Consciousness and Cognition 136 (2025) 103958

Contents lists available at ScienceDirect

Consciousness and Cognition

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

Full Length Article

The phenomenology of encoding: Experience sampling reveals thoughts associated with the retention of visual and verbal materials

Brecken Marome[a,*,1], Shivang Shelat[a,b,1], Jonathan W. Schooler[a]

[a] Department of Psychological and Brain Sciences, University of California, Santa Barbara, Santa Barbara, CA 93106, USA
[b] Institute for Collaborative Biotechnologies, University of California, Santa Barbara, Santa Barbara, CA 93106, USA

ARTICLE INFO

Keywords:
Phenomenology
Long-term memory
Mind-wandering
Experience sampling
Recognition
Memorability

ABSTRACT

Evaluating ongoing thoughts during behavioral tasks can offer valuable insight into underlying cognitive processes. Yet, despite their ubiquity, dimensions of thought are often overlooked in experimental psychology, where researchers typically prioritize the assessment of task performance and neglect the accompanying mental experience. In this study, we used experience sampling to investigate the phenomenology of task-unrelated and task-relevant thoughts during memory encoding for verbal and visual stimuli. In two experiments, participants studied words and images matched in intrinsic memorability before completing a recognition memory test. During the study phase, participants responded to several thought probes at pseudorandom intervals, rating dimensions of task-relatedness, awareness, unguidedness, inner speech, visual imagery, auditory imagery, bodily sensation, and emotion. Our results revealed a robust effect across experiments between task-unrelated thoughts (TUTs) and recognition failures, suggesting attention lapses interfere with encoding. Meta-awareness during TUTs showed a protective effect on encoding in Experiment 1 that did not replicate in Experiment 2, and modality-matched TUTs (e.g., visual TUT during image study) did not differentially impair memory. During task-oriented states, verbal stimuli evoked more auditory imagery, while visual stimuli evoked more unguidedness, visual imagery, and emotion. Importantly, certain on-task thought qualities, such as awareness and inner speech, were uniquely linked to enhanced memory performance, suggesting that task-relevant thoughts are heterogeneous in their role in processing and encoding. By emphasizing the intricate relationship between external stimuli, inner experience, and memory encoding, this work calls for a more integrative approach that incorporates phenomenological perspectives in the study of cognition.

1. Introduction

Fluctuations of the content and qualities of thought are core to our inner experience (Garg et al., 2025; Gross et al., 2020; Heavey & Hurlburt, 2008; Ortega et al., 2025) and may explain differences in how we process information and interact with our environment (Oblak et al., 2024). However, traditional methods in experimental psychology sometimes singularly use task performance data to

* Corresponding author at: UC Santa Barbara, Santa Barbara, CA 93107, USA.
E-mail address: brecken@uchicago.edu (B. Marome).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.concog.2025.103958
Received 27 June 2025; Received in revised form 12 October 2025; Accepted 29 October 2025

infer how changes in behavior represent changes in inner cognitive processes, often sidelining subjective experiences as potentially significant determinants in shaping behavior. Even when completing a simple change detection task, a myriad of complex thoughts rotate through conscious awareness, all whilst participants appear to be focused on the task at hand (Oblak et al., 2022). This may explain intra- and inter-individual variability in task performance, contributing to the generalizability of results across populations (Froese et al., 2012). Beyond these methodological implications, the experiences elicited by cognitive tasks are valuable objects of study in their own right, providing insights into how lived experience and cognition interact (Černe & Kordeš, 2023; Hurlburt et al., 2016; Laybourn et al., 2022; Lutz, 2002; Oblak et al., 2022; Slana Ozimič et al., 2023; van Vugt & Broers, 2016).

1.1. Experience sampling for phenomenology of thought

To study thoughts that co-occur with changes in task performance, researchers use experience sampling (ES; Larson & Csikszentmihalyi, 1983). ES, sometimes referred to as ecological momentary assessment (EMA; Shiffman et al., 2008), prompts participants to rate the qualities of their mental states in the past few moments. By allowing participants to quantify dimensions of thought phenomenology, ES supports a pluralistic method required to compare task performance and conscious experience at specific time points (Hurlburt, 2011) without heavily relying on memory. This approach embodies a shift in the cognitive sciences toward viewing lived experience and cognition as interconnected aspects of the same mental process (Bayne & Montague, 2011; Berkovich-Ohana et al., 2020; Rowlands, 2010; Varela, 1996).

Within this framework, thoughts during cognitive tasks can be categorized as task-related or task-unrelated. Task-unrelated thoughts (TUTs — i.e., mind-wandering; Smallwood & Schooler, 2006) are well known to disrupt the encoding of external information into memory (see Blondé et al., 2022 for a review). Recent research reveals that mind-wandering varies in both its phenomenological complexity (Fernyhough et al., 2018) and its disruptive potential (van Vugt & Broers, 2016), indicating the importance of considering other qualitative aspects of experience alongside TUT. For example, while the quality of *meta*-awareness can mitigate some negative effects of TUT on task performance, it does not appear to blunt TUT's effects on some types of memory encoding (Smallwood et al., 2007). A complementary line of work focuses on unguidedness (the extent to which thought flows without deliberate control; Irving, 2016), which is argued to be a core defining feature of mind-wandering and may possess distinct electrophysiological signatures from TUT (Kam et al., 2021). Beyond the dynamics of mind-wandering itself, its modality (whether the thought is visual, auditory, verbal) may further shape its behavioral impact. Notably, Choi et al. (2017) found that participants reported mind-wandering with more visual imagery during visual tasks and more auditory imagery during auditory tasks, highlighting how our subjective experience molds to the nature of a task. Neurophysiological work likewise indicates that cortical resources recruited for imaginative experiences (e.g., visualizing) overlap with resources engaged during perceptual experiences (e.g., seeing) for experiences in similar modalities (Villena-González et al., 2016). Based on this, it may be the case that visual or verbal qualities of TUT amplify a disruptive effect based on whether they match the stimuli being studied (e.g., verbal TUT during word study or visual TUT during image study). Choi et al. (2017) explored this idea, reporting preliminary evidence for a particularly disruptive effect of task-unrelated auditory imagery during their auditory task, but no comparable evidence was found for their visual task. Thus, the question on how TUT qualities interact with the processing of our environment remains open for exploration.

1.2. Dimensions of task-relevant thought

While research into the heterogeneity of TUT is ongoing, less is known about specific qualities of *task-relevant* thought (TRT) evoked during tasks, and whether these qualities interact with the type of material being encoded to shape memory outcomes. For example, a study by Miller et al. (2012) found that both cognitive style and task-specific strategy modulate brain activity during memory retrieval, especially in response to visual vs. abstract verbal stimuli. This suggests that TRTs are context-sensitive and shaped by task structure. In another study by Zhou et al. (2025) that compared verbal and visual production tasks for memorizing words, they found that the visual production task was associated with more spontaneous visual thinking (as opposed to deliberate verbal thinking), suggesting that thought patterns were dependent on the modality of the task at hand. Participants using the visual production task also performed significantly better during a free recall task, opening the possibility that the differences in modality-specific thought patterns played a key role in facilitating task performance.

Indeed, other research suggests that dimensions of TRTs can influence cognitive outcomes. Nedergaard et al. (2023) showed that participants who reported task-related inner speech performed better on a boring sustained attention task, indicating a possible inner verbal strategy to maintain focus. Other researchers found that simply thinking about a task during a resting period was directly correlated with increased connectivity between brain regions responsible for task performance (Gregory et al., 2016). Importantly, Jaeger et al. (2024) demonstrated that specific encoding strategies, such as interactive imagery and sentence generation, significantly enhance memory: across five experiments using word pairs, these strategies outperformed retrieval practice for associative memory, and sentence generation was especially effective for recognition and cued recall. These findings underscore that successful memory encoding is often supported by spontaneously or deliberately adopted strategies that align with the nature of the task. However, less is known about how such strategies unfold subjectively and in real time, or how they vary across different types of stimuli.

1.3. Selecting qualities of experience

Cognitive psychologists have long attempted to peer into the black box of human consciousness by prompting participants to report on the phenomenology of their inner experience. Instead of relying on participants' untrained introspection, researchers select qualities based upon established foundations (Lutz, 2002; Niikawa et al., 2020; Nishida et al., 2024). This methodological approach aligns with what Gallagher (2003) describes as "front-loaded phenomenology," where empirical research is grounded in existing phenomenological knowledge about human experience.

We selected seven specific phenomenological qualities based on both theoretical considerations and empirical precedent. Awareness (Smallwood et al., 2007) and unguidedness (Kam et al., 2021) were chosen because they represent core dimensions in mind-wandering research that likely also manifest in task-relevant states. Inner speech was included due to its role in working memory and cognitive control (Nedergaard & Lupyan, 2024). Visual and auditory imagery (Choi et al., 2017) were selected because of their modality specific relevance to encoding different types of stimuli, based on evidence that imaginative and perceptual processes recruit overlapping neural resources (Villena-González et al., 2016). Emotion was selected because affective states may modulate attention and memory encoding (Tyng et al., 2017), and bodily sensation was included since it is a commonly reported experiential quality (Bastian et al., 2017; Heavey & Hurlburt, 2008) that could potentially influence encoding performance. Rather than using multidimensional experience sampling (MDES; Smallwood et al., 2016), which captures a wide range of spontaneous thought dimensions, we focused on seven qualities to test hypotheses about their relationship to memory encoding across stimulus modalities, balancing specificity with a manageable probe structure.

1.4. The present research

While mind-wandering research has established that TUTs vary in their disruptive effects, systematic analysis of task-relevant thought qualities remains sparse; this limits a full understanding of the nature of memory. Much work in memory psychology primarily focuses on types of retrieval, stimulus properties, and strategies while largely overlooking the experiential qualities that accompany encoding processes. By examining how different dimensions of subjective experience predict memory outcomes, this research may reveal aspects of cognition that traditional behavioral measures alone cannot capture. We extend beyond previous literature and contribute to growing efforts to consider lived experience during various cognitive tasks (Laybourn et al., 2022; Oblak et al., 2022; Oblak et al., 2024; Slana Ozimič et al., 2023). Our goal was to provide a quantitative characterization of inner experience during memory encoding in response to different types of stimuli.

Across two within-subjects experiments, we investigated relationships between stimulus type (verbal and visual), various phenomenological qualities of thought during encoding, and recognition memory performance. Participants completed memory encoding tasks for words and images with an integrated ES procedure. Periodic probes prompted participants to first classify their thoughts as on-task vs. task-unrelated before rating seven qualities; the first two qualities were awareness and unguidedness, drawing on constructs from mind-wandering literature (Kane et al., 2021; Marcusson-Clavertz et al., 2023; McVay & Kane, 2009). The other five qualities have been shown to commonly emerge in daily experience: inner speech, visual imagery, auditory imagery, bodily sensation, and emotion (Bastian et al., 2017; Heavey & Hurlburt, 2008). We posed four research questions:

1. Does task-unrelated thought (TUT) during encoding impair later recognition?
2. Within TUT episodes, does *meta*-awareness shape encoding, and do inner speech and visual imagery during TUT interact with stimulus type to influence memory?
3. Do verbal and visual stimuli evoke distinct qualities of task-*relevant* thought (TRT)?
4. Do specific qualities of TRT interact with stimulus type to predict recognition performance?

1.5. Open practices statement

Experiment 1 was not preregistered. The sample size, exclusion criterion, and analysis plan for Experiment 2 were preregistered, and all additions to the preregistered analysis plan are explicitly noted. Exploratory analyses are labeled. The experimental code, materials, analysis code, and aforementioned analysis plan are publicly accessible in a repository via the Open Science Framework (osf.io/n9gk8).

2. Methods

We report the methods and results of Experiment 1 and Experiment 2 together since they were highly similar. For convenience, we provide a list of the differences between the experiments and their rationales in section 2.4 "Summary of critical differences between experiments." All procedures were approved by the Human Subjects Committee at the University of California, Santa Barbara (UCSB).

2.1. Participants

For Experiment 1, we intended for a sample of around 70 participants to double the sample used by Choi et al. (2017). We used Choi et al. (2017) as a model because of their analogous aim to study inner experience and behavioral performance across varying task parameters. For Experiment 2, we wanted to ensure having a post-exclusion sample of at least 70, so we planned to collect data from around 100 participants. Participants from both experiments were compensated for their time with class credits. They were only allowed to participate if they were at least 18 years old, had normal/corrected-to-normal vision, and reported they were in a quiet room, free from distractions, with access to a computer.

We recruited 73 pre-exclusion participants for Experiment 1 and 109 pre-exclusion participants for Experiment 2 from the online subject pool at UCSB. During post-processing, we performed several filtering steps. We removed those who failed an attention check, had an irregularly high false alarm rate during either study phase (>2.5 SDs above sample mean), or scored irregularly low on the Vividness of Visual Imagery Questionnaire (VVIQ; a measure of mental imagery ability; >2.5 SDs below sample mean). The sample size and exclusion criteria were preregistered for Experiment 2. See Table 1 for demographic information and exclusion details.

2.2. Measures and materials

2.2.1. Stimuli

To control for a well-established effect where images tend to be better remembered than words in recognition tasks (Stenberg et al., 1995) and isolate the relationship between thought, stimulus type, and memory, we sought to rule out another factor known to influence memory performance across populations: memorability. Memorability is the stimulus-inherent property that reflects the probability of remembering a given item across populations (Isola et al., 2011).

For verbal stimuli, we examined English words and corresponding memorability scores provided by Madan (2021), derived from free recall performance in the Penn Electrophysiology of Encoding and Retrieval Study (PEERS; memory.psych.upenn.edu/Penn_Electrophysiology_of_Encoding_and_Retrieval_Study). PEERS was a large-scale multi-session memory study in which 16-item lists were presented for 3,000 ms per word, followed by a free recall test (see Lohnas & Kahana, 2013; Long et al., 2015; Madan, 2021, for full procedural details). Here, we use recall data sourced by Madan (2021), who examined the performance of 147 young adults (ages 16–30) who completed 20 sessions of free recall across 1,638 non-plural nouns. Memorability was operationalized as the proportion of participants who successfully recalled a given word following the presentation of a study list. The average memorability of the 1,638 words was 0.69 (SD = 0.080). We defined a ± 0.25 SD bound around this mean value, yielding 351 candidate words with memorability values between ~ 0.67 and ~ 0.71. We selected 60 of these to be included in the experiment, varying by meaning, length, structure, animacy, and the first letter. All words were non-plural nouns. Example words include “animal” and “mask.” During the experiment, all words were printed in capital letters with a font size of 50 pixels.

For visual stimuli, we selected complex scene images from the MIT FIne-GRained Image Memorability (FIGRIM) dataset (figrim.mit.edu), which provides high-confidence recognition scores for 1,754 images (Bylinskii et al., 2015). Memorability was defined as the hit rate in a continuous recognition task administered on Amazon Mechanical Turk, in which participants viewed image streams and pressed a key upon detecting a repeat. Repeat lags were approximately 100 trials apart, and ~ 80 participants viewed each image. Hit rate — computed as hits / (hits + misses) — served as the measure of image memorability (M = 0.66, SD = 0.14; see Bylinskii et al., 2015 for full procedural details).

Although Madan’s memorability scores were based on free recall proportions and Bylinskii et al.’s scores were based on high-confidence recognition hit rates, both scores reflect stimulus-intrinsic memorability across participants. Thus, we used these datasets to match the words and images on memorability for the present study. We sampled images from FIGRIM’s full across-category collection, distributing selections across 21 semantic sub-categories (e.g., bedroom, terminal). We applied the same filtering bounds used for word selection (~0.67 to ~ 0.71), yielding 216 candidate images. From these, we selected 60 images that were approximately balanced in brightness and color vibrance. We excluded images containing human faces or figures, obvious enhancement or editing, unusual borders, or high similarity to other selected images. All images were cropped to 500 × 500 pixel squares.

An independent samples Welch’s *t*-test was conducted to compare memorability between the image and word stimuli. There was no significant difference in memorability between the stimulus sets, $t(117.86) = 0.63$, $p = 0.53$, $d = 0.11$. The 95 % confidence interval for the difference in memorability means ranged from −0.0030 to 0.0057 ($M_{images} = 0.69$; $SD_{images} = 0.012$; $M_{words} = 0.68$, $SD_{words} = 0.012$).

2.2.2. Thought probes

To categorize periods of task engagement as on-task vs. off-task, we used a validated probe method with wording adapted from Kane et al. (2021). Participants were prompted: “Just prior to the onset of this screen, I was thinking about….” Options included “the task”, “everyday things”, “my current state of being”, “personal worries”, “daydreams”, “external environment”, and “other.” During the encoding phase, thought probes were inserted at pseudorandom intervals. A total of 20 probes were presented for each participant, with 10 appearing during each study phase. Probes appeared after every 2nd or 3rd stimulus for each set of 3 stimuli. We selected this

Table 1
Participant demographics and exclusion criteria for Experiments 1 and 2.

	Experiment 1	Experiment 2
Recruited (n)	73	109
Failed attention check	4	5
High false alarm rate	3	5
Low VVIQ score	1	2
Total excluded	8	11
Final sample	65	98
Age M (SD)	19.00 (0.96)	19.34 (3.41)
Gender (M/F/other)	Not collected	11/87/0
VVIQ M (SD)	55.74 (10.11)	54.55 (10.52)

probe cadence based on prior work indicating that, when probed every 30 s, you would expect mind-wandering to occur around half of the time (Seli et al., 2013). TRT and TUT were binarized as reporting thoughts about "the task" vs. any other option, respectively.

For both TRT and TUT reports, each probe also measured seven different qualities of thought: awareness, unguidedness, inner speech, visual imagery, auditory imagery, bodily sensation, and emotion. Participants rated all other qualities on 7-point Likert scales (minimum = 0, maximum = 6). In response to "I was guiding my thoughts," participants rated their thoughts from "Completely unguided" to "Completely guided," which was reverse-scored to produce unguidedness. In response to "I was aware of my thoughts," participants rated their thoughts from "No awareness" to "Complete awareness." For inner speech, visual imagery, auditory imagery, bodily sensation, and emotion, participants responded to the prompt "My thoughts were focused on [insert thought quality]." They were rated on scales from "No [insert thought quality]" to "Lots of [insert thought quality]." Due to a concern that participants were reporting high amounts of visual imagery in response to the mere presence of images on the computer screen in Experiment 1, we inserted the word "inner" in front of each occurrence of "visual imagery" on each thought probe for Experiment 2 to emphasize the importance of reporting only the internal experience of visual imagery.

Prior to the onset of the study phase, participants were greeted with a practice probe and instructions for responding to each probe item. For task-relatedness, they were told: "Everyday things should be selected when you're thinking about routine events in the recent or distant past or future. Current state of being should be selected when you're thinking about physical or emotional states. Personal worries should be selected when you're thinking about life concerns. Daydreams should be selected when you're thinking about fantasies or thoughts disconnected from reality. External environment should be selected when you're thinking about objects or task-unrelated events in the room. Other should be selected when you're thinking about anything that was not captured by the other choices."

For the other seven qualities, they were instructed: "Thoughts with high 'guidance' are thoughts that you are controlling (i.e., not spontaneous and not freely moving). Thoughts with high 'awareness' are thoughts that you had explicit knowledge of. Thoughts with high 'inner speech' are for when you are internally talking to yourself using words that you could report. Thoughts with high 'visual imagery' are ones where you have the visual experience of a mental image. Thoughts of 'auditory imagery' have a sound quality, like a tune in your head. Thoughts on 'bodily sensation' are for when you're focusing on a pain or itch in your foot, for example. Thoughts with a lot of 'emotion' are those that are strongly valenced (e.g., sad or happy). Make sure to evaluate your thoughts on each of these categories. They can belong to multiple of these at one time!"

2.2.3. Vividness of visual imagery Questionnaire (VVIQ)

Participants were administered the VVIQ to measure their capacity for visual imagery (Marks, 1973). Responses to the 16 VVIQ items are designated along a 5-point Likert scale. A score per participant is calculated by adding the responses together, ranging from 16 to 80. From low to high, the sum represents the participant's vividness of visual imagery. The questions ask the participant to imagine a specific scene (e.g., "a rising sun") and indicate how vividly they can imagine a certain aspect of this image (e.g., "rising above the horizon into a hazy sky"). There is no agreed-upon cutoff to categorize a certain individual as aphantasic or not (Bainbridge et al., 2021), so we conservatively chose to exclude participants if their score was an outlier (see section 2.1) in order to generalize our results to individuals who can generate visual imagery.

2.2.4. Cognitive dimension of the 3D-WS

The cognitive dimension of the Three-Dimensional Wisdom Scale (3D-WS; Ardelt, 2003) was administered for another study. It will not be discussed further in the present manuscript.

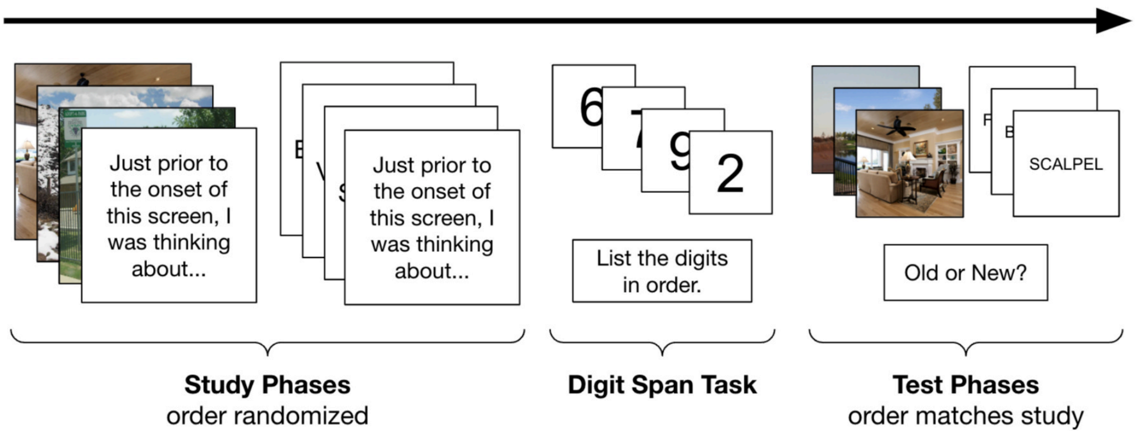

Fig. 1. A procedure flowchart of the memory experiment. Participants first studied images and words with intermittent thought probes. Then, they completed a digit span task to prevent information maintenance strategies in working memory. Finally, participants underwent a recognition memory test including previously studied stimuli and intermixed foils.

2.2.5. Exploratory questionnaire

Experiment 2 also included exploratory items intended to tease apart participant perceptions of the task. First, it inquired into the nature of the inner rehearsal mechanism and asked participants to report whether they selected auditory imagery and/or inner speech to report it. Participants were prompted: "When using inner rehearsal to remember the stimuli, what did you indicate during the thought probe?" They responded on a 5-point Likert scale with the following choices: "A lot more inner auditory imagery", "Somewhat more inner auditory imagery", "Equal auditory imagery and inner speech", "Somewhat more inner speech", "A lot more inner speech." Then, two free-response questions asked about the strategies used to remember words and images: "What strategies did you use to remember [words/images]?" (Miller et al., 2012; Slana Ozimič et al., 2023).

2.3. Procedure

The experiment took approximately 27 min to complete and consisted of three main parts: the study phases, the digit span task, and the recognition test (Fig. 1). This experiment was built using JavaScript (jsPsych plugins; de Leeuw, 2015) and hosted on cognition.run (Vidal et al., 2025).

2.3.1. Study phases

The encoding phase consisted of two blocks in a randomized order across participants. Each block included 30 stimuli (images or words) and 10 thought probes. Each stimulus remained on the screen for 10 s, with a 100 ms interstimulus interval (ISI) before the next. Participants were informed that they should study the stimuli, as they would later be tested, but were not informed of the nature of the test.

2.3.2. Digit span task

Mirroring common methods used to prevent the rehearsal of information between study and test, we administered a ladder-design digit span task to flush working memory and create a delay after memory encoding (e.g., Bainbridge et al., 2019). Methods from this task resemble those described by Luthra and Todd (2019). Participants viewed a number sequence, where each number was presented independently (stimulus time = 1 s, ISI = 0 ms), and then were asked to type the digits in the order they appeared. This version of the task adds an extra digit (up to 9) if the previous answer was correct and decreases the number of digits (down to 3) if two incorrect answers are provided in a row. All participants completed exactly 20 repetitions of the task, which typically took 3–5 min. Participants received accuracy feedback after each trial.

2.3.3. Recognition memory tests

Afterward, participants were tested on their recognition memory. The tests were presented in the same order as the study phases. For example, if the participant studied words first, they were tested on words first. The 30 stimuli originally studied were intermixed with 30 new foils of the same stimulus type, with the order completely randomized. Here, they were asked to indicate whether each stimulus presented was old (seen before) or new (not seen before). Response options were provided on a 4-point scale: 1 = high-confidence new, 2 = low-confidence new, 3 = low-confidence old, and 4 = high-confidence old. Participants indicated their responses using the number keys on their keyboards. The response options were presented below each stimulus, remaining on-screen until a selection was made, followed by a blank 500 ms trial gap.

2.4. Summary of critical differences between experiments

(1) Sample size.

In Experiment 1, our sample size was smaller than intended after applying our exclusion criteria. Thus, in Experiment 2, we aimed to collect data until we had reached approximately 100 participants. *This step was taken to ensure that our sample size would be over double that of* Choi et al. (2017) *after exclusion.*

(2) Visual imagery probe.

In Experiment 1, participants responded to probes prompting "My thoughts were focused on visual imagery." Participants rated their visual imagery from "No visual imagery" to "Lots of visual imagery" on a 7-point Likert scale. In Experiment 2, we added the word "inner" before each instance of "visual imagery." Participants responded to probes prompting "My thoughts were focused on inner visual imagery." Participants rated their inner visual imagery from "No inner visual imagery" to "Lots of inner visual imagery" on a 7-point Likert scale. *This step was taken to address a concern that participants were reporting high amounts of visual imagery in response to the mere presence of images on the computer screen. We wanted to emphasize the internal experience of visual imagery.*

(3) Retrospective questionnaire.

In both experiments, the VVIQ and the Cognitive Dimension of the 3D-WS were administered immediately after the recognition test with order randomized. In Experiment 2, we added exploratory items (see Section 2.2.5) after these questionnaires. *This step was taken to provide insight into the strategies that participants employed and each participant's interpretation of these strategies.*

2.5. Statistical analysis

We only considered an effect significant if its p-value was below 0.05 for both experiments. This minimizes the probability of spurious results since the likelihood of a p-value passing that threshold in two independent samples by chance is.05 × .05 = 0.0025, or

0.25 %. Still, we discuss all results that were at least significant in one experiment ($p < 0.05$) and trending in another experiment ($p < 0.10$).

For trial-by-trial analyses, we primarily used linear and logistic mixed-effects regression models with subjectID assigned as a random effect. All betas reported here are unstandardized. An analysis plan was preregistered at osf.io/n9gk8 for Experiment 2; a table with specific preregistered model parameters can be found in the Appendix for ease of comparison.

We calculated signal detection theory metrics of d-prime and criterion after applying a correction to hit and false alarm rates of 0 or 1. Specifically, values of 0 were replaced with 0.5 / 30, and values of 1 were replaced with (30 − 0.5) / 30, where 30 was the number of target or lure trials (Macmillan & Kaplan, 1985). D-prime and criterion were calculated as:

$$d' = z(hitrate) - z(falsealarmrate)$$

$$c = -0.5 \times [z(hitrate) + z(falsealarmrate)]$$

All data analyses were conducted in R version 4.4.2 (R Core Team, 2025). We used ggeffects to generate model predictions for visualizing the data (Lüdecke, D, 2018).

3. Results

3.1. Descriptive statistics and order effects

First, we examined overall memory performance metrics and proportion of TUT across stimuli study phases (Table 2). Correct recognitions were classified as a “hit” if participants responded with either low or high confidence “old” (i.e., response = 3 or 4) and the stimulus had been presented during the encoding phase; otherwise, the response was coded as a “miss.” If participants gave a low or high confidence “old” response to a stimulus that had not been presented during encoding, the response was classified as a “false alarm”; otherwise, it was coded as a “correct rejection.”

We found no evidence that the type of stimuli being studied affected memory performance or rates of mind-wandering (Table 3).

We fitted a series of linear mixed-effects regression models predicting memory metrics and TUT rates from study phase and whether participants saw that phase first (“words first” vs. “images first”). A main effect of study phase order or an interaction would indicate an order effect. No such effects replicated across both experiments. In Experiment 1, there was a significant interaction for hit rate ($p = 0.0037$), but this did not replicate in Experiment 2 ($p = 0.68$). Similarly, an interaction for d-prime was significant in Experiment 1 ($p = 0.022$) but absent in Experiment 2 ($p = 0.56$). Overall, the evidence did not support consistent order effects.

3.2. Does TUT during encoding impair later recognition?

Our first research question asked whether TUT during encoding impaired later recognition. We predicted that stimuli encountered during TUT would have a reduced likelihood of being recognized during test. We fitted logistic mixed-effect regressions with memory trial accuracy as the dependent variable and TUT as the predictor. In both experiments, TUT during encoding was associated with a reduced probability of correct recognition of the stimuli directly before the thought probe (Experiment 1: $b = -0.43$, $SE = 0.16$, $z = -2.74$, $p = 0.0061$; Experiment 2: $b = -0.64$, $SE = 0.13$, $z = -4.98$, $p < 0.001$; Fig. 2).

3.3. How do thought qualities interact with TUT during encoding?

Our second research question was split into two parts: Within TUT episodes, does *meta*-awareness shape encoding, and do inner speech and visual imagery during TUT interact with stimulus type to influence memory? To answer this, we filtered our data to only TUT reports to assess how different qualities of TUT shaped memory encoding. This filtering step removed two participants from Experiment 1 and four participants from Experiment 2 who did not report any TUT.

3.3.1. Does meta-awareness during TUT shape encoding?

We modeled the effects of *meta*-awareness on hit probability to test whether awareness during mind-wandering mitigated the effects of TUT on encoding (cf. Smallwood et al., 2007). There was a positive effect of awareness during TUT on recognition in Experiment 1 ($b = 0.28$, $SE = 0.063$, $z = 4.40$, $p < 0.001$) but not Experiment 2 ($b = 0.079$, $SE = 0.050$, $z = 1.57$, $p = 0.12$).

Table 2
Descriptive statistics for both experiments across study phases. Values are mean (SD).

experiment	task	hit rate	false alarm rate	d-prime	criterion	proportion TUT
exp1	image recognition	0.76 (0.17)	0.24 (0.20)	1.81 (1.32)	0.024 (0.34)	0.55 (0.30)
exp1	word recognition	0.77 (0.19)	0.23 (0.20)	1.87 (1.37)	0.021 (0.38)	0.49 (0.31)
exp2	image recognition	0.76 (0.19)	0.17 (0.18)	2.06 (1.25)	0.16 (0.40)	0.53 (0.31)
exp2	word recognition	0.78 (0.20)	0.20 (0.20)	2.02 (1.35)	0.10 (0.41)	0.52 (0.32)

Table 3
A set of paired-samples t-tests on memory metrics and TUT rates across stimuli study phases for both experiments.

experiment	metric	t-statistic	df	p	d	magnitude
exp1	hit rate	−0.57	64	0.57	−0.070	negligible
exp1	false alarm rate	0.054	64	0.96	0.0067	negligible
exp1	d-prime	−0.52	64	0.61	−0.064	negligible
exp1	criterion	0.061	64	0.95	0.0076	negligible
exp1	proportion TUT	1.71	64	0.092	0.21	small
exp2	hit rate	−0.59	97	0.56	−0.060	negligible
exp2	false alarm rate	−1.49	97	0.14	−0.15	negligible
exp2	d-prime	0.36	97	0.72	0.036	negligible
exp2	criterion	1.50	97	0.14	0.15	negligible
exp2	proportion TUT	0.40	97	0.69	0.040	negligible

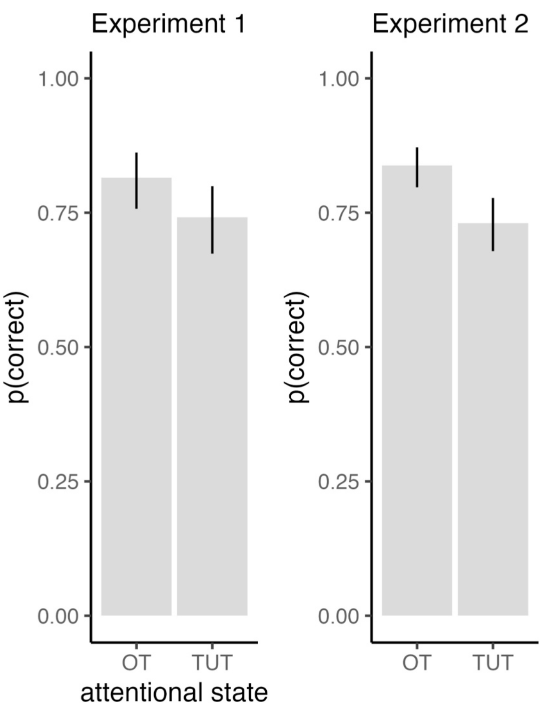

Fig. 2. Logistic mixed-effect regression predictions from a statistical test modeling memory trial accuracy by attentional state. Error bars indicate the 95% confidence interval. Stimuli encountered during TUT were less likely to be recognized during test than those encountered during on-task (OT) states.

3.3.2. Do inner speech and visual imagery during TUT interact with stimulus type to influence memory?

Another unregistered prediction was that verbal and visual mind-wandering would especially interfere with memory encoding for stimuli of the same modality (e.g., verbal TUT during word study; visual TUT during image study). Using the same filtered dataframe with only TUT reports, we fitted two models: memory hit by 1) the interaction between study phase and inner speech and 2) the interaction between study phase and visual imagery. There was no interaction with inner speech (Experiment 1: $p = 0.44$; Experiment 2: $p = 0.81$) nor visual imagery (Experiment 1: $p = 0.80$; Experiment 2: $p = 0.41$).

3.4. Do verbal and visual stimuli evoke distinct qualities of TRT?

Our third research question was how task-relevant thought qualities were shaped by the type of stimuli in the study phase. We applied linear mixed-effect regression models (Table 4) to a filtered dataframe that only had on-task thought reports; this filtering step removed three participants from Experiment 1 and four participants from Experiment 2 who reported all TUT and thus had no OT reports to consider.

The results suggested that studying words relative to images incurred greater auditory thoughts, less emotional thoughts, less imagery, and less unguidedness (Fig. 3). These findings indicate that the modality of the study stimuli modulates TRT qualities, with words eliciting more auditory processing and fewer affective, visual, or spontaneous thought experiences compared to images.

Table 4
Linear mixed-effect regressions with thought qualities as a function of study phase stimuli. Bolded rows highlight non-intercept terms that replicated across both experiments. The image condition was coded as the intercept.

experiment	dv	term	estimate	SE	t-statistic	p
exp1	awareness	(intercept)	4.43	0.16	28.07	< 0.001
exp1	awareness	study phase (words)	0.19	0.087	2.16	0.031
exp2	awareness	(intercept)	4.64	0.10	44.71	< 0.001
exp2	awareness	study phase (words)	0.088	0.070	1.25	0.21
exp1	bodily sensation	(intercept)	1.21	0.14	8.38	< 0.001
exp1	bodily sensation	study phase (words)	−0.14	0.091	−1.58	0.11
exp2	bodily sensation	(intercept)	1.33	0.13	10.38	< 0.001
exp2	bodily sensation	study phase (words)	−0.084	0.079	−1.06	0.29
exp1	auditory imagery	(intercept)	1.33	0.18	7.31	< 0.001
exp1	**auditory imagery**	**study phase (words)**	**0.64**	**0.12**	**5.24**	**< 0.001**
exp2	auditory imagery	(intercept)	1.69	0.17	10.15	< 0.001
exp2	**auditory imagery**	**study phase (words)**	**0.26**	**0.093**	**2.79**	**0.0053**
exp1	emotion	(intercept)	1.40	0.16	8.93	< 0.001
exp1	**emotion**	**study phase (words)**	**−0.34**	**0.074**	**−4.63**	**< 0.001**
exp2	emotion	(intercept)	1.63	0.13	12.45	< 0.001
exp2	**emotion**	**study phase (words)**	**−0.34**	**0.077**	**−4.42**	**< 0.001**
exp1	visual imagery	(intercept)	4.45	0.17	25.53	< 0.001
exp1	**visual imagery**	**study phase (words)**	**−0.74**	**0.12**	**−6.19**	**< 0.001**
exp2	visual imagery	(intercept)	3.57	0.17	20.58	< 0.001
exp2	**visual imagery**	**study phase (words)**	**−0.24**	**0.11**	**−2.26**	**0.024**
exp1	inner speech	(intercept)	3.70	0.22	17.15	< 0.001
exp1	inner speech	study phase (words)	0.20	0.11	1.89	0.059
exp2	inner speech	(intercept)	3.73	0.16	23.30	< 0.001
exp2	inner speech	study phase (words)	0.48	0.087	5.59	< 0.001
exp1	unguidedness	(intercept)	1.87	0.17	10.82	< 0.001
exp1	**unguidedness**	**study phase (words)**	**−0.22**	**0.097**	**−2.26**	**0.024**
exp2	unguidedness	(intercept)	1.81	0.13	14.06	< 0.001
exp2	**unguidedness**	**study phase (words)**	**−0.30**	**0.074**	**−4.02**	**< 0.001**

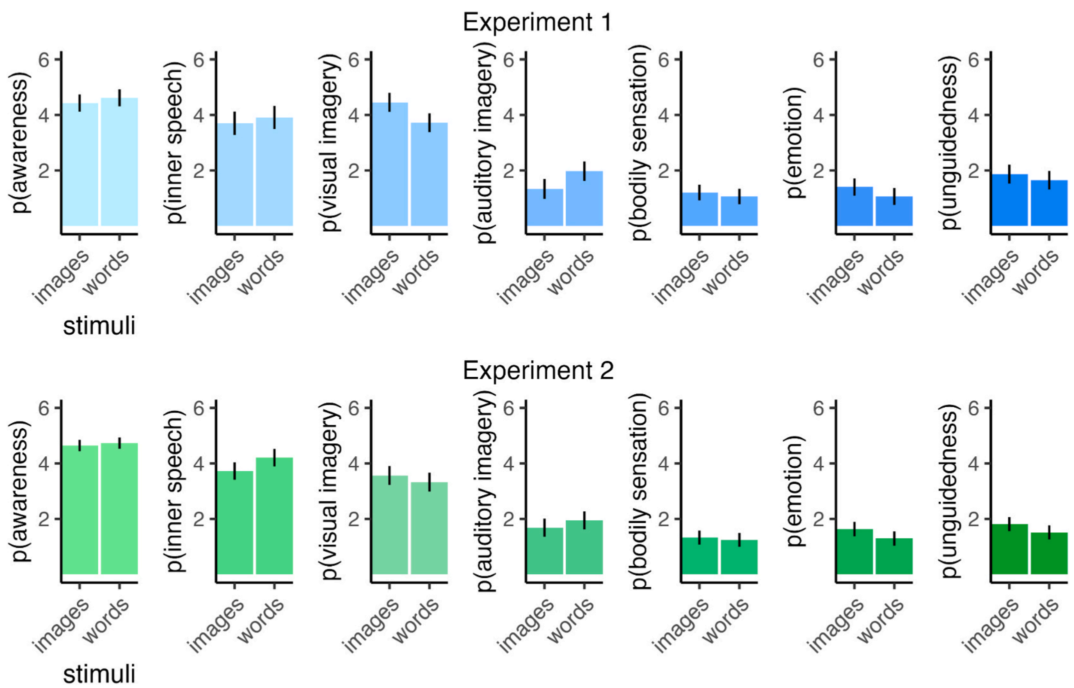

Fig. 3. Linear mixed-effect regression predictions for on-task thought qualities as a function of stimuli during the study phase. Error bars indicate the 95% confidence interval. A significant main effect of stimulus type on auditory imagery, emotion, visual imagery, and unguidedness of task-oriented thoughts replicated across both experiments.

3.5. Do specific qualities of TRT interact with stimulus type to predict recognition performance?

Our fourth research question was whether any phenomenological qualities of task-oriented thought were associated with memory performance, and if this varied by stimuli type. To test this, we specified a set of logistic mixed-effect regressions (Table 5) with memory hit as the DV and the interaction between thought quality and study phase as the predictor.

While there was no replicable evidence of any interactions between thought qualities and the type of stimuli being studied, there were two main effects: higher levels of on-task awareness and increased reports of inner speech were significantly associated with greater hit probability. These effects were consistent across the encoding of words and images (Fig. 4). This implies that when participants were more aware of their on-task state and/or engaged in inner speech during encoding, they were more likely to successfully

Table 5
Logistic mixed-effect regressions with memory hits as a function of thought qualities and study phase. Bolded rows highlight non-intercept terms that replicated across both experiments.

experiment	dv	term	estimate	SE	t-statistic	p
exp1	awareness	(intercept)	0.52	0.60	0.87	0.38
exp1	**awareness**	**awareness**	**0.26**	**0.13**	**2.04**	**0.041**
exp1	awareness	study phase (words)	0.27	0.74	0.36	0.72
exp1	awareness	awareness:study phase (words)	−0.089	0.16	−0.58	0.57
exp2	awareness	(intercept)	0.092	0.49	0.19	0.85
exp2	**awareness**	**awareness**	**0.37**	**0.10**	**3.67**	**< 0.001**
exp2	awareness	study phase (words)	0.64	0.67	0.96	0.34
exp2	awareness	awareness:study phase (words)	−0.16	0.14	−1.14	0.26
exp1	bodily sensation	(intercept)	1.76	0.28	6.35	< 0.001
exp1	bodily sensation	bodily sensation	−0.063	0.12	−0.52	0.60
exp1	bodily sensation	study phase (words)	−0.18	0.29	−0.62	0.53
exp1	bodily sensation	bodily sensation:study phase (words)	0.052	0.16	0.32	0.75
exp2	bodily sensation	(intercept)	1.84	0.24	7.64	< 0.001
exp2	bodily sensation	bodily sensation	−0.052	0.096	−0.54	0.59
exp2	bodily sensation	study phase (words)	−0.32	0.25	−1.31	0.19
exp2	bodily sensation	bodily sensation:study phase (words)	0.24	0.13	1.80	0.072
exp1	auditory imagery	(intercept)	1.85	0.29	6.46	< 0.001
exp1	auditory imagery	auditory imagery	−0.11	0.11	−1.04	0.30
exp1	auditory imagery	study phase (words)	−0.38	0.32	−1.19	0.23
exp1	auditory imagery	auditory imagery:study phase (words)	0.16	0.13	1.22	0.22
exp2	auditory imagery	(intercept)	1.66	0.24	6.84	< 0.001
exp2	auditory imagery	auditory imagery	0.071	0.087	0.81	0.42
exp2	auditory imagery	study phase (words)	−0.31	0.25	−1.24	0.21
exp2	auditory imagery	auditory imagery:study phase (words)	0.16	0.11	1.44	0.15
exp1	emotion	(intercept)	1.90	0.29	6.56	< 0.001
exp1	emotion	emotion	−0.16	0.12	−1.31	0.19
exp1	emotion	study phase (words)	−0.27	0.30	−0.89	0.38
exp1	emotion	emotion:study phase (words)	0.088	0.16	0.55	0.58
exp2	emotion	(intercept)	1.97	0.26	7.65	< 0.001
exp2	emotion	emotion	−0.12	0.089	−1.36	0.17
exp2	emotion	study phase (words)	−0.34	0.26	−1.29	0.20
exp2	emotion	emotion:study phase (words)	0.21	0.13	1.56	0.12
exp1	visual imagery	(intercept)	0.15	0.48	0.31	0.75
exp1	visual imagery	visual imagery	0.36	0.10	3.47	< 0.001
exp1	visual imagery	study phase (words)	0.81	0.53	1.52	0.13
exp1	visual imagery	visual imagery:study phase (words)	−0.19	0.12	−1.52	0.13
exp2	visual imagery	(intercept)	1.43	0.32	4.43	< 0.001
exp2	visual imagery	visual imagery	0.10	0.074	1.41	0.16
exp2	visual imagery	study phase (words)	−0.45	0.35	−1.26	0.21
exp2	visual imagery	visual imagery:study phase (words)	0.15	0.094	1.54	0.12
exp1	inner speech	(intercept)	0.59	0.38	1.54	0.12
exp1	**inner speech**	**inner speech**	**0.30**	**0.092**	**3.25**	**0.0012**
exp1	inner speech	study phase (words)	0.68	0.47	1.45	0.15
exp1	inner speech	inner speech:study phase (words)	−0.23	0.11	−2.00	0.046
exp2	inner speech	(intercept)	1.09	0.33	3.34	< 0.001
exp2	**inner speech**	**inner speech**	**0.19**	**0.074**	**2.52**	**0.012**
exp2	inner speech	study phase (words)	−0.24	0.43	−0.56	0.57
exp2	inner speech	inner speech:study phase (words)	0.026	0.10	0.25	0.80
exp1	unguidedness	(intercept)	2.05	0.32	6.45	< 0.001
exp1	unguidedness	unguidedness	−0.21	0.12	−1.79	0.073
exp1	unguidedness	study phase (words)	−0.46	0.33	−1.37	0.17
exp1	unguidedness	unguidedness:study phase (words)	0.19	0.15	1.33	0.18
exp2	unguidedness	(intercept)	2.31	0.28	8.32	< 0.001
exp2	unguidedness	unguidedness	−0.28	0.089	−3.16	0.0016
exp2	unguidedness	study phase (words)	−0.23	0.29	−0.80	0.43
exp2	unguidedness	unguidedness:study phase (words)	0.060	0.13	0.48	0.63

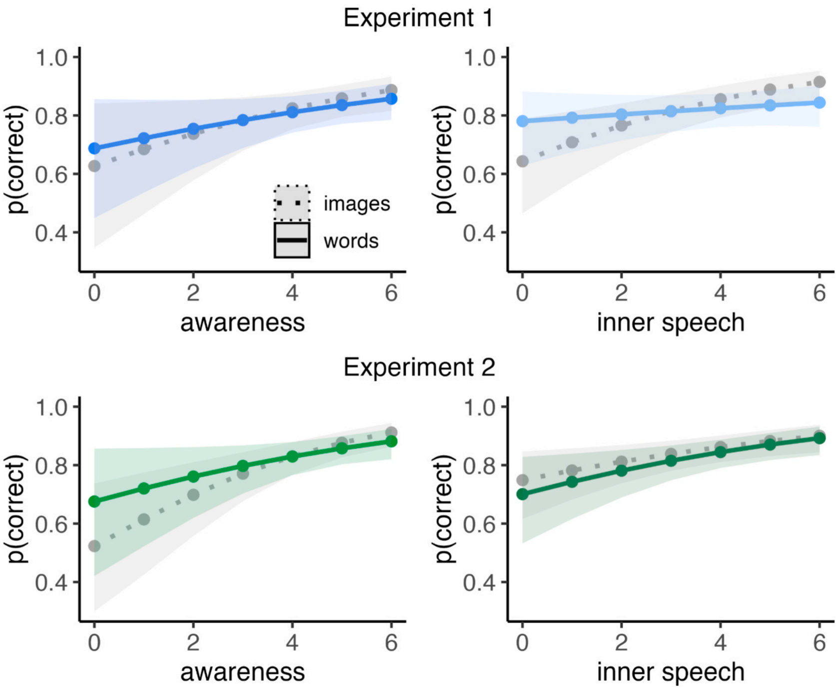

Fig. 4. Logistic mixed-effect model predictions for memory hit probability as a function of two on-task thought qualities: awareness and inner speech. Error shading represents the 95% confidence interval. Dotted lines represent memory for images, whereas solid lines represent words. Task-relevant awareness and inner speech were significantly positively associated with memory performance across both types of stimuli in both experiments.

recognize studied items later. Interestingly, on-task unguidedness was significantly associated with poorer memory performance in Experiment 2, but this effect only trended toward our p-value threshold in Experiment 1. This suggests that when participants' thoughts were less structured or deliberate during encoding, it impeded storage of stimuli into memory. However, we interpret this with caution since the effect was not significant across both experiments.

Given that awareness and inner speech are linked to one another in the literature (Bastian et al., 2017), we conducted an exploratory control analysis to test if they would stay significantly associated with memory when covarying each other. To increase the power of this exploratory check, we merged the data from both experiments. We specified a logistic mixed-effects regression with memory hit as the DV and awareness + inner speech as the predictors. On-task awareness ($b = 0.18$, $SE = 0.060$, $z = 2.87$, $p = 0.0041$) and inner speech ($b = 0.12$, $SE = 0.050$, $z = 2.38$, $p = 0.017$) both remained statistically significantly associated with recognition performance, indicating that they are uniquely linked with encoding.

In addition to analyzing subjective reports of thought qualities, we were interested in how participants evaluate their own internal memorization strategies. One addition to Experiment 2 enabled us to check how participants interpreted the concept of "inner rehearsal" during encoding. Specifically, we wanted to conduct an exploratory analysis on whether they associated a rehearsal strategy with inner auditory imagery or inner speech, as both qualities were measured by the thought probe. After the memory test, participants provided ratings on a 5-point Likert scale (0 = inner auditory, 4 = inner speech) to indicate what they meant by engaging in rehearsal. We performed a one-sample *t*-test against the scale midpoint (mu = 2) and found that participants largely reported inner speech when using a rehearsal strategy, $t(97) = 6.41$, $p < 0.001$, $d = 0.65$.

We then checked how this Likert response affected the coupling between inner on-task auditory thoughts and memory performance in Experiment 2. The logistic mixed-effects regression showed a positive main effect of inner auditory experience on hit probability ($b = 0.27$, $SE = 0.093$, $z = 2.86$, $p = 0.0042$). There was a significant interaction such that those who attributed rehearsal to inner speech had reduced coupling between inner auditory and memory ($b = -0.068$, $SE = 0.032$, $z = -2.17$, $p = 0.030$), as opposed to those who attributed rehearsal to inner auditory. This confirms that participants' conceptualization of inner rehearsal (i.e., as a verbal process vs. an auditory process) moderated how their thought reports related to memory encoding.

3.6. Trait and state visual imagery

The design of our task enabled us to run exploratory tests on VVIQ scores and visual imagery across TRT and TUT during encoding. We fitted linear mixed-effect regressions with visual imagery as the DV and the interaction between VVIQ scores and TUT as the predictor. In both experiments, VVIQ was associated with greater imagery (Experiment 1: $b = 0.045$, $SE = 0.015$, $p = 0.0039$; Experiment 2: $b = 0.036$, $SE = 0.014$, $p = 0.014$). TUTs had greater imagery than TRTs in Experiment 2 ($b = 1.47$, $SE = 0.47$, $p =$

0.0019), but not significantly in Experiment 1 ($b = 0.91$, $SE = 0.57$, $p = 0.11$). Analyses from both samples revealed a significant interaction between VVIQ and TUT (Experiment 1: $b = -0.027$, $SE = 0.010$, $p = 0.0093$; Experiment 2: $b = -0.039$, $SE = 0.0085$, $p < 0.001$). Simple slopes post-hoc analyses showed that this was driven by stronger positive coupling between VVIQ and visual imagery during on-task states (Experiment 1: TRT $b = 0.045$, $SE = 0.015$, $p = 0.0040$; Experiment 2: TRT $b = 0.036$, $SE = 0.014$, $p = 0.014$), but not TUT (Experiment 1: TUT $b = 0.018$, $SE = 0.016$, $p = 0.25$; Experiment 2: TUT $b = -0.0029$, $SE = 0.015$, $p = 0.85$). This exploratory result suggests that participants with a greater capacity to generate inner visual imagery (high VVIQ) reported increased visual processing during memory encoding, but this was reduced during mind-wandering episodes.

4. Discussion

Psychologists have long assumed that traditional memory paradigms isolate the cognitive process they intend to address, but inner experience likely shapes encoding. Researchers have attempted to measure these covert processes by using interviews and questionnaires (Miller et al., 2012; Oblak et al., 2022). Here, we applied experience sampling across two experiments to (1) replicate a detrimental effect of trial-by-trial task-unrelated thoughts (TUT) on recognition, (2) test whether *meta*-awareness, inner speech or visual imagery during TUT are linked to memory, (3) examine which task-relevant thought (TRT) qualities are evoked by verbal vs. visual materials, and (4) identify TRT correlates of memory performance.

4.1. How TUTs with specific phenomenology affect encoding

Across both experiments, stimuli encountered during TUTs were significantly less likely to be recognized, reaffirming the role of attentional state in memory formation (Blondé et al., 2022). With long stimulus durations and no motor demands (deBettencourt et al., 2018; Wakeland-Hart et al., 2022), participants still reported TUTs nearly half of the time, consistent with Seli et al. (2013). Even with 10 full seconds of presentation, the absorptive nature of TUT significantly impaired later recognition.

4.1.1. Awareness during TUTs

Meta-awareness of mind-wandering is proposed to be an executive function of the brain, allowing the mind to redirect attention away from TUTs and back toward the task at hand (Smallwood et al., 2007). Our Experiment 1 suggested that awareness during TUTs might mitigate some encoding interference; this effect was in the same direction but not statistically replicated in Experiment 2. This pattern suggests a small protective influence of *meta*-awareness on encoding, though the inconsistent replication may simply reflect sample variation or the modest size of any protective effect.

Smallwood et al. (2007) observed that participants' awareness of TUT did not prevent memory impairments for word stems, implying that awareness alone may be insufficient to counteract the costs of TUT on encoding. Our paradigm differs from theirs in both the type of memory test (recognition vs. stem-completion) and stimulus presentation duration, any of which could modulate the effect of awareness. The directionality of the effects across both of our experiments here suggests that future high-powered research is necessary; however, based on our evidence combined with Smallwood et al.'s, any potential effect is likely modest at least.

4.1.2. Intra-modal TUT bears no more harm to encoding than cross-modal TUT

We found no evidence that intra-modal TUT (e.g., verbal mind-wandering during word encoding) was more disruptive than cross-modal TUT. This analysis was inspired by exploratory work in Choi et al. (2017), who found that TUT with auditory imagery interfered more with accuracy on an auditory 2-back task than TUT with other phenomenology. Although our generally analogous analysis with long-term memory used a larger sample, as was our intent, it is important to consider the methodological differences between Choi et al. (2017)'s study and ours, and to recognize the continued need for further research on intra-modal TUT.

Previous work in neuroscience indicates that intra-modal internal and external processes recruit resources from the same brain regions (Villena-González et al., 2016), suggesting that when TUT shares the same modality as an ongoing task, it may especially interfere with task performance. If we posit this intra-modal interference, then one explanation for why we found no significant support for this effect may lie in our variable operationalization. Specifically, Choi et al. (2017) distinguished between auditory and visual modalities using auditory and visual 2-back tasks. We, however, distinguished between verbal and visual modalities by presenting two types of stimuli on a computer screen: words and images. While it may be intuitive that words pertain to the verbal modality and images pertain to the visual, one could argue that encoding words presented on a computer screen is a visual task rather than a verbal one, or at least resides in a gray area between the two. Thus, before any strong conclusions can be made, we suggest future research that definitively separates task modalities is necessary to explore potentially implicating effects of intra-modal TUT.

4.2. Task stimuli evoke specific thought qualities

By manipulating stimulus type, we showed that task-oriented experience varies with external content. These results corroborate previous research showing that task modality shapes the phenomenological characteristics of thought, and contribute to a field providing evidence for why and how internal experience can be considered during laboratory tasks (Choi et al., 2017; Laybourn et al., 2022; Oblak et al., 2022; Oblak et al., 2024; Slana Ozimič et al., 2023; Zhou et al., 2025).

Words evoked more auditory imagery (and, in Experiment 2, more inner speech) than images. This effect can likely be attributed to the reading of the presented words, which may possess an internal auditory-like sensation. It may also arise from the perception of one's own rehearsal strategies that leverage an inner voice. Images, compared with words, evoked more task-relevant visual imagery, unguidedness, and emotions. Greater visual imagery likely reflects general engagement and some deliberate encoding strategy such as mentally holding and iterating over an image's features in mind. However, the parallel rise in unguidedness and emotion is subtler. We speculate that complex scenes can spontaneously cue richer associative processing, loosening on-task focus and amplifying affect. Consistent with this, neuroimaging shows greater emotional valence responses for pictures than words regardless of tone (Kensinger & Schacter, 2006). Thus, it is plausible that the richness of our visual stimuli drove increases in unguided and emotional inner experience. Future work may probe specific valence qualities or distinguish between dimensions of thought fluidity (e.g., unguided vs. unconstrained; Christoff et al., 2016).

The consistency of these findings across experiments underscores the reliable relationship between the content of presented stimuli and the structure of thought during encoding. It further confirms that subjective experiences are not uniform but are altered by external stimuli, and that these alterations can be tracked using ES. This paves a way to empirically track fluctuations in internal experience with greater granularity in laboratory environments (cf. Garg et al., 2025; Shelat et al., 2024).

4.3. Thought qualities are linked to depth of encoding

Task-relevant awareness and inner speech reliably predicted better recognition memory, suggesting that sustained engagement and verbal strategies enhance encoding. This interpretation aligns with Nedergaard et al. (2023), who linked inner speech to attentional control by showing that task-related inner speech was associated with faster and more consistent reaction times In their study, participants were asked to detect a black dot that appeared after long intervals, and to report on the nature of their inner experience at the time of each dot's onset. Although their research question was on attentional regulation rather than memory encoding, both of our results suggest that task-relevant inner speech enhances engagement. However, given the correlational nature of this work, the directionality of the association between inner speech and enhanced memory performance remains uncertain.

Reports of on-task awareness may reflect a general readiness to engage with the task or possibly the intentional use of a strategy. For instance, an individual might intentionally chunk together sequences of complex scenes using interactive imagery or sentence generation to better memorize them (Jaeger et al., 2024); if an image of a carnival is displayed and followed by a picture of a cockpit, the participant might implement cognitive effort to remember "flying to a carnival," rather than each individual detail of both images. The conscious implementation of this strategy may facilitate high reports of task-relevant awareness. Inner speech, by contrast, may represent a more specific encoding strategy, such as verbal rehearsal during word trials or narrative generation during image trials. While we can speculate on the mechanism of these correlations, the consistency of these effects across both experiments and stimulus modalities suggests a robust link between inner experience and memory outcomes.

Previously, Bastian et al. (2017) investigated the roles of inner speech and *meta*-awareness in attentional regulation using a sustained attention to response task (SART; Robertson et al., 1997). In the second of three experiments, they primed participants to engage verbal working memory during the task. Using experience sampling, they found that verbal priming increased awareness of mind-wandering and inferred that inner speech facilitates conscious awareness of one's thoughts. Given this, we conducted a follow-up analysis and showed that awareness and inner speech each explained unique variance in memory encoding when controlling for one another. This indicates that while these qualities are tightly intertwined in prior research, they are functionally distinct in their link with encoding.

In Experiment 1, visual imagery was associated with an enhancement in memory encoding, but this effect was not present in Experiment 2. This discrepancy may have been caused by a methodological adjustment made to the wording of the probe item. Specifically, in Experiment 2, we added the term "inner" to "visual imagery" in the probe to emphasize the internal, mental aspect of visual imagery rather than the external visual stimulus presented on the screen. The disappearance of the effect in Experiment 2 could suggest that the initial enhancement observed in Experiment 1 may have partially reflected participants interpreting the original probe as referring to the physical presence of the visual stimuli rather than their mental visualization. This underscores the need for careful consideration of probe wording when aiming to study inner experience using ES.

The negative association between unguidedness and memory observed in Experiment 2 indicates that task-relevant yet directionless thoughts may reduce encoding efficiency. When inner experience lacks the structure provided by deliberate strategies such as inner rehearsal, cognitive resources may be misallocated. One could argue that spontaneous semantic/episodic associations might *help* memory, especially for rich scenes vs. words. However, our data show the opposite: unguidedness was linked with poorer recognition (significant in Experiment 2, trend in Experiment 1). Clarifying when unguided thoughts during task-oriented states help vs. hurt memory is a valuable question for future work.

4.4. Memorability as a tool for comparing cross-modal task stimuli

Mounting evidence supports the presence of a picture superiority effect in recognition tasks, which states that objects are more memorable in image form than in word form (Stenberg et al., 1995). From this, one might assume that when compared to each other, recognition test performance would be better for images than words. To isolate the effects of attention and thought on memory encoding, we aimed to equate our words and images based on their inherent memorability. In both experiments, participants scored similarly in all performance metrics (i.e., d-prime, hit rate, false alarm rate, and criterion) and TUT rate. We show no evidence of the picture superiority effect when equating intrinsic memorability in order to isolate the effects of attention on encoding. This helps to ensure that differences in thought phenomenology and memory performance can be attributed to the cognitive processes under investigation rather than extraneous factors related to the stimuli.

4.5. Attentional state shapes trait–state visual imagery

Exploratory analyses revealed that trait vividness of imagery predicted greater visual imagery only during on-task states, not during TUT. This suggests that VVIQ shapes phenomenology, but its expression hinges on attentional state. Even individuals who generally report high vividness of mental visual imagery do not experience it consistently unless they are engaged with the task.

Importantly, our inference is not that VVIQ directly predicts behavior or task performance. Instead, we believe that the VVIQ can expose individual differences in subjective cognition (i.e., how vividly people represent their inner world). Framed this way, our results emphasize the importance of considering how trait-level tendencies and situational factors interact to shape the moment-by-moment content of thought (Smallwood & Schooler, 2006). Future research should further explore how trait–state interactions like this extend to other domains of consciousness and cognition, such as verbal rehearsal, emotional processing, or metacognitive awareness.

4.6. Limitations

A limitation of this study is the lack of formal corrections for multiple comparisons. Without such corrections, there is an increased risk of false positives, as the probability of identifying significant results by chance rises with the number of statistical tests conducted. However, this concern is partially mitigated by the preregistered analysis plan in Experiment 2, which focused on testing a subset of findings from Experiment 1 in a new sample. Although the confirmatory nature of Experiment 2 strengthens the reliability of replicated findings, all of our results should be interpreted with caution and viewed as generating hypotheses for future, more targeted studies.

As noted earlier, Experiment 1 did not collect gender data. In Experiment 2, 87 of the 98 participants (88.78 %) identified as female. This disproportionate representation raises concerns about the generalizability of our findings, particularly regarding whether the observed effects hold across genders. The existing literature on potential gender differences in mind-wandering is sparse, but one study suggests that males report higher levels of mind-wandering than females in non-ADHD populations, though these gender differences disappear in individuals with ADHD (Mowlem et al., 2019). Given our sample's imbalanced gender distribution, we are unable to test gender's role on our findings. However, we believe the results still offer meaningful insight into the human mind, though future studies must recruit more gender-diverse samples to ensure broader generalizability.

We assumed a basic level of English language proficiency given that the participant pool consisted of students at an American university; however, we did not systematically assess this. It is possible that varying levels of English fluency may have affected encoding or even memorability's contributions to word stimuli. Follow-up work should include explicit language proficiency measures when using verbal materials.

Although participants reported that they were in a quiet room and free from distractions, we have no way to verify their reports. It is possible that some participants were dishonest, creating further variability in thought phenomenology and focused attention on the task.

While we were able to manipulate the task stimuli to evoke different qualities of thought, our investigation into the relationships that thought qualities and task stimuli share with memory encoding are all correlational. Future research could build upon this study and Zhou et al. (2025)'s study to further explore how altering inner experience may provide evidence for a causal relationship between specific thought qualities and memory encoding. Further, the instructions before the task may have primed participants to attend to specific dimensions of their experience or shaped the types of thoughts they reported. It may be worth exploring whether probe-naive participants show different patterns of thought qualities.

While we designed probes to balance specificity with manageability for participants, we cannot rule out that probe demands affected memory encoding or that fatigue accumulated across the session. This reflects a common issue in experience sampling: capturing intermittent phenomenology requires interrupting ongoing cognition. Future studies should investigate how the number of probe items, probe frequency, and response demands influence both the validity of phenomenological reports and task performance (cf. Wiemers & Redick, 2019; Schubert et al., 2020).

Other limitations include our limited diversity of stimulus types and aspects of experience. We intentionally chose previously researched stimuli and qualities of thought so as to build upon a preset foundation. Unfortunately, this resulted in our overlapping of verbal and visual modalities for our word stimuli, which was discussed in Section 4.1.2. Subsequent work may consider using an auditory task, similar to Choi et al. (2017), or different types of stimuli and qualities of thought to further consider diverse cognitive experiences.

While our study primarily focused on state-level fluctuations in thought phenomenology, personality traits may also shape experience. Here, we conducted trait-level exploratory analysis using the VVIQ but did not investigate how other traits such as conscientiousness could similarly interact with mind-wandering during encoding tasks. Future research should examine how personality dimensions, such as the Big Five, influence memory performance.

4.7. Conclusion

This study demonstrates that experience sampling reveals relationships between thought phenomenology and memory encoding that behavioral measures alone cannot capture. Across both experiments, we found that task-unrelated thoughts consistently impaired later recognition, different stimuli evoked distinct patterns of thought, and certain qualities of task-relevant thought (i.e., awareness and inner speech) predicted better memory regardless of stimulus type. Together, these findings suggest that external environments evoke specific conscious experiences and these experiences shape how information is encoded and retained.

Understanding thought phenomenology during memory encoding has potential applications in educational settings, where teaching students to monitor their attentional state may enhance learning. In clinical populations with attentional impairments (e.g., ADHD), training *meta*-awareness of mind-wandering could help mitigate encoding failures, though our inconsistent replication of this effect warrants caution. These applications remain speculative given our controlled laboratory paradigm and constrained sample characteristics.

Looking forward, researchers should examine whether interventions that manipulate specific thought qualities causally improve encoding, and investigate how these patterns generalize across diverse populations and naturalistic contexts. By integrating phenomenological measures into cognitive paradigms, this work demonstrates that a complete understanding of memory formation requires examining the conscious experiences that accompany encoding.

CRediT authorship contribution statement

Brecken Marome: Writing – review & editing, Writing – original draft, Visualization, Software, Project administration, Methodology, Formal analysis, Conceptualization. **Shivang Shelat:** Writing – review & editing, Visualization, Methodology, Formal analysis, Conceptualization. **Jonathan W. Schooler:** Writing – review & editing, Supervision.

Funding

SS is supported by a National Science Foundation Graduate Research Fellowship under grant 2139319.

Data Availability: All data and materials will be made available in an open-source repository.

Declaration of competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.concog.2025.103958.

Data availability

All data will be made available in an OSF repository.

References

Ardelt, M. (2003). Empirical assessment of a three-dimensional wisdom scale. *Research on Aging, 25*(3), 275–324. https://doi.org/10.1177/0164027503025003004

Bainbridge, W. A., Hall, E. H., & Baker, C. I. (2019). Drawings of real-world scenes during free recall reveal detailed object and spatial information in memory. *Nature Communications, 10*(1), 5. https://doi.org/10.1038/s41467-018-07830-6

Bainbridge, W. A., Pounder, Z., Eardley, A. F., & Baker, C. I. (2021). Quantifying aphantasia through drawing: Those without visual imagery show deficits in object but not spatial memory. *Cortex, 135*, 159–172. https://doi.org/10.1016/j.cortex.2020.11.014

Bastian, M., Lerique, S., Adam, V., Franklin, M. S., Schooler, J. W., & Sackur, J. (2017). Language facilitates introspection: Verbal mind-wandering has privileged access to consciousness. *Consciousness and Cognition, 49*, 86–97. https://doi.org/10.1016/j.concog.2017.01.002

Bayne, T., & Montague, M. (2011). *Cognitive Phenomenology*. OUP Oxford.

Berkovich-Ohana, A., Dor-Ziderman, Y., Trautwein, F.-M., Schweitzer, Y., Nave, O., Fulder, S., & Ataria, Y. (2020). The Hitchhiker's Guide to Neurophenomenology – the Case of Studying self Boundaries with Meditators. *Frontiers in Psychology, 11*. https://doi.org/10.3389/fpsyg.2020.01680

Blondé, P., Girardeau, J.-C., Sperduti, M., & Piolino, P. (2022). A wandering mind is a forgetful mind: A systematic review on the influence of mind wandering on episodic memory encoding. *Neuroscience & Biobehavioral Reviews, 132*, 774–792. https://doi.org/10.1016/j.neubiorev.2021.11.015

Bylinskii, Z., Isola, P., Bainbridge, C., Torralba, A., & Oliva, A. (2015). Intrinsic and extrinsic effects on image memorability. *Vision Research, 116*, 165–178. https://doi.org/10.1016/j.visres.2015.03.005

Černe, J., & Kordeš, U. (2023). Deconstructing Accurate and Inaccurate recall in the DRM Paradigm: A Phenomenological and Behavioral Exploration. *Constructivist Foundations, 19*(1), 38–59.

Choi, H., Geden, M., & Feng, J. (2017). More visual mind wandering occurrence during visual task performance: Modality of the concurrent task affects how the mind wanders. *PLoS One1, 12*(12), Article e0189667. https://doi.org/10.1371/journal.pone.0189667

Christoff, K., Irving, Z., Fox, K., Spreng, R. N., & Andrews-Hanna, J. R. (2016). Mind-wandering as spontaneous thought: A dynamic framework. *Nature Reviews. Neuroscience, 17*, 718–731. https://doi.org/10.1038/nrn.2016.113

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods, 47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

deBettencourt, M. T., Norman, K. A., & Turk-Browne, N. B. (2018). Forgetting from lapses of sustained attention. *Psychonomic Bulletin & Review, 25*(2), 605–611. https://doi.org/10.3758/s13423-017-1309-5

Fernyhough, C., Alderson-Day, B., Hurlburt, R. T., & Kühn, S. (2018). Investigating Multiple Streams of Consciousness: Using Descriptive Experience Sampling to Explore Internally and Externally Directed Streams of Thought. *Frontiers in Human Neuroscience, 12*. https://doi.org/10.3389/fnhum.2018.00494

Froese, T., Suzuki, K., Ogai, Y., & Ikegami, T. (2012). Using Human–Computer Interfaces to Investigate 'Mind-As-It-Could-Be' from the First-Person Perspective. *Cognitive Computation, 4*(3), 365–382. https://doi.org/10.1007/s12559-012-9153-4

Gallagher, S. (2003). Phenomenology and Experimental Design Toward a Phenomenologically Enlightened Experimental Science. *Journal of Consciousness Studies, 10*(9–10), 85–99.

Garg, A., Shelat, S., Gross, M. E., Smallwood, J., Seli, P., Taxali, A., Sripada, C. S., & Schooler, J. W. (2025). Opening the black box: Think Aloud as a method to study the spontaneous stream of consciousness. *Consciousness and Cognition, 128*, Article 103815. https://doi.org/10.1016/j.concog.2025.103815

Gregory, M. D., Robertson, E. M., Manoach, D. S., & Stickgold, R. (2016). Thinking about a Task is Associated with increased Connectivity in Regions Activated by Task Performance. *Brain Connectivity, 6*(2), 164–168. https://doi.org/10.1089/brain.2015.0386

Gross, M. E., Smith, A. P., Graveline, Y. M., Beaty, R. E., Schooler, J. W., & Seli, P. (2020). Comparing the phenomenological qualities of stimulus-independent thought, stimulus-dependent thought and dreams using experience sampling. *Philosophical Transactions of the Royal Society B: Biological Sciences, 376*(1817), Article 20190694. https://doi.org/10.1098/rstb.2019.0694

Heavey, C. L., & Hurlburt, R. T. (2008). The phenomena of inner experience. *Consciousness and Cognition, 17*(3), 798–810. https://doi.org/10.1016/j.concog.2007.12.006

Hurlburt, R. T. (2011). *Investigating Pristine inner Experience: Moments of Truth*. Cambridge University Press.

Hurlburt, R. T., Alderson-Day, B., Kühn, S., & Fernyhough, C. (2016). Exploring the Ecological Validity of Thinking on demand: Neural Correlates of Elicited vs. *Spontaneously Occurring Inner Speech. PLOS ONE, 11*(2), Article e0147932. https://doi.org/10.1371/journal.pone.0147932

Irving, Z. C. (2016). Mind-wandering is unguided attention: Accounting for the "purposeful" wanderer. *Philosophical Studies, 173*(2), 547–571. https://doi.org/10.1007/s11098-015-0506-1

Isola, P., Xiao, J., Torralba, A., & Oliva, A. (2011). What makes an image memorable? *CVPR, 2011*, 145–152. https://doi.org/10.1109/CVPR.2011.5995721

Jaeger, A., Martins, T. H. P. G., Rodrigues, J. P. P., Muniz, B. F. B., & da Silveira Fonseca, Ana Luísa Santiago, & de Oliveira Gonçalves, A.. (2024). The benefits of elaborative encoding over retrieval practice for associative learning. *Memory & Cognition*. https://doi.org/10.3758/s13421-024-01671-z

Kam, J. W. Y., Irving, Z. C., Mills, C., Patel, S., Gopnik, A., & Knight, R. T. (2021). Distinct electrophysiological signatures of task-unrelated and dynamic thoughts. *Proceedings of the National Academy of Sciences, 118*(4), Article e2011796118. https://doi.org/10.1073/pnas.2011796118

Kane, M. J., Smeekens, B. A., Meier, M. E., Welhaf, M. S., & Phillips, N. E. (2021). Testing the construct validity of competing measurement approaches to probed mind-wandering reports. *Behavior Research Methods, 53*(6), 2372–2411. https://doi.org/10.3758/s13428-021-01557-x

Kensinger, E. A., & Schacter, D. L. (2006). Processing emotional pictures and words: Effects of valence and arousal. *Cognitive, Affective & Behavioral Neuroscience, 6*(2), 110–126. https://doi.org/10.3758/CABN.6.2.110

Larson, R., & Csikszentmihalyi, M. (1983). The Experience Sampling Method. *New Directions for Methodology of Social & Behavioral Science, 15*, 41–56.

Laybourn, S., Frenzel, A. C., Constant, M., & Liesefeld, H. R. (2022). Unintended emotions in the laboratory: Emotions incidentally induced by a standard visual working memory task relate to task performance. *Journal of Experimental Psychology: General, 151*(7), 1591–1605. https://doi.org/10.1037/xge0001147

Lohnas, L. J., & Kahana, M. J. (2013). Parametric effects of word frequency in memory for mixed frequency lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(6), 1943–1946. https://doi.org/10.1037/a0033669

Long, N. M., Danoff, M. S., & Kahana, M. J. (2015). Recall dynamics reveal the retrieval of emotional context. *Psychonomic Bulletin & Review, 22*(5), 1328–1333. https://doi.org/10.3758/s13423-014-0791-2

Lüdecke, D. (2018). ggeffects: Tidy Data Frames of Marginal Effects from Regression Models. *Journal of Open Source Software, 3*(26), 772. https://doi.org/10.21105/joss.00772

Luthra, M., & Todd, P. M. (2019). Role of Working memory on strategy use in the Probability Learning Task. *Proceedings of the Annual Meeting of the Cognitive Science Society*.

Lutz, A. (2002). Toward a neurophenomenology as an account of generative passages: A first empirical case study. *Phenomenology and the Cognitive Sciences, 1*(2), 133–167. https://doi.org/10.1023/A:1020320221083

Macmillan, N. A., & Kaplan, H. L. (1985). Detection theory analysis of group data: Estimating sensitivity from average hit and false-alarm rates. *Psychological Bulletin, 98*(1), 185–199. https://doi.org/10.1037/0033-2909.98.1.185

Madan, C. R. (2021). Exploring word memorability: How well do different word properties explain item free-recall probability? *Psychonomic Bulletin & Review, 28*(2), 583–595. https://doi.org/10.3758/s13423-020-01820-w

Marcusson-Clavertz, D., Persson, S. D., Davidson, P., Kim, J., Cardeña, E., & Kuehner, C. (2023). Mind wandering and sleep in daily life: A combined actigraphy and experience sampling study. *Consciousness and Cognition, 107*, Article 103447. https://doi.org/10.1016/j.concog.2022.103447

Marks, D. F. (1973). Visual Imagery differences in the recall of Pictures. *British Journal of Psychology, 64*(1), 17–24. https://doi.org/10.1111/j.2044-8295.1973.tb01322.x

McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: Working memory capacity, goal neglect, and mind wandering in an executive-control task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(1), 196–204. https://doi.org/10.1037/a0014104

Miller, M. B., Donovan, C.-L., Bennett, C. M., Aminoff, E. M., & Mayer, R. E. (2012). Individual differences in cognitive style and strategy predict similarities in the patterns of brain activity between individuals. *NeuroImage, 59*(1), 83–93. https://doi.org/10.1016/j.neuroimage.2011.05.060

Mowlem, F. D., Agnew-Blais, J., Pingault, J.-B., & Asherson, P. (2019). Evaluating a scale of excessive mind wandering among males and females with and without attention-deficit/hyperactivity disorder from a population sample. *Scientific Reports, 9*(1), 3071. https://doi.org/10.1038/s41598-019-39227-w

Nedergaard, J. S. K., & Lupyan, G. (2024). Not everybody has an Inner Voice: Behavioral Consequences of Anendophasia. *Psychological Science, 35*(7), 780–797. https://doi.org/10.1177/09567976241243004

Nedergaard, J., Skewes, J. C., & Wallentin, M. (2023). "Stay focused!": The role of inner speech in maintaining attention during a boring task. *Journal of Experimental Psychology: Human Perception and Performance, 49*(4), 451–464. https://doi.org/10.1037/xhp0001112

Niikawa, T., Miyahara, K., Hamada, H. T., & Nishida, S. (2020). A new experimental phenomenological method to explore the subjective features of psychological phenomena: Its application to binocular rivalry. *Neuroscience of Consciousness, 2020*(1), Article niaa018. https://doi.org/10.1093/nc/niaa018

Nishida, S., Hamada, H. T., Niikawa, T., & Miyahara, K. (2024). *Neural correlates of phenomenological attitude toward perceptual experience* (p. 2024.07.07.602347). bioRxiv. Doi: 10.1101/2024.07.07.602347.

Oblak, A., Dragan, O., Slana Ozimič, A., Kordeš, U., Purg, N., Bon, J., & Repovš, G. (2024). What is it like to do a visuo-spatial working memory task: A qualitative phenomenological study of the visual span task. *Consciousness and Cognition, 118*, Article 103628. https://doi.org/10.1016/j.concog.2023.103628

Oblak, A., Slana Ozimič, A., Repovš, G., & Kordeš, U. (2022). What individuals Experience during Visuo-Spatial Working memory Task Performance: An Exploratory Phenomenological Study. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.811712

Ortega, J. R., Gross, M. E., & Schooler, J. W. (2025). When life is but a dream: Does transliminality predict continuity of thought across the sleep-wake cycle? *Philosophy and the Mind Sciences, 6*. https://doi.org/10.33735/phimisci.2025.10272

R Core Team. (2025). *R: The R Project for Statistical Computing*. Retrieved May 4, 2025, from https://www.r-project.org/.

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). Oops!': Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia, 35*(6), 747–758. https://doi.org/10.1016/S0028-3932(97)00015-8

Rowlands, M. (2010). *The new science of the mind: From extended mind to embodied phenomenology*. In *The new science of the mind: From extended mind to embodied phenomenology (p. Chapter x, 249 Pages)*. MIT Press. https://doi.org/10.7551/mitpress/9780262014557.001.0001.

Schubert, A.-L., Frischkorn, G. T., & Rummel, J. (2020). The validity of the online thought-probing procedure of mind wandering is not threatened by variations of probe rate and probe framing. *Psychological Research, 84*(7), 1846–1856. https://doi.org/10.1007/s00426-019-01194-2

Seli, P., Carriere, J. S. A., Levene, M., & Smilek, D. (2013). How few and far between? Examining the effects of probe rate on self-reported mind wandering. *Frontiers in Psychology, 4*. https://doi.org/10.3389/fpsyg.2013.00430

Shelat, S., Schooler, J. W., & Giesbrecht, B. (2024). Predicting attentional lapses using response time speed in continuous performance tasks. *Frontiers in Cognition, 3*. https://doi.org/10.3389/fcogn.2024.1460349

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological Momentary Assessment. *Annual Review of Clinical Psychology, 4*. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Slana Ozimič, A., Oblak, A., Kordeš, U., Purg, N., Bon, J., & Repovš, G. (2023). The Diversity of strategies used in Working memory for Colors, Orientations, and Positions: A Quantitative Approach to a First-Person Inquiry. *Cognitive Science, 47*(8), Article e13333. https://doi.org/10.1111/cogs.13333

Smallwood, J., Karapanagiotidis, T., Ruby, F., Medea, B., de Caso, I., Konishi, M., Wang, H.-T., Hallam, G., Margulies, D. S., & Jefferies, E. (2016). Representing Representation: Integration between the Temporal Lobe and the Posterior Cingulate Influences the Content and form of Spontaneous Thought. *PLOS ONE, 11*(4), Article e0152272. https://doi.org/10.1371/journal.pone.0152272

Smallwood, J., McSpadden, M., & Schooler, J. W. (2007). The lights are on but no one's home: Meta-awareness and the decoupling of attention when the mind wanders. *Psychonomic Bulletin & Review, 14*(3), 527–533. https://doi.org/10.3758/BF03194102

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

Stenberg, G., Radeborg, K., & Hedman, L. R. (1995). The picture superiority effect in a cross-modality recognition task. *Memory & Cognition, 23*(4), 425–441. https://doi.org/10.3758/BF03197244

Tyng, C. M., Amin, H. U., Saad, M. N. M., & Malik, A. S. (2017). The Influences of Emotion on Learning and memory. *Frontiers in Psychology, 8*. https://doi.org/10.3389/fpsyg.2017.01454

van Vugt, M. K., & Broers, N. (2016). Self-Reported Stickiness of Mind-Wandering Affects Task Performance. *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.00732

Varela, F. J. (1996). Neurophenomenology: A methodological remedy for the hard problem. *Journal of Consciousness Studies, 3*(4), 330–349.

Vidal, J., Alzueta, E., Michelmann, S., & Yebra, M. (2025). *Online behavioral experiments for researchers*. Cognition. Retrieved May 7, 2025, from https://www.cognition.run.

Villena-González, M., López, V., & Rodríguez, E. (2016). Orienting attention to visual or verbal/auditory imagery differentially impairs the processing of visual stimuli. *NeuroImage, 132*, 71–78. https://doi.org/10.1016/j.neuroimage.2016.02.013

Wakeland-Hart, C. D., Cao, S. A., deBettencourt, M. T., Bainbridge, W. A., & Rosenberg, M. D. (2022). Predicting visual memory across images and within individuals. *Cognition, 227*, Article 105201. https://doi.org/10.1016/j.cognition.2022.105201

Wiemers, E. A., & Redick, T. S. (2019). The influence of thought probes on performance: Does the mind wander more if you ask it? *Psychonomic Bulletin & Review, 26*(1), 367–373. https://doi.org/10.3758/s13423-018-1529-3

Zhou, S. S., Rowchan, K., Mckeown, B., Smallwood, J., & Wammes, J. D. (2025). Drawing behaviour influences ongoing thought patterns and subsequent memory. *Consciousness and Cognition, 127*, Article 103791. https://doi.org/10.1016/j.concog.2024.103791