Universiteit Leiden

Psychologie
Faculteit der Sociale Wetenschappen

Effect of Post-Encoding Task Difficulty on Episodic Memory Consolidation

Name: Macy Ho
Student number: 2689413
Date: 31-7-2020
Supervisor: Samarth Varma
Second reader: Wouter Boekel
Word count: 5922
Cognitive Psychology
Thesis MSc Applied Cognitive Psychology

Abstract

Recent studies showed that performing a 2-Back during post-encoding had similar facilitating effects on memory consolidation as wakeful rest, possibly due to a balance between the interreference effects of autobiographical thinking (suppressed during 2-Back) and the facilitating effects of offline replay (increased during rest) on memory consolidation. The current study investigates the relationship between autobiographical thinking and memory consolidation in a within-subjects design, by comparing two post-encoding periods differing in task difficulty: 0-Back and 2-Back. Participants (N = 22) performed two sessions of three blocks; an encoding task, a 15-minute post-encoding period, and a recognition task. Autobiographical thinking was measured via experience-sampling thought probes. We predicted autobiographical thinking to interfere with consolidation and more autobiographical thinking during the 0-Back compared to the 2-Back. We also expected equal memory consolidation across conditions, assuming increased offline replay during the 0-Back. As hypothesized, autobiographical thinking was higher during the 0-Back than during the 2-Back, and results showed no difference in memory consolidation across the post-encoding conditions. The present findings did not show a direct relationship between autobiographical thinking during post-encoding and consolidation. We argue that the failure to detect a relationship can be explained by the current study being underpowered, due to small sample size and the number of thought probes. In conclusion, results do show that manipulating task difficulty of the n-Back can be an effective method to affect autobiographical thinking. However, the relationship of autobiographical thinking and memory consolidation remains unclear, and requires further research.

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

Effect of Post-Encoding Task Difficulty on Episodic Memory Consolidation

For centuries we have been searching for methods and strategies to decrease forgetting. Newly formed memories are labile and vulnerable to interference (Lechner et al., 1999). Episodic memory consolidation is an essential cognitive function, in which labile memory content becomes stable, long-lasting memories with increased probability of recall at a later point in time (Dudai et al., 2015). Memory will become less reliable on the temporary hippocampal store and more on the permanent neocortical store, which will decrease the chance of forgetting due to integration with the existing knowledge network (Squire et al., 2015).

Consolidation begins within seconds after memory is encoded and can last for days, months or even years, depending on the memory system and task (Dudai et al., 2015). During that post-encoding period, activities can facilitate or interfere with memory consolidation. Research has shown that brief periods of wakeful rest after encoding facilitate consolidation (Dewar et al., 2007). The positive effects of wakeful rest can be explained by an increase of offline replay which is associated with memory consolidation (Carr et al., 2011; Tambini et al., 2010). Offline replay is a neural process, in which the hippocampus triggers repeated reactivation of recently encoded memory traces in the brain. During wakeful rest the hippocampus is disengaged from encoding new memories, through which brain resources become available for offline replay (Mednick et al., 2011).

Opposite to wakeful rest, performing different cognitive tasks (e.g. watching videos, spot-the-difference games, tone detection task) after encoding has shown to interfere with memory consolidation (Dewar et al., 2007). Varma et al. (2017) hypothesized that interference of a cognitive task may arise from increased demand on hippocampal and other memory resources. During cognitive demanding tasks brain resources can be limited, because of the constant metabolic consumption of the brain (Raichle & Gusnard, 2002). A demanding cognitive task can cause reallocation of brain resources towards the task, and away from consolidation-relevant processes, such as offline replay. This could lead to a decrease in memory consolidation.

Just like external cognitive tasks can interfere with memory consolidation, internal ones can too. A resting mind tends to wander (Smallwood & Schooler, 2006) and focusing on internal information during post-encoding, such as autobiographical thinking (i.e. recall of individual memories of specific events), can decrease memory consolidation. Specifically, promotion of autobiographical thinking (i.e. via familiar sound-cues during post-encoding), has shown to interfere with memory consolidation (Craig et al., 2014; Varma et al., 2018).

Furthermore, neuroimaging studies have shown that autobiographical thinking is associated with activity in the default-mode network (DMN). The DMN consists of brain regions, including the hippocampus, that remain active during wakeful rest (Mason et al., 2007). Autobiographical thinking could interfere with consolidation by reallocating hippocampal resources necessary for memory consolidation (Mednick et al., 2011). From these results it seems that anything other than wakeful rest, whether an internal task or an external task, cause interference.

Interestingly, a study by Varma et al. (2017) seemed to challenge the prevalent notion that any form of mental effort causes interference. They administered a 2-Back task – a cognitively demanding task used to measure working memory – during a post-encoding period and found that it did not interfere with memory consolidation. Varma et al. (2017) elaborate on two explanations of how engaging in a 2-Back during post-encoding does not interfere with memory consolidation. Firstly, an n-Back task might not interrupt consolidation processes due to the hippocampal independent nature of the n-Back task. Engaging in a 2-Back task even decreases hippocampal activity (Esposito et al., 2006), that could drive the facilitation of memory consolidation (Mednick et al., 2011). Secondly, difficulty of the 2-Back task is likely to suppress autobiographical thinking. Varma et al. (2017) hypothesized that autobiographical thinking interferes with memory consolidation, but conclusive results are still lacking.

The 2-Back task is a form of what is called an n-Back task (Owen et al., 2005). Usually participants have to monitor a series of stimuli and indicate whether the currently presented stimulus is the same as the defined number (‘n’) of trials before. Task difficulty can be manipulated by changing the value of ‘n’. Required brain resources differ across difficulty levels, but sustained attention, change detection, and memory of the rules are required in all task loads (Miller et al., 2009). Performing a 1-Back or 2-Back task is accompanied by deactivation of different brain regions of the DMN, but this deactivation has not been detected while performing a 0-Back (Cousijn et al., 2012; Esposito et al., 2006). Brain imaging studies show a linear brain response to increasing task load, meaning when task load increases the DMN, including the hippocampus, more strongly deactivates (Smith et al., 2018).

A later study by Varma et al. (2018) compared memory consolidation across post-encoding periods, in which autobiographical thinking was manipulated. One post-encoding period involved wakeful rest accompanied with familiar sound cues, which promoted autobiographical thinking. The other post-encoding involved executing a 2-Back task, which suppressed autobiographical thinking. However, they did not observe a direct relationship between autobiographical thinking and memory consolidation (Varma et al., 2018). Possibly due to methodological issues. Autobiographical thinking during the post-encoding periods was measured using a questionnaire at the end of the experiment. The questionnaire was unable to ascertain the content and degree of mindwandering. Moreover, the delay of administering the questionnaire possibly reduced accuracy and reliability in reporting thought proportions. Being unable to thoroughly measure participants' thoughts during the post-encoding period, the cognitive processes activated during the n-Back task still remain unclear.

Current study

As an extension of previous studies done by Varma et al. (2018) we further investigate the relationship between autobiographical thinking and memory consolidation. The procedure consisted of two sessions of three blocks; an encoding task, a post-encoding period filled with 0-Back or 2-Back in a counterbalanced order, and a recognition task. To capture resources that were engaged in memory consolidation processes, permissibility of autobiographical thinking was affected by a manipulation of task difficulty of the n-Back task during the post-encoding periods; 0-Back permits high autobiographical thinking, where 2-Back permits low autobiographical thinking. Previous research has shown a robust significant difference in cognitive load across these conditions, reaction times were faster and accuracy was higher in the 0-Back condition compared to the 2-Back condition (Cousijn et al., 2012; Qin et al., 2009).

During the post-encoding periods, autobiographical thinking was measured with the use of experience-sampling thought probes (i.e. frequently asking participants to classify their current thoughts), which has been shown to be a reliable method to investigate the frequency and content of thoughts (Csikszentmihalyi & Larson, 1987; Weinstein, 2018). We will operate under the definition of autobiographical thinking as off-task thoughts. Thoughts about the n-Back task, n-Back performance, rehearsal of encoding stimuli, or the absence of thought (mind blanking) will not be considered as autobiographical thoughts in this study, which is consistent with previous research (Unsworth & Robison, 2018).

By increasing task-load of the n-Back we expect a decrease in the proportions of autobiographical thoughts. Thus, we firstly expect participants to engage in more autobiographical thinking during the 0-Back task compared to the 2-Back task. Previous research which used experience-sampling probes during an n-Back task, has shown consistent results, namely more off-task thoughts in low-load conditions compared to high-load conditions, e.g. 0-Back vs. 2-Back (Ju & Lien, 2018), 1-Back vs. 3-Back (Robison et al., 2020; Rummel & Boywitt, 2014), 0-Back vs. 1-Back vs. 2-Back (Iijima & Tanno, 2012).

Secondly, we predict autobiographical thinking to interfere with memory consolidation based on the assumption that during post-encoding, autobiographical thinking causes reallocation of brain resources away from memory consolidation processes (Craig et al., 2014; Varma et al., 2018). We therefore expect memory performance to be negatively correlated with autobiographical thinking.

Previous research that compared post-encoding periods, involving wakeful rest or engaging in a 0-Back task, did not found a difference in memory consolidation of declarative memory (Hørlyck et al., 2019). Additionally, several experiments reported no difference in memory consolidation between wakeful rest and 2-Back (Varma et al., 2017). Following this, we thirdly expect no difference in memory performance between the 0-Back task and 2-Back task. This hypothesis seems contradictory to our previous hypotheses and requires more explanation. The hypothesis is based on the idea that engaging in a 0-Back, due to its low difficulty, would result in a similar effect as engaging in wakeful rest during post-encoding. By increasing task-load we do not only expect a decrease in autobiographical thinking but also a decrease in offline replay. It is suspected that a decrease in autobiographical thinking would facilitate consolidation. Conversely, a decrease in offline replay should interfere with consolidation. These two effects should counter-act. Despite this we did not measure offline replay; we assume that during a post-encoding task that only requires low mental effort (e.g. wakeful rest), the chances of offline replay increase (Varma et al., 2018). It is hypothesized that the balance between autobiographical thinking and offline replay stay the same across post-encoding periods, resulting in no difference in memory consolidation between post-encoding periods occupied with a 0-Back or 2-Back.

Results from this research will deepen scientific understanding of the relationship between, cognitive resources, offline replay and autobiographical thinking involved in memory consolidation. Results from the current behavioral study might pave way for future neurological research. Furthermore, it may serve as a basis for developing optimal learning strategies, by knowing what interferes with memory consolidation and how forgetting can be decreased.

Methods

Design

This study was conducted online, consisted in two sessions, with 24-72 hours between the sessions. Each session consisted of an encoding task, a ~15-minute post-encoding period (n-Back: 9 min, thought probes: max 6 min), and a recognition task, see Figure 1. A within-subjects design was applied, with a difficulty manipulation of the post-encoding period (low-load: 0-Back vs. high-load: 2-Back). The order of post-encoding conditions was counterbalanced across subjects. During the 0-Back task and 2-Back task experience-sampling probes were shown to measure autobiographical thinking. Dependent variables were the two memory performance scores associated with each condition, and the proportion of autobiographical thoughts. PsychoPy software was used to design the experiment and it was hosted on an online repository called Pavlovia.org (Peirce et al., 2019).

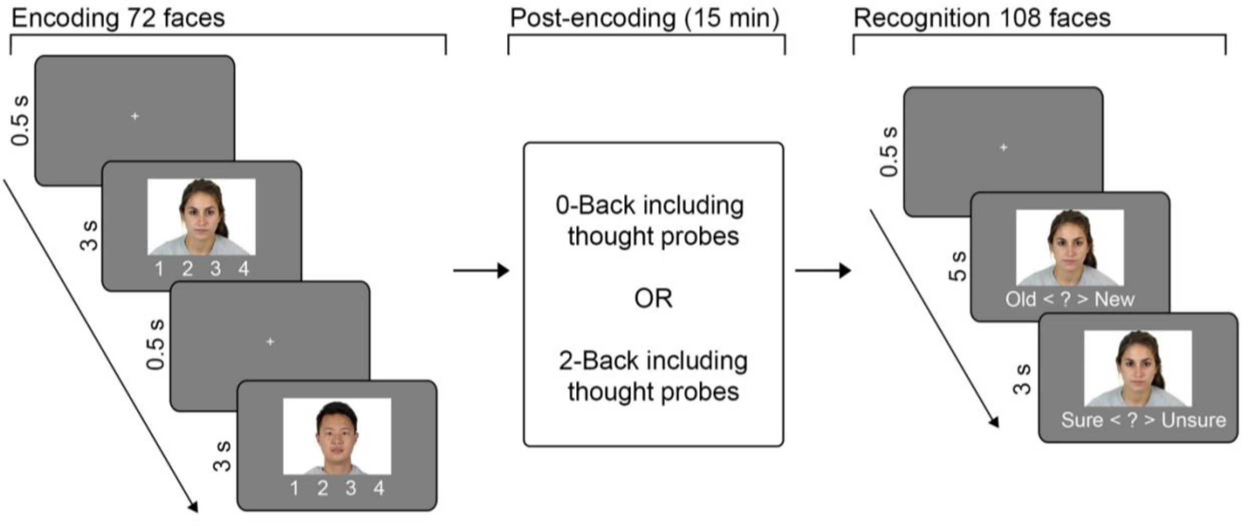

Figure 1. General experimental design per session, involved encoding of 72 faces. Followed by a post-encoding period of 15 min, occupied either by 0-Back or 2-Back, both included experience-sampling thought probes. After the post-encoding period the recognition task started, in which 108 faces (72 new, 36 old) were shown. Participants indicated if the face was old or new, and subsequently how confident they were about their response.

Participants

For this online experiment, 34 participants (23 female; M_{age} = 25.41, *SD* = 2.13) were recruited. Based on prior work, 36 participants were required for reliable statistical power, on a two condition experiment (Experiment 3; Varma et al., 2017). All participants had a high education level (i.e. college or university education) and were aged between 22 and 31 years. Participants with face recognition impairments, color-blindness or visual constraints, factors that would have made it impossible to perform the experiment, were excluded. After outlier removal, based on poor performance on the n-Back tasks (combined d-prime < 2-SD below group average), non-conformance, or technical errors, 22 participants (14 female; M_{age} = 25.59, *SD* = 2.15) were considered for analysis. This experiment was approved by the Ethics

Committee of the Faculty of Social Sciences of Leiden University. Participants filled an online consent form in accordance with the Declaration of Helsinki, after reading an instruction form and watching an instruction video. Participation was voluntarily, with a choice of either no compensation or monetary compensation (€7).

Materials

Encoding and Recognition lists

The to-be-learned images consisted of female and male faces with a neutral expression (downloaded from Chicago Face Database; Ma et al., 2015). Stimulus material was difficult to rehearse and challenging to retain, which was necessary to prevent ceiling effects. Per participant, 216 unique faces were randomly pooled, while keeping equal number of male and female faces. Each encoding list comprised of 72 faces, and each recognition list comprised of the same 72 faces and 36 new faces (total 108 faces). Outside the pooled faces, 12 faces were used for practice and instruction video.

Thought probes

During the post-encoding periods, filled with a 0-Back or 2-Back task, autobiographical thoughts were measured via experience-sampling thought probes. Participants were periodically presented with thought probes, asking them to classify the current contents of their thoughts by choosing from a list of options. Response options for the thought probes were based on prior research of mindwandering (Unsworth & Robison, 2018). Participants were asked (18 times), “Just before this interrupt were you…?”. Possible responses were (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, (7) Unknowingly thinking about personal stuff. Participants responded by pressing corresponding number on the keyboard. Responses 2, 6, and 7 were combined as off-task thoughts.

Procedure

The experiment was conducted online and consisted of two 30-minute sessions to keep participants motivated and focused during the experiment. Participants received instructions via an e-mail, with an information document and instruction video. After signing informed consent, links for opening the experiment were send to the participants on scheduled dates. Participants were able to schedule their experimental sessions with 24-72 hours between the sessions, and were allowed to start the experiment on any time point of that day. The experiment was executed on pavlova.org, at home on a personal computer or laptop. Clear instructions to perform the experiment in a quiet room without any distractions were given, to keep consistency over the experiment.

Each session started with a practice round, which consisted of a short version of all the elements of the experiment; encoding task, n-Back task (with thought probes), and recognition task. After practice participants began with an encoding task. One by one, 72 unique faces were shown for 3 seconds. Participants were instructed to memorize the faces and judge them on friendliness by using the keys 1 to 4 (1 = 'definitely unfriendly', 4 = 'definitely friendly'), with a maximum allowed response duration of 3 seconds.

The encoding task was directly followed by one of the two counterbalanced post-encoding conditions. The post-encoding periods were filled with a 0-Back or 2-Back task, representing low- and high-load tasks. A modified n-Back version similar to previous studies was used (Varma et al., 2017). During both n-Back tasks, greyscale numbers (1 to 5) appeared in a fixed pace (2 s) in the middle of a dark screen. The target in 0-Back condition was a number that matches the pre-specified number 3. In the 2-Back condition, the target was a number that was identical to the one shown two trails back (Figure 2). When a target was shown, the participants had to press "right", otherwise they had to press "left". Short feedback via coloring was shown to keep participants' attention and induce optimal performance. When participants gave a correct response in the given time the number turned green for 300 ms. If the participants were too late with responding or if they gave an incorrect response the number turned red for 300 ms. During the n-Back tasks, 18 thought probes appeared with a quasi-random timing with an average interval of 30 s (range from 18 s to 42 s). The participants had to indicate their current thought by pressing a corresponding number on the keyboard with a maximum response duration of 20 seconds.

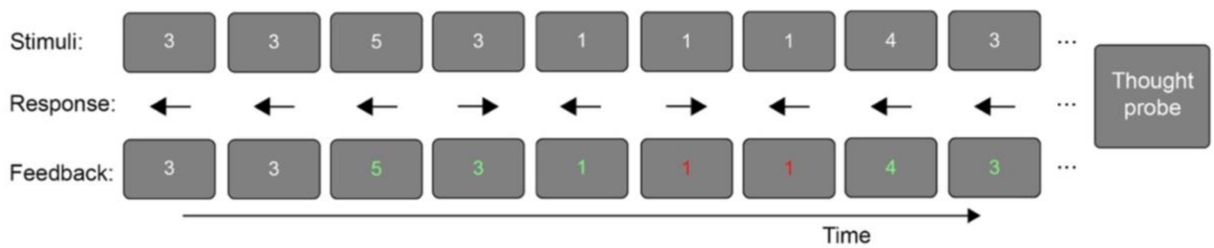

Figure 2. Sample sequence of the 2-Back task. Sequence was 18 times interrupted by a thought probe, in which participant had to classify content of current thoughts. When a target was shown, participants had to press the "right arrow" key, otherwise they had to press the "left arrow" key. Colored numbers indicate given feedback, correct response (green) or incorrect response (red).

After the ~15 minute post-encoding period, the recognition task started. Participants had 5 seconds to indicate if they recognized the presented face or not, by pressing the key “left” if they saw an ‘old’ face and “right” if they saw a ‘new’ face. In the next screen, participants had 3 seconds to indicate the confidence about their response (left = ‘sure’, right = ‘unsure’). The face remained on the screen during old/new and sure/unsure identification. During each recognition task 108 faces (72 old, 36 new) were shown. After finishing both sessions, the participants were debriefed and compensated according to their choice (no compensation or monetary compensation of €7).

Analyses

Prior to calculation of memory performance of each condition, data was cleaned by removing recognition trails of any stimuli that was not encoded correctly or not responded to during the encoding task. Furthermore, only confident responses were used for calculating d-prime, which minimalized the influence of guessing on memory performance.

N-Back task performance is represented by reaction time and calculated d-prime scores of n-Back performance. Memory performance on the recognition tasks is also represented by calculated d-prime scores, the standardized difference between hit and false alarm rates (Stanislaw & Todorov, 1999). Hit rates of each recognition task were calculated separately by dividing the number of “old” trials that were correctly identified, by the total number of “old” faces to which the participant responded. False alarm rates were similarly calculated by dividing the number of “new” trials that were incorrectly identified as “old”, by the total number of “new” faces to which the participant responded. From the experience-sampling data, the proportion of autobiographical thoughts were calculated for each condition separately, the sum of autobiographical thoughts (responses 2, 6, 7) were divided by the total amount of thought probes to which the participant responded

In order to confirm task difficulty difference between 0-Back and 2-Back, we ran a paired samples *t*-tests with reaction time and d-prime scores of n-Back task performance. To test the effect of task difficulty during the post-encoding periods on memory performance, we ran a 2 x 2 mixed repeated measures analysis of variance (RM-ANOVA), with post-encoding condition (0-Back, 2-Back) as within-participants factor, and post-encoding condition order (0-Back – 2-Back, 2-Back – 0-Back) as between-participants factor. Furthermore, we ran a paired-samples *t*-test to compare post-encoding conditions. We also ran a paired-samples *t*-test with data from the thought probes, to compare proportion off-task thoughts across the post-encoding conditions. To determine if memory performance was affected by autobiographical thinking, we carried out correlation tests (1-tailed) between memory performance and autobiographical thoughts for each of the post-encoding conditions. Spearman's rho (r_s) was used when assumptions of normality were violated (Shapiro-Wilk, $p < .05$), otherwise Pearson product-moment correlation coefficient was computed. IBM SPSS 25 was used for analyzing all results, with an alpha level of 0.05.

Results

N-Back performance

Performance on the n-Back tasks were analyzed by calculating d-prime scores per post-encoding condition. To compare post-encoding conditions, we performed paired samples *t*-tests. Performance was higher on the 0-Back task (*M* = 2.32, *SD* = 0.12) compared to performance on the 2-Back task (*M* = 1.57, *SD* = 0.45), $t(21) = 7.27$, $p < .001$. Reaction times were faster during the 0-Back task (*M* = 0.46 s, *SD* = 0.04), than during the 2-Back task (*M* = 0.65 s, *SD* = 0.09), $t(21) = -10.76$, $p < .001$. Lower d-prime score and longer reaction time during the 2-Back task shows that manipulation of difficulty was successful during the post-encoding periods.

Thought probes

For the hypothesis that autobiographical thinking (responses 2, 6, 7) is higher during the 0-Back task compared to the 2-Back task, a paired samples *t*-test was carried out. As expected, the experience-sampling thought probes show a difference in autobiographical thoughts between the task difficulty conditions, $t(21) = 4.78$, $p < .001$. Proportions of autobiographical thoughts was higher during the 0-Back task (*M* = 0.25, *SD* = 0.18) compared to the 2-Back task (*M* = 0.07, *SD* = 0.10), see Figure 3 and Table 1.

Correlation analyses were performed to investigate the relationship between autobiographical thoughts and memory performance. We expected a negative relationship between proportion of autobiographical thoughts during the post-encoding period and memory performance. However, there was no significant correlation between autobiographical thoughts and memory performance in the 0-Back task, $r_s(22) = -.07$, $p = .373$ (1-tailed) and in the 2-Back condition $r_s(22) = -.13$, $p = .279$ (1-tailed). Considering all responses (sure and unsure) for analysis, showed similar results; no significant correlation between autobiographical thoughts and memory performance in the 0-Back task, $r_s(22) = -.23$, $p = .149$ (1-tailed) and in the 2-Back condition $r_s(22) = -.02$, $p = .472$ (1-tailed).

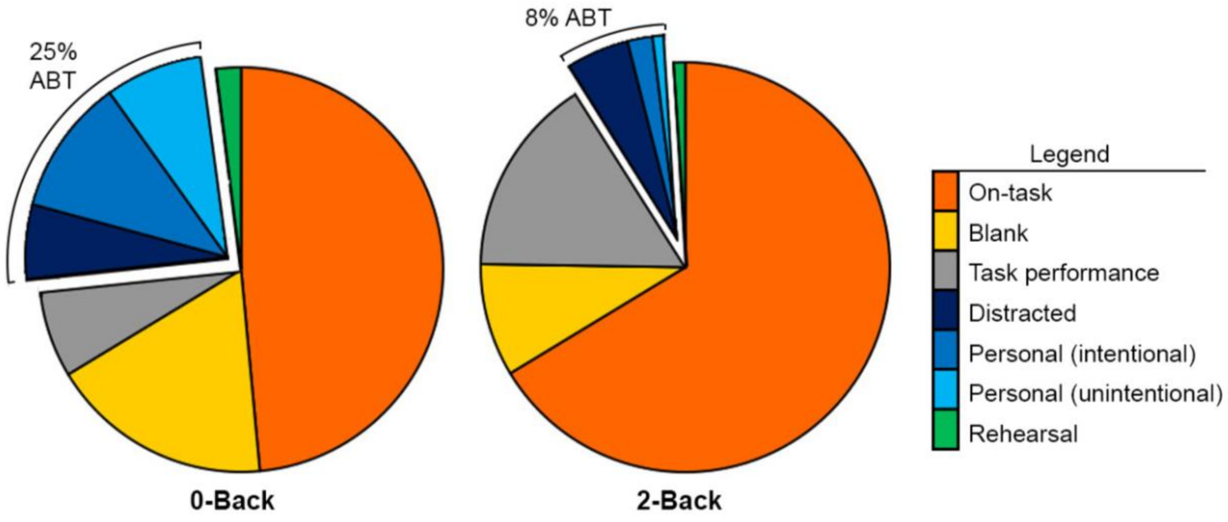

Figure 3. Results from experience-sampling thought probes, showing average proportions of thought probe categories during post-encoding periods; 0-Back (left) and 2-Back (right). As shown proportional autobiographical thinking during the 0-Back (25%) is higher compared with the 2-Back (8%). Here 'ABT' stands for autobiographical thinking.

Table 1

Thought Proportions During 0-Back and 2-Back Task

Thought probe category	0-Back condition	2-Back condition
On-task	49%	67%
Blank	18%	9%
Task performance	7%	16%
Distracted	6%	5%
Personal (intentional)	11%	2%
Personal (unintentional)	8%	1%
Rehearsal	2%	1%

Memory performance

The results from the RM-ANOVA showed, as predicted, no significant main effect of post-encoding condition on memory performance, $F(1, 20) = 2.06$, $p = .167$, $\eta_{2p} = .093$. This means that participant's memory performance did not differ when the post-encoding period consisted of a low-load task; 0-Back ($M = 2.09$, $SD = 0.82$) or a high-load task; 2-Back ($M = 2.30$, $SD = 0.80$). There was no main effect of post-encoding condition order on overall memory performance, $F(1, 20) = 3.82$, $p = .065$, $\eta^2_p = .160$. Although significance was not reached, memory performance in both conditions was higher when the experiment ended with the 2-Back task (0-Back: $M = 2.39$, $SD = 0.70$, 2-Back: $M = 2.53$, $SD = 0.82$) compared to the experiment ending with the 0-Back task (0-Back: $M = 1.74$, $SD = 0.85$, 2-Back: $M = 2.02$, $SD = 0.71$). The interaction between post-encoding condition and post-encoding order was also not significant, $F(1, 20) = 0.21$, $p = .654$, $\eta^2_p = .010$. While we counterbalanced post-encoding condition order equally across participants, due to outlier removal the groups were not equally distributed (2-Back – 0-Back $n = 10$, 0-Back – 2-Back $n = 12$). However, since we did not find an effect of order on memory performance or an interaction effect between order and task difficulty conditions, we can assume the unequal groups did not affect our results.

After ruling out the effect of order we performed a paired samples *t*-test, which showed no difference in d-prime scores across post-encoding conditions, $t(21) = -1.43$, $p = .169$, see Figure 2. We also looked at differences in the post-encoding condition memory scores by calculating memory performance while considering both ‘sure’ and ‘unsure’ trails together. In that analysis we found a difference in memory performance between post-encoding conditions, $t(21) = -2.16$, $p = .040$. Memory performance was higher when encoding was followed by a 2-Back task ($M = 1.67$, $SD = 0.43$) compared to the 0-Back task ($M = 1.46$, $SD = 0.56$).

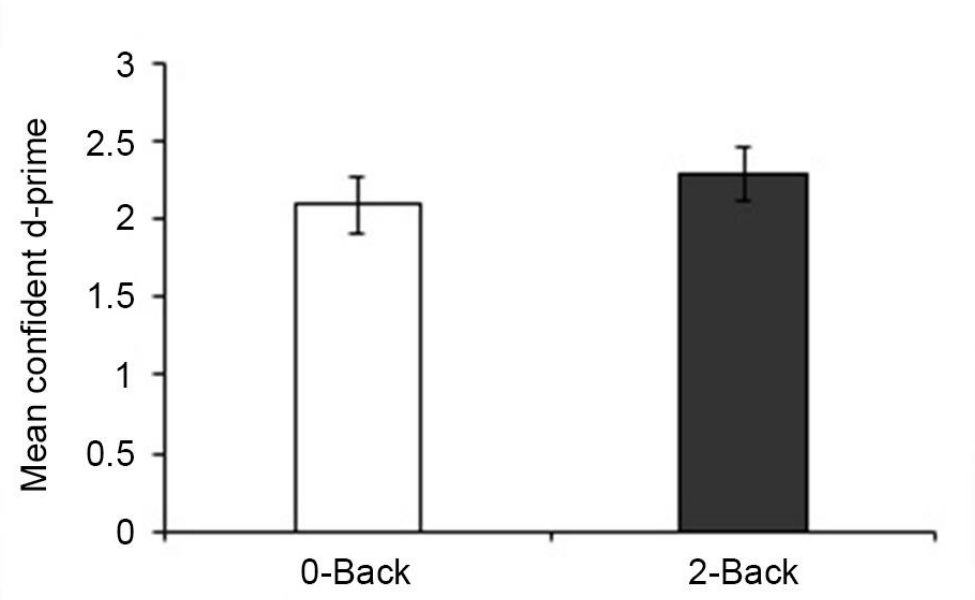

Figure 4. The effects of post-encoding tasks on mean d-prime scores (included only confident trials). No significant difference between 0-Back and 2-Back was found. Error bars represent standard errors.

Discussion

The aim of the current study was to examine the specific prerequisites for interfering with episodic memory consolidation. Previous research has challenged the notion of mental effort as the all-encompassing cause of memory inference and suspected autobiographical thinking to interfere with memory consolidation (Varma et al., 2017). The current study manipulated post-encoding task difficulty, and measured frequency and content of thoughts during the n-Back tasks (0-Back, 2-Back). The results of our analyses showed that, (a) autobiographical thinking was higher during the 0-Back task compared to the 2-Back task, (b) there was no relationship between autobiographical thinking during the post-encoding periods and memory performance, and (c) there was no difference in memory performance between post-encoding periods (0-Back vs. 2-Back). In subsequent sections, we address the findings in more detail and explain their implications.

In agreement with prior work that has shown more off-task thoughts during low-load tasks compared with high-load tasks (Seli et al., 2018), we observed more autobiographical thoughts during the 0-Back compared with the 2-Back task. Results can be explained by reallocation of limited brain resources, necessary for performing an n-Back task and autobiographical thinking. Performing a 0-Back task recruits relatively few executives resources, which leaves enough resources available for autobiographical thinking (Smallwood & Schooler, 2006). Conversely, the high task difficulty of the 2-Back leaves fewer resources available for autobiographical thinking. This finding provides evidence for the suitability of the current research design, considering that manipulating post-encoding task difficulty of the n-Back task indeed leads to different levels of autobiographical thinking.

Another finding concerns the relationship between autobiographical thinking and memory consolidation. Contrary to our expectations, we did not observe a direct correlation between autobiographical thinking and memory consolidation associated with each condition. This result implies that autobiographical thinking does not interfere with memory consolidation. However, this seems very unlikely in comparison with previous studies (Craig et al., 2014; Varma et al., 2018). Autobiographical thinking during post-encoding was expected to interfere with memory consolidation, due to the increased engagement of the hippocampus (Mednick et al., 2011). Despite our results, we still speculate that reallocation of brain resources towards autobiographical thinking away from memory consolidation processes can cause interference. Although it is possible that in the current study autobiographical thinking was too superficial to cause such reallocation. Engagement in an n-Back task may not allow for rich autobiographical thoughts to occur. Previous research has shown that vividness and the amount of detail of autobiographical thinking affects hippocampal involvement (Svoboda et al, 2006). The suspected superficial autobiographical thoughts may not involve enough engagement of the hippocampus to cause interference with memory consolidation. Furthermore, failure to detect a significant relationship between autobiographical thinking and memory consolidation can be explained by design limitations. It is possible that the low frequency of experience-sampling thought probes was insufficient to measure autobiographical thinking properly (see limitations section).

As previously explained, we expected equal memory consolidation across post-encoding conditions (0-Back, 2-Back). That may seem contradictory to the previous two hypotheses but is explained by a suspected retained balance between offline replay and autobiographical thinking across post-encoding conditions. In agreement with prior work that has shown no difference in memory consolidation between post-encoding periods involving wakeful rest or a 2-Back task (Varma et al., 2017), we observed no difference in memory performance between post-encoding periods involving a 0-Back or a 2-Back task. This finding supports our assumption that 0-Back acts similar as wakeful rest during post-encoding, possible due to its low difficulty (Hørlyck et al., 2019). No difference in memory consolidation between conditions can be explained by a retained balance between offline replay and autobiographical thinking, because offline replay is assumed to facilitate consolidation and autobiographical thinking to interfere with consolidation (Varma et al., 2018). We speculate that increasing task difficulty of the n-Back equally decreases offline replay and autobiographical thinking. However, this is stated with caution because in the current experimental design only autobiographical thinking was measured.

The findings discussed above were based only on confident d-prime scores. Analyses including non-confident trials led to a different result: higher memory performance in the post-encoding 2-Back condition compared to 0-Back condition. Analyses considering only confident trials are based on recollection, whereas considering all trials (sure and unsure) are based on both recollection and familiarity (Yonelinas et al., 2010). Recollection and familiarity are distinct processes and dependent on different subregions of the medial temporal lobe. Manipulating task difficulty of n-Back tasks affects required brain resources (Smith et al., 2018), and can affect participants' strategies to perform the n-Back task (Lovett et al., 2000). Observing different results can be explained in terms of participants using familiarity strategies as supposed to only recollection during the recognition task to identify an item as old or new. It is possible that the 2-Back task allows for greater chances of familiarity-based recognition. Since recollection and familiarity are two distinct processes it is possible that it is differently affected by post-encoding task difficulty. The additional analyses considering all trials is important to report, because during a recognition task both recollection and familiarity play a role. However, conclusions are drawn upon the confident trials, because the current study is focused on episodic memory based on recollection that reflects the retrieval of qualitative information and not familiarity (Yonelinas et al., 2010).

Limitations and Future Directions

A few weaknesses and limitations of the current study are worth mentioning. First and foremost, due to the relatively small number of participants, results need to be interpreted with a general caution. Finding no relationship between autobiographical thinking and consolidation, and no effect of post-encoding periods (0-Back vs. 2-Back) on consolidation, could have been the result of inadequate power to detect an effect. Planned sample size of 36 participants, based on prior work of Varma et al. (Experiment 3; 2017), was not met due to circumstances of the 2020 Coronavirus. It is possible that sufficient powered future studies detect a signification relationship between autobiographical thinking and memory consolidation. In order to assure sufficient power, future studies should consult papers of Humiston et al. (2019) to find true effect of post-encoding wakeful rest and experiments of Varma et al. (2017) to find true effect of post-encoding n-Back tasks.

Furthermore, it is possible that the amount of thought probes was insufficient, which may have limited our ability to measure autobiographical thinking during the n-Back tasks. In the current design, we included 18 thought probes during each post-encoding period. The amount of thought probes could have been higher by increasing the post-encoding duration. Previous research included 30 probes during a 2-Back task, which took approximately 30 minutes (Choi et al., 2017). However, in the current online study 30 minutes of the n-Back tasks was not deemed possible, as we suspected it would have been too tiring for the participants and they would lose motivation. In a lab setting, it may be more reasonable to expect participants to perform an n-Back task for such duration. Besides, to keep the duration of the post-encoding periods similar to previous interference studies, approximately 15 minutes of post-encoding was most appropriate (Dewar et al., 2012; Varma et al., 2017). The current study was unable to validate the used probing method (i.e. n-Back performance was not positively correlated with on-task thoughts and not negatively correlated with off-task thoughts), which should be prevented in future research. Since experience-sampling thought probes method is not commonly used during an n-Back task, the minimal amount of thought probes for reliable results is unclear. Based on previous research, we suggest to extent the post-encoding duration to include approximately 30 thought probes (Choi et al., 2017).

The online nature of the current study is also worth mentioning. Due to circumstances, it was not possible to conduct the current experiment in a laboratory setting. The online setting increased the chance of distraction and possibly decreased overall motivation. It is possible that the personal environment in which the experiment was conducted affected the current results. Previous research has shown that offline replay is more prevalent in a novel environment compared to a familiar one (Carr et al., 2011). Furthermore, Kane et al. (2017) showed a difference in off-task thoughts in daily life settings compared to lab settings. The degree in which offline replay and off-task thoughts were affected by the online setting is unknown. However, since we found similar effects as previous lab studies (Varma et al., 2018), it is reasonable to assume that the online setting did not have detrimental effects on the reliability of our results.

Even though we have discussed that task difficulty possibly affects the permissibility of offline replay similar as autobiographical thinking, our study design did not measure offline replay. Since our results showed no difference in memory consolidation between 0-Back and 2-Back, it is interesting to further investigate the effect of the balance between autobiographical thinking and offline replay on memory consolidation. Autobiographical thinking and offline replay can both be measured with the use of brain imaging techniques, such as functional magnetic resonance imaging (Chou et al., 2017; Levy & Wagner, 2013). Moreover, such brain imaging studies can form a neurological foundation to understand the effects of engaging in an n-Back task during a post-encoding period on memory consolidation. Due to our solely behavioral study design, we can only speculate and make assumptions about brain activations. Future neuroimaging research is necessary to substantiate our assumptions regarding resource allocation.

Conclusion

The aim of the current study was to deepen scientific understanding of the relationship between cognitive resources and autobiographical thinking involved in memory consolidation. We investigated the relationship between autobiographical thinking and memory consolidation by manipulating post-encoding task difficulty and measuring autobiographical thinking during post-encoding. Unfortunately, the present findings do not show a direct relationship between autobiographical thinking during post-encoding and memory consolidation. We argue that the failure to detect a relationship can be explained by the current study limitations, such as small sample size and insufficient amount of thought probes to measure autobiographical thinking. Nevertheless, this study can be used as a guide for future (neuroimaging) studies investigating memory consolidation. Previous research compared a 2-Back with wakeful rest (Varma et al., 2017). However, these conditions are different in many ways. Based on current findings and previous studies, 0-Back and 2-Back can be used in future research as a reliable method to manipulate the frequency and content of autobiographical thoughts. Comparing a 0-Back and 2-Back, which only differ in task difficulty, will decrease uncertainty about the interpretation of results. Building on the current findings, future research should further investigate the relationship between autobiographical thinking and memory consolidation. The use of neuroimaging studies can give clear insights of the balance between autobiographical thinking and offline replay, and how this balance affects memory consolidation. More knowledge about mental activities during post-encoding will deepen the understanding of the mechanisms behind memory consolidation.

References

Carr, M. F., Jadhav, S. P., & Frank, L. M. (2011). Hippocampal replay in the awake state: A potential substrate for memory consolidation and retrieval. *Nature Neuroscience*, *14*(2), 147–153. https://doi.org/10.1038/nn.2732

Choi, H. S., 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). https://doi.org/10.1371/journal.pone.0189667

Chou, Y. H., Sundman, M., Whitson, H. E., Gaur, P., Chu, M. L., Weingarten, C. P., Madden, D. J., Wang, L., Kirste, I., Joliot, M., Diaz., M. T., Li, Y. J., Song, A. W., & Chen, N. K. (2017). Maintenance and representation of mind wandering during resting-state fMRI. *Scientific Reports*, *7*, 40722. https://doi.org/10.1038/srep40722

Cousijn, H., Rijpkema, M., Qin, S., Van Wingen, G. A., & Fernández, G. (2012). Phasic deactivation of the medial temporal lobe enables working memory processing under stress. *NeuroImage*, *59*(2), 1161–1167. https://doi.org/10.1016/j.neuroimage.2011.09.027

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

Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the Experience-Sampling Method. *The Journal of Nervous and Mental Disease*, *175*(9), 526–536. https://doi.org/10.1097/00005053-198709000-00004

Dewar, M., Alber, J., Butler, C., Cowan, N., & Della Sala, S. (2012). Brief wakeful resting boosts new memories over the long term. *Psychological Science*, *23*(9), 955–960. https://doi.org/10.1177/0956797612441220

Dewar, M. T., Cowan, N., & Sala, S. Della. (2007). Forgetting due to retroactive interference: A fusion of Müller and Pilzecker's (1990) 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

Dudai, Y., Karni, A., & Born, J. (2015). The consolidation and transformation of memory. *Neuron*, *88*(1), 20–32. https://doi.org/10.1016/j.neuron.2015.09.004

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

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*(1), 13403. https://doi.org/10.1038/s41598-019-49634-8

Humiston, G. B., Tucker, M. A., Summer, T., & Wamsley, E. J. (2019). Resting states and memory consolidation: A preregistered replication and meta-analysis. *Scientific Reports*, *9*(1), 19345. https://doi.org/10.1038/s41598-019-56033-6

Iijima, Y., & Tanno, Y. (2012). The effect of cognitive load on the temporal focus of mind wandering. *Shinrigaku Kenkyu*, *83*(3), 232–236. https://doi.org/10.4992/jjpsy.83.232

Ju, Y. J., & Lien, Y. W. (2018). Who is prone to wander and when? Examining an integrative effect of working memory capacity and mindfulness trait on mind wandering under different task loads. *Consciousness and Cognition*, *63*, 1–10. https://doi.org/10.1016/j.concog.2018.06.006

Kane, M. J., Gross, G. M., Chun, C. A., Smeekens, B. A., Meier, M. E., Silvia, P. J., & Kwapil, T. R. (2017). For whom the mind wanders, and when, varies across laboratory and daily-life settings. *Psychological Science*, *28*(9), 1271–1289. https://doi.org/10.1177/0956797617706086

Lechner, H. A., Squire, L. R., & Byrne, J. H. (1999). 100 years of consolidation--remembering Muller and Pilzecker. *Learning and Memory*, *6*(2), 77–87. https://doi.org/10.1101/lm.6.2.77

Levy, B. J., & Wagner, A. D. (2013). Measuring memory reactivation with functional MRI: Implications for psychological Theory. *Perspectives on Psychological Science*, *8*(1), 72–78. https://doi.org/10.1177/1745691612469031

Lovett, M. C., Daily, L. Z., & Reder, L. M. (2000). A source activation theory of working memory: Cross-task prediction of performance in ACT-R. *Cognitive Systems Research*, *1*(2), 99–118. https://doi.org/10.1016/S1389-0417(99)00012-1

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*(4), 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 (New York, N.Y.)*, *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

Miller, K. M., Price, C. C., Okun, M. S., Montijo, H., & Bowers, D. (2009). Is the n-back task a valid neuropsychological measure for assessing working memory? *Archives of Clinical Neuropsychology*, *27*(7), 711–717. https://doi.org/10.1093/arclin/acp063

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping*, *25*(1), 46–59. https://doi.org/10.1002/hbm.20131

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Qin, S., Hermans, E. J., van Marle, H. J. F., Luo, J., & Fernández, 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

Raichle, M. E., & Gusnard, D. A. (2002). Appraising the brain's energy budget. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(16), 10237–10239. https://doi.org/10.1073/pnas.172399499

Robison, M. K., Miller, A. L., & Unsworth, N. (2020). A multi-faceted approach to understanding individual differences in mind-wandering. *Cognition*, *198*, 104078. https://doi.org/10.1016/j.cognition.2019.104078

Rummel, J., & Boywitt, C. D. (2014). Controlling the stream of thought: Working memory capacity predicts adjustment of mind-wandering to situational demands. *Psychonomic Bulletin and Review*, *21*(5), 1309–1315. https://doi.org/10.3758/s13423-013-0580-3

Seli, P., Konishi, M., Risko, E. F., & Smilek, D. (2018). The role of task difficulty in theoretical accounts of mind wandering. *Consciousness and Cognition*, *65*, 255–262. https://doi.org/10.1016/j.concog.2018.08.005

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

Smith, J., Browning, M., Conen, S., Smallman, R., Buchbjerg, J., Larsen, K. G., Olsen, C. K., Christensen, S. R., Dawson, G. R., Deakin, J. F., Hawkins, P., Morris, R., Goodwin, G., & Harmer, C. J. (2018). Vortioxetine reduces BOLD signal during performance of the N-back working memory task: A randomised neuroimaging trial in remitted depressed patients and healthy controls. *Molecular Psychiatry*, *23*(5), 1127–1133. https://doi.org/10.1038/mp.2017.104

Squire, L. R., Genzel, L., Wixted, J. T., & Morris, R. G. (2015). Memory consolidation. *Cold*

Spring Harbor Perspectives in Biology, *7*(8). https://doi.org/10.1101/cshperspect.a021766

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, and Computers*, *31*(1), 137–149. https://doi.org/10.3758/BF03207704

Svoboda, E., McKinnon, M. C., & Levine, B. (2006). The functional neuroanatomy of autobiographical memory: A meta-analysis. *Neuropsychologia*, *44*(12), 2189–2208. https://doi.org/10.1016/j.neuropsychologia.2006.05.023

Tambini, A., Ketz, N., & Davachi, L. (2010). Enhanced brain correlations during rest are related to memory for recent experiences. *Neuron*, *65*(2), 280–290. https://doi.org/10.1016/j.neuron.2010.01.001

Unsworth, N., & Robison, M. K. (2018). Tracking arousal state and mind wandering with pupillometry. *Cognitive, Affective and Behavioral Neuroscience*, *18*(4), 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., 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. *Behavior Research Methods*, *50*(2), 642–661. https://doi.org/10.3758/s13428-017-0891-9

Yonelinas, A. P., Aly, M., Wang, W. C., & Koen, J. D. (2010). Recollection and familiarity: Examining controversial assumptions and new directions. *Hippocampus*, *20*(11), 1178–1194. https://doi.org/10.1002/hipo.20864