Universiteit Leiden

Psychologie
Faculteit der Sociale Wetenschappen

Effect of Post-Encoding Task Difficulty on Episodic Memory Consolidation

Name: Koen van Lunteren
Student number: 1681516
Date: 29th of July 2020
Supervisor: Samarth Varma
Second reader: Dr. Wouter Boekel
Word count: 7.416 (excl. reference list)
Cognitive Psychology
Thesis Msci Applied Cognitive Psychology

Abstract

Previous research has shown that a post-encoding rest period has similar effects on ongoing memory consolidation as executing a 2-back task, due to a suspected balance between interference suppression (in 2-back) and spontaneous reactivation of memory traces (in rest). The current study investigates the relationship between post-encoding task difficulty, autobiographical thinking and episodic memory consolidation by improving methodological issues in previous work. Episodic memory consolidation in a 2-back condition is compared to a 0-back condition, while analyzing autobiographical thinking proportions. It is expected that autobiographical thinking negatively correlates with consolidation and that autobiographical thinking is higher in the 0-back condition, compared to 2-back. However, given the higher chance of spontaneous reactivation of memory traces in the 0-back condition, it is expected that both conditions show equal memory consolidation. In a within-subjects design, participants ($N = 22$) are shown face stimuli, which they later identify in a recognition test, after executing a 0-back or 2-back task. During these n-back tasks, thought probes are presented targeting autobiographical thinking (and other mindwandering instances). Results show no direct relationship between autobiographical thinking and memory consolidation, but previous studies are replicated by showing similar memory consolidation after performing the two n-back tasks. In conclusion, the results reinvigorate the idea of a post-encoding 2-back task resulting in similar consolidation outcomes as a less demanding post-encoding task. Moreover, the outcomes suggest that n-back task difficulty affects autobiographical thinking, but the direct effect of this process on memory consolidation requires further investigation, potentially in the form of a neuroimaging study.

Keywords: episodic memory consolidation, autobiographical thinking, n-back task, interference suppression

Introduction

Episodic memory, which is the memory of events that one experienced personally, is a vital component of human life. Although being an extensively researched concept in the past decades, there is still much to discover about the distinct processes that take place within episodic memory. It is important to investigate these processes thoroughly, as memory research is an ongoing quest for small pieces of information that fit the memory puzzle altogether.

An indispensable process of episodic memory is consolidation. Episodic memory consolidation can be described as the gradual strengthening of episodic memory over time, which happens in two consecutive steps (Ji & Wilson, 2007). When one experiences a personal event (e.g. having a conversation with their partner), it is encoded as an episodic memory in the hippocampus. At this stage, the memory has a ‘weak’ form and is thereby susceptible to the influence of other encoded memory traces. Over time, the hippocampus subconsciously reactivates the episodic memory trace to gradually increase its synaptic strength and resistance against other memories, relocating it to the neocortex (i.e. episodic memory consolidation) (Ji & Wilson, 2007).

Research has suggested that any mental effort (e.g. picture search, spot-the-difference tasks) performed during this process of memory consolidation might interfere with the spontaneous reactivation of memory traces, called ‘retroactive interference’. The resulting retroactive interference would lead to detrimental effects on consolidation (Dewar et al., 2007; Craig et al., 2014). Therefore, a post-encoding rest period is beneficial for episodic memory consolidation, because it enables opportunities for this spontaneous reactivation of episodic memories (Tambini & Davachi, 2013). This beneficial effect is seen even up to seven days after a learning period (Dewar et al., 2014).

However, periods of rest do not lack mental activity. In fact, a post-encoding rest period stimulates the mind to wander, shifting attention away from any on-going task, towards internally-focused thoughts (Mason et al., 2007; Carciofo et al., 2014). Within this state of mindwandering, one is likely to engage in autobiographical thinking (i.e. internal thoughts about past, present and future events that are unrelated to the memory of stimuli that is being formed). Research has shown that autobiographical thinking, in which novel episodic memory processing takes place, uses the same resources in the hippocampus as ongoing episodic memory consolidation does (Cabeza & St Jacques, 2007; Daselaar et al., 2008; Qin et al., 2009). Due to this sharing of resources, autobiographical thinking forms a potential source of interference (Mednick et al., 2011; Craig et al., 2014). Following this line of thinking, it is possible that by engaging in post-encoding mental effort that reduces the formation of new episodic memories or recall of existing episodic memories, interference from autobiographical thinking can be reduced.

Indeed, in a few recent studies, participants showed similar episodic memory performance after a post-encoding rest period as after performing a post-encoding 2-back task (Varma et al., 2017; Varma et al., 2018; Varma et al., 2019). The authors concluded that a post-encoding 2-back task can be equally beneficial for episodic memory consolidation as post-encoding rest, due to the balancing of two opposite processes. On the one hand, there is a reduction in interference from autobiographical thinking in the 2-back task (as compared to rest), because of the high cognitive load of this task. This reduced interference can be attributed to the deactivation of the default mode network (DMN), which is involved in processes related to mindwandering (Qin et al., 2009). On the other hand, there is a higher chance of spontaneous reactivation of the study material in the rest condition (as compared to the 2-back task) (Tambini & Davachi, 2013). Besides assuming this balance, it was shown that actively stimulating autobiographical thinking by presenting everyday sound cues in a post-encoding rest period was associated with significantly lower episodic memory performance. At the end of one of these experiments, Varma et al. (2018) administered a questionnaire to investigate the content and frequency of mindwandering during the post-encoding activities that involved rest, a 2-back task and stimulation of autobiographical thinking. In this study, only an indirect relationship between autobiographical thinking and episodic memory consolidation was found.

Taking the current research on post-encoding activities and episodic memory consolidation into account, it still remains unclear in what specific way the n-back task suppresses interference (Varma et al., 2017; Varma et al., 2018; Varma et al., 2019). Moreover, a clear relationship between the suppression of autobiographical thinking and memory consolidation has not been convincingly shown yet. There are two important design issues to be tackled, in order to potentially generate more conclusive results about the role of the n-back task in interference suppression.

First, the post-encoding periods of past research were generally too distinct for researchers to be able to accurately compare mindwandering data. A post-encoding rest condition might involve substantially different mental activity than engaging in a cognitive task, like a post-encoding n-back task (Fishburn et al., 2014). Due to this difficulty in comparing conditions, it is hard to pinpoint the influencing factor(s) of these post-encoding conditions on episodic memory consolidation. In order to accurately compare mindwandering data of different experimental conditions, it is important that these conditions are as similar as possible. By manipulating the difficulty of post-encoding tasks while keeping the tasks as similar as possible, it may be possible to uncover the relationship between task difficulty and episodic memory consolidation. Qin et al. (2009) have shown an increase in working memory load when the execution of the 2-back task is compared to a 0-back task, resulting in lower accuracy and higher response times in the former. By having a within-subjects design in which participants execute a post-encoding 0-back task and a 2-back task, differences in difficulty can be objectively manipulated while keeping the stimuli identical. Moreover, a relatively high task difficulty (induced by a new task, opposed to a familiar task) was shown to decrease the activation of the default mode network, thereby indicating a negative correlation between task difficulty and frequency of autobiographical thinking (Mason et al., 2007).

Secondly, the questionnaire measuring mindwandering was administered at the end and not during post-encoding activities (Varma et al., 2018). This delay between actual mindwandering and the subjective recollection of its frequency and content might have caused inaccuracies in the responses of participants. Other mindwandering questionnaires, like the imaginal process inventory (sIPI), were used to measure general tendencies of participants to mindwander in daily life (Varma et al., 2019). Such a questionnaire does not provide a reliable measure of content and frequency of mindwandering during the post-encoding tasks. To potentially uncover the role of mindwandering (and autobiographical thinking) in the effects of these n-back tasks on episodic memory consolidation, it is vital to measure the content and frequency of mindwandering during these tasks (instead of in a post-task fashion). This more accurate way of measuring mindwandering can be achieved through the use of an experience sampling method. In this method, thought probes are administered during the execution of a task (Weinstein, 2018). Participants receive a question about their current thoughts multiple times during the task (often in a random interval), thereby collecting their mindwandering processes close to real-time. The answer options of these mindwandering probes differ substantially across research, ranging from simple dichotomous answer options as ‘I was focused on-task or off-task’ (Thomson et al., 2013) to a more diverse range of options, including ‘I am totally focused on the task’, ‘my mind is blank’ and ‘I am distracted by sights/sounds/temperature or by physical sensations’ (Unsworth, 2018). Christoff et al. (2009) have examined the use of such thought probes in combination with fMRI measures, showing that both default mode network and executive network regions are active when mindwandering is reported. Moreover, activity of these brain regions is even higher when participants are unaware of their mindwandering up until the thought probe is presented, showing the importance of including thought probe response options that cover unintentional forms of mindwandering. By administering thought probes while participants execute the n-back tasks, the resulting data is presumably more representative of the mindwandering properties of participants during these tasks.

Given the need for more conclusive results about this interplay between post-encoding mental activity, autobiographical thinking and episodic memory consolidation, this study investigates two research questions:

1. What is the relationship between autobiographical thinking during post-encoding 0-back and 2-back tasks and episodic memory consolidation?
2. What is the influence of task difficulty of post-encoding 0-back and 2-back tasks on episodic memory consolidation?

It is hypothesized that the proportion of autobiographical thoughts shows a negative correlation with episodic memory performance in both the 0-back and 2-back condition, because autobiographical thinking forms a potential source of interference in episodic memory consolidation (Mednick et al., 2011; Craig et al., 2014). Since Mason et al. (2007) showed an increase in default mode network activation in relatively easy tasks, it is also expected that this proportion of autobiographical thoughts is higher in the 0-back condition, compared to the 2-back condition.

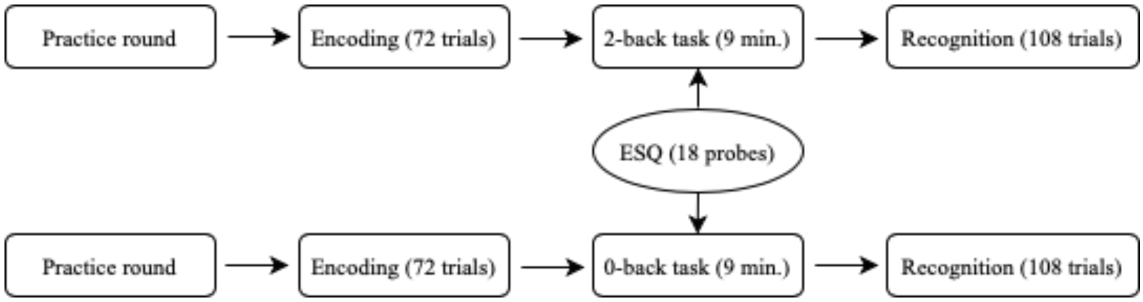
Moreover, it is hypothesized that the increased difficulty of the post-encoding 2-back condition (relative to post-encoding 0-back) results in similar levels of episodic memory consolidation as in the post-encoding 0-back condition, as measured by episodic memory performance. Previous research has suggested that there is higher suppression of autobiographical thinking in the 2-back condition, due to the continuous attentional demands needed to execute this task (Varma et al., 2018). However, there are higher chances of (automatic) episodic memory reactivation in post-encoding rest, due to the low mental effort needed in this post-encoding task (Tambini & Davachi, 2013). To resemble a rest-like state, the 0-back task is used in this study. Hørlyck et al. (2019) have shown that a post-encoding rest condition and 0-back condition do not differ in their effects on deliberate memory performance after watching a (trauma) film. Since participants in the current study are specifically asked to recall episodic stimuli in a similar fashion, post-encoding 0-back is assumed to be a good alternative for a rest-like state.

In conclusion, the proposed research involves a within-subjects design with two conditions (0-back and 2-back), in which experience sampling data is obtained during the execution of the n-back tasks. The effect of post-encoding task difficulty on episodic memory consolidation is measured.

Implications of finding support for the hypotheses may include a more thorough understanding of the role of autobiographical thinking in episodic memory consolidation during early post-encoding periods. To further progress the scientific knowledge on post-encoding factors influencing episodic memory consolidation, results of this study might pave the way for a future neuroimaging study in which neural correlates of autobiographical thinking can be directly correlated with episodic memory consolidation. On a more practical level, implications for the educational domain may include insight into how to optimally shape teaching methods to reduce interference between successive learning sessions or to stimulate retention of the study material.

Figure 1

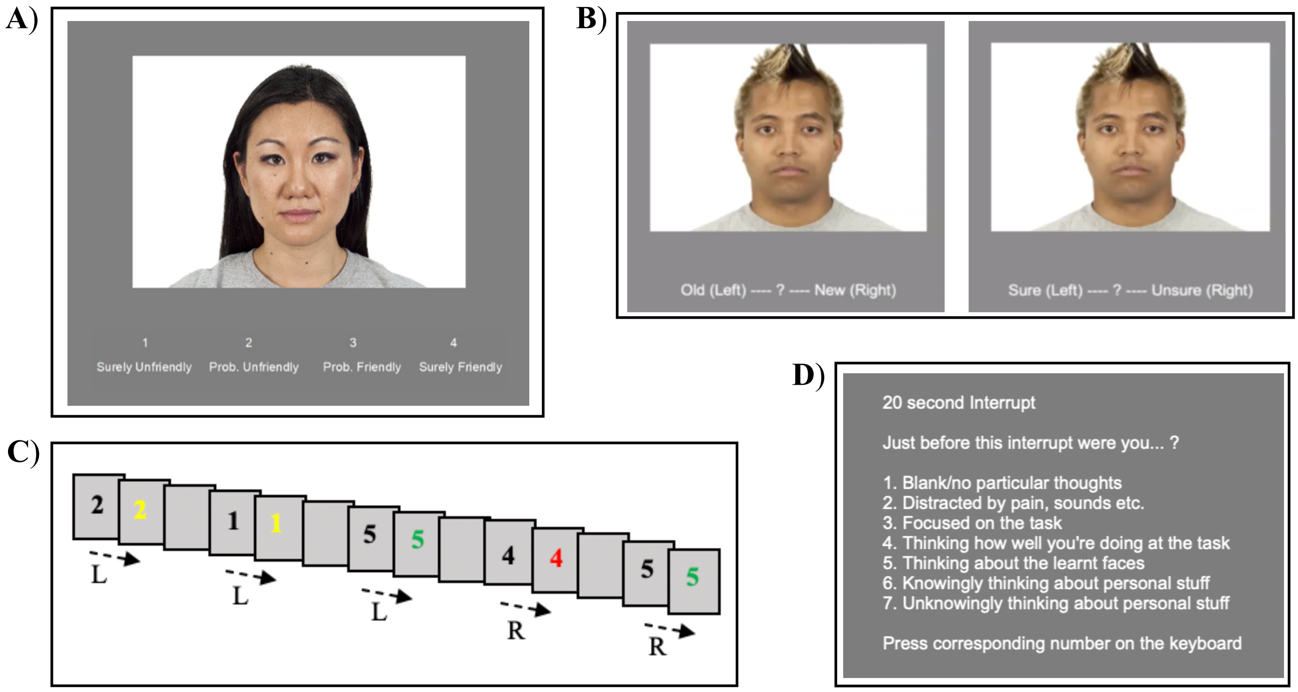
Experimental Design



*Note***.** Each of the two within-subject conditions had a practice block, followed by an encoding task, either a 2-back or 0-back task (depending on the condition) with a built-in Experience Sampling Questionnaire (ESQ), and a recognition task. Participants completed all tasks of each condition within a time gap of 24-72 hours.

Figure 2

Experimental Tasks



Note. **A**) Encoding task. Participants were asked to rate the friendliness of the face stimuli on a 4-point scale. **B)** Recognition task (two panels). Participants were instructed to indicate whether they had seen the face stimuli in the encoding task (old) or not (new), and how confident they were in their response. **C**) 2-back task. Participants indicated whether the same number stimulus was shown two trials back by pressing right (R), or not, by pressing left (L). The colors signal feedback about their response (yellow = first two trials after interruption; green = correct; red = incorrect or no response). **D**) Experience-sampling questionnaire. Participants indicated their current thoughts by choosing one out of seven options.

Methods

Design

The experiment involved a within-subjects design with two conditions (see Figure 1), spread over a time gap of 24-72 hours. The experiment took place in an online setting. The general design of the study and the tasks used (see Figure 2) were based on the design of Varma et al. (2017). Each condition started with an encoding task of face stimuli. This task was followed by a delay period of 15 minutes (maximum), which involved nine minutes of either a 0-back task or 2-back task (depending on the condition) and six minutes of experience sampling thought probes. The measurement level of the n-back manipulation is nominal, since these tasks differ in their instructions. During the execution of these post-encoding n-back tasks, an experience sampling questionnaire with thought probes was administered to measure participants' mindwandering (including autobiographical thinking). Both conditions ended with a recognition task of the face stimuli shown at the start of the experiment. The main dependent variable was the episodic memory performance, measured by d-prime memory scores (interval level). Besides measuring memory performance, proportions of mindwandering content during n-back (interval level) were measured to provide data for the hypothesis that suppression of autobiographical thinking plays a role in episodic memory consolidation. Moreover, performance in the n-back tasks (as measured by d-prime scores and reaction times, both interval level) were measured to pinpoint outliers and to investigate whether the manipulation of task difficulty was sufficient. The order of conditions was counterbalanced between participants to control for alternative explanations concerning order effects.

Participants

A total number of 34 healthy participants (23 females, $M_{age} = 25.41$, $SD = 2.13$) were recruited for the study. Being an online study, there is no clear estimate of the required sample size. However, this sample size is an approximation of the 36 participants used in experiment 3 by Varma et al. (2017), which had a similar within-subjects design with two conditions. Participants between 18 and 31 years of age and with no deficits in visual or memory processing were included in the study. Moreover, participants needed to be proficient in English and have obtained a college degree or higher. Of these 34 participants, seven participants dropped out during the study due to personal or technical reasons. Moreover, a total of five participants were excluded based on lack of adherence to the instructions or inaccurate button presses, as shown in poor performance on the 0-back and 2-back task (combined d-prime > 2-SD below combined d-prime average of 0-back and 2-back). A similar criterium was used for outlier removal in the study of Varma et al. (2018), for which they used d-prime group averages. As a result, 22 participants (14 females, $M_{age} = 25.59$, $SD = 2.15$) were considered for analysis after outlier removal. All participants gave digital informed consent in accordance with the Declaration of Helsinki. At the end of the experiment, participants received either a monetary compensation or no compensation (based on their response in the consent form stating how they would like to be compensated). The study has been reviewed and approved by the Psychology Research Ethics Committee of the Faculty of Social and Behavioral Sciences of Leiden University.

Procedure

Participants received an information letter, instruction video and informed consent via their personal email. After giving consent via an online form, participants received two separate emails on their preferred days (indicated in the informed consent, with a time period of 24-72 hours in between) with the link to the corresponding part of the experiment. These parts corresponded to the two conditions (0-back and 2-back) and the order of these conditions was counterbalanced. In total, the study took participants up to 60 minutes to complete, divided in 30 minutes on Day 1 and 30 minutes on the day on which the second part of the experiment was completed. Participants started in each condition with a short practice of all tasks involved. After this familiarization, participants executed a face-encoding task, in which they were asked to memorize a set of faces while rating their friendliness. This task was followed by a delay period involving either a 0-back or a 2-back task. During these delay periods, experience-sampling probes were presented in a pseudo-random interval with an average of 30 seconds. These experience-sampling probes required participants to categorize the content of their thoughts prior to the interruption, such as ‘I was focused on the task’ or ‘I was knowingly thinking about personal stuff’ (see ‘Tasks’ for further detail). At the end of the delay period, a face-recognition memory test was administered to test episodic memory consolidation. Participants had to indicate whether every presented face was ‘old’ (meaning it was presented earlier) or ‘new’ and how confident they were in their answer. After completing the experiment, participants received an email with the debriefing and received compensation if applicable. Participants performed the experiment on an online repository called Pavlovia.org, in which they were able to access a JavaScript version of the PsychoPy environment (Peirce, 2009). Participants used their own computer or laptop to execute the experiment.

Materials and tasks

Face-Encoding and Face-Recognition Tasks

A total of 216 unique face stimuli per participant were randomly pooled from the Chicago Face Database (CFD) (Ma et al., 2015). The choice for faces as stimuli type was made based on the fact that a string of faces is difficult to rehearse. These faces have been selected on their lack of salient facial features (e.g. tattoos, scars). In the instruction video and practice part of these tasks, all participants were shown 12 additional faces (which were identical across participants).

In the face-encoding task, a total of 72 face stimuli were consecutively presented to the participant, with a fixed duration of three seconds (taking up a total amount of 3.5 minutes). Participants were instructed to rate their friendliness on a 4-point scale ('surely unfriendly', 'probably unfriendly', 'probably friendly' and 'surely friendly'), while being aware of a later recognition test.

In the face-recognition task, a total of 108 face stimuli were presented consecutively, of which 72 face stimuli appeared in the face-encoding task and 36 faces were new. Participants were asked to correctly identify the faces shown during the face-encoding task. For every face, participants had to indicate 'old' (LEFT key) if they recognized it from the start of the experiment, and 'new' (RIGHT key) if this face seemed new to them. After their response, a second screen appeared asking participants to rate their confidence in their answer by indicating 'sure' (LEFT key) or 'unsure' (RIGHT key). These two screens were self-paced with a maximum screen time of five and three seconds, respectively. In the practice version of this task, participants received feedback on whether their response was correct. In the experimental version, they did not receive feedback. Responses on the face-recognition task allowed a conversion to hit rates (proportion of correctly identified 'old' face) and false alarm rates (proportion of erroneous claims of seeing an 'old' face), with which d' prime memory scores and thereby episodic memory performance were calculated.

0-Back Task

In this 9-minute task, number stimuli ranging from 1-5 were shown to the participants, with a fixed duration of 1.2 seconds each. Participants were instructed to indicate whether the currently displayed number was the number '3'. Participants could do so by pressing the RIGHT key. For every other number stimuli, participants had to respond with the LEFT key. Participants received feedback after the fixed 1.2 seconds per stimuli indicating the correctness of their response, by showing the color red (incorrect response or no response), green (correct response) or yellow (first two trials after every probe interruption). Performance (as measured by d-prime scores and reaction times) was measured as an indication of difficulty and the participant's focus during the task.

2-Back Task

In this 9-minute task, the presentation of stimuli, the feedback shown and the outcome measures were identical to the 0-back task. Participants were instructed to indicate whether the currently displayed number was the same number as shown two trials earlier. Participants could do so by pressing the RIGHT key. For every number that was different from the number shown two trials back, participants had to respond with the LEFT key.

Experience Sampling Questionnaire

During the 0-back and 2-back task, participants were interrupted by an experience sampling questionnaire. These thought probes were presented 18 times during the 15-minute period in a pseudo-random interval (with an average of 30 seconds in between probes). The probe was self-paced and remained on screen for a maximum of 20 seconds. This thought probe required participants to categorize the content of their thoughts prior to receiving the probe. Participants indicated their answer by choosing one out of the following seven answer options: (1) 'Blank/no particular thoughts', (2) 'Distracted by pain, sounds etc.', (3) 'Focused on the task', (4) 'Thinking how well you're doing at the task', (5) 'Thinking about the learnt faces', (6) 'Knowingly thinking about personal stuff', and (7) 'Unknowingly thinking about personal stuff'. This questionnaire was adapted from Unsworth and Robison (2018) by including two additional questions (answer options 5 and 6), with the purpose of measuring instances of stimuli rehearsal and disentangling intentional from unintentional autobiographical thinking, respectively. This questionnaire was used to measure content and frequency of mindwandering instances.

All tasks have been thoroughly tested in two pilot rounds ($N = 8$) to effectively adjust both timing and frequency of the presented stimuli during the tasks.

Analysis

To analyse the experiment data, participant data was imported from Pavlovia.org, data quality was assessed, data was combined into the relevant variables, and hypothesis testing was performed. First, all recognition trials that were lacking a response from the participant at the corresponding encoding trial (showing the same face) were removed from further analysis. Moreover, only confident responses to recognition trials were included in the analysis.

Episodic memory performance was calculated by using d-prime memory scores (Stanislaw & Todorov, 1999). To do this for each participant, hit rates of each participant were calculated separately by dividing the number of correctly identified “old” faces by the total number of old faces to which the participant responded. False alarm rates of each participant were calculated by dividing the number of “new” faces incorrectly identified as “old” faces, by the total number of new faces to which the participant responded. The standardized difference between hit and false alarm rates resulted in d-prime memory scores. Both 0-back and 2-back performance were calculated in a similar fashion by using the standardized difference between hit and false alarm rates, resulting in 0-back and 2-back d-prime scores.

The proportion of autobiographical thoughts was calculated for every participant in every n-back condition by counting the number of times a participant responded to the mindwandering probes with answer options ‘knowingly thinking about personal stuff’ (6) and ‘unknowingly thinking about personal stuff’ (7), dividing that sum by the total amount of mindwandering probes the participants responded to during the corresponding n-back condition.

To test the first hypothesis, a paired samples t-test was performed to check whether autobiographical thought proportion was higher in the 0-back condition, compared to the 2-back condition. Next, a Shapiro Wilk’s test was executed to check the normality scores of the autobiographical thought variables. Furthermore, Spearman’s rho was calculated for the two correlations between the proportion of autobiographical thoughts during the 0-back and 2-back tasks and episodic memory scores. For this hypothesis, it was expected that these correlations would be significant and negative.

To test the hypothesis that increased task difficulty in the 2-back condition results in similar levels of episodic memory consolidation as in the 0-back condition with relatively low task difficulty, a repeated-measures (RM-)ANOVA was executed with task difficulty as the variable of interest and condition order as a covariate. For this hypothesis, it was expected that d-prime memory scores would show an insignificant difference between the 0-back and 2-back condition.

To account for the potential effect of autobiographical thinking on familiarity, a planned t-test was executed involving the memory scores of all responses (both sure and unsure responses in recognition trials), instead of taking into account confident responses alone. Additionally, correlations between autobiographical thoughts and on-task thoughts during both 0-back and 2-back were calculated to check whether the first hypothesis could be addressed by looking at the ‘counterpart’ of autobiographical thinking. Lastly, correlational tests between on-task thought proportion in both conditions and confident and all memory scores were executed.

Table 1

Descriptive Statistics

Measurements	Condition: 0-back	Condition: 2-back
D-prime: n-back task	2.32 ± 0.12	1.57 ± 0.45
RT: n-back task (in seconds)	0.46 ± 0.04	0.65 ± 0.09
Hit rate: recognition (confident only)	0.82 ± 0.14	0.87 ± 0.08
False alarm rate: recognition (confident only)	0.20 ± 0.15	0.18 ± 0.14
D-prime: recognition (confident only)	2.09 ± 0.82	2.30 ± 0.80
Hit rate: recognition (all responses)	0.74 ± 0.13	0.78 ± 0.08
False alarm rate: recognition (all responses)	0.23 ± 0.11	0.21 ± 0.10
D-prime: recognition (all responses)	1.46 ± 0.56	1.67 ± 0.43

Note. Scores of participants in the 0-back and 2-back condition. The scores represent the mean and one standard deviation. Hit rates, false alarm rates and d-primes in the recognition task were calculated separately for confident responses and all responses.

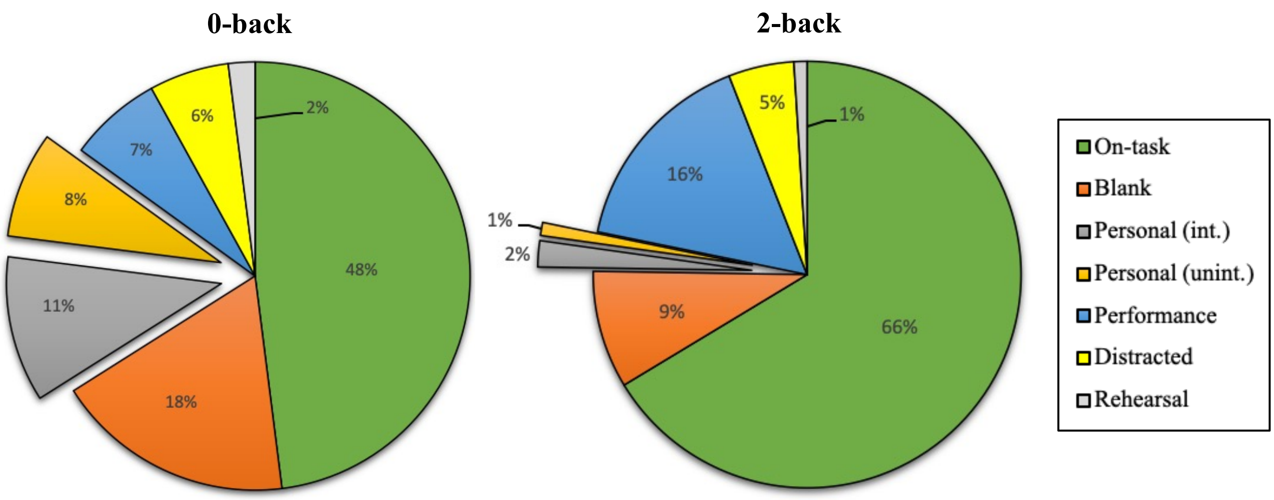
Results

The average performance of participants during the 2-back task was indicated by a mean d-prime of 1.57 (*SD* = 0.45) and an average reaction time of 0.65 seconds (*SD* = 0.09). In the 0-back task, the mean d-prime of participants reached 2.32 (*SD* = 0.12), with an average reaction time of 0.46 seconds (*SD* = 0.04) (see Table 1).

Out of a total of 7 individual mindwandering proportions (see Figure 3 and Table 2), two new variables were created. These variables indicated the proportion of autobiographical thoughts in the 0-back and 2-back condition and were based on the sum of individual proportions indicating intentional and unintentional thinking about personal events. Results of a paired t-test showed that the autobiographical thought proportion was in fact significantly higher in the 0-back condition (*M* = 0.19, *SD* = 0.16), compared to the 2-back condition (*M* = 0.02, *SD* = 0.04), $t(21) = -5.121$, $p < .001$. This suggests that the experimental manipulation of task difficulty was successful.

Figure 3

Thought Category Distribution during N-Back Tasks



Note. The mean distribution of thought categories during the 0-back (left) and 2-back task (right), indicated by responses to the experience sampling questionnaire. Autobiographical thought proportions used in the analysis are represented in the figure by the 'detached' slices.

Table 2

Thought Proportions during N-Back Tasks

Mindwandering (category)	Condition: 0-back	Condition: 2-back
On-task	0.48 ± 0.28	0.66 ± 0.25
Blank	0.18 ± 0.23	0.09 ± 0.19
Personal (intentional)	0.11 ± 0.11	0.02 ± 0.04
Personal (unintentional)	0.08 ± 0.10	0.01 ± 0.02
Performance	0.07 ± 0.09	0.16 ± 0.18
Distracted	0.06 ± 0.06	0.05 ± 0.08
Rehearsal	0.02 ± 0.03	0.01 ± 0.02

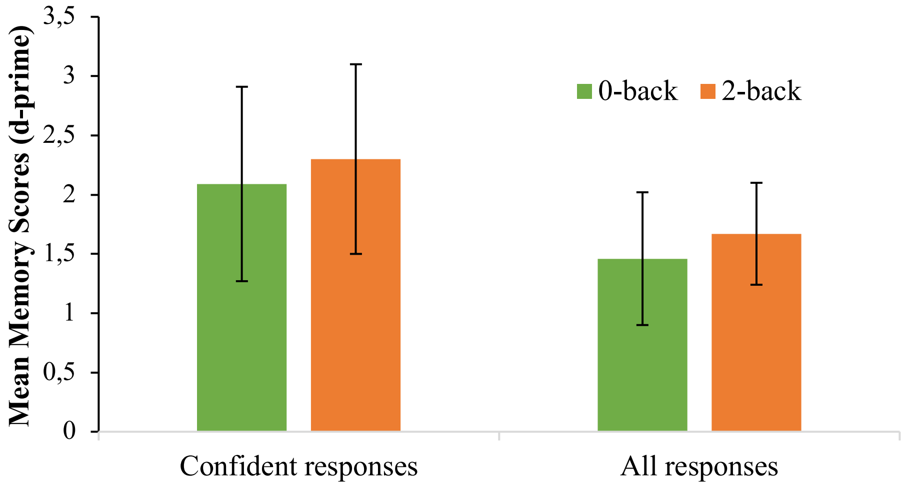
Note. Thought proportions of participants during the 0-back and 2-back task. The proportions represent the mean and one standard deviation.

In order to run correlational tests with these autobiographical thought scores, they were checked for normality. A Shapiro-Wilk test showed a significant departure from normality for both scores (0-back autobiographical thought proportion: $W(22) = 0.877$, $p = .011$; 2-back autobiographical thought proportion: $W(22), = 0.542$, $p < .001$). Correlations of these variables with the episodic memory scores of confident responses were calculated. For the 0-back condition, no significant correlation was found between proportions indicating autobiographical thoughts and confident memory scores, $r_s = .125$, $n = 22$, $p = .578$. In a similar fashion, no significant correlations were found in the 2-back condition between autobiographical thought proportions and confident memory scores, $r_s = -.175$, $n = 22$, $p = .437$. These results went against expectations, suggesting that other factors might have had a more profound impact on the episodic memory scores.

A RM-ANOVA was executed comparing episodic memory performance between the 0-back and 2-back condition, with order as a between-subject variable. Results showed that episodic memory scores based on confident responses of the participants did not significantly differ between the two n-back conditions (0-back: $M = 2.09$, $SD = 0.82$; 2-back: $M = 2.30$, $SD = 0.80$), $F(1, 20) = 2.057$, $p = .167$, $\eta_p^2 = .093$ (see Figure 4). This is in line with the second hypothesis, indicating that the level of episodic memory consolidation was similar across both conditions. Moreover, no interaction effect of the order of conditions with the episodic memory scores was found, $F(1, 20) = 0.207$, $p = .654$, $\eta_p^2 = .010$. However, a mild trend towards significance of the main effect of order was seen, $F(1, 20) = 3.820$, $p = .065$, $\eta_p^2 = .160$. Participants who ended the experiment with the 2-back condition showed higher d-prime episodic memory scores in both the 2-back and 0-back condition (2-back: $M = 2.53$, $SD = 0.22$; 0-back: $M = 2.39$, $SD = 0.22$), compared to participants who ended with 0-back (2-back: $M = 2.02$, $SD = 0.24$; 0-back: $M = 1.74$, $SD = 0.26$). This might be an indication of a practice effect. Since no significant interaction between order and episodic memory scores was found, this finding was not further analyzed. Next, a paired t-test was run with confident d-prime episodic memory scores of the 0-back and 2-back condition. This analysis showed similar results as before, $t(21) = -1.426$, $p = .169$, indicating no significant difference in episodic memory consolidation between the 2-back and 0-back condition.

Figure 4

Mean Memory Scores in N-Back Conditions



Note. Mean memory scores of participants in the 0-back and 2-back condition. Episodic memory scores are represented by d-prime scores for confident responses and all responses (both sure and unsure) separately. No significant difference between conditions was observed for confident responses. For all responses, mean d-prime memory score was higher in the 2-back condition, compared to the 0-back condition. Error bars represent one standard deviation from the mean.

Tests Involving Familiarity and On-Task Thoughts

A planned paired t-test for all recognition responses showed a significantly higher episodic memory performance in the 2-back condition (M = 1.67, SD = 0.43), compared to the 0-back condition (M = 1.46, SD = 0.56), $t(21) = -2.157$, $p = .043$ (see Figure 4). This may be an indication of a stronger familiarity effect in the 2-back condition, compared to the 0-back condition. The execution of planned correlational tests between autobiographical thought proportions in the two conditions and all episodic memory scores did not lead to different results. As predicted, a planned one-tailed Spearman's correlational test showed a negative correlation between autobiographical thought proportion and on-task proportion, in both the 0-back ($r_s = -.465$, $n = 22$, $p = .015$) and 2-back condition ($r_s = -.370$, $n = 22$, $p = .045$). A Shapiro-Wilk test showed no deviance from normality for on-task thought proportions in both conditions. The execution of a planned Pearson's correlational test revealed that on-task thought proportion in the 0-back condition was not correlated with either confident nor all memory scores. For the 2-back condition, a Pearson's correlational test showed a significant correlation between on-task thought proportion and confident memory scores, $r = .467$, $n = 22$, $p = .028$, suggesting that participants who were focused on the 2-back task showed increased episodic memory consolidation.

Discussion

The aim of this study was to investigate the interplay between post-encoding task difficulty, autobiographical thinking (as a form of mindwandering) and episodic memory consolidation. Previous studies have investigated this relationship by looking into the effects of post-encoding n-back tasks on content and frequency of mindwandering and episodic memory performance, but a convincing relationship hasn't been shown yet (Varma et al., 2017; Varma et al., 2018; Varma et al., 2019). By improving design issues through the implementation of an experience sampling questionnaire and comparable experimental conditions (0-back condition versus 2-back condition), this study focused on generating more conclusive results about the role of the n-back task in interference suppression and consolidation.

The study did not show any direct relationship between autobiographical thinking during the 0-back or 2-back task and episodic memory consolidation. This suggests that autobiographical thinking does not interfere with episodic memory consolidation, which contradicts tentative conclusions of previous studies about the interfering effects of autobiographical thinking (Craig et al., 2014; Varma et al., 2018). These results went against expectations. A possible explanation for this unexpected result is the suboptimal distribution of autobiographical thought proportions, that deviated from normality in both conditions (see 'Limitations and Future Directions'). Although no clear relationship was shown, the 0-back condition did show an increased level of autobiographical thinking relative to the 2-back condition. This is likely the result of the relatively low task difficulty of 0-back, which increases DMN activity and therefore stimulates mindwandering and autobiographical thinking (Mason et al., 2009; Qin et al., 2009). This finding shows that manipulating post-encoding task difficulty might be a good method for modulating autobiographical thinking in this line of research.

It was hypothesized that episodic memory consolidation would be similar in the 0-back and 2-back condition, due to the balancing of two opposing processes: more interference suppression in the more difficult 2-back task (opposed to the 0-back task) through deactivation of the default mode network (Qin et al., 2009), but more spontaneous reactivation of the study material in the easier 0-back task (Tambini & Davachi, 2013). Indeed, results of this study did not show a significant difference in confident episodic memory scores after executing a 0-back or 2-back task, indicating that consolidation in these two conditions is similar. Taking the 0-back condition as an approximation of a rest-like state (Hørlyck et al., 2019), these results replicate earlier findings showing similar consolidation after a post-encoding 2-back task or rest (Varma et al., 2018).

However, there is a possibility that post-encoding task difficulty and/or autobiographical thinking affect not only pure memory recognition (as indicated by confident responses in the recognition task), but tap into other memory processes as well, such as familiarity. In an extensive meta-review, Yonelinas (2002) examined the contribution of recollection and familiarity in recognition memory tests, stating that these two processes are independent, but both important for retrieval. Neurological evidence for this dissociation between recollection and familiarity comes from the distinct event-related brain potentials (ERP) that these processes show during memory tasks; temporally, topographically, and functionally (Friedmann & Johnson, 2000). Moreover, activity changes in the perirhinal cortex are consistently associated with familiarity, while the hippocampus and parahippocampal cortex are associated with recollection (Ranganath et al., 2003). Both recollection and familiarity play a role in recognition tasks and it might be hard for participants to disentangle the ‘source’ of their recognition response accurately by stating whether they were sure or unsure. Following this line of thinking, the potential influence of the experimental tasks on familiarity effects were taken into account by looking at all responses to the recognition task (both sure and unsure). Indeed, higher episodic memory consolidation was shown in the 2-back condition, relative to the 0-back condition, when including unsure, but correct responses from participants to recognition stimuli. This result implies that executing a post-encoding 2-back task seems to boost familiarity effects more than a 0-back task does. A potential explanation for this finding lies in the suppression of the DMN during the execution of the 2-back task (Mason et al., 2007). During memory retrieval, the hippocampus is strongly coupled with the DMN (Huijbers et al., 2011). By suppressing the DMN during 2-back, the hippocampus also shows a strong deactivation (Esposito et al., 2006). If this effect (temporarily) carries over to the recognition task, pure recollection as a hippocampal process might also be partially hindered. On the contrary, the perirhinal cortex might be relatively unaffected by the 2-back task, making judgements based on familiarity likely to be overrepresented after executing this 2-back task. Further study is needed to investigate this assumption (see ‘Limitations and Future Directions’).

To address the relationship between autobiographical thinking and consolidation from a different angle, proportions of on-task thoughts as a counterpart of autobiographical thinking were investigated. On-task thought proportion showed an anticorrelation with autobiographical thought proportion in both conditions, which was intuitively expected. This indicates that a higher frequency of on-task thoughts during the n-back task is associated with less autobiographical thinking. Now, on-task thoughts were positively correlated with episodic memory performance, but only in the 2-back condition. These results suggest that a sustained focus on the 2-back task leads to better episodic memory consolidation. Given the attentional demands needed to remain focused on the 2-back task, it is likely that the beneficial effect of this sustained focus on consolidation lies in the active suppression of mindwandering and therefore, a decrease in resource sharing (Mednick et al., 2011). Since a decreased level of autobiographical thinking in the 2-back task does not seem to be the reason for the increased memory consolidation shown in this study, it might be that the suppression of other mindwandering instances are the primary cause of improved memory consolidation (e.g. moments of 'blank' and/or distraction). For example, Dewar et al. (2007) found that common post-encoding distractors, such as listening to the radio, were a source of interference while consolidating, leading to a weaker consolidation and more forgetting. Participants might have been able to successfully ignore these types of distractors (which are likely to occur due to the online nature of the experiment) by attentively focusing on the demanding 2-back task, explaining the improved consolidation in the 2-back condition. Moreover, a lack of general mindwandering (as indicated by a high amount of on-task thoughts) in a 2-back condition improves consolidation, but leads to similar consolidation outcomes as a 0-back condition. Therefore, it is likely that chances of spontaneous reactivation of memory traces (which is beneficial for consolidation) are indeed higher in a post-encoding 0-back condition, as was shown for post-encoding rest (Tambini & Davachi, 2013).

In summary, these results suggest that increased post-encoding task difficulty decreases autobiographical thinking frequency, as shown in the low frequency of autobiographical thinking in the 2-back condition due to high task demands. However, autobiographical thinking does not show a direct relationship with episodic memory consolidation in this study, which might be explained by either experimental limitations and/or the influence of other more prominent factors (such as the frequency of on-task thoughts). Despite the lack of a clear relationship between autobiographical thinking and episodic memory consolidation, similar consolidation is shown after engaging in a 2-back task or 0-back task when looking at confident recognition, replicating similar studies (Varma et al., 2018).

Limitations and Future Directions

Although the aim of this study was to improve the design of earlier studies to find concrete evidence for a relationship between post-encoding task difficulty, autobiographical thinking and episodic memory consolidation, there are a couple of important limitations to note in order to point towards future directions in this line of research.

First, the final sample size did not meet the number of participants as required by the standards of similar studies (Varma et al., 2017). The removal of outliers based on relevant criteria has depleted the final number of participants substantially, leading to a low statistical power. Although this study includes some important alterations relative to previous designs, such as an online setting, the sample size should have been higher to generate more convincing conclusions about the interplay between variables. It is advised to replicate this study with a final sample size of at least 36 participants (Varma et al., 2017) to match earlier work in this scientific field.

Secondly, the study might have involved a suboptimal design of thought probe presentation. The study used a post-encoding n-back tasks of nine minutes, adding up to a maximum delay period of 15 minutes in total when including the experience sampling questionnaire. Having participants executing these n-back tasks for a relatively short period of time might not have induced autobiographical thinking to a level that potentially shows more pronounced effects on episodic memory consolidation. For example, Choi et al. (2017) used a similar experimental design in which participants had to execute visual and auditory n-back tasks while responding to mindwandering thought probes, for 30 minutes per task. The relatively low delay period duration of 15 minutes is expected to have caused the deviance from normality for the distribution of autobiographical thought proportions, which led to skewed results. On a more conceptual level, the definition of 'autobiographical thinking' might not have been fully captured by only including concrete thinking about personal events. As pointed out in a meta-review by Weinstein (2018), there is a considerable lack of consensus in the field of investigating mindwandering by using experience sampling methods. The existence of autobiographical thinking during the n-back tasks in this study might have been more pronounced, if other experience sampling responses would have been included. For example, distraction might be another indication of autobiographical thinking, if the specific form of distraction were to include everyday sounds (which is likely given the online nature of the experiment). Relatedly, it was shown that the presence of familiar sound cues in the environment led to increased autobiographical thinking (Craig et al., 2014; Varma et al., 2018). Besides distraction, indications of a 'blank' might also be partially related to autobiographical thinking (without awareness), since participants themselves do not know what the content of their thoughts is when responding with 'blank'. In this study, if participants responded to such events by indicating 'I am distracted' or 'blank/no particular thoughts', the response was not recorded as autobiographical thinking and therefore not included in the final analysis. To ensure that autobiographical thinking proportion is normally distributed and that the concept of autobiographical thinking is reliably captured in a post-encoding period, it is important to examine the optimal frequency, timing and wording of probe presentation. To investigate this, a future study could include four conditions of a 2-back task of approximately 25 minutes, in which the frequency and timing of presented probes increases in each condition (condition one: one probe/min., condition two: two probes/min., condition three: three probes/min., condition four: four probes/min.). Moreover, probes in these conditions can be phrased differently, alternating between more concrete ("I am thinking about events in the past") and abstract probes ("I am daydreaming"). Following such a design, the timing, frequency and probe phrasing in the conditions can be manipulated in various combinations. Moreover, this study should employ concurrent fMRI measurements of the default mode network and executive network regions while participants engage in the 2-back task, similar to the work of Christoff et al. (2009) (which included a go/no-go task). This way, it is possible to investigate which frequency, timing and phrasing of probes are able to predict autobiographical thinking best in a post-encoding task. Results of this study could be directly implemented in the current line of research, in order to employ the most valid and reliable experience sampling method to investigate autobiographical thinking.

Lastly, a concrete explanation of the familiarity effects (indicated by a high amount of unsure, but correct responses) in the 2-back condition, relative to the 0-back condition, cannot be provided based on the current design. In order to investigate these effects more thoroughly, it would be worthwhile to design a study that compares brain activity patterns in the perirhinal cortex and (para)hippocampal regions using fMRI. This study could have two conditions similar to the current study: an encoding task, a post-encoding 2-back or 0-back task, and a final recognition task in both conditions. Since perirhinal cortex activity is associated with familiarity processes (Ranganath et al., 2003) and familiarity effects in this study were more prevalent in the recognition task of the 2-back condition, it would be interesting to examine whether this finding is reflected in neural correlates in the brain. Moreover, this study should compare perirhinal cortex activity with measurements of (para)hippocampal activity, which correlates with recollection processes (Ranganath et al., 2003), in both conditions during n-back performance and recognition. This way, inferences can be made about the influence of and interplay between familiarity and recollection during a recognition task after post-encoding n-back task efforts.

Conclusion

This study investigated the relationship between post-encoding task difficulty, autobiographical thinking and episodic memory consolidation by improving methodological issues found in previous work (Varma et al., 2017; Varma et al., 2018). Executing a post-encoding 0-back task leads to more autobiographical thinking overall, comparing it to a more cognitively demanding 2-back task that suppresses these thoughts. Although this study shows a successful manipulation of task difficulty to alter autobiographical thinking frequency, the decrease in autobiographical thoughts in the 2-back task does not directly predict episodic memory consolidation outcomes, which contradicts suggestions made in previous research (Craig et al., 2014; Varma et al., 2018). However, being able to concentrate on the 2-back task without engaging in any form of mindwandering does lead to better memory consolidation overall, indicating that suppression of other forms of mindwandering might play a more prominent role in enhancing consolidation. Similar to the outcomes of previous research (Varma et al., 2018), episodic memory is equally consolidated after performing a 2-back task or 0-back task, when looking at pure recollection processes. The results of this study reveal the need for a neuroimaging study, in which the mechanisms behind the influence of n-back task difficulty on episodic memory consolidation through suppression of autobiographical thinking (and other forms of mindwandering) can be more thoroughly investigated. Further research is needed on post-encoding tasks and memory consolidation before employing these methods in an educational setting.

References

Cabeza, R., & St Jacques, P. (2007). Functional neuroimaging of autobiographical memory. *Trends in Cognitive Sciences, 11*(5), 219-227. https://doi.org/10.1016/j.tics.2007.02.005

Carciofo, R., Du, F., Song, N., & Zhang, K. (2014). Mind wandering, sleep quality, affect and chronotype: An exploratory study. *PLoS One, 9*(3), e91285. https://doi.org/10.1371/journal.pone.0091285

Choi, H., Geden, M., & Feng, J. (2017). More visual mind wandering occurrence during visual task performance: Modality of the concurrent task affects how the mind wanders. *PLoS ONE, 12*(12), e0189667. https://doi.org/10.1371/journal.pone.0189667

Christoff, K., Gordon, A. M., Smallwood, J., Smith, R., & Schooler, J. W. (2009). Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proceedings of the National Academy of Sciences, 106*(21), 8719-8724. https://doi.org/10.1073/pnas.0900234106

Craig, M., Della Sala, S., & Dewar, M. (2014). Autobiographical thinking interferes with episodic memory consolidation. *PLoS ONE, 9*(4), e93915. https://doi.org/10.1371/journal.pone.0093915

Daselaar, S. M., Rice, H. J., Greenberg, D. L., Cabeza, R., LaBar, K. S., & Rubin, D. C. (2008). The spatiotemporal dynamics of autobiographical memory: Neural correlates of recall, emotional intensity, and reliving. *Cerebral Cortex, 18*(1), 217-229. https://doi.org/10.1093/cercor/bhm048

Dewar, M., Alber, J., Cowan, N., & Della Sala, S. (2014). Boosting long-term memory via wakeful rest: Intentional rehearsal is not necessary, consolidation is sufficient. *PLoS ONE, 9*(10), e109542. https://doi.org/0.1371/journal.pone.0109542

Dewar, M., Cowan, N., & Della Sala, S. (2007). Forgetting due to retroactive interference: A fusion of Muller and Pilzecker's (1900) early insights into everyday forgetting and recent research on anterograde amnesia. *Cortex, 43*(5), 616-634. https://doi.org/10.1016/s0010-9452(08)70492-1

Esposito, F., Bertolino, A., Scarabino, T., Latorre, V., Blasi, G., Popolizio, T., Tedeschi, G., Cirillo, S., Goebel, R., & Di Salle, F. (2006). Independent component model of the default-mode brain function: Assessing the impact of active thinking. *Brain Research Bulletin, 70*(4-6), 263-269. https://doi.org/10.1016/j.brainresbull.2006.06.012

Fishburn, F. A., Norr, M. E., Medvedev, A. V., & Vaidya, C. J. (2014). Sensitivity of fNIRS to cognitive state and load. *Frontiers in Human Neuroscience, 8,* 76. https://doi.org/10.3389/fnhum.2014.00076

Friedmann, D., & Johnson, R. J. (2000). Event-related potential (ERP) studies of memory encoding and retrieval: A selective review. *Microscopy Research and Technique, 51*, 6-28. https://doi.org/10.1002/1097-0029(20001001)51:1<6::AID-JEMT2>3.0.CO;2-R

Hørlyck, L. D., Bisby, J. A., King, J. A., & Burgess, N. (2019). Wakeful rest compared to vigilance reduces intrusive but not deliberate memory for traumatic videos. *Scientific Reports, 9*, 13403. https://doi.org/10.1038/s41598-019-49634-8

Huijbers, W., Pennartz, C. M. A., Cabeza, R., & Daselaar, S. M. (2011). The hippocampus is coupled with the default mode network during memory retrieval but not during memory encoding. *PLoS ONE, 6*(4): e17463. https://doi.org/10.1371/journal.pone.0017463

Ji, D., & Wilson, M. A. (2007). Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nature Neuroscience, 10*, 100-107. https://doi.org/10.1038/nn1825

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavioral Research Methods, 47,* 1122-1135. https://doi.org/10.3758/s13428-014-0532-5

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

Mednick, S. C., Cai, D. J., Shuman, T., Anagnostaras, S., & Wixted, J. T. (2011). An opportunistic theory of cellular and systems consolidation. *Trends in Neurosciences, 34*(10), 504-514. https://doi.org/10.1016/j.tins.2011.06.003

Peirce, J. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics, 2*, 10. https://doi.org/10.3389/neuro.11.010.2008

Qin, S., Hermans, E. J., van Marle, H. J. F., Luo, J., & Férnandez, G. (2009). Acute psychological stress reduces working memory-related activity in the dorsolateral prefrontal cortex. *Biological Psychiatry, 66*(1), 25-32. https://doi.org/10.1016/j.biopsych.2009.03.006

Ranganath, C., Yonelinas, A. P., Cohen, M. X., Dy, C. J., Tom, S. M., & D'Esposito, M. (2013). Dissociable correlates of recollection and familiarity within the medial temporal lobes. *Neuropsychologia, 42,* 2-13. https://doi.org/10.1016/j.neuropsychologia.2003.07.006

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers, 31,* 137-149. https://doi.org/10.3758/BF03207704

Tambini, A., & Davachi, L. (2013). Persistence of hippocampal multivoxel patterns into postencoding rest is related to memory. *Proceedings of the National Academy of Sciences, 110*(48), 19591-19596. https://doi.org./10.1073/pnas.1308499110

Thomson, D. R., Besner, D., & Smilek, D. (2013). In pursuit of off-task thought: Mind wandering-performance trade-offs while reading aloud and color naming. *Frontiers in Psychology, 4,* 360. https://doi.org/10.3389/fpsyg.2013.00360

Unsworth, N., & Robison, M. K. (2018). Tracking arousal state and mind wandering with pupillometry. *Cognitive, Affective, & Behavioral Neuroscience, 18,* 638-664. https://doi.org/10.3758/s13415-018-0594-4

Varma, S., Daselaar, S. M., Kessels, R. P. C., & Takashima, A. (2018). Promotion and suppression of autobiographical thinking differentially affect episodic memory consolidation. *PLoS ONE, 13*(8), e0201780. https://doi.org/10.1371/journal.pone.0201780

Varma, S., Takashima, A., Fu, L., & Kessels, R. P. C. (2019). Mindwandering propensity modulates episodic memory consolidation. *Aging Clinical and Experimental Research, 31,* 1601-1607. https://doi.org/10.1007/s40520-019-01251-1

Varma, S., Takashima, A., Krewinkel, S., van Kooten, M., Fu, L., Medendorp, W. P., Kessels, R. P. C, & Daselaar, S. M. (2017). Non-interfering effects of active post-encoding tasks on episodic memory consolidation in humans. *Frontiers in Behavioral Neuroscience, 11*, 54. https://doi.org/10.3389/fnbeh.2017.00054

Weinstein, Y. (2018). Mind-wandering, how do I measure thee with probes? Let me count the ways. *Behavioral Research Methods, 50,* 642-661. https://doi.org/10.3758/s13428-017-0891-9

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46*(3), 441-517. https://doi.org/10.1006/jmla.2002.2864