = 1.56 η^2 = .01, p = .213
Correct response reaction time (in section) to vigilance target	0.79 (0.19)	0.84 (0.17)	0.80 (0.15)	0.77 (0.18)	0.94 (0.19)	0.90 (0.18)	F(1, 218) = 5.39 η^2 = .02, p = .021	F(2, 218) = 7.22 η^2 = .06, p = .001	F(2, 218) = 3.11 η^2 = .03, p = .047

(*table continues*)

Table 6 (*continued*)

	The type of working memory load								
	Verbal working memory load			Spatial working memory load					
	Working memory load condition								
	Control: No working memory load ($N = 40$)	Low working memory load ($N = 39$)	High working memory load ($N = 36$)	Control: No working memory load ($N = 40$)	Low working memory load ($N = 38$)	High working memory load ($N = 32$)	Effect of type of working memory load (verbal, spatial)	Effect of working memory load (none, low, high)	Interaction
Outcome measure	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	Test		
Concentration rating	4.33 (0.97)	4.57 (0.97)	4.50 (0.83)	4.06 (1.25)	4.27 (0.88)	4.61 (1.12)	$F(1, 218) = 1.35$ $\eta^2 = .01, p = .246$	$F(2, 218) = 2.39$ $\eta^2 = .02, p = .094$	$F(2, 218) = 0.89$ $\eta^2 = .01, p = .413$
Correct *N*-back response		0.94 (0.07)	0.78 (0.12)		0.94 (0.08)	0.72 (0.12)	$F(1, 141) = 3.01$ $\eta^2 = .02, p = .085$	$F(1, 141) = 132.20$ $\eta^2 = .48, p = .001$	$F(2, 141) = 2.23$ $\eta^2 = .02, p = .138$
Correct *N*-back response time		0.77 (0.15)	1.01 (0.16)		0.77 (0.12)	0.99 (0.16)	$F(1, 141) = 0.06$ $\eta^2 = .01, p = .811$	$F(1, 141) = 90.32$ $\eta^2 = .39, p = .001$	$F(1, 141) = 0.12$ $\eta^2 = .01, p = .730$
Successful cue recognition	0.65 (0.09)	0.65 (0.09)	0.60 (0.07)	0.66 (0.09)	0.66 (0.07)	0.61 (0.08)	$F(2, 218) = 0.45$ $\eta^2 = .01, p = .502$	$F(2, 218) = 8.28$ $\eta^2 = .07, p = .001$	$F(2, 218) = 0.02$ $\eta^2 = .01, p = .977$
Attention checks	0.84 (0.12)	0.81 (0.16)	0.79 (0.18)	0.81 (0.18)	0.82 (0.15)	0.76 (0.21)	$F(1, 218) = 0.53$ $\eta^2 = .01, p = .469$	$F(2, 218) = 2.02$ $\eta^2 = .02, p = .135$	$F(2, 218) = 0.45$ $\eta^2 = .01, p = .641$
Number of IAMs	$r(40) = .38$ $p = .016$	$r(39) = .57$ $p = .001$	$r(35) = .34$ $p = .048$	$r(40) = .31$ $p = .052$	$r(38) = .37$ $p = .022$	$r(28) = .28$ $p = .143$			
Number of IFTs	$r(40) = -.06$ $p = .716$	$r(39) = .20$ $p = .221$	$r(35) = -.03$ $p = .875$	$r(40) = .10$ $p = .544$	$r(38) = .15$ $p = .369$	$r(28) = .49$ $p = .009$			

Note. Study 2: Means and standard deviations for variables measuring mood, fatigue, concentration, motivation, and performance on the vigilance task, working memory task as a function of the type of working memory load (verbal, spatial), and the level of working memory load (none, low, high). Low working memory conditions included 1-back task; high working memory load included 3-back task. Significant results are highlighted in grey. Additional and not preregistered findings: Correlations (Pearson's correlation coefficient) between successful cue-recognition and the frequency of IAMs and IFTs across groups and conditions in Study 2. PANAS = Positive and Negative Affect Schedule; IAMs = involuntary autobiographical memories; IFTs = involuntary future thoughts.

Table 7
Study 2: Frequency of Involuntary Thoughts as a Function of Working Memory Load and Cue Presentation

Involuntary thought category	The type of working memory load												Effect of type of working memory task (verbal, spatial)	Effect of working memory load (none, low, high)	Interaction
	Verbal working memory load						Spatial working memory load								
	Working memory load condition												Test		
	Control: No working memory load ($N = 40$)		Low working memory load ($N = 39$)		High working memory load ($N = 36$)		Control: No working memory load ($N = 40$)		Low working memory load ($N = 38$)		High working memory load ($N = 32$)				
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*			
All involuntary task-unrelated thoughts	12.33	5.21	12.41	5.89	9.97	5.11	12.63	4.32	13.03	4.87	11.82	4.97	$F(1, 214) = 1.78$, $\eta^2 = .01$, $p = .184$	$F(2, 214) = 2.53$, $\eta^2 = .01$, $p = .082$	$F(2, 214) = 0.44$, $\eta^2 = .01$, $p = .646$
Involuntary task-oriented and task-interference thoughts	6.60	4.42	5.74	3.82	9.00	4.99	6.85	3.70	5.89	4.40	6.69	5.18	$F(1, 219) = 1.16$, $\eta^2 = .01$, $p = .282$	$F(2, 219) = 3.79$, $\eta^2 = .03$, $p = .024$	$F(2, 219) = 1.92$, $\eta^2 = .01$, $p = .149$
Involuntary autobiographical memories	5.40	3.99	6.36	3.97	5.17	3.98	5.23	3.23	6.21	3.94	5.29	3.65	$F(1, 214) = 0.02$, $\eta^2 = .01$, $p = .893$	$F(2, 214) = 1.76$, $\eta^2 = .02$, $p = .174$	$F(2, 214) = 0.03$, $\eta^2 = .01$, $p = .971$
Involuntary future-oriented thoughts	3.65	3.63	2.59	1.98	1.74	1.74	4.47	3.51	2.82	2.40	2.75	2.30	$F(1, 214) = 3.40$, $\eta^2 = .02$, $p = .066$	$F(2, 214) = 8.77$, $\eta^2 = .08$, $p = .001$	$F(2, 214) = 0.40$, $\eta^2 = .01$, $p = .669$

Note. Study 2: Means and standard deviations for the number of involuntary task-unrelated and task-related thoughts, involuntary autobiographical memories and future-oriented thoughts as a function of the type of working memory task (verbal, spatial), and the level of working memory load (none, low, high). Low working memory conditions included 1-back task; high working memory load included 3-back task. Significant results are highlighted in grey.

memory load objectively increased task difficulty. Finally, a significant *interaction between type of working memory task and working memory load* was found for response times. Specifically, while there were no differences between verbal and spatial conditions in the control group ($p = .497$), participants in both low and high working memory load conditions responded faster in the verbal than in the spatial condition (*p*s $< .028$), possibly suggesting that the spatial version of the task imposed slightly greater processing demands under load.

Levels of Probe-Based Concentration Ratings. Analysis of probe-based concentration ratings revealed no significant main or interaction effects. Participants' self-rated concentration remained comparable across the different types and loads of working memory tasks.

Performance on the N-Back Task

The results are presented in Table 6. A series of 2 (type of working memory task: verbal, spatial) × 2 (working memory load: low, high) ANOVAs were conducted on the mean proportion of *N*-back responses and the correct *N*-back response times. As presented in Table 6, no significant main effects of the type of working memory load were observed suggesting the comparability between the two types of *N*-back tasks. The significant main effect of working memory load was observed for both correct *N*-back responses and response times. More precisely, high working memory load compared to low working memory load led to lower accuracy and higher response time. These results suggest the higher cognitive load of conditions in 3-back tasks. None of the two-way interaction effects were significant.

Frequency and Type of Recorded Task-Unrelated Thoughts

In terms of data preparation and coding, we repeated the procedure described in Study 1 (see the Frequency and Type of Recorded Thoughts section). To examine possible effects of the type of working memory task and working memory load on the number of all ITUTs reported during the vigilance task, we entered the number of these thoughts into a 2 (type of working memory task) × 3 (working memory load) ANOVA (see Table 7). This analysis resulted in nonsignificant main and interaction effects (*ps* $> .082$).

In addition to analyses of task-unrelated spontaneous thoughts, we examined task-oriented and task-interference thoughts to better understand how cognitive load reshapes the overall mental landscape during task performance. There was a significant main effect of working memory load on the number of these task-related thoughts, with more of them occurring in the high working memory load condition compared to the low working memory load condition ($p = .005$).

We followed the same data preparation and coding procedure as described in Study 1 (see The Positive and Negative Affect Schedule section and Barzykowski et al., 2024b). In total, participants reported 5,175 thoughts during the task. Of these, 1,908 were coded as task-related or task-related interference thoughts and 274 as "nothing," which were excluded from further analyses. Out of the remaining task-unrelated thoughts ($n = 2{,}993$), 336 were classified by participants as occurring deliberately rather than coming to mind involuntarily. The exclusion of these deliberate thoughts resulted in 2,657 spontaneous task-unrelated thoughts that served as the focus of the analyses presented below. The exact number of reevaluated entries was 644 out of the total 2,657 ITUTs.

Temporal Focus of Thoughts

The number of IAMs and IFTs was then entered into two separate two-way ANOVAs with the type of working memory task (verbal vs. spatial) and working memory load (none, low, high) as between-subjects factors. The means are provided in Table 7 (see also Figure 4), which also reports the corresponding *F* statistics and *p* values for the main effects on IAMs and IFTs.

The Effects of Working Memory Type (Verbal, Spatial). There were no significant main effects of the type of working memory task for either IAMs or IFTs, although for IFTs the effect approached significance ($p = .066$), with more IFTs occurring in the spatial than in the verbal conditions.

The Effects of Working Memory Load (Conditions: Control, Low, High). While there were no differences across working memory load conditions in the frequency of IAMs (all $ps > .101$), the working memory load significantly influenced the frequency of IFTs. Specifically, the highest frequency of IFTs was observed in the control condition compared to both 1-back and 3-back task conditions (both $ps < .002$) that did not differ from each other ($p = .273$). Additionally, no interaction effects were observed for either IAMs or IFTs (see Table 7).

Since we observed a null result for IAMs, we supplemented these analyses by reporting BFs. The Bayesian ANOVA provided a BF_{10} of 0.148 for the main effect of type of working memory task and 0.237 for the main effect of working memory load on the frequency of IAMs. Following Wetzels et al. (2011), for the interpretation of the BF, this suggests substantial evidence in favor of the null hypothesis.

The Comparison of IAMs and IFTs Across Conditions. To examine whether different working memory components differentially affected IAM and IFT retrieval, the number of IAMs and IFTs was entered into a 2 (type of working memory task: verbal vs. spatial) × 3 (working memory load: none, low, high) × 2 (type of thought: IAMs vs. IFTs) mixed ANOVA with repeated measures on the last factor.

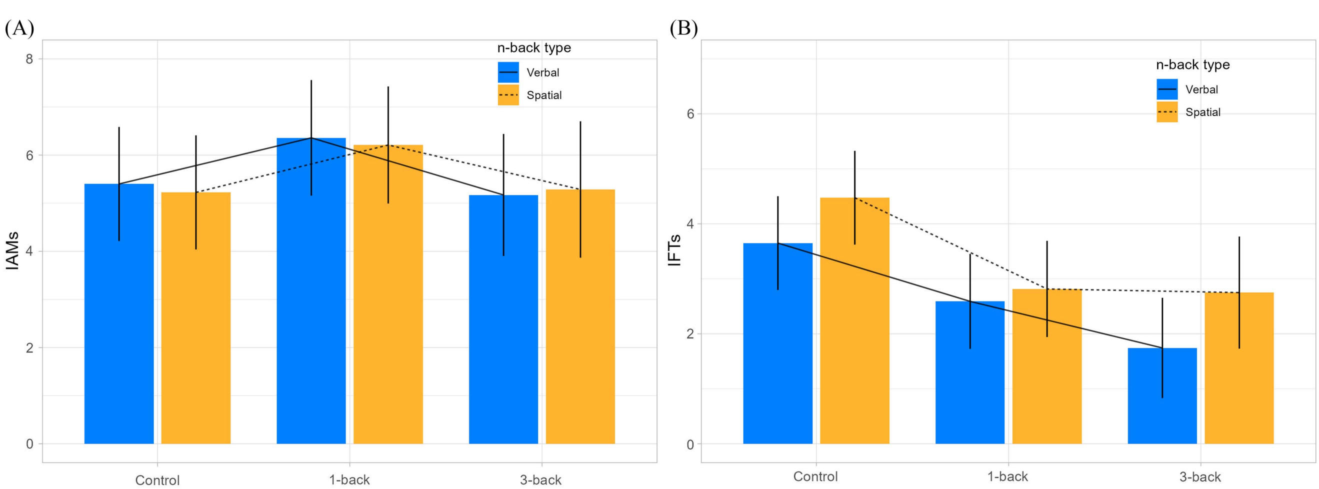
Main Effects. First, there was a significant main effect of working memory load, $F(2, 214) = 3.12$, $p = .046$, $\eta^2 = .03$. Follow-up comparisons indicated that the high working memory load condition was associated with a lower overall number of involuntary thoughts (collapsed across IAMs and IFTs) compared to both the control ($p = .049$) and low working memory load conditions ($p = .013$), which did not differ from each other ($p = .603$). Second, there was a significant main effect of type of thought, $F(1, 214) = 66.44$, $p = .001$, $\eta^2 = .23$, indicating that participants reported more IAMs ($M = 5.63$, $SD = 3.79$) than IFTs ($M = 3.05$, $SD = 2.84$). There was no significant main effect of type of working memory task (letter vs. visuospatial) on the overall number of involuntary thoughts (collapsed across IAMs and IFTs), $F(1, 214) = 0.93$, $p = .333$, $\eta^2 = .01$.

Interaction Effects. Importantly, the main effect of type of thought was qualified by a significant interaction between working memory load and type of thought, $F(2, 214) = 5.17$, $p = .006$, $\eta^2 = .05$. Follow-up analyses revealed that working memory load significantly affected IFTs but not IAMs. Specifically, while the number of IAMs did not differ significantly across working memory load conditions ($ps > .060$), participants reported more IFTs in the control condition compared to both the low ($p = .010$) and high ($p = .001$) working memory load conditions, with no difference between the latter two ($p = .365$).

Cue-Recognition Ratio Across Participants and Conditions

The mean cue-recognition ratio was entered into a two-way ANOVA with type of working memory task (verbal vs. visual) and working memory load (none, low, high) as a between-subjects variable (see Table 6; see also Figure 5). There was a significant main effect of working memory load with participants in the control and

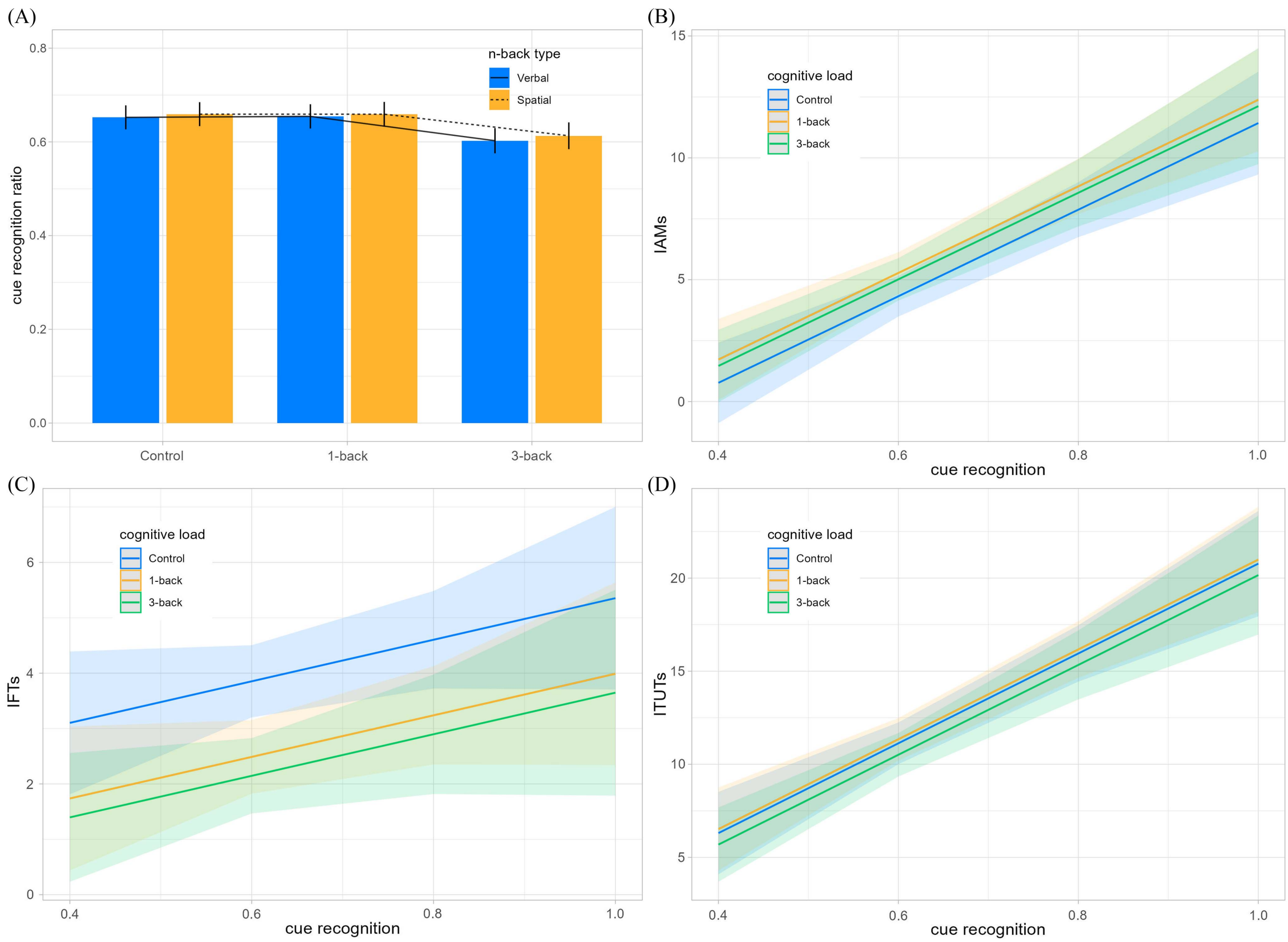
Figure 4
Study 2: The Number of (Panel A) IAMs and (Panel B) Future-Oriented Thoughts (IFTs) as a Function of Working Memory Task (Verbal, Spatial) and Working Memory Load (None, Low, High)



Note. Shaded areas and erros bars represents 95% confidence intervals. IAMs = involuntary autobiographical memories; IFTs = involuntary future thoughts. See the online article for the color version of this figure.

Figure 5

Cue Recognition Across Conditions and Its Relationship With IAMs and IFTs in Study 2



Note. Study 2: Successful cue recognition as (Panel A) function of the type of working memory load (verbal, spatial) and the level of working memory load (none, low, high), predictor of the number of (Panel B) involuntary autobiographical memories (IAMs), (Panel C) involuntary future thoughts (IFTs), and (Panel D) involuntary task-unrelated thoughts (ITUTs). Shaded areas and erros bars represents 95% confidence intervals. See the online article for the color version of this figure.

low working memory load conditions showing superior cue recognition compared to the 3-back task condition (all *ps* < .001).

Finally, a more detailed analysis of cue-recognition performance as a function of cue valence (positive, neutral, negative) is provided in Appendix E. Briefly, neutral cues were recognized less accurately than both positive and negative cues, with no differences between the latter two.

Additional and Not-Pre-Registered Findings: Cue Recognition and Frequency of IAMs and IFTs

As in Study 1, we investigated the relationship between successful cue-recognition ratio and the frequency of IAMs and IFTs. First, across all groups and conditions, the recognition ratio was positively and significantly correlated with the number of IAMs, $r(232) = .39$, $p = .001$, and with the number of IFTs, $r(232) = .15$, $p = .025$, indicating that participants who recognized more cues correctly were also reporting higher number of IAMs and IFTs than participants who recognized fewer cues presented in the vigilance task (for correlations across groups and conditions, see Table 6). Second, we used the linear regression model to further examine the relationship between the frequency of IAMs and IFTs and cue recognition scores while controlling for cognitive load condition. The relationship between IAMs and cue-recognition scores remained statistically significant and positive, $t(216) = 6.14$, $p < .001$, while it was no longer significant for IFTs, $t(216) = 1.66$, $p = .099$. Importantly, models including the interaction of cue recognition and cognitive load were not significantly better than simpler models explaining both IAMs, $F(2, 214) = 0.73$, $p = .481$, and IFTs, $F(2, 214) = 0.64$, $p = .529$, which suggest that the relationship between cue recognition and IAMs/IFTs frequency is similar regardless of the level of manipulation of cognitive load. Therefore, the visualizations present models without interaction effects (see Figure 5B–5D).

Discussion

In Study 2, we aimed to further investigate the effects of working memory load by comparing two types of working memory task. In particular, in addition to a verbal *N*-back task (letters varying in identity but presented in the same position) used in Study 1, we also used a visuospatial *N*-back task (the same letter presented in varying positions). Importantly, we employed only sequential cue presentation to maximize the likelihood of cue noticing, independently of the type of working memory load. This design allowed us to examine whether the retrieval of IAMs and IFTs depends on different components of working memory. With respect to background variables, no significant differences emerged between the verbal and spatial working memory conditions, except that in the spatial condition participants demonstrated significantly slower RTs to vigilance task targets (vertical lines) compared to the verbal condition. This may be explained by differences in eye movements: In the spatial version, participants likely needed to scan the entire screen to locate the target, which increased processing time. Overall, however, the findings suggest that the two types of working memory load were fairly comparable.

As in Study 1, high working memory load conditions—regardless of type of working memory task—were reported as more difficult than control conditions. They were also associated with lower concentration on verbal phrases, stronger interference with task performance, and consequently a greater tendency to ignore cues and suppress involuntary thoughts. Again, similar to Study 1, an interesting possibility emerges of an interplay between deliberate cognitive strategies adopted by participants and processes that operate beyond conscious control. In summary, the findings demonstrate that our manipulation of working memory load—independent of its type—successfully created conditions that differed in the extent to which participants were engaged in precisely tailored mental activity. As in Study 1, we next discuss the results by considering (a) the main effects of working memory type, (b) the main effects of working memory load level, and (c) their interaction.

Main Effect of Working Memory Type

Surprisingly, we did not observe significant effects of the type of working memory type on the number of IAMs or IFTs. Based on our predictions (see Table 2), we expected the lower frequency of IAMs in the spatial condition compared to the verbal condition and the reverse pattern for IFTs (i.e., the higher frequency in the spatial condition). Contrary to these expectations, the type of task did not result in differences in the overall number of involuntary task-unrelated thoughts, IAMs, or IFTs. The difference was only marginally significant ($p = .066$) in the expected direction, with more IFTs in the visuospatial condition. This finding suggests that possible effects of visual versus verbal working memory load on IAMs and IFTs warrant further investigation using a stronger manipulation of these types of loads.

Main Effect of Working Memory Load

Unlike Study 1, we did not observe a significant main effect of working memory load on the overall number of involuntary task-unrelated thoughts. However, we did replicate the effect for task-oriented and task-interference thoughts, although its magnitude was smaller ($\eta^2 = .03$) than what was observed in Study 1 ($\eta^2 = .05$). This weaker effect may suggest that in Study 2 task-related and task-related interference thoughts were less prevalent even under higher cognitive load, thereby reducing the extent to which they suppressed other types of thoughts. This pattern raises the possibility that more demanding tasks may actually increase the dominance of task-related thoughts, further limiting the cognitive space available for involuntary memories—a question that future research should address. Importantly, we replicated the effect of cognitive load on IFTs, showing that more demanding tasks were associated with fewer IFTs. Surprisingly, and perhaps for the first time in the literature, we did not observe the main effect of cognitive load on IAMs. In addition to the abovementioned possibility, a not mutually exclusive explanation may also be that participants in Study 2 were less engaged in actively suppressing thoughts ($F = 3.04$, $\eta^2 = .03$, $p = .050$) and in ignoring verbal phrases ($F = 9.97$, $\eta^2 = .08$, $p = .001$) compared to Study 1 ($F = 7.86$, $\eta^2 = .07$, $p = .001$ for thought suppression; $F = 17.54$, $\eta^2 = .14$, $p < .001$ for ignoring verbal phrases). These differences may have further attenuated the overall taxing impact of cognitive load on IAMs, highlighting the role of deliberate strategies such as suppression and cue avoidance in shaping spontaneous thought retrieval. Future studies are needed to systematically examine how these deliberate strategies interact with cognitive load to influence the occurrence of IAMs.

Interaction of Working Memory Type, Load, and Temporal Focus (IAMs vs. IFTs)

Consistent with the omnibus analysis reported in the Results section, we observed a main effect of working memory load on the overall number of involuntary thoughts, such that high working memory load was associated with fewer spontaneous thoughts when IAMs and IFTs were considered together. However, this effect was qualified by an interaction with temporal focus. Specifically, the reduction under load was driven primarily by IFTs, whereas IAMs remained relatively stable across load conditions. In addition, replicating previous findings (e.g., Barzykowski et al., 2021; Branch, 2023; Plimpton et al., 2015), participants across all conditions reported more thoughts about their personal past than about their future. This persistent asymmetry raises an important question regarding its underlying source. One possibility is that the use of verbal cues preferentially facilitates access to past events. Future studies could directly test this by systematically manipulating the presence versus absence of cues, or by varying cue characteristics, to determine whether the past–future imbalance reflects cue-driven processes or more fundamental differences in spontaneous retrieval dynamics.

Cue Recognition

Finally, cue recognition results revealed a consistent pattern across both studies and in line with previous work (e.g., Barzykowski et al., 2022). Specifically, high working memory load impaired cue processing, whereas no differences emerged between the verbal and spatial tasks. This suggests that the overall level of working memory load, rather than its type, drives reductions in cue recognition. As in Study 1, cue recognition was significantly related to the number of IAMs. These findings again highlight the important role of cue processing in the reduction of IAMs under working memory load.

Interestingly, and for the first time, we also observed a similar relationship for IFTs. Rather than suggesting a simple trade-off between IAMs and IFTs, this pattern may indicate that under certain cognitive conditions, efficient cue processing supports the emergence of both types of spontaneous thoughts. This possibility warrants further investigation.

General Discussion

The present research set out to investigate the mechanisms through which cognitive load affects the occurrence of involuntary autobiographical memories and involuntary future thoughts (IAMs and IFTs, respectively). Across two studies, we systematically manipulated the level (Studies 1 and 2) and type (Study 2) of working memory load and the conditions of cue presentation (Study 1: sequential vs. simultaneous cue presentation) to disentangle the role of limited cognitive resources and cue processing in the generation of spontaneous thoughts. To summarize, in Study 1, we demonstrated that higher working memory load reduced the overall frequency of spontaneous thoughts, with IAMs being more strongly affected than IFTs, and that cue recognition was impaired under high load and in simultaneous cue-presentation conditions. We did not observe a significant effect of cue presentation on the overall frequency of spontaneous thoughts (collapsed across IAMs and IFTs). In Study 2, we extended this work by comparing verbal and visuospatial working memory load while holding cue presentation constant (sequential only). Here, we replicated the effect of cognitive load on IFTs but not for IAMs or the total number of spontaneous task-unrelated thoughts. Surprisingly, we found no robust differences between verbal and spatial tasks and observed again that the highest working memory load impaired cue recognition.

Importantly, while investigating the somewhat "pure" cognitive mechanisms of cognitive load, we were also attentive to the "human factor"—namely, participants' individually applied strategies—which may additionally influence the frequency of spontaneous thoughts. Our findings clearly demonstrated that when engaged in cognitively demanding tasks, people also suppress and ignore thoughts and triggers, a factor that should not be overlooked. We discuss this possibility in more detail below. Taken together, the two studies provide converging evidence that both working memory demands and cue processing contribute to the occurrence of involuntary memories.

Our findings carry several implications for understanding the underlying mechanisms of cognitive load dependency of IAMs and IFTs retrieval, offering both theoretical insights and methodological considerations for why these experiences are limited under cognitive demand. We frame our discussion in terms of previously proposed explanations, which are not mutually exclusive, including impairments in cue noticing and the dependency on working memory capacity. Importantly, the present findings also highlight a factor that warrants further systematic investigation: participants' use of strategic cognitive activities. That is, beyond automatic cognitive constraints, individuals may intentionally suppress thoughts or ignore cues, further modulating the occurrence of spontaneous mental content. We discuss these implications below.

First, as to the *human factor*, it seems that previous research—including work by the authors of the present study—has somewhat overlooked an important factor operating during engagement in highly demanding cognitive tasks: namely, the deliberate strategies participants may use to meet the requirements of the ongoing task. To our best knowledge, this is the first time the "human factor" is explicitly discussed in the context of cognitive load dependence of IAMs and IFTs retrieval. For instance, in a highly demanding task such as the 3-back, a straightforward strategy might be to focus only on the letters presented on the screen while deliberately avoiding verbal phrases. This would manifest as ignoring the phrases and any other stimuli that could distract attention. A simple answer to the question of why we do not experience involuntary memories during demanding tasks could therefore be because we actively prevent them in order to maintain focus on the task. To further examine this possibility, we conducted exploratory analysis by examining correlations between IAM frequency and participants' self-reports of how much they suppressed involuntary thoughts and how much they ignored verbal phrases. As expected, both correlations were negative, Study 1: thought suppression, $r(219) = -.15$, $p = .026$; phrase ignoring, $r(219) = -.15$, $p = .028$. This explanation is not mutually exclusive with other cognitive mechanisms—cognitive processing is rarely so simple—but it may play a significant role in the effects of cognitive load on the diminished frequency of spontaneous thoughts. We suggest that systematically examining this factor in future research will help clarify how deliberate strategies interact with cognitive load to influence the occurrence of involuntary memories and future thoughts. In a broader sense, this suggestion aligns with the idea of closely examining the metacognitive processes that accompany IAMs and IFTs retrieval, as recently highlighted by Carreras et al. (2023, 2024, in press, Carreras, Tales, et al., 2026).

In fact, in an unpublished study from our lab, we manipulated the extent to which participants were actively and explicitly instructed to treat spontaneous thoughts as unwanted during a cognitively demanding task. Specifically, we induced a "human factor" by providing instructions that motivated participants to perform the task as well as possible and to avoid making mistakes. These instructions included (a) emphasizing the importance of performing both the main and additional tasks as accurately as possible, (b) informing participants that any additional thoughts could negatively affect task performance, (c) explaining that reporting such thoughts was critical to evaluate their impact on attention, (d) providing on-task reminders to maintain focus and avoid mistakes ("ATTENTION. Work with concentration, as quickly as you can. Both speed and accuracy are important. Do your best. Do not make mistakes"), and (e) raising the status of the task by requiring repetition of trial series if errors occurred in previous trials. In a sample of 51 participants ($M_{age} = 14.94$, $SD = 1.81$), although all groups experienced the same cognitive load as in Barzykowski and Niedźwieńska (2018a), Study 1),[7] participants who received the "human factor" instructions reported significantly fewer involuntary memories ($M = 2.73$, $SD = 2.81$) compared to the standard cognitive load condition, $M = 6.36$, $SD = 5.86$; $t(49) = 2.84$, $p = .007$, Cohen's $d = 0.79$, large effect. This pattern suggests that participants' deliberate strategies—motivated by task-related instructions—can

[7] Briefly, the cognitive load task required participants to perform a demanding decision-making exercise. Each time the central square on the screen turned green, they had to decide whether the number displayed inside it matched the number of blue lines (ranging from 2 to 8) shown elsewhere on the screen. Participants responded by pressing the green button ("z" key) for yes or the black button ("c" key) for no. Out of 800 trials, 180 critical trials with green squares were randomly distributed. These appeared roughly once every 8 s (approximately every fourth card) and included an equal number of yes and no responses.

substantially reduce the occurrence of IAMs. In terms of causal mechanism, this reduction may occur because participants both allocate more working memory resources to the primary task and actively ignore potential cues, limiting opportunities for spontaneous retrieval. It would not be the first time that the frequency of IAMs has been shown to vary as a function of participant instructions. For example, previous studies (e.g., Barzykowski & Niedźwieńska, 2016; Vannucci et al., 2014) demonstrated that instructing participants to report only memory-oriented thoughts increased their frequency, and a recent study by Safati et al. (2024) found that simple instructions to either increase or decrease mind wandering were sufficient to influence its occurrence. Taken together, these findings provide strong preliminary evidence that the "human factor" plays an important role in spontaneous memory retrieval and should not be overlooked. They also highlight the need for future studies to systematically examine how intentional cognitive strategies interact with cognitive load and cue processing to shape involuntary thought processes.

Second, as to the *cue noticing* and their role in IAMs and IFTs retrieval, our findings also support the explanation suggested by Kvavilashvili and Mandler (2004), according to which increasing engagement in cognitively demanding tasks reduces the likelihood of noticing cues that could otherwise serve as triggers for involuntary thoughts. This possibility was first examined by Barzykowski et al. (2022), and our results provide further support for that account. Specifically, we observed that the higher the working memory load, the lower the cue-recognition ratio. Cue recognition was also lower in the simultaneous compared to the sequential presentation condition in Study 1. Moreover, cue recognition was positively correlated with the frequency of IAMs, Study 1: $r(220) = .21, p = .002$; Study 2: $r(232) = .39, p = .001$; see also regression analyses in the Cue-Recognition Ratio Across Participants and Conditions section. While these results seem straightforward, the underlying mechanisms remain an open question. One possibility is that participants are simply less likely to notice cues while engaged in an attentionally demanding task—especially given that they may also be motivated to ignore verbal phrases on the screen (see the "human factor" account discussed above). This explanation suggests a primarily perceptual, mechanical process: When fewer cues are noticed, fewer opportunities arise to trigger IAMs and IFTs. This idea aligns with recent claims by Vannucci and Hanczakowski (2023) advocating for more systematic investigations of attentional factors in the retrieval of spontaneous cognition phenomena. The present study fits within this framework, specifically examining how attentional engagement with cues influences IAMs and IFTs retrieval. Consistent with this interpretation, we experimentally manipulated cue presentation to test whether noticing was driving the effect. Although the main effect was not statistically significant, the pattern in Figure 2A indicates a trend (difference between simultaneous and sequential conditions in control groups: $p = .067$) suggesting that cue presentation plays some role, though clearly not the only one. Thus, retrieval of IAMs does not depend solely on cue detection but also on cue processing. Beyond merely noticing a cue, it may be necessary to engage with it at a deeper cognitive level (Craik & Lockhart, 1972). Even when a cue is registered, high cognitive load may leave insufficient working memory resources to elaborate its meaning, connect it to autobiographical memory, launch spreading activation, and eventually achieve successful retrieval. Future studies should therefore aim to disentangle these processes more systematically—for instance, by combining manipulations of working memory load with manipulations of the depth of cue processing—an approach we are currently pursuing in our laboratory. Promising avenue within this direction may be the selection of cues that are highly efficient in triggering autobiographical memories—that is, cues that, once noticed, require minimal further elaboration. For example, recent studies by Carreras et al. (2024, Carreras, Staugaard, et al., 2026, also Matsumoto, 2024) identified fluent-access cues, which allow rapid retrieval of autobiographical memories, and limited-access cues, which provide insufficient retrieval support and result in slower or unsuccessful memory access. By combining fluent-access cues with manipulations of working memory load, future research could disentangle the extent to which cognitive load operates at a mechanistic level of cue noticing versus deeper cue processing.

Finally, our findings also support the idea that IFTs, being less dependent on external cues, may arise especially when frequency of IAMs is somewhat limited. This interpretation aligns with recent naturalistic experience-sampling results by Warden et al. (2019), who showed that IFTs in general—and especially involuntary thoughts about upcoming prospective memory tasks—can emerge without any external or internal triggers. It is also possible that IFT retrieval is cognitively more demanding than IAM retrieval, as suggested by Branch (2023), who found that IFTs more often occur voluntarily than involuntarily. This stands in contrast to accounts by Cole and Kvavilashvili (2021) and Kvavilashvili and Rummel (2020) but may explain why IFTs occur less frequently, particularly when IAMs are especially prevalent. Future studies using self-caught methods could directly test this possibility.

Third, as to the *working memory load*, our results underscore that the type and intensity of working memory load matter, but their impact on IAMs and IFTs may vary. Across both studies, a robust finding was that the more cognitively demanding the task, the more frequently participants reported task-related and task-interfering thoughts. While this might seem intuitive, its implications have rarely been considered in previous work. One possibility is that such task-oriented thoughts themselves consume working memory resources, thereby reducing the capacity available for retrieving IAMs and IFTs—resources that may function as the "oxygen" required for these involuntary processes and/or cue processing to ignite. In this sense, demanding tasks may constrain spontaneous cognition not only by taxing resources directly but also by crowding mental space with task-related mentation. Future work could examine how this dynamic between non-IAM thoughts, cognitive load, and spontaneous memories shapes the balance between IAMs and IFTs.

This perspective may also shed light on the puzzling absence of a main effect of cognitive load on IAMs in Study 2.[8] Notably, IAM frequency was descriptively higher in Study 2 compared to Study 1. One plausible explanation is that the balance between working memory resource depletion and the generation of task-related and task-interference thoughts varied across studies.

In Study 2, task-related thoughts were less frequent, and participants appeared less engaged in active suppression and cue ignoring, potentially leaving more cognitive capacity available for spontaneous

[8] Inspired by Nosek et al. (2012), who reflected on their own experience of running two identical experiments and failing to replicate the original finding, we believe that reporting this null result contributes positively to science by demonstrating transparency and adherence to open scientific practices.

retrieval. Consequently, the taxing effect of cognitive load on IAMs may have been attenuated. Together, these findings suggest that the interplay between working memory resource taxation, task-related and task-interference mentation, and deliberate regulation strategies (e.g., suppression, cue avoidance) jointly shape whether IAMs and IFTs emerge under demanding conditions. Future studies will need to disentangle these mechanisms more systematically.

Fourth, the last but not the least, as to the *type of working memory task*, while we had expected differences between verbal and visuospatial working memory, the main effect was not statistically significant. However, a trend emerged, with IFTs being experienced more frequently under visuospatial working memory load, suggesting that this effect warrants further investigation. This pattern encourages the use of visuospatial cognitive load when studying IFTs in laboratory settings. This approach is particularly promising because, given "the centrality of imagery to autobiographical memory" (M. A. Conway & Pleydell-Pearce, 2000, p. 263) and recent study by Congleton et al. (2021) showing that visuospatial cues increase the frequency of IAMs, high visuospatial load should disproportionately limit the retrieval of IAMs relative to IFTs, thereby leaving more cognitive resources available to selectively examine IFTs. This idea is both intriguing and deserving of systematic testing in future research.

Methodological Considerations and Avenues for Future Research

Despite the insights gained from the present studies, several methodological recommendations should be noted. First, the type of working memory task—verbal or visuospatial—should be carefully chosen depending on the type of spontaneous cognition under investigation, as visuospatial load may differentially impact IAMs versus IFTs. Separating cues from ongoing tasks can further minimize interference and better isolate the processes that lead to involuntary retrieval of task-unrelated thoughts. Second, the assessment of cue noticing and processing could be refined. In the present studies, we used an unexpected cue-recognition task, which provides a clear and objective metric but may capture only a limited aspect of the cue influence on involuntary thought retrieval. Future studies could explore alternative approaches, such as self-caught methods, where participants report spontaneously triggered thoughts, allowing analysis of RTs and thought frequency avoiding the restrictions in timings of stop trials or thought probes imposed by experimental design. Similarly, experience-sampling or probe-caught methods could also be used to examine real-time cue noticing, as well as perceived distinctiveness or salience of cues. For instance, the program could occasionally pause to ask participants about the cues presented a few trials earlier or to assess specific properties of the cues at the moment of their presentation, rather than relying on retrospective reports. Finally, it is crucial to control for potential overlap or interference between ongoing task demands and cue processing. As noted by Barzykowski and Niedźwieńska (2018a), additional task not only taxes cognitive resources but may also reduce exposure to cues that could directly trigger IAMs and IFTs. To capture the "pure" cognitive mechanisms underlying spontaneous thought retrieval, future studies should carefully minimize such confounds. This could be achieved by separating task-related stimuli from cues across conditions, ensuring that the materials used in ongoing tasks do not interfere with the presentation or perception of cues, and systematically monitoring participants' attention to both task and cue materials. These precautions will allow a clearer understanding of how cognitive load affects involuntary thoughts, in general, and IAMs and IFTs in particular.

Final Conclusions

The present article contributes to a broader line of research examining why involuntary thoughts are more likely to arise during undemanding ongoing tasks. Alternatively, this can be framed in reverse: Why do we experience fewer involuntary thoughts when engaged in cognitively demanding activities? Previous research has proposed several explanations, including impairments in cue noticing, the role of inhibitory control, and the dependency on working memory capacity. Importantly, these mechanisms are not mutually exclusive; multiple processes are likely to operate simultaneously, each influencing the frequency of IAMs and IFTs to varying degrees. Our studies further highlight an additional factor that warrants systematic investigation: participants' use of strategic cognitive activities. Beyond automatic cognitive constraints, individuals may actively suppress thoughts or ignore cues, thereby modulating the occurrence of spontaneous mental contents. In summary, our findings support the idea that cognitive load impairs the retrieval of spontaneous thoughts primarily by interfering with the noticing and processing of cues. Moreover, Study 2 explored the extent to which the retrieval of IAMs and IFTs depends on available different types of working memory resources. Although different types of working memory may contribute to IAM and IFT retrieval to varying degrees, their overall impact appears weaker than theoretical accounts might suggest. Taken together, these studies open promising avenues for future research particularly regarding how working memory load and its various types, cue processing, and strategies actively used by participants interact to shape the occurrence of involuntary past and future thoughts. Our findings suggest that the occurrence of IAMs and IFTs reflects the cumulative influence of multiple factors operating during their retrieval. Based on our work, we believe that future studies will further illuminate these factors by more precisely controlling, isolating, and disentangling their individual contributions.

References

Andrews-Hanna, J. R., Reidler, J. S., Sepulcre, J., Poulin, R., & Buckner, R. L. (2010). Functional-anatomic fractionation of the brain's default network. *Neuron*, *65*(4), 550–562. https://doi.org/10.1016/j.neuron.2010.02.005

Baddeley, A. (2007). *Working memory, thought, and action*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198528012.001.0001

Baddeley, A. D. (1986). *Working memory*. Oxford University Press.

Baddeley, A. D., & Hitch, G. J. (1994). Developments in the concept of working memory. *Neuropsychology*, *8*(4), 485–493. https://doi.org/10.1037/0894-4105.8.4.485

Ball, C. T. (2007). Can we elicit involuntary autobiographical memories in the laboratory? In J. H. Mace (Ed.), *Involuntary memory* (pp. 127–152). Blackwell Publishing. https://doi.org/10.1002/9780470774069

Barzykowski, K., Hajdas, S., Radel, R., & Kvavilashvili, L. (2022). Effects of inhibitory control capacity and cognitive load on involuntary past and future thoughts: A laboratory study. *Consciousness and Cognition*, *102*, Article 103353. https://doi.org/10.1016/j.concog.2022.103353

Barzykowski, K., Hajdas, S., Radel, R., Niedźwieńska, A., & Kvavilashvili, L. (2021). The role of inhibitory control and ADHD symptoms in the occurrence of involuntary thoughts about the past and future: An individual differences study. *Consciousness and Cognition*, *95*, Article 103208. https://doi.org/10.1016/j.concog.2021.103208

Barzykowski, K., Ilczuk, E., & Kvavilashvili, L. (2024a). *Autobiographical and future-related thoughts recorded during laboratory vigilance task with working memory load manipulation* (Version V1) [Data set]. Jagiellonian University. https://doi.org/10.57903/UJ/UXMKX6

Barzykowski, K., Ilczuk, E., & Kvavilashvili, L. (2024b). A comprehensive guide to research protocols for collecting and coding involuntary past and future thoughts. *MethodsX*, *12*, Article 102732. https://doi.org/10.1016/j.mex.2024.102732

Barzykowski, K., Ilczuk, E., & Kvavilashvili, L. (2024c). Dataset of autobiographical and future-related thoughts recorded during a laboratory vigilance task with working memory load manipulation. *Data in Brief*, *56*, Article 110831. https://doi.org/10.1016/j.dib.2024.110831

Barzykowski, K., & Kvavilashvili, L. (2024). Involuntary autobiographical memory. In L. M. Bietti & M. Pogacar (Eds.), *The Palgrave encyclopedia of memory studies*. Palgrave Macmillan. https://doi.org/10.1007/978-3-030-93789-8_73-1

Barzykowski, K., & Mazzoni, G. (2022). Do intuitive ideas of the qualities that should characterize involuntary and voluntary memories affect their classification? *Psychological Research*, *86*(1), 170–195. https://doi.org/10.1007/s00426-020-01465-3

Barzykowski, K., & Niedźwieńska, A. (2016). The effects of instruction on the frequency and characteristics of involuntary autobiographical memories. *PLOS ONE*, *11*(6), Article e0157121. https://doi.org/10.1371/journal.pone.0157121

Barzykowski, K., & Niedźwieńska, A. (2018a). Involuntary autobiographical memories are relatively more often reported during high cognitive load tasks. *Acta Psychologica*, *182*, 119–128. https://doi.org/10.1016/j.actpsy.2017.11.014

Barzykowski, K., & Niedźwieńska, A. (2018b). Priming involuntary autobiographical memories in the lab. *Memory*, *26*(2), 277–289. https://doi.org/10.1080/09658211.2017.1353102

Barzykowski, K., Niedźwieńska, A., & Mazzoni, G. (2019). How intention to retrieve a memory and expectation that a memory will come to mind influence the retrieval of autobiographical memories. *Consciousness and Cognition*, *72*, 31–48. https://doi.org/10.1016/j.concog.2019.03.011

Barzykowski, K., Radel, R., Niedźwieńska, A., & Kvavilashvili, L. (2019). Why are we not flooded by involuntary thoughts about the past and future? Testing the cognitive inhibition dependency hypothesis. *Psychological Research*, *83*(4), 666–683. https://doi.org/10.1007/s00426-018-1120-6

Barzykowski, K., & Staugaard, S. R. (2016). Does retrieval intentionality really matter? Similarities and differences between involuntary memories and directly and generatively retrieved voluntary memories. *British Journal of Psychology*, *107*(3), 519–536. https://doi.org/10.1111/bjop.12160

Barzykowski, K., & Staugaard, S. R. (2018). How intention and monitoring your thoughts influence characteristics of autobiographical memories. *British Journal of Psychology*, *109*(2), 321–340. https://doi.org/10.1111/bjop.12259

Berntsen, D. (1996). Involuntary autobiographical memories. *Applied Cognitive Psychology*, *10*(5), 435–454. https://doi.org/10.1002/(sici)1099-0720(199610)10:5<435::aid-acp408>3.0.co;2-l

Berntsen, D. (1998). Voluntary and involuntary access to autobiographical memory. *Memory*, *6*(2), 113–141. https://doi.org/10.1080/741942071

Berntsen, D. (2009). *Involuntary autobiographical memories: An introduction to the unbidden past*. Cambridge University Press. https://doi.org/10.1017/cbo9780511575921

Berntsen, D., & Bohn, A. (2010). Remembering and forecasting: The relation between autobiographical memory and episodic future thinking. *Memory & Cognition*, *38*(3), 265–278. https://doi.org/10.3758/MC.38.3.265

Berntsen, D., & Jacobsen, A. S. (2008). Involuntary (spontaneous) mental time travel into the past and future. *Consciousness and Cognition*, *17*(4), 1093–1104. https://doi.org/10.1016/j.concog.2008.03.001

Branch, J. G. (2023). Individual differences in the frequency of voluntary & involuntary episodic memories, future thoughts, and counterfactual thoughts. *Psychological Research*, *87*(7), 2171–2182. https://doi.org/10.1007/s00426-023-01802-2

Brewer, W. F. (1996). What is recollective memory? In D. C. Rubin (Ed.), *Remembering our past: Studies in autobiographical memory* (pp. 19–66). Cambridge University Press. https://doi.org/10.1017/cbo9780511527913.002

Brzozowski, P. (2010). *Skala Uczuć Pozytywnych i Negatywnych (SUPIN). Polska adaptacja skali PANAS Dawida Watsona i Lee Anny Clark. Podręcznik* [The manual of the Polish adaptation of SUPIN. Positive and Negative Affect Schedule (PANAS)]. Pracownia Testów Psychologicznych.

Carreras, F., & Moulin, C. J. A. (2023). Evidence for a metacognitive awareness of autobiographical memory organisation. *Scientific Reports*, *13*(1), Article 15624. https://doi.org/10.1038/s41598-023-34389-0

Carreras, F., Moulin, C. J. A., Tales, A., Barnes, C. M., & Souchay, C. (2024). Metacognitive processes accompanying the first stages of autobiographical retrieval in the self-memory system. *Memory*, *32*(6), 776–789. https://doi.org/10.1080/09658211.2024.2370532

Carreras, F., Staugaard, S. R., Moulin, C. J. A., & Barzykowski, K. (2026). Unravelling autobiographical memory retrieval: Processes and phenomenology [Manuscript submitted for publication]. In M. El Haj, S. M. J. Janssen, & M. Irish (Eds.), *The handbook of autobiographical memory: From basic to applied perspectives*. Elsevier.

Carreras, F., Tales, A., Barnes, C. M., Souchay, C., & Barzykowski, K. (2026). Autobiographical memory and metacognition in aging: A preserved ability to monitor memory retrieval. *Psychology and Aging*. Advance online publication. https://doi.org/10.1037/pag0000983

Carreras, F., Tales, A., Barnes, C., Souchay, C., & Barzykowski, K. (in press). Metacognitive abilities related to autobiographical memory retrieval in aging. *Psychology and Aging*.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. Academic Press.

Cole, S., & Kvavilashvili, L. (2019). Spontaneous future cognition: The past, present and future of an emerging topic. *Psychological Research*, *83*(4), 631–650. https://doi.org/10.1007/s00426-019-01193-3

Cole, S., & Kvavilashvili, L. (2021). Spontaneous and deliberate future thinking: A dual process account. *Psychological Research*, *85*(2), 464–479. https://doi.org/10.1007/s00426-019-01262-7

Cole, S. N., Staugaard, S. R., & Berntsen, D. (2016). Inducing involuntary and voluntary mental time travel using a laboratory paradigm. *Memory & Cognition*, *44*(3). 376–389. https://doi.org/10.3758/s13421-015-0564-9

Congleton, A. R., Nielsen, N. P., & Berntsen, D. (2021). Through the gateway of the senses: Investigating the influence of sensory modality-specific retrieval cues on involuntary episodic memory. *Psychological Research*, *85*(3), 1292–1306. https://doi.org/10.1007/s00426-020-01304-5

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*(5), 769–786. https://doi.org/10.3758/bf03196772

Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*, *107*(2), 261–288. https://doi.org/10.1037/0033-295X.107.2.261

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

D'Argembeau, A., & Van der Linden, M. (2004). Phenomenal characteristics associated with projecting oneself back into the past and forward into the future: Influence of valence and temporal distance. *Consciousness and Cognition*, *13*(4), 844–858. https://doi.org/10.1016/j.concog.2004.07.007

D'Argembeau, A., & Van der Linden, M. (2006). Individual differences in the phenomenology of mental time travel: The effect of vivid visual imagery and emotion regulation strategies. *Consciousness and Cognition*, *15*(2), 342–350. https://doi.org/10.1016/j.concog.2005.09.001

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Finnbogadóttir, H., & Berntsen, D. (2013). Involuntary future projections are as frequent as involuntary memories, but more positive. *Consciousness and Cognition*, *22*(1), 272–280. https://doi.org/10.1016/j.concog.2012.06.014

Gajewski, P. D., Hanisch, E., Falkenstein, M., Thönes, S., & Wascher, E. (2018). What does the *n*-back task measure as we get older? Relations between working-memory measures and other cognitive functions across the lifespan. *Frontiers in Psychology*, *9*, Article 2208. https://doi.org/10.3389/fpsyg.2018.02208

Gamboz, N., Brandimonte, M. A., & de Vito, S. (2010). The role of past in the simulation of autobiographical future episodes. *Experimental Psychology*, *57*(6), 419–428. https://doi.org/10.1027/1618-3169/a000052

Hall, N. M., Gjedde, A., & Kupers, R. (2008). Neural mechanisms of voluntary and involuntary recall: A PET study. *Behavioural Brain Research*, *186*(2), 261–272. https://doi.org/10.1016/j.bbr.2007.08.026

Hall, S. A., Rubin, D. C., Miles, A., Davis, S. W., Wing, E. A., Cabeza, R., & Berntsen, D. (2014). The neural basis of involuntary episodic memories. *Journal of Cognitive Neuroscience*, *26*(10), 2385–2399. https://doi.org/10.1162/jocn_a_00633

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the *N*-back task as a working memory measure. *Memory*, *18*(4), 394–412. https://doi.org/10.1080/09658211003702171

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review*, *9*(4), 637–671. https://doi.org/10.3758/BF03196323

Kvavilashvili, L., & Mandler, G. (2004). Out of one's mind: A study of involuntary semantic memories. *Cognitive Psychology*, *48*(1), 47–94. https://doi.org/10.1016/S0010-0285(03)00115-4

Kvavilashvili, L., & Rummel, J. (2020). On the nature of everyday prospection: A review and theoretical integration of research on mind-wandering, future thinking, and prospective memory. *Review of General Psychology*, *24*(3), 210–237. https://doi.org/10.1177/1089268020918843

Mace, J. H. (2004). Involuntary autobiographical memories are highly dependent on abstract cuing: The Proustian view is incorrect. *Applied Cognitive Psychology*, *18*(7), 893–899. https://doi.org/10.1002/acp.1020

Mandler, G. (1994). Hypermnesia, incubation, and mind popping: On remembering without really trying. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 3–33). MIT Press.

Matsumoto, N. (2024). Meta-memory (prediction) of specific autobiographical recall: An experimental approach using a modified autobiographical memory test. *Memory & Cognition*, *52*(6), 1263–1274. https://doi.org/10.3758/s13421-024-01541-8

Mazzoni, G. (2019). Involuntary memories and involuntary future thinking differently tax cognitive resources. *Psychological Research*, *83*(4), 684–697. https://doi.org/10.1007/s00426-018-1123-3

Mazzoni, G., Vannucci, M., & Batool, I. (2014). Manipulating cues in involuntary autobiographical memory: Verbal cues are more effective than pictorial cues. *Memory & Cognition*, *42*(7), 1076–1085. https://doi.org/10.3758/s13421-014-0420-3

Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific utopia: II. Restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science*, *7*(6), 615–631. https://doi.org/10.1177/1745691612459058

Philippi, C. L., Bruss, J., Boes, A. D., Albazron, F. M., Deifelt Streese, C., Ciaramelli, E., Rudrauf, D., & Tranel, D. (2021). Lesion network mapping demonstrates that mind-wandering is associated with the default mode network. *Journal of Neuroscience Research*, *99*(1), 361–373. https://doi.org/10.1002/jnr.24648

Plimpton, B., Patel, P., & Kvavilashvili, L. (2015). Role of triggers and dysphoria in mind-wandering about past, present and future: A laboratory study. *Consciousness and Cognition*, *33*, 261–276. https://doi.org/10.1016/j.concog.2015.01.014

Rac-Lubashevsky, R., & Kessler, Y. (2016). Decomposing the *n*-back task: An individual differences study using the reference-back paradigm. *Neuropsychologia*, *90*, 190–199. https://doi.org/10.1016/j.neuropsychologia.2016.07.013

Raichle, M. E. (2015). The brain's default mode network. *Annual Review of Neuroscience*, *38*, 433–447. https://doi.org/10.1146/annurev-neuro-071013-014030

Rasmussen, A., Zucca, R., Johansson, F., Jirenhed, D.-A., & Hesslow, G. (2015). Purkinje cell activity during classical conditioning with different conditional stimuli explains central tenet of Rescorla–Wagner model. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(45), 14060–14065. https://doi.org/10.1073/pnas.1516986112

Rubin, D. C. (2006). The basic-systems model of episodic memory. *Perspectives on Psychological Science*, *1*(4), 277–311. https://doi.org/10.1111/j.1745-6916.2006.00017.x

Safati, A. B., Carr, T. H., Lowe, C. J., & Smilek, D. (2024). Mind wandering on command. *Frontiers in Psychology*, *15*, Article 1448226. https://doi.org/10.3389/fpsyg.2024.1448226

Schlagman, S., Kliegel, M., Schulz, J., & Kvavilashvili, L. (2009). Differential effects of age on involuntary and voluntary autobiographical memory. *Psychology and Aging*, *24*(2), 397–411. https://doi.org/10.1037/a0015785

Schlagman, S., & Kvavilashvili, L. (2008). Involuntary autobiographical memories in and outside the laboratory: How different are they from voluntary autobiographical memories? *Memory & Cognition*, *36*(5), 920–932. https://doi.org/10.3758/MC.36.5.920

Schlagman, S., Kvavilashvili, L., & Schulz, J. (2007). Effects of age on involuntary autobiographical memories. In J. H. Mace (Ed.), *Involuntary memory* (pp. 87–112). Blackwell Publishing. https://doi.org/10.1002/9780470774069.ch5

Smallwood, J., Obonsawin, M., & Reid, H. (2003). The effects of block duration and task demands on the experience of task unrelated thought. *Imagination, Cognition and Personality*, *22*(1), 13–31. https://doi.org/10.2190/TBML-N8JN-W5YB-4L9R

Smallwood, J., Obsonsawin, M., Baracaia, S. F., Reid, H., O'Connor, R., & Heim, D. (2003). The relationship between rumination, dysphoria, and self-referent thinking: Some preliminary findings. *Imagination, Cognition and Personality*, *22*(4), 317–342. https://doi.org/10.2190/2n80-avm3-4a23-leaj

Smith, E. E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science*, *283*(5408), 1657–1661. https://doi.org/10.1126/science.283.5408.1657

Stawarczyk, D., Cassol, H., & D'Argembeau, A. (2013). Phenomenology of future-oriented mind-wandering episodes. *Frontiers in Psychology*, *4*, Article 425. https://doi.org/10.3389/fpsyg.2013.00425

Van Calster, L., D'Argembeau, A., Salmon, E., Peters, F., & Majerus, S. (2017). Fluctuations of attentional networks and default mode network during the resting state reflect variations in cognitive states: Evidence from a novel resting-state experience sampling method. *Journal of Cognitive Neuroscience*, *29*(1), 95–113. https://doi.org/10.1162/jocn_a_01025

Vannucci, M., Batool, I., Pelagatti, C., & Mazzoni, G. (2014). Modifying the frequency and characteristics of involuntary autobiographical memories. *PLOS ONE*, *9*(4), Article e89582. https://doi.org/10.1371/journal.pone.0089582

Vannucci, M., & Hanczakowski, M. (2023). Involuntary autobiographical memories and déjà vu: When and why attention makes a difference. *Behavioral and Brain Sciences*, *46*, Article e379. https://doi.org/10.1017/S0140525X23000031

Vannucci, M., Pelagatti, C., Chiorri, C., & Mazzoni, G. (2016). Visual object imagery and autobiographical memory: Object imagers are better at remembering their personal past. *Memory*, *24*(4), 455–470. https://doi.org/10.1080/09658211.2015.1018277

Vannucci, M., Pelagatti, C., Hanczakowski, M., & Chiorri, C. (2019). Visual attentional load affects the frequency of involuntary autobiographical memories and their level of meta-awareness. *Memory & Cognition*, *47*(1), 117–129. https://doi.org/10.3758/s13421-018-0854-0

Vannucci, M., Pelagatti, C., Hanczakowski, M., Mazzoni, G., & Paccani, C. R. (2015). Why are we not flooded by involuntary autobiographical memories? Few cues are more effective than many. *Psychological Research*, *79*(6), 1077–1085. https://doi.org/10.1007/s00426-014-0632-y

Vannucci, M., Pelagatti, C., & Marchetti, I. (2017). Manipulating cues in mind wandering: Verbal cues affect the frequency and the temporal focus of mind wandering. *Consciousness and Cognition*, *53*, 61–69. https://doi.org/10.1016/j.concog.2017.06.004

Warden, E. A., Plimpton, B., & Kvavilashvili, L. (2019). Absence of age effects on spontaneous past and future thinking in daily life. *Psychological Research*, *83*(4), 727–746. https://doi.org/10.1007/s00426-018-1103-7

Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E.-J. (2011). Statistical evidence in experimental psychology: An empirical comparison using 855 *t* tests. *Perspectives on Psychological Science*, *6*(3), 291–298. https://doi.org/10.1177/1745691611406923

Whelan, R. (2008). Effective analysis of reaction time data. *The Psychological Record*, *58*(3), 475–482. https://doi.org/10.1007/bf03395630

Zareen, G., Souchay, C., Barzykowski, K., & Moulin, C. J. A. (2026). Spontaneous metacognitive experiences and involuntary memories in the laboratory. *Consciousness and Cognition*, *137*(1), Article 103976. https://doi.org/10.1016/j.concog.2025.103976

Appendix A

Study 1: Equivalence of Experimental Groups Before, During, and After the Vigilance Task

Table 4 provides the mean ratings of various cognitive and noncognitive variables provided by participants before, during, or after the vigilance task. To test the comparability of conditions before and after the vigilance task in terms of participants' mood ratings, the overall mean positive and negative PANAS scores were entered into two separate 3 (working memory load: none, low, high) × 2 (cue presentation: simultaneous, sequential) × 2 (time: before, after vigilance task) mixed ANOVAs with repeated measures on the last factor.

The only significant main effect for positive PANAS scores emerged for the factor of time, revealing lower positive affect after the vigilance task compared to before, $F(1, 219) = 146.20$, $p = .001$; $\eta^2 = .40$. Regarding negative PANAS scores, there was a significant main effect of time, with lower negative affect reported after the vigilance task than before its completion, $F(1, 219) = 7.69$, $p = .006$; $\eta^2 = .03$. However, this main effect was qualified by a significant Time × Working Memory Load interaction, $F(1, 219) = 7.88$, $p = .001$; $\eta^2 = .07$. Tests of simple main effects showed that the reduction in negative affect occurred only in the control ($p = .001$) and the low working memory load conditions ($p = .004$) but not in the high working memory load condition ($p = .110$). The observed pattern of results is consistent with previous findings (e.g., Barzykowski et al., 2021, 2022; Barzykowski & Niedźwieńska, 2018b; Barzykowski, Radel, et al., 2019), suggesting that performing a monotonous vigilance task across conditions reliably influenced participants' mood in a more neutral direction (i.e., becoming less positive but also less negative).

We repeated this mixed ANOVA for physical and mental fatigue. For mental fatigue, the only significant main effect was observed for time, indicating higher mental fatigue after the vigilance task compared to before starting the task, $F(1, 219) = 58.38$, $p = .001$; $\eta^2 = .21$. For physical fatigue, there was a significant main effect of time, indicating higher physical fatigue after the vigilance task compared to before, $F(1, 219) = 91.23$, $p = .001$; $\eta^2 = .29$; a two-way interaction between time and working memory load, $F(1, 219) = 3.93$, $p = .021$; $\eta^2 = .03$; and a three-way interaction between time, working memory load, and cue presentation, $F(1, 219) = 3.10$, $p = .047$; $\eta^2 = .03$. Follow-up tests revealed that the increase in physical fatigue (i.e., the difference between before and after vigilance task) was greater in both the low ($p = .0001$) and high ($p = .0001$) working memory load conditions compared to the control condition ($p = .001$). Additionally, there were no significant differences in physical fatigue between sequential and simultaneous cue presentation measured before and after the vigilance task in the control and high working memory conditions. While in the low working memory condition, physical fatigue did not differ before the vigilance task for sequential and simultaneous cue presentation, still, after the vigilance task, it was higher for sequential than simultaneous cue presentation.

Next, we ran a series of three working memory load condition (none, low, high) by two cue presentation (simultaneous, sequential) between-subjects ANOVAs on the mean self-reported ratings of various cognitive and noncognitive variables (see Table 4) provided by participants before, during, or after the vigilance task.[A1] These analyses allowed us to examine any possible differences between the conditions in terms of their levels of retrospectively self-reported concentration, motivation, fatigue, or the way participants interacted with the vigilance task and its stimuli.

The Effects of Working Memory Load (Control, 1-Back, 3-Back)

As presented in Table 4, significant main effects of working memory load were observed for the following variables: (a) concentration on vertical lines, (b) concentration on verbal phrases and the extent to which verbal phrases were ignored, (c) the extent to which verbal phrases were perceived as interfering with the vigilance task, (d) the extent to which involuntary thoughts were being suppressed, and (e) the perceived task difficulty. These significant main effects indicated that participants in the high working memory load condition reported the lowest concentration ratings on vertical lines compared to the control condition ($p = .001$) but not the low working memory load condition ($p = .082$). These ratings were also lower in the low load condition compared to the control condition

[A1] Due to a program error, we were unable to report analyses on questions relating to the extent to which involuntary thoughts were perceived as interfering with the vigilance task.

(*Appendices continue*)

(p = .003). Importantly, in terms of self-reported attention to incidental verbal phrases presented on some of the slides, participants reported the lowest concentration ratings on phrases in the high working memory load condition compared to the low working memory (p = .011) and control conditions (p = .001). This difference was also significant between the low working memory and control conditions (p = .019). Similarly, verbal phrases were rated as least interfering with the vigilance task in the control condition compared to both the low (p = .001) and high (p = .001) working memory load conditions. In addition, participants' ratings of ignoring verbal phrases were lower in the control condition compared to the low and high working memory load conditions (all ps < .001), and the difference between low and high working memory load conditions was approaching significance in the expected direction (p = .049). In line with these findings, participants also reported the highest extent of thought suppression in the high working memory load condition compared to both the low working memory load (p = .001) and control conditions (p = .001). Finally, the vigilance task in the high working memory load (3-back) was rated as more difficult than in the control (p = .001) or low working memory load conditions (p = .001), and the low working memory condition (1-back) was also perceived as more difficult than the control condition (p = .048).

In summary, the high working memory load condition was consistently associated with greater perceived task difficulty and lower concentration ratings on both vertical lines and phrases compared to the control condition. Verbal phrases were perceived as least interfering with the vigilance task and were ignored to the least extent in the control condition. Additionally, participants reported the highest extent of thought suppression in the high working memory load condition.

The Effects of Cue Presentation (Sequential, Simultaneous)

The only significant main effect of cue presentation type (simultaneous or sequential) was found for concentration on vertical lines, with higher ratings observed for sequential presentation compared to simultaneous presentation (p = .053).

The Interaction Effects (Working Memory Load × Cue Presentation)

Across the control and low working memory load conditions, no differences were observed in the extent to which verbal phrases were ignored between sequential and simultaneous cue-presentation conditions (p = .343). However, in the high working memory load condition, participants reported a higher extent of ignoring verbal phrases in simultaneous compared to sequential conditions (p = .023).

Appendix B

Study 1: Analysis of Spontaneous Task-Unrelated Thoughts

To examine the possible effects of working memory load and cue presentation on the number of involuntary task-unrelated thoughts (ITUTs) reported during the vigilance task, the number of these thoughts was entered into a 3 (working memory load condition: none, low, high) × 2 (cue presentation: simultaneous, sequential) between-subjects ANOVA (Table 5). This analysis resulted in a significant main effect of the working memory load with the lowest number of ITUTs being reported in the high working memory load condition (M = 9.63, SD = 4.83) compared to both the control condition (M = 12.87, SD = 4.84; p = .001) and the low working memory load condition (M = 11.63, SD = 4.81; p = .014), which did not differ from each other (p = .116). The main effect of cue presentation and the Working Memory Load × Cue Presentation interaction were not significant.

We also repeated this analysis on the mean number of involuntary task-oriented and task-interference thoughts combined together. We observed a significant main effect of the working memory load. Specifically, while the control and 1-back task conditions did not differ from each other, there were more task-oriented and task-interference thoughts in the 3-back task compared to these two groups (all ps < .015). Neither the main effect of cue presentation nor the Working Memory Load × Cue Presentation interaction was significant.

Appendix C

Study 1: Cue-Recognition Ratio Across Participants and Conditions as a Function of a Cue Valence

To further inspect possible differences in terms of the type of cue being successfully recognized, the overall mean cue recognition was entered into a separate mixed 3 (working memory load: none, low, high) × 2 (cue presentation: simultaneous, sequential) × 3 (cue type: positive, neutral, negative) ANOVA with repeated measures on the last factor. This analysis resulted in significant main effects of working memory load, $F(2, 219) = 5.85$, $\eta^2 = .05$, $p = .003$, and cue presentation, $F(2, 219) = 5.85$, $\eta^2 = .09$, $p = .001$. Post hoc analysis revealed that cue recognition was poorest in the 3-back task and in simultaneous cue-presentation conditions. Additionally, a significant main effect of cue type was observed, $F(2, 438) = 84.92$, $\eta^2 = .28$, $p = .001$, indicating poorer cue recognition for neutral cues compared to both positive and negative cues (ps < .001). Furthermore, a significant interaction between cue type and cue presentation was identified, $F(2, 438) = 4.55$, $\eta^2 = .02$, $p = .011$, suggesting that the differences in cue recognition between simultaneous and sequential conditions were smaller for neutral cues compared to positive and negative cues. These results overall suggest that neutral cues may be more poorly processed than positive and negative cue words.

(*Appendices continue*)

Appendix D

Study 2: Equivalence of Experimental Groups Before, During, and After the Vigilance Task

Table 6 provides the mean ratings of various cognitive and noncognitive variables provided by participants before, during, or after the vigilance task. The overall mean positive and negative PANAS scores were entered into two separate mixed 2 (type of working memory: verbal vs. spatial) × 3 (working memory load: none, low, high) × 2 (time: before vs. after vigilance task) mixed ANOVAs with repeated measures on the last factor.

The only significant main effect for positive PANAS scores emerged for the factor of time, revealing lower positive affect after the vigilance task compared to before, $F(1, 219) = 137.63$, $p = .001$; $\eta^2 = .39$. Regarding negative PANAS scores, there was a significant main effect of time, with lower negative affect after the vigilance task compared to before, $F(1, 219) = 9.21$, $p = .003$; $\eta^2 = .04$, and the Time × Working Memory Load interaction, $F(1, 219) = 3.00$, $p = .052$; $\eta^2 = .03$. Tests of simple main effects showed the significant reduction in negative affect in the control ($p = .002$) and the low working memory load conditions ($p = .012$) but not in the high working memory load condition ($p = .823$). These results are similar to findings of Study 1 and previous research (e.g., Barzykowski et al., 2021, 2022; Barzykowski & Niedźwieńska, 2018b; Barzykowski, Radel, et al., 2019) and suggest that engaging in a repetitive vigilance task consistently impacted participants' mood across conditions, leading to a decrease in positivity while also reducing negativity.

Furthermore, we conducted mixed ANOVAs to examine physical and mental fatigue. Regarding physical fatigue, significant main effects were evident for time, with higher physical fatigue after the vigilance task compared to before, $F(1, 219) = 66.64$, $p = .001$; $\eta^2 = .23$, and a three-way Time × Type of Working Memory × Working Memory Load interaction, $F(1, 219) = 3.23$, $p = .041$; $\eta^2 = .03$. Follow-up tests revealed that while before the vigilance task there were no differences in physical fatigue between spatial and verbal working memory tasks implying comparability of experimental conditions, after the vigilance task, there were no significant differences only in the low working memory load condition (i.e., 1-back). In the control condition, the vigilance task in the visual condition was rated as slightly more physically fatiguing compared to the verbal condition ($p = .060$), but this trend was reversed in the high working memory load (i.e., 3-back), where participants rated the visual *N*-back version of the vigilance task as slightly less physically fatiguing ($p = .066$).

For mental fatigue, there was a significant main effect of time with higher mental fatigue after the vigilance task compared to before, $F(1, 219) = 58.38$, $p = .001$; $\eta^2 = .21$, and significant Time × Working Memory Load, $F(1, 219) = 3.59$, $p = .029$; $\eta^2 = .03$, and Type of Working Memory Task × Working Memory Load interactions, $F(1, 219) = 4.57$, $p = .011$; $\eta^2 = .04$. Concerning the first interaction, tests of simple main effects indicated that the differences in mental fatigue between before (lower mental fatigue ratings) and after (higher mental fatigue ratings) vigilance task performance were smallest in the control condition ($p = .004$) and highest in the high working memory load condition ($p = .001$). For the second interaction, follow-up tests revealed that while in the control condition with no working memory load, the vigilance task with the visual *N*-back task was rated as more mentally fatiguing than the verbal *N*-back task ($p = .058$), in the working memory load conditions (both 1- and 3-back), it was rated as less mentally fatiguing ($ps = .048$).

Next, we ran a series of 2 (type of working memory: verbal, spatial) × 3 (working memory load: none, low, high) ANOVAs on the mean self-reported ratings of various cognitive and noncognitive variables (see Table 6) provided by participants before, during, or after the vigilance task. As presented in Table 6, no significant main effects of the type of *N*-back task were observed for any of the self-reported background task performance variables.

The Effects of Working Memory Load (Conditions: Control, 1-Back, 3-Back)

We observed significant main effects of working memory load on levels of perceived task difficulty, concentration on the task, vertical lines, and phrases, and the extent to which verbal phrases were ignored, perceived as interfering with the vigilance task, and involuntary thoughts were being suppressed. Follow-up tests revealed that, as expected, the high working memory load condition was perceived as more difficult than the low working memory ($p = .001$) and control conditions ($p = .001$), which did not differ from each other ($p = .349$). Participants in the control condition also reported being least concentrated on the task and most concentrated on vertical lines compared to those in the low ($p = .038$, $p = .001$, respectively) and high ($ps = .001$) working memory load conditions, with no significant differences in the latter two groups ($p = .229$, $p = .924$, respectively). Additionally, participants in the high working memory load condition reported being least concentrated on verbal phrases compared to both the control ($p = .001$) and low working memory load conditions ($p = .028$), while the latter two did not differ ($p = .127$). They also reported verbal phrases interfering with the vigilance task to a higher degree compared to the control ($p = .001$) and the low working memory conditions ($p = .015$) who also differed significantly from each other ($p = .001$). Finally, participants in the high working memory load condition reported higher levels of ignoring verbal phrases and suppressing involuntary thoughts compared to those in the control and low working memory load conditions ($p = .001$, $p = .016$), who significantly differed from each other only in terms of ignoring verbal phrases ($p = .003$, $p = .320$, respectively).

The Interaction Effects of Type of Working Memory Task and Working Memory Load

The only significant interaction effects were observed for the ratings of concentration on vertical lines and the suppression of involuntary thoughts. Regarding the former, while there were no differences between the verbal and visual conditions in the control condition ($p = .95$), participants in the low working memory condition reported higher concentration on lines in the verbal than visual working memory task condition ($p = .020$), and the reverse pattern was observed in the high working memory load condition ($p = .066$). Concerning the latter, while there were no differences in the extent to which involuntary thoughts were suppressed between verbal and visual working memory conditions in both the control ($p = .329$) and high working memory conditions ($p = .565$), in the low working memory condition, participants declared a higher extent of suppression in the verbal than visual *N*-back task conditions ($p = .008$).

(*Appendices continue*)

Appendix E

Study 2: Cue-Recognition Ratio Across Participants and Conditions as a Function of a Cue Valence

To further examine potential differences in terms of the type of cue successfully recognized, the cue-recognition scores were entered into a separate 2 (type of working memory load: verbal, spatial) × 3 (working memory load: none, low, high) × 3 (cue type: positive, neutral, negative) mixed ANOVA with repeated measures on the last factor. There was a significant main effect of cue type, $F(2, 438) = 75.87$, $\eta^2 = .26$, $p = .001$, indicating the lowest cue recognition for neutral cues compared to both positive and negative cues (ps < .001), with no differences between positive and negative cues ($p = .290$).

Received September 28, 2021
Revision received March 4, 2026
Accepted March 6, 2026 ■