OXFORD

*Cerebral Cortex*, 2022, 1–14
https://doi.org/10.1093/cercor/bhac270
**Original Article**

# The influence of imagery vividness and internally-directed attention on the neural mechanisms underlying the encoding of visual mental images into episodic memory

Eva Gjorgieva, MA[1,3,*], Benjamin R. Geib, PhD[1,3], Roberto Cabeza, PhD[1,3], Marty G. Woldorff, PhD[1,2,3]

[1]Department of Psychology and Neuroscience, Duke University, Durham, NC 27708, United States,
[2]Department of Psychiatry, Duke University, Durham, NC 27708, United States,
[3]Center for Cognitive Neuroscience, Duke Institute for Brain Sciences, Duke University, Durham, NC 27708, United States
*Corresponding author: Center for Cognitive Neuroscience, Box 90999, Duke University, Durham, NC 27708, United States. Email: eva.gjorgieva@duke.edu

Attention can be directed externally toward sensory information or internally toward self-generated information. Using electroencephalography (EEG), we investigated the attentional processes underlying the formation and encoding of self-generated mental images into episodic memory. Participants viewed flickering words referring to common objects and were tasked with forming visual mental images of the objects and rating their vividness. Subsequent memory for the presented object words was assessed using an old–new recognition task. Internally-directed attention during image generation was indexed as a reduction in steady-state visual evoked potentials (SSVEPs), oscillatory EEG responses at the frequency of a flickering stimulus. The results yielded 3 main findings. First, SSVEP power driven by the flickering word stimuli decreased as subjects directed attention internally to form the corresponding mental image. Second, SSVEP power returned to pre-imagery baseline more slowly for low- than high-vividness later remembered items, suggesting that longer internally-directed attention is required to generate subsequently remembered low-vividness images. Finally, the event-related-potential difference due to memory was more sustained for subsequently remembered low- versus high-vividness items, suggesting that additional conceptual processing may have been needed to remember the low-vividness visual images. Taken together, the results clarify the neural mechanisms supporting the encoding of self-generated information.

*Key words*: internally-directed attention; episodic memory; mental imagery; electroencephalography; SSVEP.

## Introduction

While sitting in a history class, your attention may be directed externally to the professor giving a lecture about a war, but you may also regularly shift your attention internally to form a mental image of a battle scene that is being discussed. Both externally- and internally-directed attention can determine your ability to learn, as both the external and internal attentional processes contribute to being able to successfully later remember the contents of the lecture. Chun et al. (2011) proposed a taxonomy distinguishing between externally- and internally-directed attentions in terms of the target of the attentional focus. More specifically, they defined externally-directed attention as the selection and modulation of sensory information (including objects, features, and spatial locations), and internally-directed attention as attention directed toward internally-generated information (such as task rules, mental imagery, or the contents of working memory and long-term memory). Although this distinction is fundamental, the vast majority of cognitive neuroscience research on attention has focused on externally-directed attention (for review, see Chun and Turk-Browne 2007; Aly and Turk-Browne 2017), and very little is known regarding the neural mechanisms of internally-directed attention and its impact on other cognitive processes. The overarching goal of the current study was to investigate the neural mechanisms by which internally-directed attention contributes to episodic memory encoding.

We employed a form of the hallmark "subsequent memory paradigm," which compares encoding activity for subsequently remembered versus subsequently forgotten items (Friedman and Johnson Jr 2000; Paller and Wagner 2002). In electroencephalographic (EEG) and event-related-potential (ERP) studies, the most investigated subsequent memory effect has been the difference due to memory (Dm), a broadly distributed, centro-parietal ERP during encoding occurring at ~400–800 ms after stimulus onset (Paller et al. 1987, 1988; Johnson 1995; Wagner et al. 1999). Previous research has shown that the Dm can be modulated by directing attention toward the conceptual or semantic meaning of an item, with a greater Dm observed for deeper levels of encoding (Paller et al. 1987; Guo et al. 2004). This finding has led to the suggestion that the Dm increases with

**Received:** December 2, 2021. **Revised:** June 10, 2022. **Accepted:** June 12, 2022
© The Author(s) 2022. Published by Oxford University Press. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

Downloaded from https://academic.oup.com/cercor/advance-article/doi/10.1093/cercor/bhac270/6659085 by Duke University user on 06 January 2023

conceptual processing, which is well known to enhance memory encoding (Craik and Lockhart 1972; Craik and Tulving 1975).

To date, however, research investigating interactions between internally-directed attention and memory has been limited. One of the few exceptions is the study by Yi et al. (2008). In this study, each trial consisted of 2 frames. In the first frame, participants viewed a scene, and in the second frame, they either viewed the scene again or generated a mental image of the scene. In comparison to viewing scenes only once, memory was improved, and to a similar degree, by viewing scenes twice and by viewing the scene once and then imagining it. These results suggest that internally-directed attention (seeing the image in your "mind's eye") improved memory as much as externally-directed attention could.

Relatedly, studies investigating the generation of mental images (a goal-directed attention task) have reported that the use of imagery during encoding improves memory performance (Gupton and Frincke 1970; Mueller and Jablonski 1970; Bower 1972; McCauley et al. 1996; Foley 2012; D'Angiulli et al. 2013). The quality of the imagery can vary from item to item, with some mental images being highly detailed and vivid and others being very poor (Cui et al. 2007). Vividness has been defined as a construct that expresses the self-reported degree of richness, level of detail, and clarity of a mental image in comparison to actually seeing the imagined stimulus (D'Angiulli and Reeves 2007). A functional magnetic resonance imaging (fMRI) study investigating vividness in mental imagery confirmed the link between imagined and perceived stimuli with results showing that the neural overlap between imagery and perception in the visual system correlated with experienced imagery vividness (Dijkstra et al. 2017). Studies have also shown improved memory for mental images that had been visualized with high-vividness (De Beni and Pazzaglia 1995; D'Angiulli et al. 2013), suggesting that the vividness influences subsequent memory. However, the neural mechanisms by which mental images of high- versus low-vividness are successfully encoded into episodic memory have been little studied.

To investigate the neural mechanisms that support the successful encoding of internally-generated information into episodic memory, we designed an incidental encoding task during which participants formed visual mental images, a task which requires internally-directed attention. Although an external stimulus may be present during internally-directed attention, the processing seems to be generally decoupled from the external environment (Schooler et al. 2011; Dixon et al. 2014). Previous electrophysiology studies have demonstrated that both early sensory-level processing (as reflected, e.g., by the visual P1 and auditory N1 ERP sensory components) and late higher-order processing (as captured by the hallmark P3 ERP wave) are reduced during periods of internally-directed attention (Smallwood et al. 2008; Baird et al. 2014). However, shifts from externally- to internally- directed attention are difficult to measure, as they occur covertly. Our novel alternative to the difficult task of measuring the neural correlates of internally-directed attention was to use steady-state visual evoked potentials (SSVEPs), typically used as a measure of externally-directed attention, to index internally-directed attention instead. The SSVEP is an oscillatory EEG response that resonates at the same frequency as a flickering stimulus and is modulated by externally-directed attention to that stimulus, with an increase in SSVEP power at that frequency reflecting increased selective attention to that stimulus (Morgan et al. 1996; Muller et al. 2006; Vialatte et al. 2010). Accordingly, it stands to reason that a "decrease" in SSVEP power to an externally flickering stimulus, which reflects a shift of externally-directed attention away from that stimulus, could correspondingly, but inversely, reflect at least in part a shift toward internally-directed attention to form the mental image as instructed. Thus, SSVEPs could serve as a high-temporal-resolution index of the external versus internal orientation of attention. More specifically, in the present study we made such use of SSVEPs to index internally-directed attention during visual mental imagery in an incidental encoding task.

Participants were presented with a series of flickering object-word stimuli and tasked with forming a mental image of the object denoted by each of the words and rating the image vividness (Fig. 1). They then completed a surprise memory retrieval task, in which they were presented with both old and new words, and instructed to identify which items were old or new, including rating their confidence of that assessment. Our hypotheses were focused on the attention-related processes during the visual mental imagery, indexed by the SSVEPs, and the high-level encoding-related processes, indexed by the Dm. We hypothesized that our behavioral results would show improved memory for the high-vividness items, but we in particular wanted to examine the neural mechanisms underlying the successful encoding of high-vividness versus low-vividness items into episodic memory.

The study had 3 main goals. The first was to investigate the use of SSVEPs to track shifts between externally- and internally-directed attention. We hypothesized that after reading the flickering word stimulus, the process of forming a mental image of the object denoted by the word would result in a decrease of the SSVEP driven by the externally presented flickering word stimulus, which would then return to baseline as the imagery process was completed and attention returned to the visual screen.

Our second goal was to investigate the relationship between SSVEPs and the quality of mental images. Given the paucity of evidence regarding mental images and SSVEPs, we entertained 2 hypotheses regarding vividness. One hypothesis was that if participants deployed greater internally-directed attention and/or spent more time forming the mental images, they would be able to form images of better quality. If so, compared with

Downloaded from https://academic.oup.com/cercor/advance-article/doi/10.1093/cercor/bhac270/6659085 by Duke University user on 06 January 2023

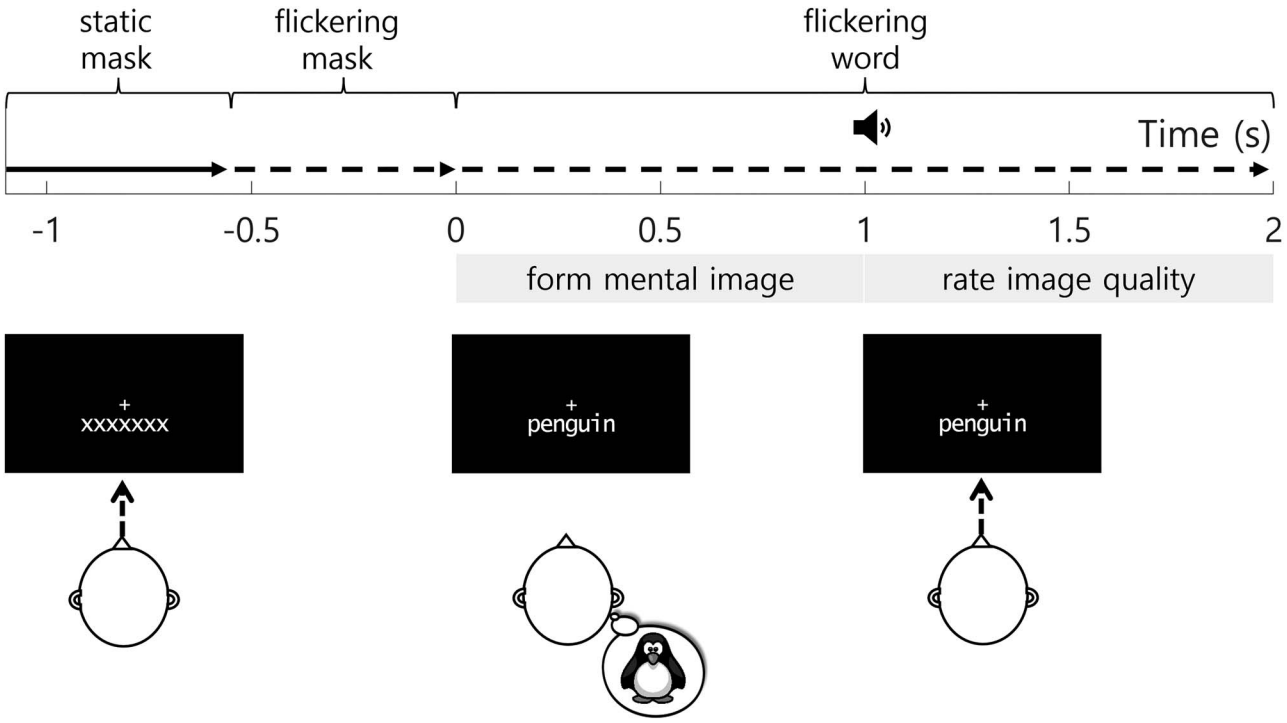


**Fig. 1.** Sequence of events over the course of a word stimulus trial. Participants were presented with a mask, which was static for half of the time and flickering for the other half. This was followed by a flickering object-word stimulus that was presented for 2 s and flickered during the entire presentation. Participants were tasked with forming a mental image of the object and, after hearing a tone, rating the quality of the mental image.

images rated as low-vividness, images rated as high-vividness would be associated with a larger and/or longer decrease in the SSVEP. The alternative hypothesis was that because a rating of low-vividness would indicate that the participant struggled with forming the image, the effect would be exactly the opposite, such that the low-vividness images would instead be associated with a larger and/or longer-duration SSVEP decrease. Given the prior research demonstrating faster RTs for more vivid mental images (D'Angiulli et al. 2013), however, and assuming that the internally-directed attention indexed by the SSVEP would be in line with RTs, we predicted that the second hypothesis would be more likely.

Finally, our third goal was to examine the influence of image vividness on the conceptually-related Dm effect during the image-generation period. We hypothesized that the elicitation of an enhanced Dm effect for the subsequently remembered high-vividness items could reflect greater encoding due to increased perceptual saliency. Alternatively, if there was a larger or more sustained Dm effect for the remembered low-vividness items, a plausible interpretation could be that enhanced conceptual processing is needed to compensate for the poor mental image in order to still be remembered.

## Methods

### Participants

Thirty-one healthy participants between the ages of 18 and 35 years ($M = 23.0$, $SD = 3.39$) were recruited through the Duke University Psychology Subject Pool and Interdisciplinary Behavioral Research Center. All participants were right-handed with normal or corrected-to-normal vision and no history of neurological or psychiatric disease. Informed consent was obtained for all participants for their credited (1 per hour) or paid ($15 per hour) participation in accordance with a protocol approved by the Duke University Medical Center Institutional Review Board. The study length was ~90 min, preceded by 1 h of preparation.

### Stimuli and procedure

The experiment was conducted in a dimly lit, electrically shielded room in which participants were seated 80 cm in front of a 24-inch stimulus presentation monitor (144 Hz refresh rate). Over the entire experimental session, participants completed an encoding phase, consisting of 600 trials, and a retrieval phase, consisting of 900 trials, with a 5-min break in between. In both the encoding and retrieval phases, participants were instructed to continuously fixate on a white cross present in the middle of a black screen. Each trial began with a static mask stimulus, which consisted of a row of 5–7 lowercase Xs, presented just below the fixation cross for 555.6 ms. This was followed immediately by the same mask stimulus presented flickering at 18 Hz, with 4 frames (27.8 ms) on and 4 frames off, for another 555.6 ms. The flickering mask was then replaced by a flickering 5–7 letter word stimulus of an object (95% of trials) or a 5–7 letter nonword stimulus (5% of trials), which flickered at 18 Hz for a full 2 s. The word and nonword stimuli were presented using a monospaced font (Lucida Console, size 24). To maintain a consistent visual input from trial to trial, the

Downloaded from https://academic.oup.com/cercor/advance-article/doi/10.1093/cercor/bhac270/6659085 by Duke University user on 06 January 2023

stimuli were grouped so that all the 5-letter words would appear in a row, followed by a group of the 6-letter words, and then a group of the 7-letter words. There were equal numbers of 5-, 6-, and 7-letter words across all stimuli. The number of Xs in the mask stimulus were matched to the number of letters in each word or nonword stimulus on that trial and presented in the same font and size. This was done so that the low-level sensory stimulation of the individual visual stimuli would not change on each trial. Presentation of the stimuli was controlled using Presentation 20.1 Software (Neurobehavioral Systems, Inc., Berkeley, CA). The encoding phase was broken up into 5 blocks of 120 trials each, and participants were able to determine the length of their breaks between blocks, with the entire encoding phase taking ~35 min to complete.

In terms of task, on trials in which a word was presented, participants were instructed to form a mental image of the word (image-generation period), adding as much imagery detail as possible. After 1 s, they heard a tone and had the remaining 1 s in the trial to rate the vividness of the image they had formed on a scale of 1–4, with 1 indicating poor/no image, 2 indicating fair image, 3 indicating good image, and 4 indicating excellent image (vividness-rating period). They were instructed to respond as quickly and accurately as possible after the tone on a keyboard using their left hand. To ensure that participants were rating the quality of the image they formed and not other features of the words, we gave the following examples: "A button response of 1 will mean you formed a very poor image or no image at all. If, for example. the word was 'carburetor' and you could not really picture what that would look like, you would rate that image as a 1. If the word was bicycle, and you were able to imagine only the shape of the bicycle, you would rate that image as 'ok', which is a 2. If you were able to imagine the pedals and wheels, you would rate that image as 'good', or a 3. If you were able to imagine the bicycle in detail - for example, the color or a basket - you would rate that image as 'excellent' or a 4". On trials in which a nonword stimulus was presented, participants were instructed to just press the space bar as quickly as possible. These catch trials were included to ensure that participants were attending during the presentation of the mask stimulus and could respond quickly as soon as the mask was replaced with the nonword stimulus.

The surprise retrieval phase was broken up into 5 blocks of 180 trials each and took a total of ~50 min to complete. Once again, participants were able to determine the length of their breaks in between each block. In this phase, all 570 words from the encoding phase were presented again, along with an additional 285 new words, as well as 45 new nonwords for catch trials (5% of total). Each word was presented for 2 s and followed by a 1.1 s inter-trial interval. On trials in which a word was presented, participants were instructed to determine whether the word was old or new and rate their confidence of that response from 1, indicating definitely new, to 6, indicating definitely old, as quickly and accurately as possible. On trials in which a nonword stimulus was presented, participants were again instructed to press the space bar as quickly as possible. Responses were made on a keyboard using their left hand. There was no tone during the retrieval phase, so participants had a full 2 s to make their response.

**Table 1.** Trial numbers per condition.

| Trial type | Minimum | Median | Maximum | Mean |
|---|---|---|---|---|
| Poor | 7 | 39 | 104 | 49 |
| Fair | 16 | 93 | 158 | 91 |
| Good | 36 | 108 | 186 | 105 |
| Excellent | 24 | 126 | 340 | 151 |
| LowVivRem | 14 | 103 | 205 | 112 |
| HighVivRem | 18 | 108 | 293 | 133 |
| LowVivForg | 2 | 25 | 81 | 29 |
| HighVivForg | 2 | 8 | 62 | 18 |

## Behavioral data and analysis

Overall memory performance was evaluated by comparing the hit rate and false alarm rate using a paired sample *t*-test. All accurate judgments were categorized as hits for old items and correct rejections for new items, regardless of confidence level. Similarly, all inaccurate judgments were categorized as misses for old items and false alarms for new items, regardless of confidence level. Mental imagery was evaluated behaviorally based on the vividness-rating reaction time (RT) at encoding and on subsequent memory for the word. The RTs are relative to the onset of the word stimuli. To create 2 well-separated levels of image vividness that also had sufficient and nearly equal numbers of trials, we binned trials into a *High Vividness* (*HighViv*) level, which included items rated at the highest vividness rating (4), and a *Low Vividness* (*LowViv*) level, consisting of trials with the two lowest vividness ratings (1 and 2). We compared the word stimuli within these two categories (*HighViv* and *LowViv*) with regard to frequency of use ($t(20) = 1.01$, $p > 0.05$) and imageability ($t(20) = 1.12$, $p > 0.05$) and found no significant differences. The percent of items subsequently remembered for these two categories were compared using paired sample *t*-tests. In addition, the vividness-rating RTs at encoding for the subsequently remembered items within these two categories (*HighVivRem* and *LowVivRem*) were compared using paired-sample *t*-tests. Trials with no behavioral responses were excluded from analysis. There were not enough forgotten trials to compare high- and low-vividness within this category (see Table 1 for the trial counts of each condition). However, our main interest was in comparing the mechanisms by which high- versus low-vividness items are successfully encoded into episodic memory.

## EEG data acquisition and preprocessing

Online EEG data were recorded across both the encoding and retrieval phases from a custom, 64-channel,

Downloaded from https://academic.oup.com/cercor/advance-article/doi/10.1093/cercor/bhac270/6659085 by Duke University user on 06 January 2023

extended-coverage cap (Woldorff et al. 2002) with active electrodes (actiCAP, Brain Products GmbH) and an online right-mastoid reference. The impedances of the ground and mastoid electrodes were maintained below 5 kOhms, and the remaining electrodes were maintained below 15 kOhms. Data were digitized at a rate of 500 Hz per channel, with a 3-stage cascaded integrator low-pass comb filter with a 130-Hz corner frequency. To monitor for blinks and horizontal eye movements, 1 electrode positioned below the left eye was used for the vertical electrooculogram (EOG) channel and 2 electrodes lateral to the outer canthus of each eye were used to compute the horizontal EOG channel. Offline, the EEG data were filtered using a 40 Hz low-pass filter, downsampled to 250 Hz, and then filtered using a 0.1 Hz high-pass causal FIR filter. In addition, the data were re-referenced to the algebraic average of the left and right mastoid electrodes. Channels that appeared excessively noisy during data collection were interpolated using a spherical spline procedure (Perrin et al. 1989).

The EEG data were epoched from −2,750 to 3,750 ms, relative to the presentation of the word or nonword stimuli, to allow sufficient buffer zones on the ends for the time-frequency analysis. In addition, the data were baseline-corrected based on the −200 to 0 ms prestimulus period, during which time the mask was flickering at 18 Hz. To detect eyeblinks occurring near the onset of the stimulus, which could influence the initiation of the neural processing cascade, the data were submitted to an algorithm using a 150-ms wide window moving across the epoch from −100 to 300 ms in 50-ms steps. Epochs with peak-to-peak voltage differences exceeding 48 $\mu$V in the vertical EOG channel were marked for rejection from the analysis. A copy of the EEG data was epoched from −1,100 to 2,000 ms, relative to the presentation of the word or nonword stimuli, to avoid overlapping trials. This copy was submitted to an infomax independent component (IC) analysis, and the resulting component weights were transferred back to the original data set. Ocular components of the IC analysis reflecting eyeblinks were identified via visual inspection and removed from the data (at most 4 ICs per participant).

To detect horizontal eye movements, data were submitted to an algorithm that used a 150-ms wide window moving across the epoch from −100 to 2,000 ms in 50-ms steps. Epochs with peak-to-peak voltage differences exceeding 24 $\mu$V in the vertical EOG channel, corresponding to ∼1.5° of movement, were marked for rejection. Additional data-cleaning procedures were then performed. To remove high-amplitude noise or excess muscle activity, if 3 or fewer electrodes (excluding the vertical and horizontal EOG channels) exceeded ±100 $\mu$V for one epoch, each of these electrodes were replaced using the spherical-spline interpolation procedure. If >3 electrodes exceeded the threshold, the epoch was marked for rejection from the analysis.

Ten participants were excluded due to >55% of trials being rejected, either because of poor behavioral performance (i.e. making responses prior to the onset of the tone or failing to respond; $n = 2$), poor EEG data quality ($n = 5$), or a combination of both ($n = 3$). This left 21 participants (12 female, mean age 22.4 years) in the final analyses. For the remaining participants, 30.5% of trials were excluded on average. The nonword-stimulus catch trials were not included in the final analysis. Data were preprocessed and analyzed in MATLAB using a combination of the EEGLAB (Delorme and Makeig 2004), ERPLAB (Lopez-Calderon and Luck 2014), and FieldTrip (Oostenveld et al. 2011) toolboxes. The EEG cap featured a modified 10–10, equidistant electrode montage with extended inferior occipital coverage (Woldorff et al. 2002). We report electrode sites based on standard 10–10 electrode names if the electrode in our montage was within 5 mm of the corresponding 10–10 site. If the electrodes were 5–10 mm from the 10–10 site, the name has a prime appended to it. A subscript 'a' means that the electrode was 10–15 mm anterior to the listed 10–10 site, and a subscript 'p' means that the electrode was 10–15 mm posterior to the listed 10–10 site. No electrodes were further than ∼1.5 cm from their closest standard 10–10 site.

## EEG and ERP analysis

Frequency decomposition of the EEG was performed using multitaper methods based on discrete prolate spheroidal sequences to estimate the power in logarithmically spaced frequencies from 3 to 40 Hz. The window widths for the tapers were 3 cycles for 3 Hz, 4 cycles for 4–7 Hz, 5 cycles for 8–14 Hz, 7 cycles for 15–20 Hz, and 10 cycles for 21–40 Hz. Spectral smoothing through multitapers was specified as $5 \times \log_{10}$ of each frequency. A decibel conversion was performed to normalize and baseline the data relative to the static period during the presentation of the mask stimulus (from −848 to −600 ms relative to the word stimulus onset).

The steady-state visual evoked response was calculated as the average magnitude of the oscillatory activity between 17 and 19 Hz across electrodes over occipital channels: Oz, O1′, O2′, PO3′, PO4′, PO7′, and PO8′. We selected the electrodes for the SSVEP analysis based on previous studies that have shown the SSVEP is observed over medial occipital sites, around Oz (for review, see Norcia et al. 2015), which was the scalp region where the SSVEPs in the current study numerically had their highest power. This measure served as an index of the direction of attention, with an increase in SSVEP power corresponding to increased externally-directed attention to the driving flickering stimuli and a decrease in SSVEP power corresponding to increased internally-directed attention and away from the flickering stimuli. To validate its use as an index of both externally- and internally-directed attention, SSVEP power data were submitted to a 1-way rANOVA with 3 levels reflecting the 3 different time periods over the course of a trial: flickering-mask period (baseline), image-generation period, and vividness-rating period. The flickering-mask

Downloaded from https://academic.oup.com/cercor/advance-article/doi/10.1093/cercor/bhac270/6659085 by Duke University user on 06 January 2023

period was averaged from −555.6 to 0 ms relative to the onset of the word stimulus, reflecting the pre-stimulus flickering of the mask stimulus, in which participants were attending to the mask stimulus in anticipation of the word onset. The image-generation period was derived from the average magnitude from 0 to 1,000 ms, covering the entire period in which participants were attending internally to form the mental image. The vividness-rating period was averaged from 1,000 to 2,000 ms, during which participants were attending externally to rate the quality of the mental image. We expected the SSVEP power to decrease while participants' attention was directed internally, away from the flickering stimulus, to form the mental image. After validating the SSVEP, we evaluated the attention-related processes during the image-generation period for subsequently remembered items. Mean SSVEP power data were submitted to a 2 (vividness: HighVivRem/LowVivRem) × 5 (time: 100–300ms/300–500 ms/500–700 ms/700–900 ms/900–1,100 ms) repeated-measures analysis of variance (rANOVA) in 200-ms time bins. The first time bin starts at 100 ms to account for time spent perceiving the word. Post-hoc paired sample *t*-tests were performed to identify the time bins that significantly differed between conditions.

For the ERP analysis, we examined the Dm effect by selectively averaging trials time-locked to the onset of the word stimuli at encoding as a function of subsequent memory (Remembered vs. Forgotten). Based on previous literature, we selected a cluster of electrodes centered on electrodes Cz and CPz (Luck 2014), including Cz, C1a, C2a, CPz, CP1′, CP2′, C1p, and C2p. Mean amplitude data were submitted to a 2 (memory: Remembered/Forgotten) × 5 (time: 400–500 ms/500–600 ms/600–700 ms/700–800 ms/800–900 ms) rANOVA to identify the Dm effect in 100-ms time bins. Post-hoc paired sample *t*-tests were conducted to identify the time bins in which the 2 conditions differed significantly. Then, to compare the encoding-related processes reflected by the Dm based on vividness, mean amplitude data of the remember-minus-forgotten difference waves were submitted to a 2 (vividness: HighVivRem-minus-Forg/LowVivRem-minus-Forg) × 5 (time: 400–500 ms/500–600 ms/600–700 ms/700–800/800–900 ms) rANOVA. Again, where the ANOVAs showed significant interactions, specific contrasts using post-hoc paired sample *t*-tests were conducted.

# Results

## Behavioral results

Overall memory performance at retrieval was well above chance with a mean hit rate of 82.9% ($SD = 10.8$) and false alarm rate of 30.2% ($SD = 16.5$), $t(20) = 12.54$, $p < 0.001$, $d = 2.74$. The data showed higher subsequent memory for the high-vividness items (HighViv; $M = 88.9\%$, $SD = 6.7\%$) in comparison to the low-vividness items (LowViv; $M = 80.0\%$, $SD = 10.5\%$), $t(20) = 5.21$, $p < 0.001$, $d = 1.14$ (Fig. 2A), which were still remembered substantially higher than chance, $t(20) = 12.21$, $p < 0.001$, $d = 2.67$. These results indicate that forming more vivid mental images facilitated better encoding of those items for later retrieval.

To behaviorally assess the time spent at encoding forming the subsequently remembered mental images, we compared the vividness-rating relative to the flickering-word onset time. Participants displayed faster RTs at encoding for the HighVivRem items ($M = 1{,}464$ ms, $SD = 86$) compared with the LowVivRem items ($M = 1{,}519$ ms, $SD = 98$), $t(20) = 4.13$, $p < 0.001$, $d = 0.90$ (Fig. 2B). These results suggest that, for the successfully encoded items (those later remembered), participants spent less time at encoding forming the mental image for the items they rated as having excellent vividness, presumably because those items were easier and thus faster to imagine.

## Goal 1: Using SSVEPs to track shifts between externally- versus internally-directed attention

The first goal of this study was to investigate the use of SSVEPs as an index of the direction of attention (external vs. internal). We hypothesized that after reading the flickering word stimulus, participants would direct their attention internally to form the mental image of the item, resulting in a decrease of the externally-driven SSVEP. The SSVEP would then return to pre-imagery baseline once participants heard the tone and re-directed attention to the visual screen to rate the vividness of the mental image. To test this hypothesis, we controlled for memory and evaluated the time-course of all subsequently remembered trials by comparing the SSVEP amplitude across the flickering-mask period (−555.6 to 0 ms), the image-generation period (0–1,000 ms), and the vividness-rating period (1,000–2,000 ms). SSVEP power data were submitted to a 1-way rANOVA with 3 time-period levels: flickering-mask period, image-generation period, and vividness-rating period. Note that a flickering stimulus was present during these 3 periods so changes in SSVEP power can be safely attributed to shifts between externally- and internally-directed attention rather than to perceptual changes.

We observed a significant effect of time-period, $F(2, 40) = 16.44$, $p < 0.001$, $n_p^2 = 0.45$ (Fig. 3A). Post-hoc paired sample *t*-tests showed that SSVEP power was significantly reduced, $t(20) = 5.33$, $p < 0.001$, $d = 1.16$, from the externally-oriented, flickering-mask period ($M = 0.51$, $SD = 0.73$), to the internally-oriented, image-generation period ($M = -0.03$, $SD = 0.67$), and then it significantly increased, $t(20) = -4.95$, $p < 0.001$, $d = -1.08$, to the externally-oriented, vividness-rating period ($M = 0.76$, $SD = 1.24$). There was no significant difference between the flickering-mask and vividness-rating periods, $t(20) = -1.62$, $p > 0.05$, $d = -0.35$.

In sum, consistent with our first hypothesis, the results validate the use of SSVEP power as a means to index shifts between externally- and internally-directed attention: (i) when participants directed attention externally during the flickering-mask period, SSVEP power was

Downloaded from https://academic.oup.com/cercor/advance-article/doi/10.1093/cercor/bhac270/6659085 by Duke University user on 06 January 2023

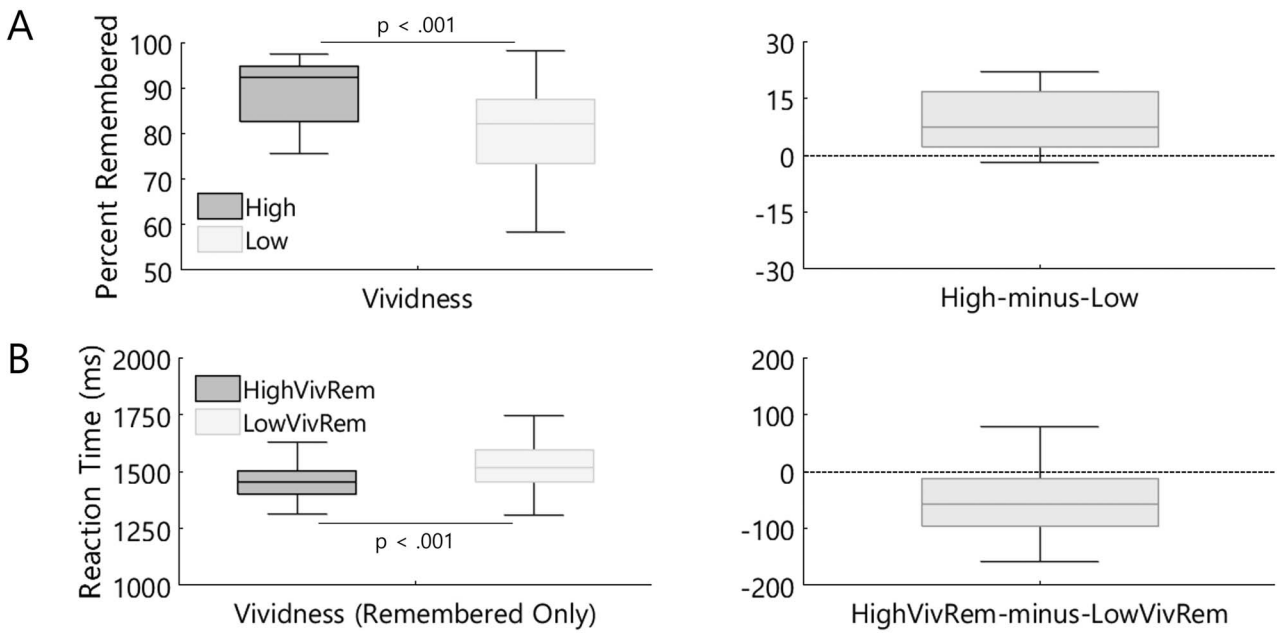


**Fig. 2.** Behavioral results. A) The left panel shows the mean and distribution for the percent of items remembered that were rated as high-vividness compared with low-vividness. The right panel shows the within-subject difference between the high- and low-vividness categories. Results showed that there was better subsequent memory for high-vividness items. B) The left panel shows the mean and distribution for the vividness-rating reaction times at encoding between the subsequently remembered high-vividness compared with low-vividness items. The right panel shows the within-subject RT difference between the 2 categories. Participants were faster to make vividness ratings for high-vividness items.

high, (ii) when they directed attention internally during the image-generation period, SSVEP power decreased significantly, and (iii) finally, when they directed their attention externally again during the vividness-rating period, SSVEP power increased reliably, returning to baseline.

## Goal 2: Investigate the relationship between SSVEPs and the quality of mental images

For our second goal, we tested 2 alternative hypotheses. According to one hypothesis, more intense and/or longer internally-directed attention would lead to higher image vividness, whereas according to the other hypothesis, more intense and/or longer internally-directed attention would be necessary when images were more difficult to generate. To avoid confounding the effects of imagery with the effects of successful encoding processes, these hypotheses were tested using only subsequently remembered items. The former hypothesis predicts that the SSVEP decrease during the image-generation period should be greater and/or more sustained for HighVivRem compared with LowVivRem items, whereas the latter hypothesis predicts the opposite result. SSVEP power data were submitted to a $2 \times 5$ rANOVA, with a 2-level factor of vividness (HighVivRem/LowVivRem) and a 5-level factor of time (100–300 ms/300–500 ms/500–700 ms/700–900 ms/900–1,100 ms). Figure 3B displays the time-course of SSVEP power for the HighVivRem and LowVivRem items, and Fig. 5A displays the mean power and within-subject standard error for each time bin. Table 2 displays means and standard deviations for the 5 time windows during the image-generation period.

**Table 2.** SSVEP mean power (dB) and standard deviation across time.

| | HighVivRem | LowVivRem |
|---|---|---|
| 100–300 ms | −.11 (.46) | −.19 (.58) |
| 300–500 ms | −.45 (.57) | −.51 (.73) |
| 500–700 ms | −.21 (.76) | −.35 (.83) |
| 700–900 ms | .32 (1.19) | −.08 (1.03) |
| 900–1100 ms | .70 (1.37) | .26 (1.10) |

These analyses indicated that there was no main effect of imagery vividness in this analysis, $F(1, 20) = 2.50$, $p = 0.13$, $n_p^2 = 0.11$, indicating no significant difference in the mean SSVEP power integrated across the entire examined time period. There was, however, a main effect of time, $F(4, 80) = 13.98$, $p < 0.001$, $n_p^2 = 0.41$, indicating that SSVEP power was modulated over the course of the trial, as shown in both Fig. 3A and B. Most importantly, there was an interaction between vividness and time, $F(4, 80) = 4.78$, $p = 0.002$, $n_p^2 = 0.19$. Post-hoc paired sample $t$-tests comparing the HighVivRem and LowVivRem items over the course of the trial showed significant differences during the 700–900 ms, $t(20) = 2.51$, $p = 0.02$, $d = 0.55$, and 900–1,100 ms time bins, $t(20) = 2.69$, $p = 0.01$, $d = 0.59$.

These results indicate that SSVEP power was reduced for a longer time for the LowVivRem compared with HighVivRem items. This finding is consistent with the alternative hypothesis that the shift to internally-directed attention is more intense and/or longer for LowVivRem than HighVivRem items, presumably because the mental images were harder to form, thus fitting with the second hypothesis above. It is worth noting that the vividness-rating RTs were also slower for LowVivRem

Downloaded from https://academic.oup.com/cercor/advance-article/doi/10.1093/cercor/bhac270/6659085 by Duke University user on 06 January 2023

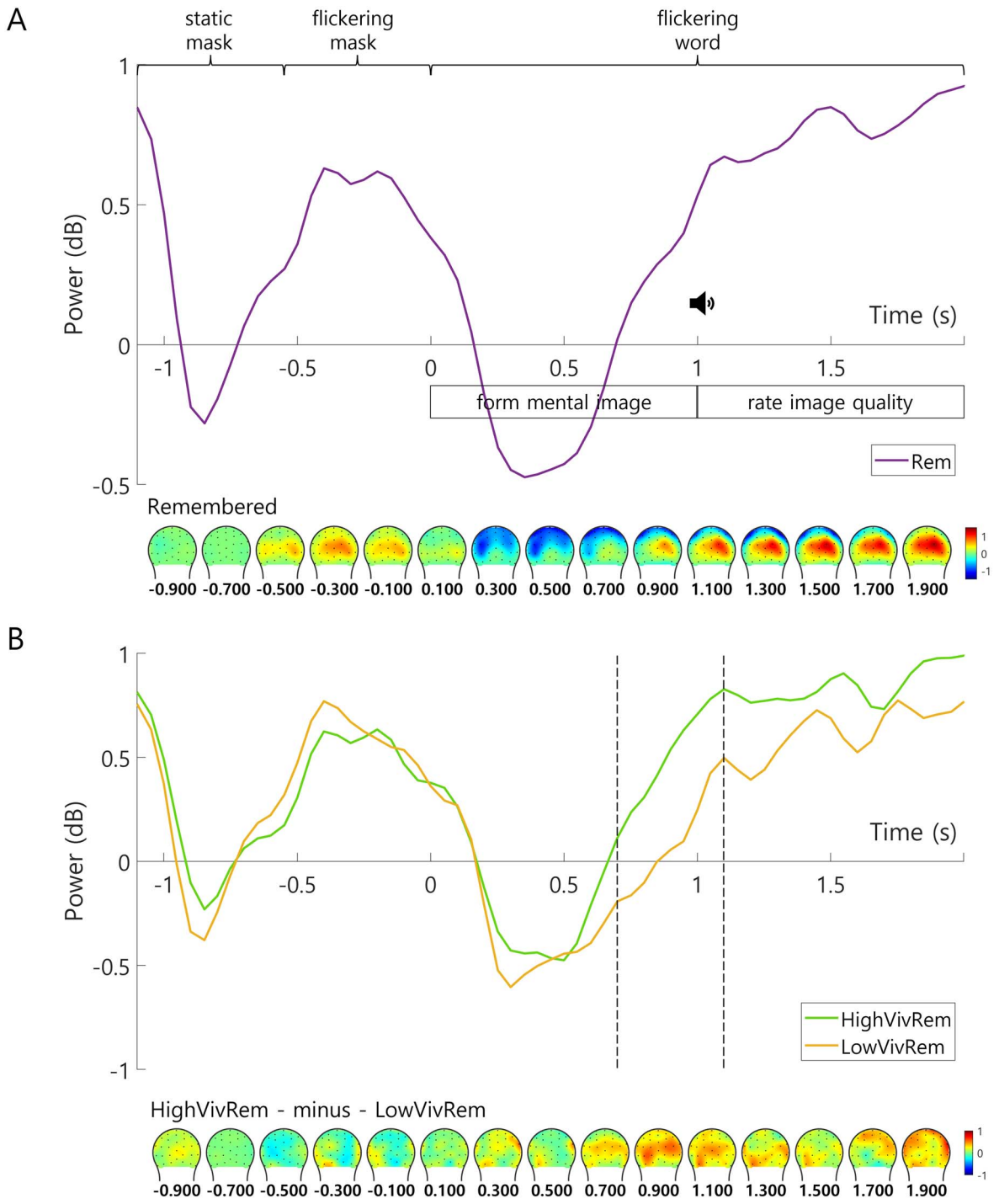


**Fig. 3.** SSVEP Results. A) Timelock-averaged occipital SSVEP power plotted over the time-course of a trial, with the onset of the flickering word stimulus at time 0. During the flickering-mask period (−555 to 0 ms) and the vividness-rating period (1,000–2,000 ms), SSVEP power was high as participants directed their attention externally toward the flickering visual stimulus on the screen. During the image-generation period in between (0–1,000 ms), however, there was a substantial decrease in SSVEP power as participants directed their attention internally to form the mental image, after which it returned to a high level as attention was redirected externally back to the screen during the vividness-rating period. The topographic plots show the time-course of the trial in 200-ms time bins, with each bin labeled with the middle time point in each 200-ms period. B) Timelock-averaged occipital SSVEP power across the time-course of a trial, plotted separately for the subsequently remembered high-vividness and low-vividness items. The SSVEP power returned more slowly to baseline when forming low-vividness mental images, consistent with participants being slower to redirect their attention externally. The dashed lines represent the time-periods in which a significant difference was observed. The topographic plots show the difference between the high- and low-vividness trials across time, plotted in 200-ms time bins, labeled again with the middle time point of each bin displayed.

Downloaded from https://academic.oup.com/cercor/advance-article/doi/10.1093/cercor/bhac270/6659085 by Duke University user on 06 January 2023

**Table 3.** Post-hoc paired sample *t*-tests for the Dm comparisons in each time window.

**A.** Remembered vs. Forgotten

| | Rem Mean (SD) | Forg Mean (SD) | t | df | p | d |
|---|---|---|---|---|---|---|
| 400–500 ms | .93 (3.42) | −.16 (2.86) | 3.68 | 20 | .001 | .80 |
| 500–600 ms | 2.16 (3.44) | .75 (3.01) | 4.10 | 20 | <.001 | .89 |
| 600–700 ms | 2.21 (2.96) | .93 (2.44) | 3.40 | 20 | .003 | .74 |
| 700–800 ms | 2.03 (2.34) | .78 (2.16) | 3.33 | 20 | .003 | .73 |
| 800–900 ms | 1.49 (2.07) | .46 (2.25) | 2.33 | 20 | .03 | .51 |

**B.** HighVivRem-minus-Forg vs. LowVivRem-minus-Forg

| | HighVivRem-minus-Forg Mean (SD) | LowVivRem-minus-Forg Mean (SD) | t | df | p | d |
|---|---|---|---|---|---|---|
| 400–500 ms | .99 (3.28) | .77 (3.75) | .60 | 20 | .56 | .13 |
| 500–600 ms | 2.20 (3.58) | 2.29 (3.77) | −.25 | 20 | .80 | −.06 |
| 600–700 ms | 2.10 (3.15) | 2.64 (3.46) | −1.13 | 20 | .27 | −.25 |
| 700–800 ms | 1.59 (2.48) | 2.74 (2.87) | −2.42 | 20 | .03 | −.53 |
| 800–900 ms | 1.02 (2.00) | 2.22 (2.83) | −2.48 | 20 | .02 | −.54 |

than HighVivRem items, providing converging evidence for the second hypothesis.

### Goal 3: Examine the influence of image vividness on the Dm effect

Beside imagery, a major factor contributing to successful encoding is thought to be related to conceptual processing (Craik and Lockhart 1972), which in ERP studies, has been often linked to the Dm ERP effect (Paller et al. 1987; Otten and Rugg 2001). Thus, it is possible that one of the reasons why LowViv items could be subsequently remembered despite the poor images that were generated for them is because they received greater conceptual processing, perhaps due to the difficulty and time required to generate the images. If so, the Dm effect should be greater or more sustained for LowViv than HighViv items. To examine differences in the encoding-related processes for the HighVivRem and LowVivRem items, Dm amplitude data during the image-generation period were submitted to a 2 (memory: remembered/forgotten) × 5 rANOVA (time: 400–500 ms/500–600 ms/600–700 ms/700–800 ms/800–900 ms). Table 3A displays the means and standard deviations for each condition and the results of the *t*-tests for each time bin.

First, we identified the Dm by comparing subsequently remembered to forgotten items. Results showed a main effect of memory, $F(1, 20) = 13.46$, $p = 0.001$, $n_p^2 = 0.40$ and a main effect of time, $F(4, 80) = 2.55$, $p < 0.05$, $n_p^2 = 0.11$, with no interaction of these 2 factors, $F(4, 80) = 0.70$, $p > 0.05$, $n_p^2 = 0.03$. Post-hoc paired sample *t*-tests showed significant differences between subsequently remembered and forgotten items in each time bin from 400 to 900 ms (Fig. 4A). After identifying the Dm *main* effect for later remembered versus later forgotten items, we compared amplitudes for the HighVivRem-minus-Forg and LowVivRem-minus-Forg Dm difference waves to investigate differences in encoding processes for high- versus low-vividness items. Table 3B displays the means and standard deviations, as well as the results of the *t*-tests for each time bin. Results showed no main effect of vividness, $F(1, 20) = 1.25$, $p > 0.05$, $n_p^2 = 0.06$, but a main effect of time, $F(4, 80) = 2.53$, $p < 0.05$, $n_p^2 = 0.11$. Most importantly, however, the results showed a significant interaction of time and vividness rating, $F(4, 80) = 6.87$, $p < 0.001$, $n_p^2 = 0.26$ (Fig. 4B), indicating the Dm effect did not follow the same time-course for each vividness condition. Figure 5B displays the mean amplitude and within-subject standard error for each time bin. Post-hoc paired sample *t*-tests showed significant differences between the HighVivRem-minus-Forg and LowVivRem-minus-Forg Dm difference waves from 700 to 900 ms. These results suggest that participants were able to successfully encode the HighViv items into episodic memory more quickly, whereas the LowViv items required additional or more sustained conceptual processing for successful subsequent retrieval.

## Discussion

Using a visual mental imagery task in combination with a subsequent-memory paradigm, we investigated the neural mechanisms supporting the successful encoding of internally-generated information into episodic memory and the modulation of those neural mechanisms by the quality of the information, which was indexed by the self-reported vividness of mental images in this paradigm. Participants were presented with a series of flickering object-word stimuli and instructed to form a visual mental image of the object referent for each word and then rate the quality of the image in terms of its vividness. In our analyses, we focused on the processing at encoding of high- versus low-vividness items that were subsequently remembered.

Our behavioral results were consistent with previous research and theories. In line with prior studies (Reisberg et al. 1986; De Beni and Pazzaglia 1995; D'Angiulli et al. 2013; Marcotti and St Jacques 2018), we observed improved subsequent memory for high-vividness items—i.e. those where participants were able

Downloaded from https://academic.oup.com/cercor/advance-article/doi/10.1093/cercor/bhac270/6659085 by Duke University user on 06 January 2023

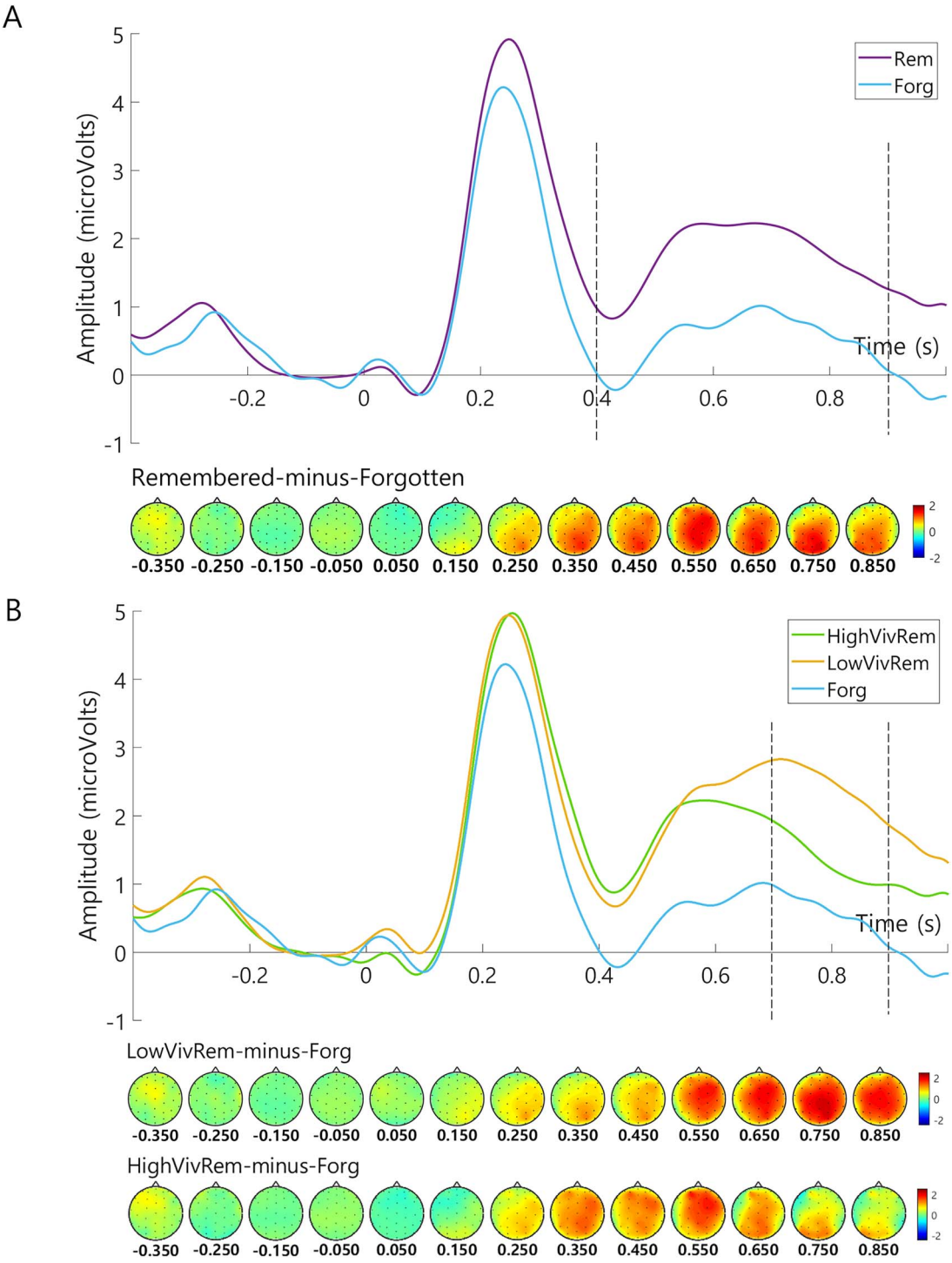


**Fig. 4.** Dm results. A) The Dm difference at encoding for subsequently remembered vs. subsequently forgotten items was observed over centro-parietal regions starting at around 400 ms. The dashed lines represent the time-periods in which a significant difference was observed. The topographic plots show the difference between the subsequently remembered and forgotten trials plotted in 100-ms time bins, with the middle time of each bin displayed. B) The Dm amplitudes were compared between the high-vividness subsequently remembered items and low-vividness subsequently remembered items. Results showed that the Dm effect persisted longer for the low-vividness items, with the dashed lines delimiting the time-periods in which significant differences were observed. The topographic plots show the remembered minus forgotten differences for both the low- and high-vividness items over time in 100-ms time bins, with the middle time displayed.

to form vivid visual mental images—as compared with low-vividness items. Our finding of shorter vividness rating RTs for high- versus low-vividness items are in line with a D'Angiulli and colleagues (2013) hypothesis—based on multi-trace memory theory (Moscovitch et al. 2005)—that vividness ratings reflect the availability of multiple sensory traces in long-term memory, and therefore, faster image generation is observed for higher-vividness items (D'Angiulli and Reeves 2002).

The neural results showed 3 main findings. First, we observed reduced SSVEP power as subjects directed attention internally to form a mental image of the word's

Downloaded from https://academic.oup.com/cercor/advance-article/doi/10.1093/cercor/bhac270/6659085 by Duke University user on 06 January 2023

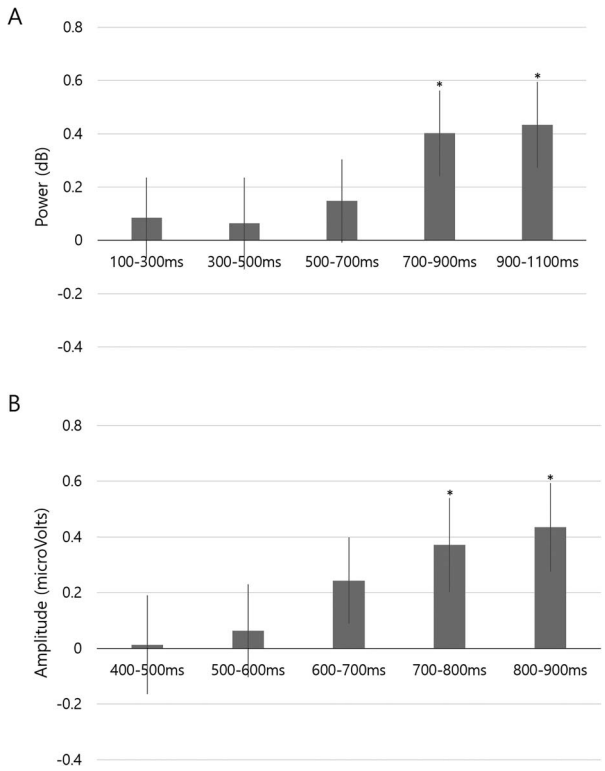


**Fig. 5.** Comparisons of within-subject differences between conditions at each time window for the main SSVEP and Dm analyses. A) The HighVivRem minus LowVivRem calculation for SSVEP power within each time bin, with error bars indicating within-subject standard error. The asterisks indicate $p < 0.05$. B) The HighVivRem minus LowVivRem calculation for the ERP Dm amplitude within each time bin, with error bars indicating within-subject standard error. The asterisks indicate $p < 0.05$.

referent. Second, objects rated as low-vividness showed a slower return of SSVEP power to pre-imagery baseline compared with high-vividness items. In line with the behavioral vividness-rating RT results, this suggests that longer internally-directed attention was needed to generate low-vividness images. Finally, we observed a more sustained Dm for the subsequently remembered low-vividness items compared with subsequently remembered high-vividness ones, suggesting additional conceptual processing was necessary to compensate for the low-vividness of the mental image and still remember the item. These 3 findings are discussed in separate sections below.

## SSVEPs can index shifts between externally- and internally-directed attention

Most SSVEP studies have used this measure to index externally-directed attention. For example, SSVEPs have been used to track visuospatial attention (Morgan et al. 1996; Andersen et al. 2008, 2011), sustained attention (Silberstein et al. 1990), working memory (Silberstein et al. 2001; Van Rooy et al. 2001; Ellis et al. 2006), and emotional stimulus processing (Kemp et al. 2002, 2004; Hindi Attar et al. 2010). In a novel use of this technique here, we employed SSVEPs to track internally-directed attention in a visual mental imagery task. We hypothesized that power would be highest during the flickering-mask period (externally-directed attention), would drop during the image-generation period (internally-directed attention), and then would return to baseline as attention was redirected to the screen during the vividness-rating period (externally-directed attention). The results confirmed this hypothesis, indicating SSVEPs can be used to track shifts between externally- and internally-directed attention during cognitive tasks. These results are in line with a recent study published by Kritzman et al. (2022), which used the SSVEP as an index of internally-directed attention in an interoceptive processing task. Their results showed that SSVEP magnitude and phase synchronization decreased when participants directed their attention toward their heartbeats.

Our finding has implications for the use of SSVEPs in research on the cognitive and neural mechanisms of shifts between externally- and internally-directed attention, as well as studies on the impact of these shifts on other cognitive processes. For example, one fruitful line of research would be to use SSVEPs to examine lapses in sustained attention, which may reflect both external distractions from the environment or from the person's own thoughts, as in the case of mind-wandering (Stawarczyk et al. 2011; Unsworth and McMillan 2014; Unsworth and Robison 2016; deBettencourt et al. 2018). Additionally, the SSVEP technique can be effectively used to measure attention and memory processes that occur internally, such as the impact of visual imagery on subsequent memory, as in the current study.

## SSVEP power returned to baseline more slowly for low- than high-vividness images

We tested 2 alternative hypotheses regarding the relationship between internally-directed attention and the quality of mental images. One hypothesis was that greater and/or longer internally-directed attention contributes to more vivid images, whereas the alternative hypothesis was that greater and/or longer internally-directed attention is required when images are difficult to form. Our second main neural finding was that a more prolonged dip in SSVEP power during the object imagery process was observed for low- compared with high-vividness remembered items. This finding is consistent with the alternative hypothesis, and it suggests that participants spent more time with attention directed internally as they tried to form the mental images of the low-vividness items.

This idea is also consistent with the vividness-rating RTs being longer for low- than high-vividness remembered items, providing converging evidence for the alternative hypothesis. Participants were not asked to press the vividness rating key as quickly as possible after they finished generating each image, so rating RTs provide only an indirect measure of the duration of the image-generation process. However, we would expect participants to be even faster at making the ratings for the high-vividness items if they did not have to wait for a

Downloaded from https://academic.oup.com/cercor/advance-article/doi/10.1093/cercor/bhac270/6659085 by Duke University user on 06 January 2023

tone, thereby increasing the difference in RTs between the high- and low-vividness items. Since participants were trying to generate vivid images, it is reasonable to infer that once they were satisfied with the quality of the mental image they generated, they began shifting to the task of rating the image more quickly, but when they were not satisfied with the image quality, they continued trying to improving their mental image and shifted to the rating task later. We were able to directly confirm that there were no differences between the items rated as high-vividness compared with those rated as low-vividness in terms of frequency of use or imageability of the words. However, further research is warranted to directly examine the role of other possible factors.

Increases in SSVEP power can be used to link externally-directed attention toward the flickering stimuli to specific brain regions, within the spatial resolution of EEG. In the current study, we observed SSVEP power over occipital cortex (Fig. 3A), consistent with externally-directed attention being focused on a central flickering visual stimulus (Andersen et al. 2011; Norcia et al. 2015). In contrast, decreases in SSVEP power suggest that a shift toward internally-directed attention occurred. However, the decrease in SSVEP power does not give us insight into the specific cognitive operation that was the focus of internally-directed attention or the specific brain regions mediating attention to these internal targets. Functional MRI studies have linked internally-directed attention during visual mental imagery to the default mode network (DMN), operating in collaboration with frontal areas (Ishai et al. 2000; Yomogida et al. 2004; Schlegel et al. 2013; Pearson 2019). Thus, a future combined EEG-fMRI study, may be useful to directly link the SSVEP power reduction to modulations of activity in the DMN and its interactions with frontal regions. Each of these areas could contribute differently to subsequent memory of internally-generated information.

### The Dm was more sustained for low- versus high-vividness remembered items

The Dm neural ERP effect showed a prolonged duration for the low-vividness versus high-vividness items that were later remembered. Previous research has shown that the Dm ERP is modulated by conceptual processing (for a review, see Friedman and Johnson Jr 2000; Wagner et al. 1999). For example, enhanced Dm amplitude was observed for tasks requiring semantic decisions compared with non-semantic decisions (Paller et al. 1987). In the current study, the Dm effect was more sustained for low- than for high-vividness items. Taken together with the SSVEP results, this finding suggests that one of the factors contributing to later memory for the more poorly imagined items could be that they received greater higher-level conceptual processing.

This idea is consistent with the results of an early behavioral study (Wiseman and Neisser 1974) in which participants viewed incomplete pictures, either providing a meaningful interpretation or just attending to the meaningless pictures. Results showed better recognition memory for the pictures that were given meaningful interpretations at encoding. Similarly, a study in which words and pictures were either named or categorized showed that pictures were later better remembered than words (picture-superiority effect) in the naming condition but equally well-remembered in the categorization condition (Vaidya and Gabrieli 2000). The results of these studies suggest that conceptual processing can compensate for poorer imagery. Furthermore, a study comparing deep versus shallow encoding showed that deep encoding was associated with a larger and longer-lasting Dm (Guo et al. 2004). Although our findings robustly show differences in Dm duration for low- versus high-vividness items that are later remembered, future studies are needed to directly investigate the compensatory strategies that may be employed when participants are unable to form a highly vivid visual mental image.

Our findings reveal important differences in the time-course for the processing of high- versus low-vividness items that are subsequently remembered. Not only did the low-vividness items require more time spent with attention directed internally (as reflected by a slower return to baseline from the dip in SSVEP power during the imagery process), but they also showed a prolonged Dm ERP effect, suggesting that there may have been additional higher-level conceptual processing for the successful encoding of these items that enabled them to be subsequently remembered. To further investigate the neuroanatomical underpinnings of such processes, future research using functional MRI could compare the perceptual versus conceptual representations for high-versus low-vividness items. Previous fMRI studies have used representational similarity analysis to investigate the different brain regions where visual and semantic representations predict perceptual memory, conceptual memory, or both (Davis et al. 2021). Although fMRI imaging measures would not be able to provide the temporal dynamics of the encoding-related processes revealed by the EEG/ERP effects presented here, they could offer insight into the specific neural regions activated at encoding that support subsequent memory for the low-vividness compared with high-vividness items.

## Conclusion

In conclusion, this study shows that differences in the vividness of a mental image are related to changes in internally-directed attentional processes and encoding-related processes, as well as their time-courses. When subjects encoded subsequently remembered images with low-vividness, we observed a slower return of the SSVEP dip back to the pre-imagery baseline and a more sustained Dm ERP positivity, relative to subsequently remembered images that had high-vividness. This pattern of results suggests that more sustained internally-directed attention, in conjunction with more prolonged encoding-related processing, is needed to successfully remember less salient mental representations.

Downloaded from https://academic.oup.com/cercor/advance-article/doi/10.1093/cercor/bhac270/6659085 by Duke University user on 06 January 2023

*Conflict of interest statement:* None declared.

## References


Aly M, Turk-Browne NB. 2017. How hippocampal memory shapes, and is shaped by, attention. In: Hannula DE, Duff MC, editors. *The hippocampus from cells to systems*. Cham: Springer International Publishing. p. 369–403. [accessed 2019 Jul 26]. http://link.springer.com/10.1007/978-3-319-50406-3_12.

Andersen SK, Hillyard SA, Müller MM. Attention facilitates multiple stimulus features in parallel in human visual cortex. *Curr Biol*. 2008:18(13):1006–1009.

Andersen SK, Müller MM, Hillyard SA. Tracking the allocation of attention in visual scenes with steady-state evoked potentials. *Cogn Neurosci Atten*. 2011:197–216.

Baird B, Smallwood J, Lutz A, Schooler JW. The decoupled mind: mind-wandering disrupts cortical phase-locking to perceptual events. *J Cogn Neurosci*. 2014:26(11):2596–2607.

Bower GH. Mental imagery and associative learning. In: *Cognition in learning and memory*. Oxford, England: John Wiley & Sons; 1972. pp. vii, 263–vii, 263

Chun MM, Turk-Browne NB. Interactions between attention and memory. *Curr Opin Neurobiol*. 2007:17(2):177–184.

Chun MM, Golomb JD, Turk-Browne NB. A taxonomy of external and internal attention. *Annu Rev Psychol*. 2011:62(1):73–101.

Craik FIM, Lockhart RS. Levels of processing: a framework for memory research. *J Verbal Learn Verbal Behav*. 1972:11(6):671–684.

Craik FIM, Tulving E. Depth of processing and the retention of words in episodic memory. *J Exp Psychol Gen*. 1975:104(3):268–294.

Cui X, Jeter CB, Yang D, Montague PR, Eagleman DM. Vividness of mental imagery: individual variability can be measured objectively. *Vis Res*. 2007:47(4):474–478.

D'Angiulli A, Reeves A. Generating visual mental images: latency and vividness are inversely related. *Mem Cogn*. 2002:30(8):1179–1188.

D'Angiulli A, Reeves A. The relationship between self-reported vividness and latency during mental size scaling of everyday items: phenomenological evidence of different types of imagery. *Am J Psychol*. 2007:120(4):521–551.

D'Angiulli A, Runge M, Faulkner A, Zakizadeh J, Chan A, Morcos S. Vividness of visual imagery and incidental recall of verbal cues, when phenomenological availability reflects long-term memory accessibility. *Front Psychol*. 2013:4:1.

Davis SW, Geib BR, Wing EA, Wang W-C, Hovhannisyan M, Monge ZA, Cabeza R. Visual and semantic representations predict subsequent memory in perceptual and conceptual memory tests. *Cereb Cortex*. 2021:31(2):974–992.

De Beni R, Pazzaglia F. Memory for different kinds of mental images: role of contextual and autobiographic variables. *Neuropsychologia*. 1995:33(11):1359–1371.

deBettencourt MT, Norman KA, Turk-Browne NB. Forgetting from lapses of sustained attention. *Psychon Bull Rev*. 2018:25(2):605–611.

Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods*. 2004:134(1):9–21.

Dijkstra N, Bosch SE, Gerven MAJ v. Vividness of Visual Imagery Depends on the Neural Overlap with Perception in Visual Areas. *J Neurosci*. 2017:37(5):1367–1373.

Dixon ML, Fox KCR, Christoff K. A framework for understanding the relationship between externally and internally directed cognition. *Neuropsychologia*. 2014:62:321–330.

Ellis KA, Silberstein RB, Nathan PJ. Exploring the temporal dynamics of the spatial working memory n-back task using steady state visual evoked potentials (SSVEP). *NeuroImage*. 2006:31(4):1741–1751.

Foley MA. Imagery encoding and false recognition errors: exploring boundary conditions of imagery's enhancing effects. *Memory*. 2012:20(7):700–716.

Friedman D, Johnson R Jr. Event-related potential (ERP) studies of memory encoding and retrieval: a selective review. *Microsc Res Tech*. 2000:51(1):6–28.

Guo C, Zhu Y, Ding J, Fan S, Paller KA. An electrophysiological investigation of memory encoding, depth of processing, and word frequency in humans. *Neurosci Lett*. 2004:356(2):79–82.

Gupton T, Frincke G. Imagery, mediational instructions, and noun position in free recall of noun-verb pairs. *J Exp Psychol*. 1970:86(3):461–462.

Hindi Attar C, Andersen SK, Müller MM. Time course of affective bias in visual attention: convergent evidence from steady-state visual evoked potentials and behavioral data. *NeuroImage*. 2010:53(4):1326–1333.

Ishai A, Ungerleider LG, Haxby JV. Distributed neural systems for the generation of visual images. *Neuron*. 2000:28(3):979–990.

Johnson R. Event-related potential insights into the neurobiology of memory systems. *Handb Neuropsychol*. 1995:10:135–164.

Kemp AH, Gray MA, Eide P, Silberstein RB, Nathan PJ. Steady-state visually evoked potential topography during processing of emotional valence in healthy subjects. *NeuroImage*. 2002:17(4):1684–1692.

Kemp AH, Silberstein RB, Armstrong SM, Nathan PJ. Gender differences in the cortical electrophysiological processing of visual emotional stimuli. *NeuroImage*. 2004:21(2):632–646.

Kritzman L, Eidelman-Rothman M, Keil A, Freche D, Sheppes G, Levit-Binnun N. Steady-state visual evoked potentials differentiate between internally and externally directed attention. *NeuroImage*. 2022:254:119133.

Lopez-Calderon J, Luck SJ. ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Front Hum Neurosci*. 2014:8:213.

Luck SJ. *An introduction to the event-related potential technique*. 2nd ed. Cambridge, MA, USA: A Bradford Book; 2014

Marcotti P, St Jacques PL. Shifting visual perspective during memory retrieval reduces the accuracy of subsequent memories. *Mem Hove Engl*. 2018:26(3):330–341.

McCauley ME, Eskes G, Moscovitch M. The effect of imagery on explicit and implicit tests of memory in young and old people: a double dissociation. *Can J Exp Psychol Can Psychol Expérimentale*. 1996:50(1):34–41.

Morgan ST, Hansen JC, Hillyard SA. Selective attention to stimulus location modulates the steady-state visual evoked potential. *Proc Natl Acad Sci U S A*. 1996:93(10):4770–4774.

Moscovitch M, Rosenbaum RS, Gilboa A, Addis DR, Westmacott R, Grady C, McAndrews MP, Levine B, Black S, Winocur G, et al. Functional neuroanatomy of remote episodic, semantic and spatial memory: a unified account based on multiple trace theory. *J Anat*. 2005:207(1):35–66.

Mueller JH, Jablonski EM. Instructions, noun imagery, and priority in free recall. *Psychol Rep*. 1970:27(2):559–566.

Muller MM, Andersen S, Trujillo NJ, Valdes-Sosa P, Malinowski P, Hillyard SA. Feature-selective attention enhances color signals in early visual areas of the human brain. *Proc Natl Acad Sci*. 2006:103(38):14250–14254.

Norcia AM, Appelbaum LG, Ales JM, Cottereau BR, Rossion B. The steady-state visual evoked potential in vision research: a review. *J Vis*. 2015:15(6):4.


Downloaded from https://academic.oup.com/cercor/advance-article/doi/10.1093/cercor/bhac270/6659085 by Duke University user on 06 January 2023

Oostenveld R, Fries P, Maris E, Schoffelen J-M. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci*. 2011:2011: 156869.

Otten LJ, Rugg MD. Task-dependency of the neural correlates of episodic encoding as measured by fMRI. *Cereb Cortex*. 2001:11(12): 1150–1160.

Paller KA, Wagner AD. Observing the transformation of experience into memory. *Trends Cogn Sci*. 2002:6(2):93–102.

Paller KA, Kutas M, Mayes AR. Neural correlates of encoding in an incidental learning paradigm. *Electroencephalogr Clin Neurophysiol*. 1987:67(4):360–371.

Paller KA, McCarthy G, Wood CC. ERPs predictive of subsequent recall and recognition performance. *Biol Psychol*. 1988:26(1): 269–276.

Pearson J. The human imagination: the cognitive neuroscience of visual mental imagery. *Nat Rev Neurosci*. 2019:20(10):624–634.

Perrin F, Pernier J, Bertrand O, Echallier JF. Spherical splines for scalp potential and current density mapping. *Electroencephalogr Clin Neurophysiol*. 1989:72(2):184–187.

Reisberg D, Culver LC, Heuer F, Fischman D. Visual memory: when imagery vividness makes a difference. *J Ment Imag*. 1986:10(4): 51–74.

Schlegel A, Kohler PJ, Fogelson SV, Alexander P, Konuthula D, Tse PU. Network structure and dynamics of the mental workspace. *Proc Natl Acad Sci*. 2013:110(40):16277–16282.

Schooler JW, Smallwood J, Christoff K, Handy TC, Reichle ED, Sayette MA. Meta-awareness, perceptual decoupling and the wandering mind. *Trends Cogn Sci*. 2011. doi: 10.1016/j.tics.2011.05.006. [accessed 2018 Apr 18]. http://linkinghub.elsevier.com/retrieve/pii/S1364661311000878.

Silberstein RB, Schier MA, Pipingas A, Ciorciari J, Wood SR, Simpson DG. Steady-State Visually Evoked Potential topography associated with a visual vigilance task. *Brain Topogr*. 1990:3(2):337–347.

Silberstein RB, Nunez PL, Pipingas A, Harris P, Danieli F. Steady state visually evoked potential (SSVEP) topography in a graded working memory task. *Int J Psychophysiol*. 2001:42(2):219–232.

Smallwood J, Beach E, Schooler JW, Handy TC. Going AWOL in the brain: mind wandering reduces cortical analysis of external events. *J Cogn Neurosci*. 2008:20(3):458–469.

Stawarczyk D, Majerus S, Maj M, Van der Linden M, D'Argembeau A. Mind-wandering: phenomenology and function as assessed with a novel experience sampling method. *Acta Psychol*. 2011:136(3): 370–381.

Unsworth N, McMillan BD. Similarities and differences between mind-wandering and external distraction: a latent variable analysis of lapses of attention and their relation to cognitive abilities. *Acta Psychol*. 2014:150:14–25.

Unsworth N, Robison MK. Pupillary correlates of lapses of sustained attention. *Cogn Affect Behav Neurosci*. 2016:16(4):601–615.

Vaidya CJ, Gabrieli JDE. Picture superiority in conceptual memory: dissociative effects of encoding and retrieval tasks. *Mem Cogn*. 2000:28(7):1165–1172.

Van Rooy C, Stough C, Pipingas A, Hocking C, Silberstein RB. Spatial working memory and intelligence: biological correlates. *Intelligence*. 2001:29(4):275–292.

Vialatte F-B, Maurice M, Dauwels J, Cichocki A. Steady-state visually evoked potentials: focus on essential paradigms and future perspectives. *Prog Neurobiol*. 2010:90(4):418–438.

Wagner AD, Koutstaal W, Schacter DL. When encodong yields remembering: insights from event-related neuroimaging. *Philos Trans R Soc B Biol Sci*. 1999:354(1387):1307–1324.

Wiseman S, Neisser U. Perceptual organization as a determinant of visual recognition memory. *Am J Psychol*. 1974:87(4):675–681.

Woldorff MG, Liotti M, Seabolt M, Busse L, Lancaster JL, Fox PT. The temporal dynamics of the effects in occipital cortex of visual-spatial selective attention. *Cogn Brain Res*. 2002:15(1):1–15.

Yi D-J, Turk-Browne NB, Chun MM, Johnson MK. When a thought equals a look: refreshing enhances perceptual memory. *J Cogn Neurosci*. 2008:20(8):1371–1380.

Yomogida Y, Sugiura M, Watanabe J, Akitsuki Y, Sassa Y, Sato T, Matsue Y, Kawashima R. Mental visual synthesis is originated in the fronto-temporal network of the left hemisphere. *Cereb Cortex*. 2004:14(12):1376–1383.

Downloaded from https://academic.oup.com/cercor/advance-article/doi/10.1093/cercor/bhac270/6659085 by Duke University user on 06 January 2023