Investigating the Pupil Response as a Physiological Correlate of Encoding and Retrieval

Name: Gloria Charite

Candidate Number: 1078303

Supervisor: Dr. Bernhard Staresina

Date: June 2025

Thesis submitted to the University of Oxford in fulfilment of the requirements for the degree of M.Sc. by Research in Experimental Psychology

Abstract

The question of how we encode and retrieve previously learned information has been explored for millennia. However, many unanswered questions remain. For instance, in daily life these two processes are constantly interleaved, and yet research has shown that they require opposing neural mechanisms. Therefore, what facilitates this ability to rapidly switch back and forth between encoding and retrieval? An influential model by Hasselmo (1995) suggests that acetylcholine is a key candidate for modulating this process. This theory has been extensively studied in animal and computational models but work using human participants has been scarce. The first goal of this thesis project was to address this gap.

To do so, we employed pupillometry in a continuous associative memory paradigm with alternating encoding and retrieval tasks. Changes in the pupil size are indirect markers of neuromodulatory activity especially that from the locus coeruleus - norepinephrine (LC- NE) and acetylcholine. Additionally, pupillometry has increasingly become touted as a reliable tool for measuring cognitive processing. However, there is debate as to what exactly pupil dilation means in the context of recognition memory. Does it reflect memory strength or the cognitive effort exerted to retrieve relevant traces? Furthermore, there is also debate as to whether pupillary changes do in fact index cognitive processes or if they simply capture mere task engagement. The second and third goal of the current study was to address these debates, respectively.

So, what did we find? First, even when task requirements were matched, encoding and retrieval exhibited distinct pupil patterns. This provided evidence for the differential impact of acetylcholine and norepinephrine on these processes which is consistent with Hasselmo’s predictions. Second, the pupil time course for retrieval was comprised of two phases: an early and later component whose characteristics could be best explained by the memory strength and retrieval effort account, respectively. Thus, allowing for a reconciliation of the mixed findings in the literature. Lastly, the pupil pattern replicated three of the most consistent findings in memory studies, namely, the old/new effect, the encoding specificity principle, and the observation that memories decay as time passes. The fact that we could replicate these well-established memory effects suggests that the pupil response is indeed sensitive to cognitive processing.

Acknowledgments

I would like to extend heartfelt gratitude to my supervisor, Bernhard Staresina, for his invaluable guidance since the inception of the project. I would also like to the thank all the members of the Staresina Lab, especially Marcel Kehl, who provided encouragement and mentorship essential for accomplishing this work. Finally, my sincere gratitude to my family, partner, David R. Salmon, and best friend, Kathia Masimbi Natete, for their unending support.

Table of Contents

1. Introduction

> “Let us, then, make in our minds a sort of waxen block... Let us say that this is a gift of Memory... and that whenever we wish to remember something we see or hear or think, we press the perception into this wax, as we might a signet ring into sealing wax, and that we remember and know what has been imprinted as long as the image remains in the wax, but whatever is rubbed out or cannot be imprinted we forget and do not know.”
>
> Plato’s Theaetetus (191c–d)

The question of how we form memories (encoding) and access previously learned information (retrieval) has been central to philosophical and scientific discourse for a long time. Indeed, debates on memory mechanisms go back as far as 2000 years ago, even predating psychology as a formal science. However, it was not until the 19th century that systematic and empirical approaches to studying memory were introduced. In the 1870s, Hermann Ebbinghaus conducted the first memory experiments on record. Using himself as a subject, Ebbinghaus meticulously memorised a long list of three-letter nonsense syllables and measured how long it took him to forget and relearn them. This work resulted in the well-known forgetting curve which provided empirical evidence that memories fade over time unless there is an effort to retain them such as through spaced repetitions (Ebbinghaus, 1885).

Beyond these important theoretical and empirical contributions, Ebbinghaus' work demonstrated that memory followed predictable and measurable patterns and thus could be studied experimentally. This would prove critical in propelling the field forward especially after the push in the 1950s to move beyond observable behaviour (behaviourism) and incorporate the study of internal mental processes (cognition) (Miller, 2003). Consequently, the scientific investigation of memory became widespread with some of the most influential findings in the field emerging in the few decades that followed. For instance, Miller (1956) defined the capacity of short-term memory as seven, plus or minus 2, Scoville and Milner (1957) studied patient HM, and Endel Tulving (1972) distinguished episodic from semantic memories which emphasised that memory is not a monolithic concept. Since then, there have been a plethora of models aimed at mapping out the functioning of the memory system and how information is processed in the brain (Lutz & Huitt, 2003; Squire, 2004).

Coincidentally, the rise of computers occurred around the same time, and this had an outsized influence on the memory field. For instance, inspired by how computers convert input into storable formats (encoding), save and combine information from multiple sources into one destination (consolidation), and access stored information (retrieval), researchers like Atkinson and Shiffrin (1968) proposed that memory was a discontinuous, multi-staged process. Specifically, borrowing from the computer science vocabulary, memory has since been described as comprising three distinct stages namely: encoding, consolidation, and retrieval (Straube, 2012; Guskjolen & Cembrowski, 2023).

As a result, encoding and retrieval have typically been studied as separate, independent processes (Dolan & Fletcher, 1999; Rugg et al., 2015). However, unlike in most experimental paradigms, in everyday life these processes are constantly interleaved. Imagine, for example, that you are meeting an old friend in a new cafe. In this scene, like many we encounter in real life, you have to switch rapidly between encoding and retrieval states. In other words, you are in an encoding state as you process the cafe's features (e.g., the location of the counter) and almost at the same time in a retrieval state as you recall what your friend looks like or the colour of the jacket that they mentioned they would be wearing.

1.1. Research Question

The main goal of this project is to examine the neurobiological mechanisms that enable us to flexibly switch back and forth between encoding and retrieval. This is especially puzzling because these two processes have been shown to require opposing mechanisms (O'Reilly & McClelland, 1994). That is, successful retrieval involves re-accessing traces of past events, but doing so at the time of encoding may result in proactive interference (Wixted & Rohrer, 1993).

Animal and computational models suggest that the ability to flexibly interleave encoding and retrieval may be regulated by acetylcholine. Hasselmo (1995) proposed a hippocampal model of episodic memory function and the role of cholinergic modulation therein (Figure 1). According to this model, henceforth referred to as Hasselmo's model, acetylcholine plays a critical role in setting appropriate dynamics for processing input from the entorhinal cortex (EC). The EC is the relay station for information from the rest of the brain into the CA1 and CA3, subregions of the hippocampus critical for memory processes. It receives the input to be encoded and the cues to be used during retrieval (Jones, 1993). The CA1 is the main output centre from the hippocampus back to the EC. It is sparse as opposed to CA3 which possess strong recurrent connections that provide a mechanism for associative memory (Treves and Rolls, 1992; Daumas et al., 2022).

There are two different pathways of how information gets into CA1 (Kesner & Rolls, 2014). The first being the perforant pathway with input coming directly from the EC (i.e., EC→CA1) and the second being the Schaffer collateral pathway where input comes from CA3 (i.e., CA3→CA1). When the perforant pathway is activated, sensory input from the EC is enhanced thus creating ideal conditions for encoding (Hasselmo & Bower, 1993). On the other hand, when the Schaffer collateral pathway is activated, it allows the internal hippocampal circuitry to engage in pattern completion which, in turn, supports retrieval (Staresina et al., 2016).

Acetylcholine plays an important role in determining which pathway is activated. Hasselmo & McGaughy (2004) proposed that it does so through its seemingly paradoxical effects. Acetylcholine, for example, increases the cell spiking of pyramidal neurons through depolarisation (Krnjević et al., 1971; Cole & Nicoll, 1984) while at the same time suppressing excitatory synaptic transmission at recurrent synapses in the hippocampus (Hasselmo & Bower, 1993; Hasselmo & Schnell, 1994). In a similar manner, it depolarises inhibitory interneurons in the hippocampus while simultaneously suppressing GABA release at their terminal (Patil and Hasselmo, 1999). Together, these effects work to reduce interference from internally generated activity and enhance responsiveness to external input.

There has been much evidence in support of this model's predictions in the animal literature. For example, De Rosa & Hasselmo (2000) used an odour learning task in rats to study the impact of blocking acetylcholine muscarinic receptors with scopolamine. The task involved learning a baseline odour pair (e.g., A+B). Following, the animals were given either scopolamine, normal saline, or methylscopolamine which served as a control for the effects of scopolamine in the peripheral nervous system as it does not pass through the blood-brain barrier (Cosquer et al., 2005). They then had to learn a novel odour pair with an overlapping component with the baseline pair (e.g., A+C) or an entirely new pair with no overlapping component (e.g., D+E). The results revealed that rats that were under the influence of scopolamine found it more difficult to learn overlapping than nonoverlapping odour pairs compared to the other groups. Therefore, supporting the hypothesis that acetylcholine suppresses interference from existing memories and facilitates flexible learning of new associations.

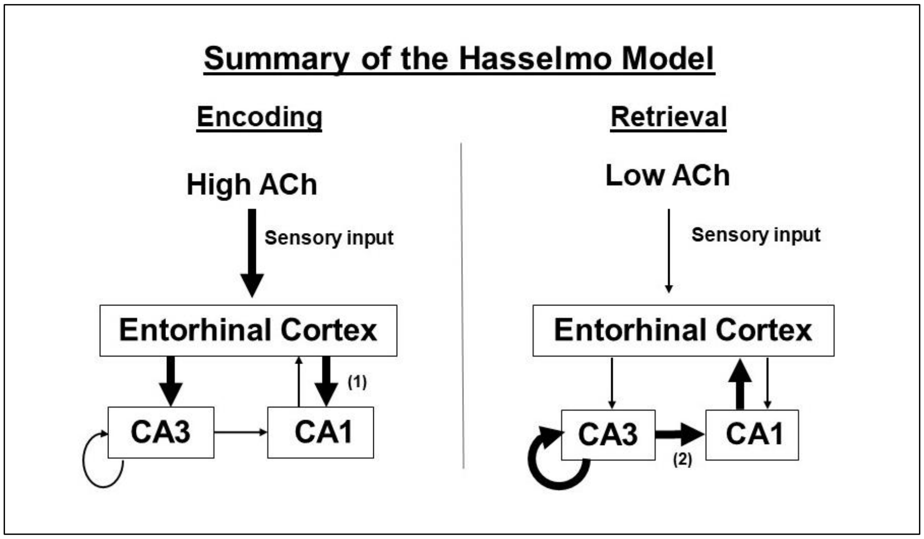

Figure 1. *Summary of the Hasselmo model adapted from figure 2 in Hasselmo & McGaughy, 2004 and figure 1 in Hasselmo et al., 2002. The left panel shows the mechanism that facilitate encoding. High levels of acetylcholine (ACh), activate the perforant pathway (1) allowing more external input from the EC to reach the hippocampus. The right panel depicts the mechanism that facilitate retrieval. Low levels of acetylcholine release the suppression of recurrent activity in CA3, thus activating the Schaffer collateral pathway (2).*

Similar findings were obtained in non-human primates by Aigner & Mishkin (1986) in a visual recognition task. In this task, rhesus monkeys were shown different objects and were later tested on their ability to distinguish between old and new stimuli. Also using scopolamine, they found that blocking acetylcholine receptors impaired the ability to learn new objects but had no effect on recognising objects that had been learned before the administration of scopolamine. In a follow-up study, Tang & Aigner (1996) did a microdialysis in which they analysed the extracellular levels of acetylcholine during the above task. They reported a 41% and 24% increase in acetylcholine in the perirhinal cortex and the dentate gyrus of the hippocampus relative to baseline levels.

More recently, using optogenetic stimulation of cholinergic fibres in rat hippocampal slices, Palacios-Filardo et al. (2021) showed that indeed acetylcholine biases hippocampal dynamics toward encoding by enhancing sensory input from the entorhinal cortex and suppressing recurrent input from CA3. Specifically, they found that acetylcholine equally depresses excitatory input from both the EC and CA3. What was different is that the Schaffer collateral (CA3→CA1) pathway was more inhibited than the perforant (EC→CA1) pathway. The net result being more room for input from the EC to CA1, thus providing enhanced conditions for successful encoding.

In humans, some studies have replicated the modulatory effects of acetylcholine on memory using drugs that block muscarinic acetylcholine receptors (e.g., Ghoneim and Mewaldt, 1975, 1977; Petersen, 1977). Besides these studies, however, investigations of Hasselmo model's predictions in humans have been scant. This is because of the many methodological constraints of probing the human brain and the ethical conundrums of utilising drugs. Nevertheless, methods like pupillometry, tracking changes in the pupil size over time, provide a promising avenue to further investigate the effects of neuromodulators like acetylcholine and noradrenaline on memory processes. Changes in pupil size correlate with their activity and have consequently been used as indirect measures of noradrenergic and cholinergic influence (Larsen & Waters, 2018). In this thesis project, we used pupillometry to test Hasselmo's proposal that neuromodulators like acetylcholine have differential effects on encoding and retrieval.

1.2. Literature Review

1.2.1. Pupillary changes as an indirect marker of neuromodulatory activity

The pupil size is controlled by two muscles, namely the iris sphincter and the iris dilator which promote constriction and dilation, respectively (Loewenfeld, 1958; May et al., 2019). The iris sphincter is regulated by the parasympathetic system via acetylcholine released by cholinergic motoneurons in the Edinger–Westphal nucleus (EWN). That is, acetylcholine binds to muscarinic receptors on the iris sphincter muscle, causing it to contract and thereby constricting the pupil (Viglione et al., 2023). Conversely, when the cholinergic neurons in the EWN are inhibited, the iris sphincter muscle relaxes, resulting in pupil dilation. For example, Naicker et al. (2016) tested the effects of blocking acetylcholine muscarinic receptors on the pupil diameter by using anticholinergic medications. They found an increase in the pupil diameter after ingesting hyoscine hydrobromide, a moderate centrally acting anticholinergic medication, compared to the placebo. This provided causal evidence that acetylcholine helps to control pupil constriction.

It is important to highlight that while acetylcholine plays a key role, other neuromodulators like norepinephrine also make major contributions in regulating the pupil size. To be specific, whereas acetylcholine primarily controls pupil constriction, norepinephrine regulates pupil dilation via the sympathetic system (Larsen & Waters, 2018; Viglione et al., 2023). This sympathetic dilation pathway originates from the locus coeruleus (LC) which is the main source of norepinephrine in the brain (Poe et al., 2020). The exact mechanism is yet to be elucidated, but the LC has broad projections throughout the brain which allows for the released norepinephrine to act on the iris dilator muscle and in turn cause the pupil to dilate. Oppositely, inhibition of neurons in the LC leads to pupil constriction (Hayat et al., 2020).

There has been substantial evidence in both animal and human studies linking activity in the LC with changes in pupil size (Larsen & Waters, 2018). In rodents, using simultaneous pupillometry and photon imaging to monitor the noradrenergic axons in the LC, Reimer et al. (2016), found that brief bursts in their activity were strongly correlated with rapid increases in pupil diameter even without locomotion. In non-human primates, Joshi et al. (2016) conducted a study in which they combined pupillometry and extracellular recordings in rhesus monkeys, finding that activation of LC units reliably preceded increases in pupil size. In humans, simultaneous fMRI and pupillometry have revealed that blood-oxygen level dependent (BOLD) activity localised to the LC increases in concert with pupil diameter at rest (Murphy et al., 2014).

Moreover, the relationship between neuromodulatory activity and pupillometry exists beyond these natural fluctuations that occur during resting conditions or as a consequence of changes in lighting conditions (Grujic et al., 2024). As reviewed, pupil size is directly regulated by the autonomic nervous system, with norepinephrine promoting dilation via the sympathetic system and acetylcholine promoting constriction via the parasympathetic system. However, during cognitive tasks, cholinergic and noradrenergic activity in the central nervous system also influence pupil dynamics (Grujic et al., 2024). Therein, these systems exhibit two modes of activity, namely phasic and tonic. Phasic activation refers to short and quick bursts of neuronal firing in response to a specific event. Whereas tonic activation is characterised by steady firing at a relatively constant frequency and reflects the general brain state (e.g., arousal level) during a given task (Aston-Jones & Cohen, 2005). Notably, the phasic activity is influenced by tonic firing, in that, they exhibit an inverted-U shape relationship whereby the former is optimal at intermediate levels of the latter (Hasselmo & McGaughy, 2004; Aston-Jones & Cohen, 2005).

The pupil is sensitive to these distinct aspects of the neuromodulatory system in cognitive tasks. For instance, Murphy et al. (2014), employed an oddball task, a classic attentional paradigm in which participants had to identify unusual stimuli in a sequence of standard frequent stimuli. They found that rapid, phasic pupil dilations closely monitored transient bursts of norepinephrine activity that arose around the onset of the unusual targets. Additionally, these targets evoked stronger pupil dilations compared to the standard stimuli. This is consistent with the well documented role of the phasic LC-NE activity in orienting attention to behaviourally relevant events (Aston-Jones & Cohen, 2005; Sara, 2009). Complementing these findings, Reimer et al. (2016) found that ongoing task engagement was associated with tonic cholinergic activity and evoked sustained pupil dilations. As a whole, these findings suggest that pupillometry is a reliable index for the neuromodulation of cognitive tasks, reflecting the combined contributions of the cholinergic and noradrenergic systems.

1.2.2. Pupillometry and memory processes

Pupillometry is well suited for investigating memory processes. It has a high temporal resolution (i.e., at the millisecond level) given the quick reactivity of pupils in responding to changes in attentional demands and mental effort (Mathôt, 2018). Thus, providing a continuous measure of these two variables that are crucial for memory encoding and retrieval. Both memory functions are covert, however, they have been theorised to involve processes that are modulated by acetylcholine and norepinephrine (Sarter et al., 2004). Therefore, pupillometry provides a framework for these processes to be monitored experimentally by tracking moment to moment changes.

Encoding

Encoding involves bottom-up and top-down activity that facilitate sensory processing. It typically begins with the registration of the target stimuli by the sensory organs, from which the information is transferred to the thalamus and then sent to the primary sensory cortices for processing (Atkinson & Shiffrin, 1968). At this sensory registration stage, there is an increase in the activity of the basal forebrain corticopetal (oriented towards the cortex) cholinergic projections. This enhances thalamocortical inputs but also suppresses intracortical associative activity (Sarter et al., 2004). Thus, resulting in an improved signal to noise ratio, which in turn, creates optimal conditions for sensory processing. Following the established relationship between pupillometry and neuromodulatory activity, we hypothesise that there will be a corresponding increase in the pupil size after the stimulus onset.

Sensory registration is also mediated by the noradrenergic system. Novel, salient and affective stimuli trigger autonomic body responses that are detected by the regions in the brainstem responsible for monitoring the body’s internal state (Poe et al., 2020). These include the LC which releases norepinephrine, thus activating the widespread ascending noradrenergic projections in the cortex (Sarter et al., 2004). This burst in norepinephrine leads to heightened cortical alertness which primes the attentional systems to respond effectively.

Importantly, the noradrenergic system has projections to the cholinergic system and their activity often correlate (Collins et al., 2023). This provides a mechanism for information on the saliency of the stimuli to be passed on to the cholinergic system which plays a significant role in encoding. Several studies have shown that unexpected stimuli, uncertainty and surprise lead to larger pupil dilations, thus, cementing the relationship between pupillary changes and neuromodulatory activity (Preuschoff et al., 2011; Nassar et al., 2012; Kloosterman et al., 2015). Accordingly, we hypothesise that stimuli deemed as less likely to occur in life or nature (i.e., low plausible stimuli) will result in more pupil dilation than those identified as more likely to occur (i.e., high plausible stimuli).

Other factors that influence sensory processing are top-down attentional processes that are informed by prior knowledge, current rules, or expectations (Desimone & Duncan, 1995; Sarter et al., 2001; Egner & Hirsch, 2005). Purely signal-driven sensory processing is rare as even attribution of novelty is informed by what we already know. The prefrontal cortex (PFC) plays a key role in orchestrating how existing information shape sensory processing and the interpretation of the input (Nelson et al., 2005). The PFC has direct projections to posterior sensory areas, which provides a pathway for input from higher-order regions to guide attention to relevant stimuli (Miller & Cohen, 2001). Here too, cholinergic modulation plays a crucial role.

Indeed, there is accumulating evidence that the PFC has the ability to regulate cholinergic activity in other cortical areas (St Peters et al., 2011; Bloem et al., 2014; Kozak et al., 2005; Nelson et al., 2005). It does this through its direct projections to the basal forebrain where its input activates corticopetal cholinergic projections which then proliferate the signals to relevant cortices (Sarter et al., 2001). For instance, Nelson et al. (2005) showed that stimulating cholinergic receptors in the PFC led to an increase in acetylcholine release in the posterior parietal cortex, one of the core regions in the attention network. This top-down modulation increases signal-to-noise by focusing attention on relevant stimuli. It also helps with making sense of the incoming data, which together work to enhance input processing.

One of the cases where the PFC-driven top-down modulation plays a critical role is during the processing of ambiguous stimuli (Kuperberg et al., 2003; Bach et al., 2015; Ishibashi et al., 2020). This is because the process requires consulting prior knowledge and using it to resolve the ambiguity. Therefore, we hypothesise that processing stimuli judged by the participants to neither have a low nor high plausibility to happen (i.e., medium plausible) will elicit greater PFC engagement. This will, in turn, increase the cortical cholinergic levels reflected by increases in pupil size.

Retrieval

There have been multiple attempts to identify and define the processes involved in the retrieval of long-term memories (Shiffrin & Atkinson, 1969; Norman & Bobrow, 1979; Tulving, 1985; Hasselmo et al., 1995; Schacter & Addis, 2007). While there are theoretical differences and predictions of the resulting models, they all converge on the idea that retrieval involves: (1) an attempt to re-access the original experience, (2) an interaction with the obtained information, and (3) response-related decisions. For example, in the influential memory model by Shiffrin and Atkinson (1969), retrieval involves three stages: search, recovery and response generation. That is, the retrieval process starts with a search in the memory repository triggered by a cue (e.g., a stimulus presented on the screen). Next, relevant information from this search is recovered and placed in the working memory. There is then an evaluation of the current content in the working memory. This is to determine whether the desired information has been obtained and give a response, and if not, whether to continue the search or render it unsuccessful and terminate it.

Later, Hasselmo (1995) would propose a memory model which, similar to Shiffrin and Atkinson (1969), involves search and response-related decision elements. Hasselmo theorised that the hippocampal CA1 acts as a "comparator" in that it identifies whether the incoming sensory input (from the EC) matches or mismatches the memory representations in the existing repository (in

CA3). When there is a high degree of mismatch (e.g., a new stimuli), there is an increase in acetylcholine. Acetylcholine suppresses intrinsic recurrent connections, so its increase promotes encoding by encouraging external input and guarding against proactive interference from existing memories (Sarter et al., 2004). By contrast, when there is a high degree of match (e.g., old stimuli), there is a reduction in acetylcholine. In turn, the suppression of the recurrent feedback connections is released, allowing the memory circuitry to engage in pattern completion and associative recall (Hasselmo, 1995; Hasselmo & McGaughy, 2004). We hypothesise that these differential effects of acetylcholine on encoding and retrieval will be reflected in pupil dynamics.

Furthermore, we expect to see differences within the retrieval process. Specifically, we employed a paradigm with retrieval trials that differed in the intensity of the search needed to access the target memory traces. Our stimuli were word-image pairs and during retrieval participants had to identify whether: there was a new word, old word-same image, or old word-different image. These trials will be referred to as new, intact, and rearranged, respectively.

For new trials, success entailed recognising that the stimuli had not been seen before. A key characteristic of the human memory system is its ability to swiftly identify whether or not it has knowledge of a given event (Headley, 1981). This is in part mediated by the noradrenergic system whose main function is to monitor changes in the environment. It has been dubbed a “novelty detector” because novel stimuli evoke phasic bursts in the LC-NE activity (Sara, 2009). The LC has strong direct connections to the hippocampus, allowing the novelty signals to reach the memory system quickly (Hansen, 2017). This makes it easier to detect that the stimulus does not match any stored memory representations and terminate the search process. In the light of this, we hypothesise that there will be a slight increase in the pupil size following the stimulus onset after which it will return to pre-stimulus levels.

The second type of trials were “intact.” Here, success entailed recognising that the word-image pair presented during retrieval was the exact copy of what was learned during encoding. The word being an “old word” will lead to CA1 marking it as a “match,” which will trigger a search for the original trace to confirm that it was paired with the current image during encoding. Research has shown that people remember better when retrieval cues match the encoding contexts in accordance with the encoding specificity principle (Tulving & Thomson, 1973; Choi et al., 2025). Therefore, although intact trials will expend more cognitive resources than new trials, they will require less mental effort compared to rearranged trials where the retrieval cue does not match the encoding context. To this end, we hypothesise that after stimulus onset, there will be a large and sustained increase in the pupil size for intact trials. But we expect an even more pronounced increase for rearranged trials.

The hypothesis that rearranged trials will evoke the largest pupil dilations is also consistent with past literature showing that during retrieval the increases in pupil size are proportionate to the interference from other information concurrently active in memory (Johansson et al., 2018). For rearranged trials, the retrieval cue is an “old word” that is paired with a different image. Thus, successful retrieval necessitates overcoming interference from the presented information to access the correct pair.

The exact neurobiological mechanism linking pupillary changes and resolution of mnemonic interference is yet to be mapped out. Nevertheless, activity in the anterior cingulate cortex (ACC), a region that monitors conflict (Botvinick et al., 2004), has been shown to trigger a release of norepinephrine in the LC and increased activity in the PFC (Aston-Jones & Cohen, 2005). The PFC supports the cognitive control needed for conflict resolution, in particular, interference from competing memories (Kuhl et al., 2007). Together, these studies suggest that norepinephrine plays a critical role in resolving interference and we expect its contribution to be reflected in the pupil response during rearranged trials.

1.3. Hypotheses

To summarise, we expect to see differences in the pupil response between and within encoding and retrieval. Firstly, we hypothesise that while both processes will evoke increases in pupil diameter, retrieval will result in larger dilations. This would reflect the underlying differential effects of cholinergic and noradrenergic activity on encoding and retrieval. Secondly, within encoding we hypothesised that unexpected stimuli will result in more pupil size increases than expected stimuli. Additionally, ambiguous stimuli will also result in pupil dilation. However, we do not have strong predictions on whether these stimuli will evoke larger or smaller increases compared to the unexpected stimuli. Finally, within retrieval, we expect that the pupil response will be sensitive to the differences in the intensity of the search required to retrieve relevant memory traces. Specifically, rearranged trials will have the largest increases in the pupil diameter, followed by intact trials and then new trials.

2. Materials and Methods

2.1. Ethics Statement

This research was approved by the University of Oxford's Central University Research Ethics committee under the reference number of R90987/RE001.

2.2. Participants

To test our hypotheses, we recruited 27 participants (22 females, 5 males), aged 18–40 years (M = 23.4). All participants were pre-screened for psychological or neurological disorders, handedness, English fluency, and vision acuity. In the final sample, all participants self-reported no history of psychological or neurological disorders, were English-speaking, right-handed, and had normal or corrected-to-normal vision with contact lenses or LASIK surgery. We did not recruit participants with glasses as it was revealed during pilot tests that glasses could interfere with pupillary responses as well as eye-tracking calibration and validation procedures. We ended up with a sample size of N = 26 for analysis. This was after excluding one participant who reported that they did not understand some words in the study stimuli.

Data collection typically occurred in the morning from 9am-1pm to ensure that participants were alert and well rested. To the same end, the night before the study, participants received a reminder to get ample sleep (7+ hours). As stipulated in the ethics protocol, the study commenced with obtaining written consent. This was after explaining to the participants the general objectives of the study and the voluntary nature of their participation. Participants received monetary compensation (£15/hour) or course credit for their time.

2.3. Data Collection

This study was a part of a larger project that aimed to investigate the neural and physiological correlates of encoding and retrieval. As a result, in addition to eye tracking, we recorded EEG, respiration using a nasal thermistor, and heart rate using peripheral EKG electrodes. For this thesis, we will be focusing on the eye tracking data. Data collection involved three main components, namely participant preparation, task instructions and practice, and the memory task.

2.3.1. Participant preparation

The quality of eye-tracking data can be affected by several factors including posture and body movements (Gotardi et al., 2020; Park et al., 2024). Before the experimental tasks, therefore, we made sure participants were well seated, relaxed, and comfortable. To familiarise them with the eye tracking procedure they performed a baseline task where they had to look at a grey screen with a fixation cross in the centre for 2 minutes. Once confirmed that there were no issues, we moved on to the task instructions. It is important to note that in the broader project, this baseline task worked as a measure of resting state. It was repeated at the end of the session to see whether physiological states shifted over the course of the study.

2.3.2. Task instruction and practice

Participants were given detailed instructions on the memory task and did a practice run to make sure they understood what the task entailed. This also helped in ensuring that they were comfortable with using the response keys on the keyboard.

2.4. Experimental Task

The experimental task was a continuous memory recognition paradigm. It consisted of 6 runs, with 60 trials each. The maximum run duration was 8.5 minutes. After each run participants were given a self-paced break. The entire task took approximately 1 hour. The memory task consisted of three main components: task prompt, encoding, and retrieval (Figure 2A).

2.4.1. Prompt/ramp-up

The task employed in this study aimed to mimic the quick transitions between encoding and retrieval we see in daily life. Therefore, encoding and retrieval were interspersed together. Prior to each trial, there was a prompt saying either “encoding” or “retrieval,” indicating which task to perform on the upcoming stimuli. There was a moving bar under the prompt to indicate when the stimuli would come up. We refer to this part of the paradigm as the ramp-up period because participants were instructed to use this period to get ready to learn or retrieve already learned associations. In the broader project, this period will be used to investigate preparatory states which is beyond the scope of this thesis. The ramp-up period was 3.5 s long.

2.4.2. Encoding

On each trial, the participant saw a trial-unique verb together with one of three repeating images: a car, a face, or a house (see section 2.5 for details on how the stimuli were chosen). Upon the image-word pair appearing on the screen, participants were invited to vividly imagine scenarios that involve the action word together with the image. For instance, if the action word was “lift” and the image was a “car”, they could imagine lifting a very heavy car. They then had to indicate how likely the scenario they imagined is to occur in real life or nature. There were three response options: “LOW plausibility”, “MEDIUM plausibility”, and “HIGH plausibility” (Figure 2B). These corresponded to the left arrow, down arrow, and right arrow keyboard keys, respectively. The maximum length of the stimuli presentation and response query was 5 s but participants could respond whenever they were ready to give a response (see section 2.9.1 for discussion on reaction times).

Structuring encoding trials in this manner served two main purposes. First, it fostered deep processing through the usage of semantic associations and visual imagery to come up with personalised scenarios. Deep processing has been shown to increase memory retention (Craik & Tulving, 1975) and task engagement (Biggs, 1987). Second, the plausibility ratings were used to explore how the pupil responds to different types of novelty (i.e., expected vs unexpected novelty) (see section 3.2.2).

2.4.3. Retrieval

During retrieval trials, participants were also presented with an action word and one of three repeating images: a car, a face, or a house (Figure 2B). The response options were “OLD word, SAME image” (we refer to this as “intact”), “OLD word, DIFFERENT image” (we refer to this as “rearranged”) and “NEW word” (we refer to this as “new”). Similar to encoding, the maximum length of stimuli presentation and response query was 5 s.

2.5. Stimuli

Three images of common objects were randomly chosen from the functional localizer (floc) image database created by Stigliani et al. (2015). These images were of equal luminance and same colour which made them fit for our study given that the pupil is very sensitive to changes in luminance and colour (Oster et al., 2022). The action words used were randomly selected from a large set of commonly used verbs in the English language.

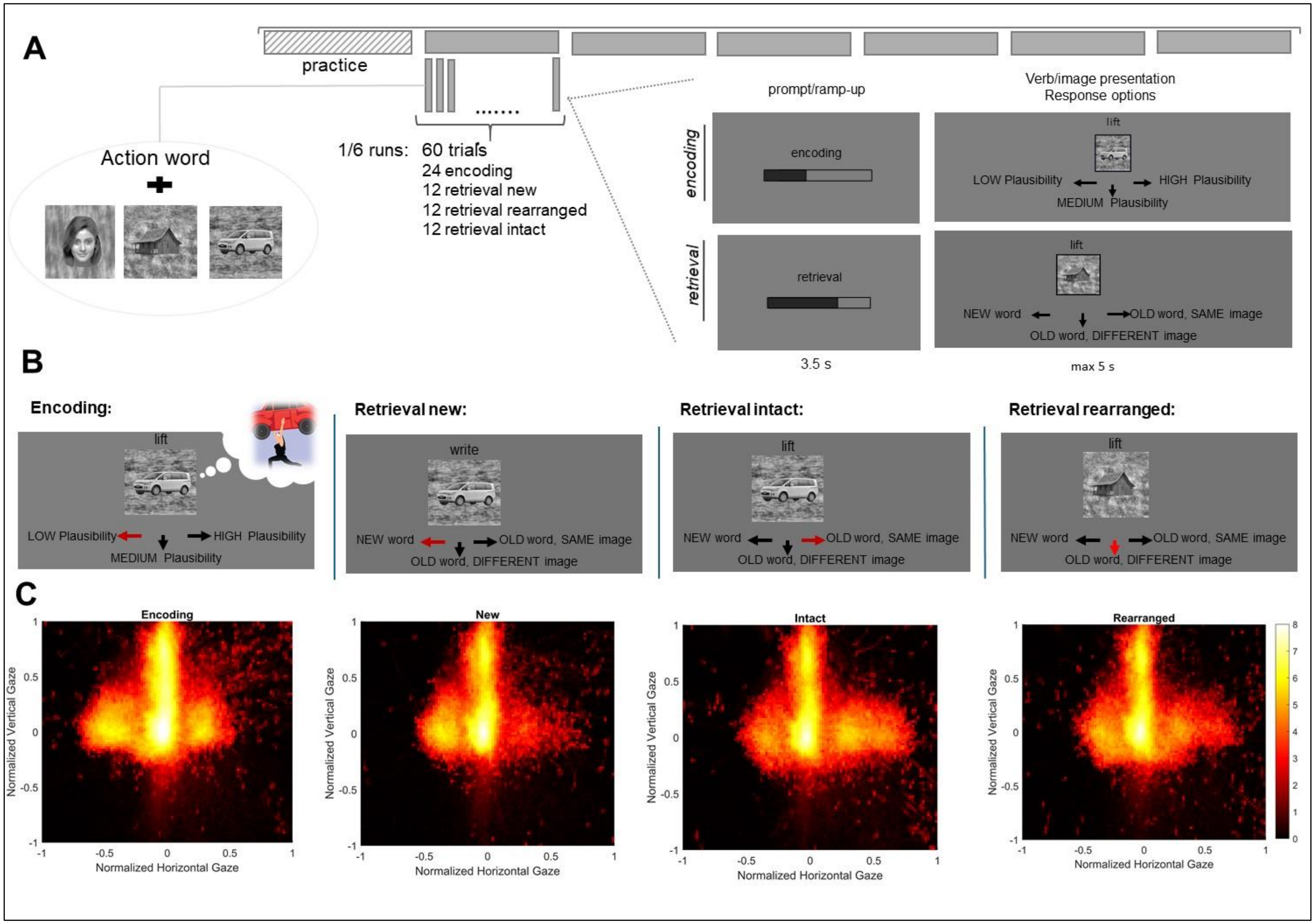

Figure 2. A *- Experimental design: Timings per trial: fixation (not shown) = 1.5 s (+/- 250 ms jitter) → ramp-up (task cue) = 3.5 s → verb/image presentation + response query = max 5 s → max trial duration = 8.5 s.* ***B*** *– An example trial of encoding and all three retrieval trial types.* ***C*** *- Fixation density from stimulus onset till trial end (0 - 5 s).*

2.6. Pupil Data Acquisition

2.6.1. Eye tracker, experimental room, and display settings

To acquire the pupil data, we used an EyeLink 1000 Plus eye tracker by SR Research. The experimental room was windowless and shielded to block all light from outside and we used dim light settings during the experimental tasks. We employed dim lighting because it engenders better tracking accuracy for infrared-based eye trackers like the EyeLink. It allows the eye tracker's infrared light to stand out more which, in turn, enhances the relative contrast between the dark pupil and more reflective structures like the iris (Hansen & Hammoud, 2006). This makes it easier to detect changes in the pupil size and reliably track other eye movements.

We tracked both eyes at 1000 Hz. For the analysis, however, we only used the dominant eye (see section 2.6.2 for the dominant eye test) except in one case where we were only able to obtain data from one eye. Stimuli were displayed on a 1920 px x 1080 px monitor whose screen brightness was kept at 30 cd/m2. We chose this intermediate level of luminance to match the dim lighting in the experimental room and after subjective reports of suitability from volunteers in pilot tests. This is important because when there is a mismatch between the luminance of the display screen and the surrounding environment (e.g., a bright screen in a dark room), it strains the eyes which results in drastic decreases in the pupil size (Tyukhova & Waters, 2018).

Participants were seated 95 cm away from the screen in a locked-in chair and head-stabilised mount to minimise movements. In addition, participants received instructions to remain as still and relaxed as possible as they would be given time to stretch and rest during a self-paced break at the end of each run. Prior to each run, participants performed a randomised 9-point eye tracking calibration and validation procedure.

2.6.2. Dominant eye test

To determine the dominant eye, we used the Miles test also known as the thumb test (Miles, 1930). It involves the following steps:

1. Find an object in a distance;
2. While both eyes are open and focused on the identified object, stretch one arm out and superimpose the thumb on the object;

3. Alternately close one eye at a time while the other remains focused on the object. For one eye it seems as though the thumb has moved, while for the other the thumb is directly in front of the object.

The eye that keeps the thumb aligned with the object is the dominant eye.

2.7. Data Pre-processing

Our pre-processing procedure followed the steps in Urai et al. (2016) and recommendations by Mathôt & Vilotijević (2023).

2.7.1. Parsing raw data into MATLAB and Fieldtrip friendly data structure

The output files from the EyeLink are in EDF format. To convert these EDF files to asc text files, we used the EDF2ASC converter which was developed by SR Research. After this conversion, we used a function, *read_eyelink_ascNK_AU*, from Urai et al. (2016) to read the asc files, namely, the header information, input triggers, messages, pupil data, gaze-position and other eye movement data such as saccades and blinks. Then, we used the *asc2dat* function, also from Urai et al. (2016), to create events and data structures compatible with MATLAB and Fieldtrip. Once this step was completed, we proceeded to clean the pupil signal.

2.7.2. Removing missing data and identifying invalid data

Missing data originates from the eye tracker being unable to locate the pupil, such as when the eyes are closed (Mathôt & Vilotijević, 2023). The EyeLink denotes these data with the value zero (0). Therefore, to avoid confusion in downstream computations we changed them into NaNs.

In contrast, invalid data from blinks where the recorded pupil size does not reflect the actual pupil size are not denoted by any special value. This meant that we had to detect them indirectly using blinks defined by the eye-tracker software. However, built-in blink detection software in eye-trackers do not always capture all the blinks (Culemann et al., 2023). As recommended by Mathôt & Vilotijević (2023) and employed in Urai et al. (2016), the remaining blinks were identified based on changes in the velocity of the pupil signal characteristic of blinks. The beginning of the eyes blinking is marked by a sudden and rapid decrease in pupil size due to the eyelid closing while the end is denoted by a rapid increase in the velocity as the eyelid opens (Culemann et al., 2023; Mathôt & Vilotijević, 2023). After all the blinks were identified, we changed their values into NaNs, essentially denoting them as missing data.

2.7.3. Interpolating the blinks

Missing data can significantly decrease statistical power (Mathôt & Vilotijević, 2023). In light of this, we used linear interpolation to estimate the data in the identified blinks and other brief periods of missing data. We started interpolating 150 ms before and after the blinks given that data surrounding blinks is unreliable (Knapen et al., 2016). We did not interpolate blink-like periods that were longer than 500 ms. Instead, we left them as missing data because physiologically blinks are generally shorter than 500 ms (Nyström et al., 2024). It is also advisable to only interpolate short periods during which pupil size is confidently predictable (Mathôt et al., 2018).

2.7.4. Estimating and removing blink and saccade effects

Interpolating blinks estimates the pupil size in the blink period but does not remove all blink driven changes in the signal, also known as blink effects (Knapten et al., 2016). Additionally, saccades also introduce changes to the pupil time course, known as saccade effects. Notably, these blink and saccade effects have been shown to last for as long as 5s meaning that they would not be eliminated by interpolation (Knapten et al., 2016). Yet, they can be easily misinterpreted as physiological responses to the stimuli or task in question (Jainta, 2011; Yoo et al., 2021). Hence, to avoid this, they need to be removed. To estimate the blink and saccade effects we used the deconvolution approach detailed in Knapten et al. (2016). Then, using the function *blink_regressout*, provided by Urai et al. (2016), we regressed these effects out.

2.7.5. Normalisation of the data

The residual pupil data was normalised by z-scoring done independently for each participant and for each run. We used this method for three reasons. Firstly, after each of the 6 runs, participants took a break such that normalising across the overall pupil time course would not have accounted for these variations. Secondly, we normalised separately for each participant to minimise inter-subject variability as there are individual differences in the pupil response (Sibley et al., 2020). In addition, the EyeLink measures the pupil diameter in arbitrary units (AU). While these AU values correlate linearly with the actual pupil diameter (Einhäuser et al., 2008), absolute values for different participants may mean different things. Hence, the need for within participant normalisation.

Lastly, we preferred this method over the common baseline corrections (Mathôt et al., 2018), because the nature of our paradigm meant that the stimulus onset was not immediately preceded by a period with no experimental manipulation to use as a baseline. This being said, we repeated our analyses using trial wise baseline correction, subtracting the mean pupil size in the 500 ms before the ramp-up period. There were no significant differences from the findings reported here.

2.7.6. Epoching and enriching trials with behavioural information

Next, we epoched the data for the different encoding and retrieval trials with stimulus-locked segments of 6 s in length (i.e., from −1 s before stimulus onset to the 5 s period with stimulus display and response query). The behavioural data was recorded in a separate log file. Thus, after epoching the data and defining trials, we added relevant information and experimental variables from the log files.

2.8. Independent Variables

To investigate whether the pupil is sensitive to retrieval operations, we varied the intensity of the search required to access the target memory traces. In the extant literature this is known as retrieval effort, defined as “the mobilisation of processing resources in service of a retrieval attempt” (Robb & Rugg, 2002, p. 583). Retrieval effort is operationalised in terms of relative difficulty. This is because the more difficult to access the information needed to execute a given retrieval task, the greater the amount of effort expended (Buckner et al., 1998; Robb & Rugg, 2002; Van Rijn et al., 2012). Therefore, for our independent variables, we used two variables that are known to require different amounts of retrieval effort, (1) alignment of encoding and retrieval cues and (2) retention interval. Additionally, we reasoned that if the pupil response indexes memory processes, it would replicate the classical ERP old/new effect (Rugg & Curran, 2007). Hence, our third independent variable was new vs old stimuli.

2.8.1. Alignment of encoding and retrieval cues

Research has shown that retrieval cues that match the encoding context are more effective at improving recall (Tulving & Thomson, 1973; Choi et al., 2025). In other words, it is easier to access desired information when encoding and retrieval cues match versus when they do not match. Therefore, to vary the retrieval effort, we had retrieval trials where the encoding and retrieval cues either matched (intact trials) or did not match (rearranged trials) (Figure 2B).

2.8.2. Retention interval

Another factor that influences retrieval effort is retention interval, which describes how long after learning an item you are tested on your ability to remember it (Hockley, 1982; Oliveira et al., 2021). Continuous recognition memory paradigms are a good fit for exploring this given that in these paradigms encoding and retrieval are intertwined.

Specifically, in our task, the action words (paired with either a car, house or face) were presented two times in a study-test procedure with an intermingling of new words that were only presented once. Retention intervals were varied through the number of words between their corresponding encoding and retrieval phases. Some words were tested shortly after while others were tested long after they were learned. For example, words that were tested three trials after being learned had a retention interval of 3. Retention intervals were randomly generated for each run and the longest possible interval was 58 given that there were 60 trials per run.

It is well established that it is easier to access recently experienced events than events in the distant past. This was first observed by Ebbinghaus (1885) and more recently demonstrated by proponents of the temporal context model (TCM) (Howard & Kahana, 2002). According to the TCM, when we encode events, we not only register the content but also the time in which the events in question occurred. This means that recalling recent events is easier as their temporal context is still active. On the other hand, recalling events further in the past requires reconstructing the temporal context, and hence involves more effort (Polyn et al., 2009). As a result, we hypothesised that the increased cognitive effort required to access memories in the longer study-test retention intervals will lead to more increases in pupil size than in short study-test retention intervals. This will be the case for both intact and rearranged trials.

2.8.3. New vs old stimuli

In recognition memory, the old/new effect originates from comparing event-related potentials (ERPs) exhibited by correct judgement of previously encountered items as old (hits) and ERPs exhibited by correct identification of new items as new (correct rejections) (Vo et al., 2008). These comparisons have repeatedly found that ERPs for hits are marked by a greater positivity, emerging at around 300-800 ms after stimulus onset, than ERPs elicited by correct rejections (Rugg & Curran, 2007). Therefore, we expect that comparisons between the pupil time course for intact and new trials will follow a similar pattern. Specifically, we hypothesise that correct judgements of already encoded word-image pairs as “old word, same image” will result in bigger pupil diameter increases than correct identification of new trials as containing “new words.”

2.9. Dependent Variables

To assess the effects of our independent variables, we employed two dependent variables: memory accuracy and reaction time. This is because retrieval effort is typically deduced from measures of reaction times and memory performance as they have been theorised to index the strength of memory traces and by extension their accessibility (Gimbel & Brewer, 2010).

2.9.1. Reaction time

Reaction times (RTs) have long been used as a measure of memory strength with faster responses associated with stronger memory traces. For example, in paradigms that test if participants feel familiar versus remember the specific details of an event, it has been shown that remembered items engender shorter RTs (Rotello & Zeng, 2008; Wixted & Mickes, 2010). Additionally, retrieval of recently experienced events evokes shorter RTs than events further in the past (Hockley, 1982). Hence, we expect that long study-test retention intervals will lead to longer RTs relative to shorter study-test retention intervals. Similarly, rearranged trials will exhibit longer RTs compared to intact and new. In addition, we predict that the same pattern will be observed for the respective pupil response with longer RTs correlating with increased pupil dilations.

Furthermore, there is ongoing debate about whether pupillary responses index cognitive processes or if they just monitor mere task engagement (Mathôt, 2018; Grujic et al., 2024). Proponents of the latter have cited the relationship between RTs and changes in pupil size as evidence for their position (Gross & Dobbins, 2020). This is because reaction times do indeed track time on task and this covaries with pupil dynamics (Hopstaken et al., 2015). Nevertheless, there is accumulating evidence showing differential pupillary responses even in conditions where reaction time is controlled for (Porter et al., 2006; Kafkas & Montaldi, 2012). The aim of the present study is to contribute to this debate.

Thus, in addition to retrieval, we recorded reaction times for encoding trials. We hypothesised that if pupil responses index the underlying mechanisms supporting these two processes, we would still see differences in the pupil signal regardless of the length of their respective RTs. The RTs for encoding trials were also used to evaluate whether there are any behaviour differences incurred during the processing of stimuli that differed in novelty expectations. Specifically, we tested if low, medium and high plausible stimuli resulted in different reaction times and if this correlates with the pupil response.

For both encoding and retrieval, reaction times were obtained by computing the difference between the onset of the stimulus + response query screen and the first keyboard press when participants gave an answer. Trials where participants did not press any of the answer keys within 5 s were marked as invalid and excluded from analysis.

2.9.2. Memory accuracy

Another factor that has been used to measure memory strength is accuracy. For instance, literature on the signal detection theory (SDT) of recognition memory has revealed that the more difficult it is to access memory events the lower the memory performance (Wixted, 2007). According to the SDT, test items elicit different levels of memory strength signal. This signal is then compared to a decision criterion, whereby if it passes a certain threshold it is classified as “old” and below that threshold it is seen as “new” (Yonelinas et al., 2010; Rotello, 2017).

Additionally, investigations into different methods used to vary the accessibility of desired memory traces have shown that conditions that facilitate easy access to previously learned information result in better memory performance (Craik & Lockhart, 1972; Glanzer et al., 1999; Wixted, 2007). Such methods include encoding techniques like shallow versus deep processing and the alignment of retrieval and encoding cues as employed in this study. Therefore, to test whether our manipulation of retrieval effort was successful, we recorded memory accuracy of different retrieval trials, namely rearranged, intact and new. Memory accuracy was measured as a binary value (1-correct or 0-incorrect) per trial.

Furthermore, there is long standing scholarship on the relationship between pupillary changes and memory performance (Goldinger & Papesh, 2012). However, while there have been extensive explorations on which pupil response during encoding is associated with better memory performance (Unsworth & Miller, 2020), few studies have investigated the pattern of pupil response that support retrieval. Although, in the limited studies that have looked at pupillometry during retrieval, there have been mixed results. Some studies have found that successful retrieval is associated with a larger increase in pupil diameter than unsuccessful retrieval (e.g., Kucewicz et al., 2018). Whereas others have found the opposite (e.g., Kafkas et al., 2023).

Both directions are plausible. On one hand, the bigger increase in the pupil size could reflect the memory strength and the neuromodulatory activity facilitating the successful retrieval of desired information (Goldinger & Papesh, 2012). On the other hand, the increase in pupil sizes could reflect the longer search that often characterise unsuccessful retrieval (Rotello & Zeng, 2008; Wixted & Mickes, 2010). This is emblematic of the broader debate in the field between the memory strength and retrieval effort account of pupil dilation in recognition memory (Taikh & Bodner, 2022). Accordingly, the present project aims to shed light on the nature of the relationship between the pupil response and retrieval operations.

3. Results

3.1. Retrieval

3.1.1. Behaviour results

We first explored the behavioural data to evaluate whether our manipulation of search intensity and operationalisation of retrieval effort was successful. For these analyses, we aggregated the RTs and memory accuracy scores to produce a single value for each participant per retrieval trial type. Specifically, for each trial type we concatenated the data across all the 6 runs and then averaged them to get a mean RT and memory score for each participant.

Invalid values, resulting from trials where participants did not give a response, and outliers, in the case of reaction times, were excluded prior to averaging. Outliers were identified using the MATLAB function *isoutlier* which removes values that are 3 median absolute deviations (MAD) away from the median. Lastly, to get a clearer picture of how variables including RT, retention interval, and later the pupil response, relate to memory operations, we only used correct trials, unless when explicitly stated otherwise.

Main effect of trial type on RT and memory accuracy

Prior to examining the influence of the specific independent variables, we first tested whether there was a main effect of trial type on RT and memory accuracy. To do so, we conducted a one-way repeated-measures ANOVA which revealed a significant main effect of trial type on RT, $F(2, 50) = 59.17$, $p < 0.001$, $\eta^2_p = 0.70$, and memory accuracy, $F(2, 50) = 41.98$, $p < 0.001$, $\eta^2_p = 0.63$. This indicated that RTs and memory performance differed significantly between the three conditions. Following that, we performed pairwise comparisons, using paired-samples *t*-tests, to further investigate where these effects arose from.

Alignment between encoding and retrieval cues

First, we assessed whether the alignment between encoding and retrieval cues was successful in varying the retrieval effort. To achieve this, we compared rearranged and intact trials. For reaction times (Figure 3B), rearranged trials ($M = 2.178$ s, $SD = 0.328$ s) had significantly longer reaction times than intact trials ($M = 1.983$ s, $SD = 0.304$ s), $t(25) = 6.944$, $p < 0.001$, Cohen's dz = 1.362, 95% CI [0.137, 0.253].[1] For memory accuracy (Figure 3A), rearranged trials ($M = 0.806$, $SD =$

---

[1] Confidence intervals used in this study were calculated using mean differences of the conditions in question.

0.107) had significantly lower scores than intact trials (*M* = 0.849, *SD* = 0.125), *t*(25) = 3.64, $p <$ 0.001, Cohen's dz = 0.714, 95% CI [0.019, 0.069]. As hypothesised, matching encoding and retrieval cues resulted in faster RTs and better memory performance.

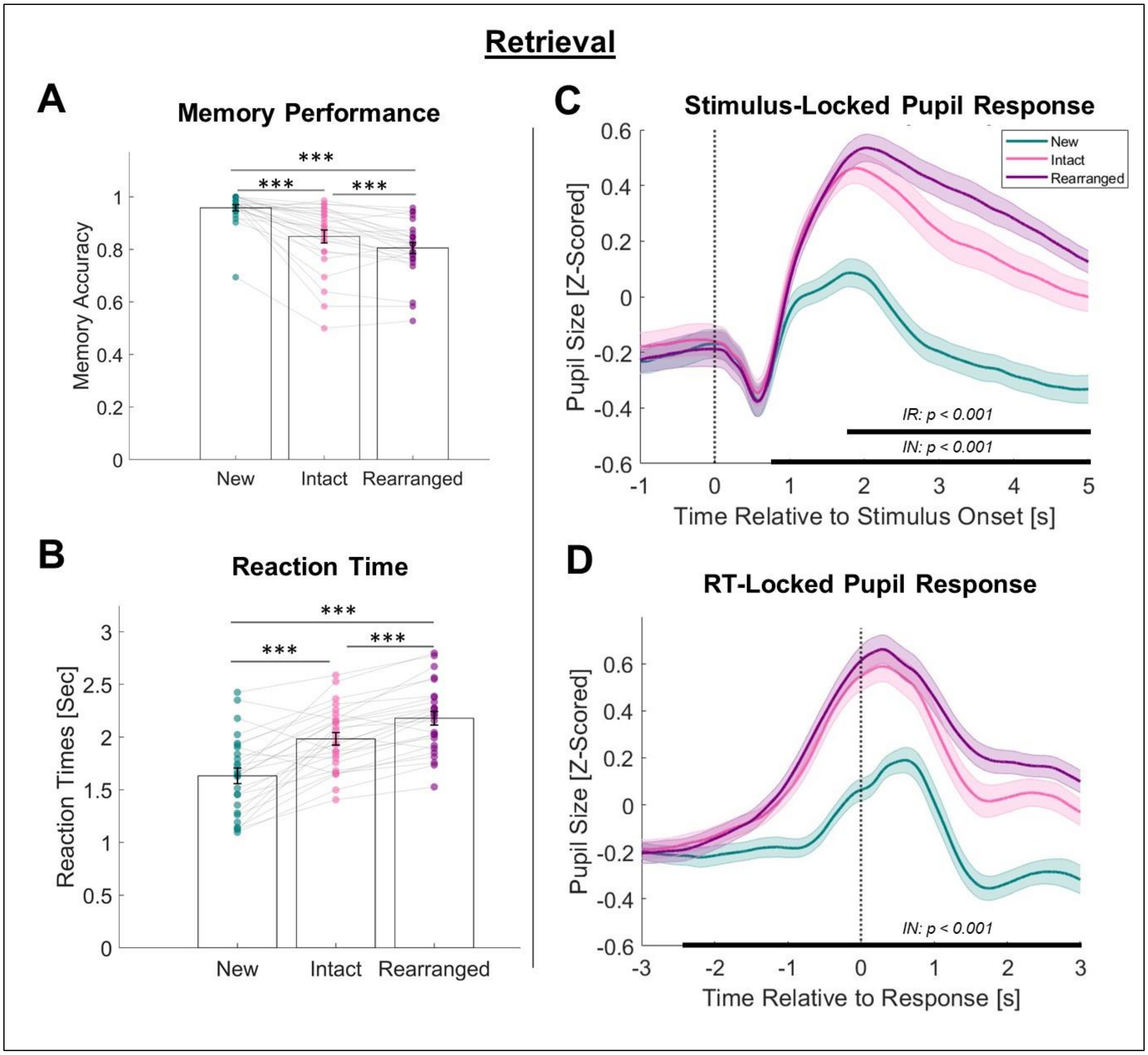

Figure 3**: Retrieval-behavioural and pupillometry results.* ***(A) *Memory performance was significantly different across all conditions, New>intact> rearranged, p < 0.001.* ***(B)*** *RTs were also significantly different across retrieval trial types, rearranged>intact>New, p < 0.001. Stimulus-locked* ***(C)*** *and RT-locked* ***(D)*** *demonstrate the pupil old/new effect (intact vs new-IN) and the encoding specificity principle (intact vs rearranged-IR). Error bars in this and subsequent figures were calculated by mean ± SEM for both the behaviour and pupillometry results.*

Old vs new stimuli

Second, we evaluated the behavioural correlates of the old/new effect, by comparing intact and new trials. RTs (Figure 3B) were significantly shorter for new trials ($M = 1.632$ s, $SD = 0.379$ s) than for intact trials ($M = 1.983$ s, $SD = 0.304$ s), $t(25) = -5.639$, $p < 0.001$, Cohen's dz = -1.106, 95% CI [-0.501, -0.247]. For memory accuracy (Figure 3A), new trials ($M = 0.958$, $SD = 0.060$) resulted in better scores than intact trials ($M = 2.027$ s, $SD = 0.319$s), $t(25) = 5.32$, $p < 0.001$, Cohen's dz = 1.043, 95% CI [0.067, 0.151]. These results, faster RTs and better memory performance for new trials compared to intact trials, suggest that correctly identifying new stimuli as "new" requires less cognitive effort than correct judgement of previously seen items as "old."

Short vs. long retention intervals

Third, we investigated the effects of study-test retention intervals. To do so, we performed a median split for rearranged ($M = 8$, $SD = 9.5$) and intact trials ($M = 8$, $SD = 9.7$). Trials above the median were subsequently considered to be a part of "long retention" intervals and those below to belong to "short retention" intervals.

To begin, we ran a 2 × 2 repeated-measures ANOVA to assess the effects of retention interval (short vs. long) and trial type (intact vs. rearranged) on reaction time. There were significant main effects of retention interval, $F(1, 25) = 20.59$, $p < 0.001$, $\eta^2_p = 0.452$, and trial type, $F(1, 25) = 63.32$, $p < 0.001$, $\eta^2_p = 0.717$. However, there was no significant interaction, $F(1, 25) = 0.04$, $p = 0.843$, $\eta^2_p = 0.002$.

Follow-up paired-samples *t*-tests revealed that in both trial types, long retention intervals were accompanied by slower response times than short intervals. Rearranged (Figure 4C): long intervals (M = 2.320 s, SD = 0.325 s), short intervals (M = 2.168 s, SD = 0.311 s), $t(25) = 2.995$, $p = 0.006$, Cohen's dz = 0.587, 95% CI [0.047, 0.257]. Intact (Figure 4A): long intervals ($M = 2.124$ s, $SD = 0.261$ s), short retention intervals ($M = 1.983$ s, $SD = 0.289$ s), $t(25) = 4.457$, $p < 0.001$, Cohen's dz = 0.874, 95% CI [0.076, 0.206]. As indicated by the lack of a significant retention interval × trial type interaction, there was no difference in the effect of retention interval on reaction time between intact and rearranged trials.

Next, a 2 × 2 repeated-measures ANOVA was conducted to examine effects of retention interval (short vs. long) and trial type (intact vs. rearranged) on memory accuracy. We found: a main effect of retention interval, $F(1, 25) = 31.05$, $p < 0.001$, $\eta^2_p = 0.554$, a main effect of trial type, $F(1, 25)$

= 14.16, $p = 0.001$, $\eta^2_p = 0.362$, and a significant retention interval × trial type interaction, $F(1, 25) = 7.54$, $p = 0.011$, $\eta^2_p = 0.232$.

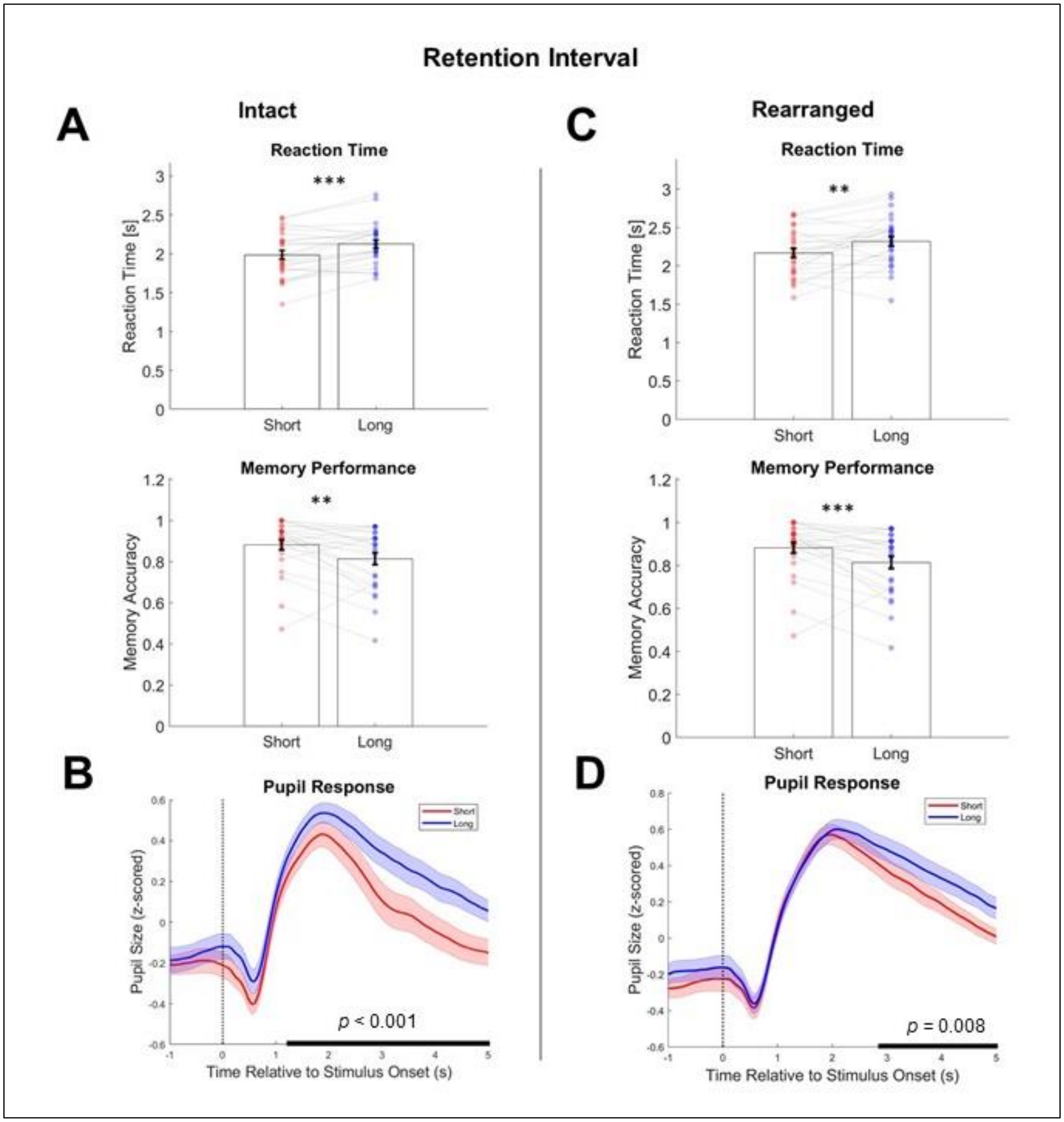

***Figure 4**: Comparison between behavioural outcomes and the pupil response in short and long retention intervals for intact (left - panel **A** and **B**) and rearranged (right - panel **C** and **D**) trials. In both conditions, long was more effortful as reflected in the RTs, accuracy and pupil response. $p < 0.001$ and $p < 0.01$ are signified by *** and ** respectively.*

Post-hoc pairwise comparisons, confirmed that for both rearranged and intact, there was a decline in memory accuracy from short to long retention intervals. Rearranged (Figure 4C): short retention (M = 0.868, SD = 0.113), long retention (M = 0.739, SD = 0.122), $t(25) = 6.678$, $p < 0.001$, Cohen's dz = 1.310, 95% CI [0.089, 0.169]. Intact (Figure 4A): short retention (M = 0.882, SD = 0.127), long retention (M = 0.814, SD = 0.147), $t(25) = 3.06$, $p = 0.005$, Cohen's dz = 0.60, 95% CI [0.022, 0.114]. However, the decline in memory performance was more pronounced for rearranged than intact trials as highlighted by the significant interaction.

In summary, short study-test retention intervals elicited faster response times and better memory performance than long intervals.

Unsuccessful vs. successful retrieval

Lastly, it has been theorised that unsuccessful retrieval is often accompanied by a longer search (Rotello & Zeng, 2008; Wixted & Mickes, 2010). To test this, we looked at whether there was a difference in reaction time for successful and unsuccessful retrieval. Specifically, we conducted a 2×2 repeated-measures ANOVA with factors accuracy (correct vs. incorrect) and trial type (intact vs. rearranged). Memory performance was generally high, but was especially close to ceiling for new trials (M = 0.958, SD = 0.060) with some participants (n = 6) having a perfect score. Therefore, new trials were not included in this analysis. Additionally, in intact, we excluded one participant who had a perfect score.

The test revealed: a significant main effect of accuracy, $F(1, 24) = 29.11$, $p < 0.001$, $\eta^2_p = 0.548$, and a significant accuracy × trial type interaction, $F(1, 24) = 21.34$, $p < 0.001$, $\eta^2_p = 0.471$. Whereas the main effect of trial type was not significant, $F(1, 24) = 1.67$, $p = 0.208$, $\eta^2_p = 0.065$. Together these results indicate that reaction times were significantly different between successful and unsuccessful retrieval, and the difference varied depending on the retrieval task.

To further understand the direction of these effects, we ran paired-samples t-tests. For intact trials (Figure 5A), incorrect responses (M = 2.8105 s, SD = 0.7250 s) took significantly longer than correct responses (M = 1.9949 s, SD = 0.3034 s), $t(24) = -5.50$, $p < 0.001$, 95% CI [-1.031, -0.469], Cohen's dz = 1.10. For rearranged trials (Figure 5B), incorrect (M = 2.4083 s, SD = 0.4306 s) were slower than correct (M = 2.1781 s, SD = 0.3280 s) trials. Similarly, this difference was significant, $t(25) = -2.97$, $p < 0.01$, 95% CI [-0.346, -0.063], Cohen's dz = 0.58, but smaller than the difference observed in intact trials. This explains the ANOVA interaction by showing that retrieval success influenced reaction times more strongly for intact than for rearranged trials.

3.1.2. Pupillometry

Having confirmed that the retrieval effort manipulation produced the predicted behavioural effects, we turned to the main question of whether it influenced the pupil response.

Data quality checks

First, we conducted data quality checks to confirm that everything was recorded as expected. Specifically, we used the gaze position data to trace where participants were looking during the different retrieval trials. As illustrated in Figure 2C, there was an area of high gaze density at the top of the screen which reflects the movements from the word to the picture and vice versa. Expectedly, for different retrieval trials, participants fixated more on the location of the respective response. For example, there were more fixations on the left for new trials as compared to intact trials that show more fixations on the right.

Comparison with the typical pupil time course

Next, we examined whether the pupil trace followed a similar pattern to that typically observed in cognitive pupillometry studies. Cognitive pupillometry uses changes in the pupil size, in response to a given stimulus, to investigate cognitive processes (Mathôt & Vilotijević, 2023). For investigations that employ visual stimuli, the pupil response has been shown to follow a characteristic pattern. This involves an increase in pupil diameter that starts to rise between 300 – 500 ms, peaking at around 1 s – 2 s after the stimulus onset. The pupil size then gradually returns to pre-stimulus levels or remain elevated for several seconds depending on task requirements (Wierda et al., 2012; Isabella et al., 2021).

Indeed, as will be revealed by the present study, the specific behaviour of the pupil signal varies depending on the experimental design and cognitive demands of the task. For example, for all the retrieval trial types, the pupil time course started with an initial constriction that occurred from 200 ms to 500 ms after stimulus onset (Figure 3C). This transient change in the pupil response is common in cases, such as in the paradigm used here, where the stimulus includes a change in luminance from the pre-stimulus period (Wang et al., 2014; Schwetlick et al., 2025). It is important to note that this change is not cognitively driven but is a defining feature of the pupillary light reflex (PLR). The PLR is a brief change in pupil size in response to a visual trigger (Mathot et al., 2014). After, the pupil signal for all the trial types followed the typical pattern described above, albeit steeper for some trial types than others. In particular, there was an increase in the diameter that emerged at ~500 ms, peaked at ~2 s, and was generally sustained until the start of the next trial at 5 s.

To plot the pupil time course, we used a custom MATLAB function that applied the *shadedErrorBar* function to indicate the standard error of the mean (SEM) through a continuous shading (Campbell, 2024). Statistical significance was assessed using cluster-based permutation tests (see below).

Using cluster-based permutation tests to establish statistical significance

From there, we assessed the effects of our independent variables on the pupil response. To achieve this, we compared the pupil signal of the conditions of interest using cluster-based permutation tests. We chose these tests to formally assess significance because they have been identified to guard against the problem of multiple comparisons in time series data (Maris & Oostenveld, 2007). Specifically, for each contrast, we computed paired-samples t-tests at every time point and identified clusters of adjacent time points for which the uncorrected p-value was below the predefined cluster-forming alpha of 0.05. We then calculated the cluster-level t statistic using the sum of all t-values within the cluster. This cluster-level t-value was then compared against a null distribution generated via 2000 random permutations. Clusters whose p-value was < 0.05 were considered significant.

The pupil response indexes retrieval operations

Alignment of encoding and retrieval cues

To start, we tested whether the differences in retrieval effort, as reflected by RTs and memory performance, between rearranged and intact trials were also seen in the pupil response. See Figure 3C for the pupil trace corresponding to these two trial types. As expected, rearranged trials exhibited greater and more sustained pupil dilation than intact, especially in the later phase of the time course. There was a significant positive cluster starting from 1.88 s to 5 s, post-stimulus, $p < 0.001$. Prior to this, namely between 0.50 s – 1.88 s, both intact and rearranged experienced a similar increase in pupil size.

Rearranged trials involve recollection-based judgements in that the participant needs to re-access the original pair to reject the familiar rule (i.e., old word) and notice the recombination. This is a relatively slow and effortful process, which might explain the more sustained pupil dilation. Intact trials, on the other hand, do not involve any interference resolution. This makes the retrieval process more fluent, leading to a lower dilation peak and a quicker return toward pre-stimulus levels. Together, these results likely reflect the greater retrieval effort needed when retrieval cues do not match the encoding context and provide a physiological correlate of the well-established encoding specificity principle.

Long vs. short retention intervals

Another consistent finding in the memory literature is that recent events are easier to access than events far in the past. In the current study, this was replicated behaviourally with longer RTs and lower memory performance associated with trials in long retention intervals compared to those in short retention intervals. But, would this be reflected in the pupil response? To answer this question, we compared the pupil trace for both intervals in intact and rearranged trials (Figure 4).

The pupil time course had two visually distinct phases. That is, an early phase where both short and long intervals elicited an equal and sharp increase in the pupil diameter. This phase was longer for rearranged trials than intact. And, a later phase where long intervals exhibited significantly larger pupil dilations. Specifically, for intact (Figure 4B) there was a significant difference from 1.23 s to 5 s post-stimulus, $p < 0.001$ and from 2.87 s to 5 s post-stimulus, $p = 0.008$ for rearranged trials (Figure 4D).

Testing the pupil old/new effect

Next, we compared the traces for intact and new trials to see whether the pupil captured the old/new effect. Traditionally, this is marked by post-stimulus ERPs of greater positivity that emerge between 300-800 ms for old relative to new stimuli (Rugg & Curran, 2007). The pupil response followed a similar pattern, with larger increases in pupil size for intact than new trials emerging around 700 ms after stimulus onset. To be specific, there was a significant positive cluster from 0.73 s to 5 s post-stimulus, $p < 0.001$(Figure 3C). As hypothesised, accurate judgments of word-image pairs as old were accompanied by more increases in pupil diameter than correct identification of new trials as new.

Stimulus-locked comparison - successful vs. unsuccessful retrieval

Lastly, to elucidate whether pupillary changes are best explained by the retrieval effort or memory strength account, we compared the pupil response for correct and incorrect trials in intact and rearranged trials (Figure 5). More pupil dilation for correct trials would support memory strength whereas more for incorrect trials would support retrieval effort. Here too, there were two phases, namely an early phase where the pupil size for correct and incorrect trials was equal (intact) or numerically higher (rearranged) for correct trials. And, a later phase (approximately between 2.5 s – 5 s) where incorrect trials were accompanied by significantly larger pupil dilations than correct trials in both conditions.

Specifically, for intact (Figure 5B), a significant positive cluster emerged between 2.49 s to 5 s post-stimulus, $p < 0.001$. Prior to this, the pupil response was similar for correct and incorrect trials, with both conditions showing an initial luminance-evoked constriction, and then dilation 0.50 s after stimulus onset. The divergence occurred ~2 s later when correct trials experienced a sharp peak and began to return to pre-stimulus levels. Thereat, incorrect trials continued to increase in dilation and reached a higher peak 1s later that was gradually maintained until the end of the trial.

For rearranged (Figure 5F), the early phase followed a relatively similar trend to what was seen in intact trials. There was a more visible difference between incorrect and correct trials, although this difference was not significant, $p = 0.1$. Subsequently, incorrect trials continued to rise and while they did not reach a peak higher than correct trials, they maintained a sustained plateau till the end of the trial. In this period, there was a significant cluster from 3.44 s to 5 s, post-stimulus, $p < 0.001$.

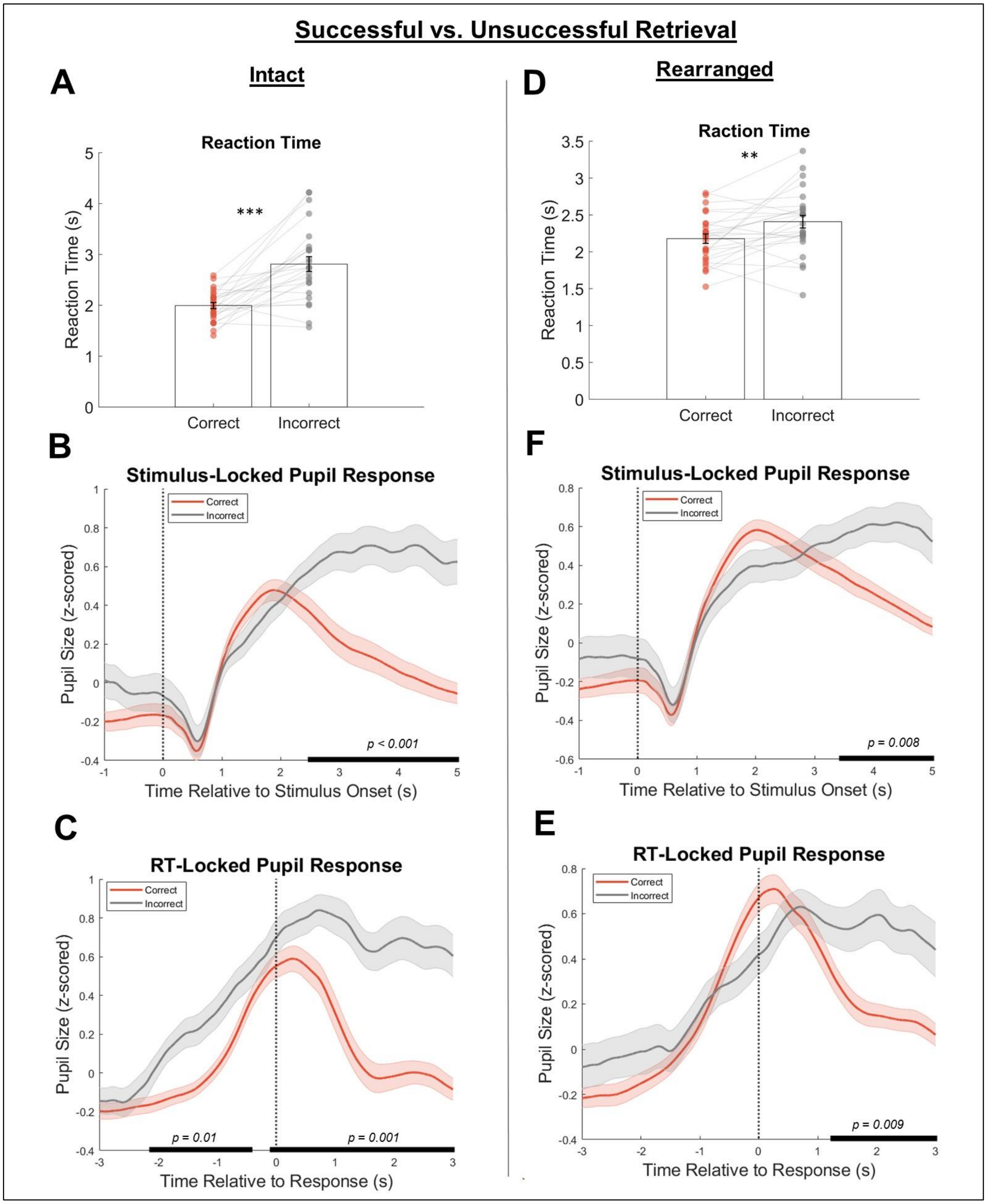

***Figure 5**: Successful vs unsuccessful retrieval. In both intact **(A)** and rearranged **(D)**, incorrect trials had longer RTs suggesting more effort needed. In both conditions the stimulus **(B, F)** and RT-locked response **(C, E)** show two distinct phases in the pupil time course, an early and late component that are best explained by the memory strength and effort accounts, respectively.*

RT- Locked comparison - successful vs. unsuccessful retrieval

To further explore these temporal dynamics, we conducted response-locked comparisons between incorrect and correct trials (Figure 5). In this analysis, point zero corresponded to the time when the participant pressed the response key. We analysed 3 s before and after to allow enough time to capture how the differences in the signal unfolded.

Starting with intact (Figure 5C), pre-response, both incorrect and correct trials elicited a steep increase in the pupil size. But, incorrect trials had significantly larger and more sustained increases than correct ones, as evidenced by a significant cluster, from -2.13 s to -0.43 s, pre-response, p = 0.01. This corroborates the behavioural results and reflects the effortful search that is characteristic of unsuccessful retrieval attempts. Post-response, correct trials saw a steep decrease in the pupil size following their dilation peak, indicating that once relevant traces were successfully retrieved, cognitive effort in service of that task subsided.

Conversely, incorrect trials continued to increase and remained elevated long after the response time, with a significant cluster emerging between -0.08 s to 3 s, post-response, p-value = 0.001. This suggests that in the absence of a satisfying and/or confident response participants continued to engage with the unresolved search. These results align with previous research showing that the pupil response is sensitive to decision uncertainty (Urai et al., 2016) and metacognitive conflict (Ryals et al., 2021). For instance, Ryals et al. (2021), found that in tip-of-the-tongue (TOT) experiences, whereby individuals feel confident that they know the answer but fail to retrieve the target items, there is greater pupil dilation than retrieval failure without TOT.

Moving on to rearranged trials (Figure 5E), the pre-response signal showed a somewhat different pattern than seen for intact. The difference in the pupil response between correct and incorrect was less pronounced and non-significant, albeit with correct trials achieving a sharper dilation peak than incorrect trials. These differences are consistent with those seen in RTs. For intact trials, incorrect had significantly longer RTs (M = 2.8105 s, SD = 0.7250 s) than correct trials (M = 1.9949 s, SD = 0.3034 s). This highlights that the former involved significantly greater search effort which was reflected in the pupil response accordingly. Oppositely, RTs in rearranged were generally slower across both incorrect (M = 2.4083 s, SD = 0.4306 s), and correct (M = 2.1781 s, SD = 0.3280 s) trials with the difference significant but much smaller than in intact. That is, even successful retrieval was effortful in the rearranged condition, given the mismatch in encoding and retrieval cues. Therefore, this might explain the diminished differentiation in the pupil response.

Post-response, the pattern in both retrieval conditions was comparable. Similar to intact, pupil dilation in incorrect trials remained elevated while correct trials returned toward pre-stimulus levels. There was a significant cluster between 1.25 s to 3 s, post-response, $p = 0.009$.

Slow vs. fast responses – successful retrieval

To further explore whether pupil dilation reflects retrieval effort or memory strength, we used a median split for RTs to divide correct trials into slow and fast responses. Then we compared the corresponding pupil signal for all the three trial types (Figure 6). The stronger the memory traces, the easily more accessible they are and the faster the response time (Rotello & Zeng, 2008; Wixted & Mickes, 2010). In other words, if faster responses were associated with larger pupil dilations, there would be support for the memory strength account. By contrast, if slow responses, yielded larger pupil dilations, the retrieval effort would be supported.

For all conditions, we observed the divide in the pupil pattern between earlier and later phases of the time course. In intact (Figure 6A) and rearranged (Figure 6B), the early phase lasted for around 1.5 s with both slow and fast exhibiting the same increases in the pupil size until fast trials reached a sharp peak and commenced a steep return towards pre-stimulus levels. Simultaneously, slow trials continued to rise reaching a higher peak which was sustained for a little over 1 s before gradually returning towards pre-stimulus levels. Notably, in this phase, there was a significant difference between 2.06 s to 5 s post-stimulus, $p < 0.001$ for intact and between 2.18 to 5 s, $p < 0.001$, for rearranged.

In new trials (Figure 6C), the early phase also lasted for approximately 1.5 s. However, fast trials elicited bigger pupil increases and achieved a sharper peak earlier than slow trials, before declining towards pre-stimulus levels. Here, there was a significant cluster from 0.54 s to 1.83 s, post-stimulus, $p = 0.01$. Interestingly, while slow trials continued to rise, and thus in this period were more dilated, their peak was lower than that for fast trials. This was marked by two significant clusters: one from 2.17 s to 3.54 s, post-stimulus, $p = 0.006$ and another from 4.08 s to 5 s, post-stimulus, $p = 0.035$.

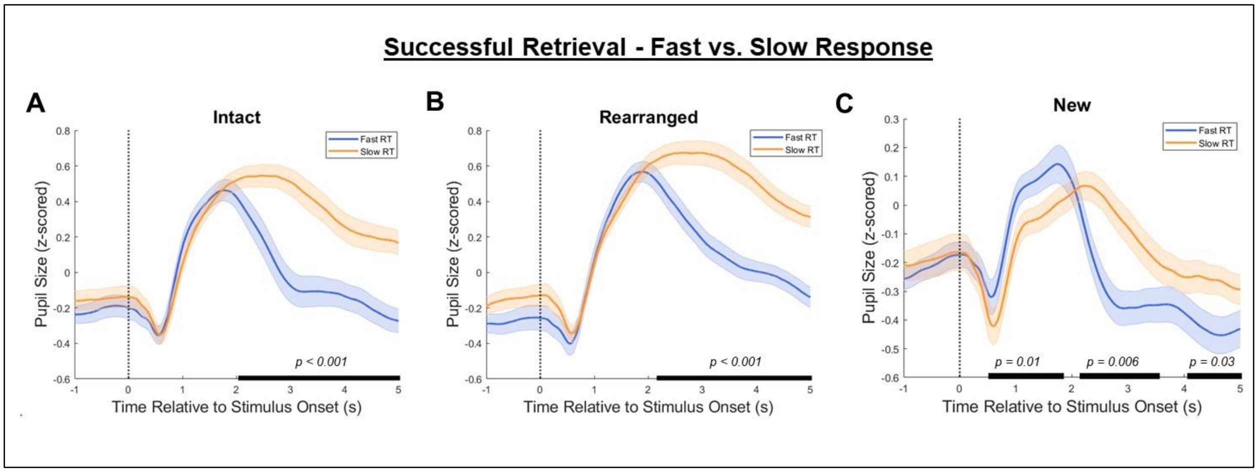

Figure 6*: Successful retrieval – fast vs. slow response for all three retrieval trial types. The pupil time course was characterised by two components. An early component where the pupil size for fast is equal to (****A**** and ****B****) or bigger than ****(C)**** that of slow responses and a later component where slow responses have significantly bigger pupil sizes.*

3.2. Encoding

The processing and statistical analyses of the behaviour and pupillometry data followed the same steps as those described in the retrieval section.

3.2.1. Behaviour results

In the encoding phase, only reaction times were recorded. Nevertheless, the paradigm was structured such that encoding trials (i.e., low, medium, and high plausible stimuli) could be retrospectively linked to memory accuracy. That is, encoding trials could be classified as later remembered or forgotten. This allowed us to examine whether the pupil response at encoding could predict subsequent memory performance (see section 3.2.2).

Reaction times

A repeated-measures ANOVA revealed a significant main effect of plausibility level on reaction time, $F(2, 50) = 24.77$, $p < 0.001$, $\eta^2_p = 0.498$. Post-hoc pairwise comparisons using paired-sample *t*-tests showed that medium plausible combinations (M = 2.495 s, SD = 0.700 s) had longer RTs compared to low (M = 2.182 s, SD = 0.771 s), $t(25) = 3.66$, $p = 0.001$, dz = 0.718, 95% CI [0.084, 0.299], and high (M = 1.994 s, SD = 0.621 s), $t(25) = 7.28$, $p < 0.0001$, dz = 1.428, 95% CI [0.263, 0.471]. Low was also significantly slower than high, $t(25) = 3.27$, $p = 0.003$, dz = 0.641, 95% CI [0.065, 0.286].

These results indicate that, as expected, stimuli rated as having medium plausibility elicited the longest response times (Figure 7A). This reflects the increased decision time due to the ambiguity of the stimuli. We did not have a hypothesised direction on the relationship between high and low plausibility as there are mixed results in the existing literature. Some studies find that the unexpected stimuli produce faster response times (e.g., Parmentier, 2024) as they stand out and quickly capture attention. While others find that stimuli that meet expectations are processed faster (e.g., Neely, 1977; Bar et al., 2006) due to top-down modulation and quick activation of relevant traces. Our finding that highly plausibility stimuli had faster RTs than low is consistent with the latter.

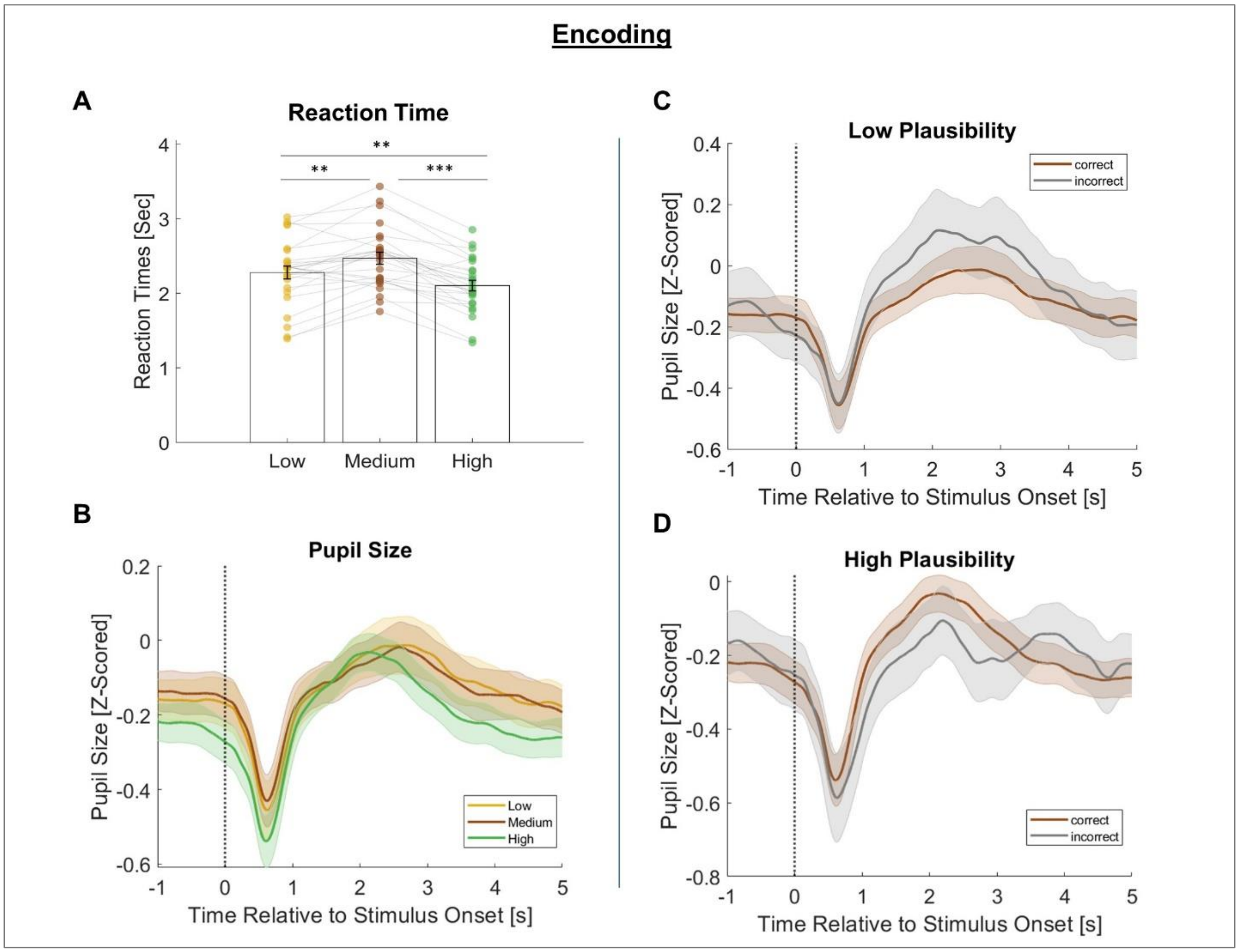

***Figure 7:** Encoding – behavioural and pupillometry results. Medium plausible stimuli had the longest RTs followed by low and high plausible stimuli **(A)**. There were no differences in the pupil size for expected (high) vs. unexpected (low) stimuli **(B)**. Pupil size at encoding did not predict memory performance in low **(C),** high **(D),** and medium (not shown) plausibility.*

3.2.2. Pupillometry

The pupil response did not differentiate between expected and unexpected stimuli

The processing of unexpected novel stimuli has been shown to rely on noradrenergic phasic activity. Therefore, we hypothesised that low plausible stimuli would engender more pupil dilation. However, we found no significant difference in the pupil size between these two conditions. Similarly, ambiguous stimuli in the form of medium plausible did not have any significant pupil size difference from the other conditions (Figure 7B).

Pupil response at encoding did not predict memory performance

Previous research has shown that the pupil response at encoding can predict later memory performance. However, there have been inconsistencies in the pupil response patterns (during encoding) that support memory. For example, some studies have found that pupil dilation enhances memory performance (e.g., Papesh et al., 2012; Bergt et al., 2018). While others have identified that pupil constriction enhances memory performance (Kafkas & Montaldi, 2011; Naber et al., 2013; Wetzel et al., 2020).

One factor that has been shown to modulate the pupil response, that support successful memory retrieval, is the expectedness of the stimuli (Kafkas, 2021; Kafkas et al. 2024). That is, how the stimuli either adheres to or violates the participant's expectation. Specifically, it has been hypothesised that increased pupil dilation is predictive of the memory performance for unexpected stimuli while for expected stimuli pupil constriction predicts memory performance.

To test this, we looked at whether the pupil size change differentially modulated memory performance based on the plausibility rating of the stimuli. In particular, we expected pupil dilation to enhance memory performance for "low plausible" stimuli whereas pupil constriction to enhance memory performance for "high plausible" stimuli. To do so, we compared the pupil signal for correct and incorrect trials in both conditions (Figure 7C & 7D). A cluster-based permutation comparison revealed no significant clusters in either condition. Interestingly, the pupil was numerically bigger for correct than incorrect trials in the low plausibility condition and smaller for correct than incorrect trials in high plausibility. To follow up on this, we conducted a repeated-measures ANOVA to test whether there was a significant interaction. No significant effect was found, $F(3, 66) = 0.80$, $p = 0.497$, $\eta^2_p = 0.035$.

3.3. Comparing Encoding and Retrieval

3.3.1. Stimulus-locked pupil response

Hasselmo's model posits that encoding and retrieval are supported by different neuromodulatory mechanisms. Given that the pupil response is an indirect measure of the underlying neuromodulation activity, we expected to see differences in the pupillometry signal for encoding and retrieval.

To test this, we compared encoding and retrieval in the new condition (Figure 8). For retrieval, we chose new trials for comparison because the task requirements for encoding and retrieval new are relatively similar. In both cases, the participant has to process new information. The only difference being that for retrieval, they then have to indicate that the information is new whereas for encoding, they need to process the information in accordance to the task's instructions. In other words, the difference in the pupil signal for these tasks would be attributed to being in a "retrieval" versus an "encoding" state. A cluster permutation test revealed that there was a more significant increase in the pupil size for retrieval compared to encoding (Figure 8A). To be specific, there was a positive significant cluster from 0.40 s to 2.08 sec, $p = 0.01$.

3.3.2. RT-locked pupil response

To further understand when exactly in the pupil time course this difference arose, we did a response-locked analysis (Figure 8B). This indicated that while both conditions elicited increases in the pupil size, retrieval resulted in greater pupil dilation for both pre- and post-response than encoding, with the most differences occurring ~500 ms post-response. Interestingly, encoding showed an earlier but more sustained pupil dilation peak and a relatively prolonged return to baseline. While, retrieval peaked slightly later with a sharper rise and fall.

3.3.3. Reaction times

To confirm that the observed differences were not simply indexing task engagement, we compared the reaction time for encoding and retrieval new (Figure 8C). Notably, encoding had significantly longer reaction times (M = 2.27 s, SD = 0.37 s) compared to retrieval new (M = 1.71 s, SD = 0.35 s), $t(25) = 9.53$, $p < 0.001$, Cohen's dz = 1.87, 95% CI [0.432, 0.671]. If the pupil response was only capturing task engagement, longer reaction times would have been associated with more pupil increase. However, we found the opposite pattern, which supports the argument that the pupil response is able to capture the underlying differences in the mechanisms that support encoding and retrieval.

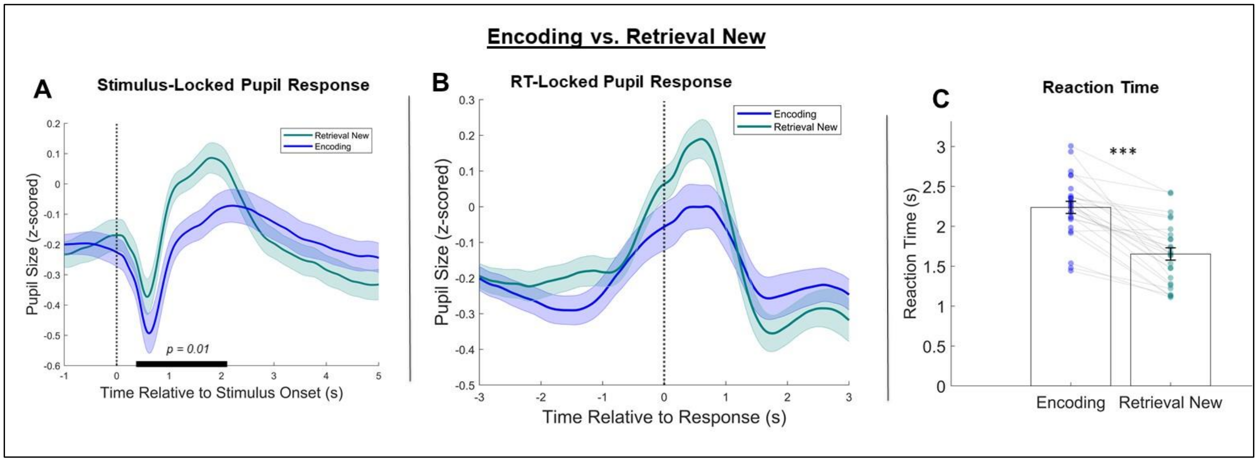

Figure 8: *Encoding vs. Retrieval New. Both stimulus-locked* ***(A)*** *and RT-locked* ***(B)*** *pupil time course revealed that retrieval was accompanied by larger dilations than encoding. This is despite encoding having a longer reaction time* ***(C)****.*

4. Discussion

Pupil dynamics have emerged as a promising measure of cognitive processing. This is especially the case for processes, such as memory encoding and retrieval, that are thought to be modulated by acetylcholine and norepinephrine (Sarter et al., 2004). These neuromodulators also control the pupil size (Aston-Jones & Cohen, 2005), and thus pupillary changes provide an indirect readout of their activity. The present study set out to interrogate whether the pupil response can (i) discriminate encoding from retrieval and (ii) differentiate long-term memory retrieval operations.

The goals of the study were threefold. Firstly, we wanted to test Hasselmo's hypothesis that the ability to switch back and forth between encoding and retrieval is supported by neuromodulators which exert differential effects on these two memory processes. Secondly, there have been competing theories as to whether pupil dilation reflects retrieval effort or memory strength. We wanted to examine which of these two accounts best explains the pupil response in recognition memory. Thirdly, we wanted to assess if pupillary changes do indeed index memory operations beyond mere task engagement. This is in light of criticisms which challenge the proposition that the pupil response can serve as a reliable proxy for cognitive processing.

In this section, we will address these goals in turn. Moreover, we will discuss the study's main findings, highlight how they relate and build on the existing literature, and provide suggestions for future research.

4.1. Interpretation and Implication of Findings

4.1.1. What mechanism supports the ability to switch between encoding and retrieval?

To tackle our first goal, we compared the pupil response for encoding and retrieval (Figure 8). We found that retrieval elicits greater pupil dilation than encoding, even when the task requirements are fairly matched across conditions. But, how do these results relate to Hasselmo's proposed cholinergic modulation of memory processes?

As reviewed in the introduction, Hasselmo (1995) and Hasselmo & McGaughy (2004) theorised that encoding is supported by high levels of acetylcholine which enhance external sensory inputs and suppress recurrent activity. Whereas, low acetylcholine levels facilitate retrieval by allowing internal associative processes (Figure 1). This has been largely tested in and supported by animal and computational models, but it is yet to be confirmed whether the same mechanism holds for the human memory system. Given that the pupil response is an indirect measure of neuromodulatory activity, whether this model generalises to humans cannot be uncovered from the pupil signal alone (see section 4.2 for further discussion). Nevertheless, below we present a proposal based on the results of the current study and other well-established findings in the extant literature.

In the central nervous system, cholinergic activity has been shown to correlate with pupillary changes (Reimer et al., 2016). However, pupil dilation is thought to be more closely linked to activity in the locus coeruleus-norepinephrine (LC-NE) system which monitors arousal and cognitive effort (Murphy et al., 2014). Compared to encoding, retrieval tasks require more top-down and internal associative information processing (Hasselmo & McGaughy, 2004). Therefore, we propose that when the brain shifts into a retrieval state, enabled by low acetylcholine, it expends cognitive energy to engage in the search for relevant memory traces and response-related decision making. This elevated demand recruits the LC-NE system more strongly, resulting in larger pupil dilations.

On the other hand, when the brain shifts into an encoding state, enabled by high acetylcholine, attention is oriented to external input whose processing may be less effortful. This would be especially the case when the target stimuli are commonly encountered items as was in our paradigm. As a result, encoding may elicit comparatively lower LC-NE activation, leading to smaller pupil dilations than retrieval.

Furthermore, there is accumulating evidence showing that pupil dilation in cognitive tasks is often accompanied by decreased activation in the visual and sensorimotor cortices (Murphy et al., 2014; Yellin et al., 2015; Schneider et al., 2016, DiNuzzo et al., 2019). This allows for attentional resources to be redirected toward higher-order cortical areas needed for the search of relevant information, interference resolution, and metacognitive monitoring. Notably, the above studies attributed the deactivation of these sensory cortices to cholinergic activity. In particular, DiNuzzo et al. (2019) suggested that it is due to a reduction in cholinergic input to the visual and sensorimotor areas, a view that is consistent with Hasselmo’s prediction of low acetylcholine levels during retrieval.

The findings of this study add to this growing literature showcasing the role of cholinergic activity in memory processes. Our observation that retrieval elicits a greater pupil dilation than encoding, even when task requirements are matched across conditions, indicates that the underlying difference in attention focus (external vs. internal) can trigger different neuromodulatory patterns. These are indexed by distinct pupil responses. Hence, this study suggests that Hasselmo's model may also apply in humans.

4.1.2. Does pupil dilation index effort or memory strength?

Next, we explored whether pupil dilation reflects retrieval effort or memory strength. To reiterate, the retrieval effort account of pupil dilation posits that the more difficult it is to retrieve a given event (regardless of success), the greater the increase in pupil size (Kahneman & Beatty, 1966; Kafkas et al., 2023). Traditionally, this has been supported by studies finding larger pupil dilations for more cognitively demanding conditions such as long study lists and study-test intervals in comparison to their short counterparts (Magliero, 1983; Piquado et al., 2010; van Rijn et al., 2012).

This is in contrast with the memory strength account which predicts more pupil dilation for stronger and easily accessible memories (Kucewicz et al., 2018). Evidence for this has typically come from dual process model (i.e., familiarity vs. recollection) studies showing that high-confidence hits and "remember" judgments exhibit larger increases in the pupil diameter than low-confidence hits, "know" responses, and misses (Papesh et al., 2011; Goldinger & Papesh, 2012; Kafkas & Montaldi, 2015, 2023).

Given these divergent results in the literature, the current study sought to investigate under what conditions each account holds. Specifically, we compared the pupil signal for: (i) conditions with varying difficulty levels namely short and long retention intervals and retrieval cues that either matched (intact) or did not match (rearranged) the encoding context, (ii) successful and unsuccessful retrieval, and (iii) fast and slow responses within successful retrieval. Additionally, we conducted stimulus-locked and response-locked analyses. Stimulus-locked comparisons are optimal for capturing memory strength driven changes following stimulus presentation. While, response-locked comparisons are best at tracking effort-related signals leading up to and after the response.

A consistent finding across the different comparisons was an emergence of two distinct components in the pupil time course. That is, an early phase that occurred roughly between 0 s – 2.5 s after stimulus onset and a later phase around 2.5 s – 5 s. In the early phase, conditions with the highest memory performance (i.e., correct, fast responses, and short retention intervals) showed a similar increase in the pupil size or numerically bigger, and in some cases significantly so, than that of their more cognitively demanding pairs. Response-locked analyses revealed that the early phase ended shortly after the response time with a sharp peak followed by a steep decrease toward pre-stimulus levels. The later phase, on the other hand, was characterised by continued increases in the pupil dilation and sustained peaks for the more cognitively demanding trials (i.e., incorrect, slow responses, and long retention intervals). This emerged post-response, resulting in significant differences in this period.

These temporal dynamics suggest that neither memory strength nor retrieval effort is sufficient to account for the pupil response in recognition memory. Hence, the mixed results in the literature. Rather, the pupil response reflects both, with the timing and amplitude signalling their relative contributions. To be more specific, the early component monitors memory strength with the increases in pupil size indexing the accessibility of target traces. On the other hand, the later component tracks how cognitively demanding a task is and post-response evaluative processes, with increases in pupil diameter proportionate to the cognitive effort expended.

Our findings add to a growing body of literature showing that the pupil response is comprised of temporally dissociable components that can index distinct memory operations (e.g., Kafkas & Montaldi, 2015; Mill et al., 2016; Oliveira et al., 2021). Simultaneously, there have been efforts to elucidate the neural correlates of memory strength and retrieval effort. These investigations have revealed that these two related but independent retrieval operations are subserved by distinct regions. Strong memories typically elicit greater activation in the hippocampus, posterior parietal cortex, and ventrolateral prefrontal regions (Liang & Preston, 2016; Ferreira et al., 2019). Whereas, retrieval effort is associated with activity in the frontoparietal regions, including the anterior prefrontal cortex and the anterior cingulate cortex (Schacter et al., 1996; Ranganath et al., 2000; Botvinick et al., 2004). Therefore, evidence that pupil measures can disentangle these processes, highlights the potential of pupillometry as a reliable proxy of neural activity.

4.1.3. How do our results relate to previous studies?

In encoding, unlike retrieval, the pupil response did not differentiate between trial types despite differences in behavioural outcomes. Stimuli rated as having medium plausibility elicited the longest RTs and high plausible stimuli exhibited significantly longer RTs than low plausible stimuli. Yet, there were no differences in the pupil response. The lack of pupillometric difference in the latter was surprising given the extensive literature showing differences in the pupil signal of expected and unexpected novelty (Kafkas, 2021; Kafkas et al. 2024).

On a closer look, this could have been due to the nature of our experimental design. In most previous studies, there was a direct manipulation of expectedness using various methods. Examples include inserting rare deviants in an ongoing pattern (Mill et al., 2016), omitting expected stimuli in a series (Dercksen et al., 2023), and changing stimulus–reward as well as cue-outcome contingencies (Pajkossy et al., 2023; Becker et al., 2024). In other words, the expectedness, or lack thereof, was directly linked to the stimuli.

Meanwhile, in our study, the plausibility ratings were tied to the imagined combination of the word and image pair. This might have reduced the magnitude of surprise. Additionally, even within the same plausibility category, different people could have imagined different scenarios with varying levels of expectedness. Together, these factors may have muted the typical phasic pupil response elicited by salient and unexpected stimuli. Future research can further parse apart the differences in the pupil response for real versus imagined events and explore appropriate paradigms to answer this question.

Furthermore, the pupil response at encoding did not predict later memory performance at retrieval. This is similar to null results reported in several other studies (e.g., Hämmerer et al., 2017, Gross and Dobbins, 2021, Unsworth and Miller, 2021, Lloyd & Nieuwenhuis, 2024). Yet, many more studies have found that pupil dynamics at encoding predict successful retrieval. This is known as the subsequent memory pupil effect and has been associated with both pupil constriction (Kafkas & Montaldi, 2011, Naber et al., 2013, Wetzel et al., 2020) and dilation (Papesh et al., 2012; Bergt et al., 2018; Kucewicz et al., 2018; Pajkossy & Racsmány, 2019; Cronin et al., 2023). Therefore, more research is needed to resolve the present inconsistencies in the literature. Specifically, there is need to explore under what conditions the subsequent memory pupil effect emerges and what determines the pupil pattern at encoding that predicts later memory performance.

4.2. Limitations and Future Directions

One of the major limitations of pupillometry, and by extension our study, is that the pupil signal is an indirect marker of neural activity. This makes it difficult to delineate the exact mechanisms driving the observed changes. For instance, we can infer that the later phase of the pupil time course was reflective of continued effort. However, the pupil alone would not inform us of the relative contributions of memory processing, decision uncertainty, and metacognitive monitoring. Thus, we encourage future studies to combine eye-tracking with brain imaging and/or electrophysiological measures for a more complete picture.

Additionally, one of the strengths of this study was that, unlike most recognition memory paradigms, we used a continuous associative task that could mimic the quick transitions between encoding and retrieval in daily life. Nevertheless, it is also far from replicating the complex multimodal contexts in real life. The increasing development of naturalist paradigms (Baldassano et al., 2017) and advancements in virtual reality (Thurley, 2022) present a prime opportunity for this endeavour. Therefore, future studies can continue to explore the potential of employing simultaneous virtual reality and eye-tracking to understand memory mechanisms.

4.3. Contributions to the Literature

In summarising our contributions to the literature, we will concurrently answer the question of whether the pupil indexes cognitive processing beyond task engagement.

Our results support the proposition that the pupil response does indeed reflect cognitive processes. In addition to differentiating between overlapping operations such as memory strength and retrieval effort, our pupillometry measures replicated three of the most well-supported findings in memory studies. First, we found that the pupil is sensitive to the differences in retrieval effort engendered by retrieval cues that match the encoding context versus those that do not. In thus doing, we showed a physiological correlate of the encoding specificity principle.

Second, we replicated the old/new effect showing that old stimuli elicited greater pupil dilations than new stimuli. The pupil old/new effect has been a consistent finding since it was first introduced by Vo et al. (2008) to the repertoire of the memory literature. Our results build on this prior work by expanding it beyond item recognition tasks used in previous studies (see Lapteva & Martarelli (2024) for review and meta-analysis of the extant studies on this effect). Instead, we employed an associative recognition task, demonstrating that the pupil old/new effect extends to relational binding.

Third, by varying the retention interval, we showed that the pupil could discriminate recently experienced events from those that were experienced further in the past. In particular, long retention intervals had slower RTs, less memory accuracy, and elicited larger increases in the pupil diameter, indicating that its more effortful to access memories as time passes. This replicated the classic effect of time delay on memory strength which was first introduced by Ebbinghaus (1885).

Finally, it can be argued that in all these three instances one condition requires more cognitive effort than the other, and by extension, requires more task engagement. This was in fact supported by our behavioural results showing longer RTs for rearranged (vs. intact), intact (vs. new), and long (vs. short) retention intervals. Thus, making it difficult to use the ensuing differences to fully disregard the criticism that perhaps pupil is just capturing task engagement. The differences observed for encoding and retrieval-new provide more decisive evidence. Encoding trials had longer RTs and thus participants were on task for much longer than for retrieval-new trials. Yet, we still see more pupil dilation for retrieval than encoding, highlighting that the pupil is sensitive to some of the underlying neurophysiological mechanisms that differentiates these two mechanisms.

5. Conclusion

To conclude, the current project explored the potential of the pupil response to act as a physiological correlate of memory encoding and retrieval. We showed that the pupil response can successfully discriminate between encoding and retrieval as well as parse apart different retrieval operations. This research is important because eye-tracking has become increasingly advanced and accessible, especially with the development of portable and affordable trackers (Fu et al., 2024; Niehorster et al., 2025). In this way, establishing pupillometric measures as reliable proxies of cognitive activity holds potential to broaden the reach and scope of experimental psychology. It would provide means to conduct ecological studies such as memory research in classrooms. In addition, it would make it more possible to carry out cognitive research in low-resource settings where access to tools like fMRI and EEG is limited and technical expertise is scarce.

References

Aigner, T. G., & Mishkin, M. (1986). The effects of physostigmine and scopolamine on recognition memory in monkeys. *Behavioral and Neural Biology*, *45*(1), 81–87. https://doi.org/10.1016/s0163-1047(86)80008-5

Aston-Jones, G., & Cohen, J. D. (2005). Adaptive gain and the role of the locus coeruleus–norepinephrine system in optimal performance. *The Journal of Comparative Neurology*, *493*(1), 99–110. https://doi.org/10.1002/cne.20723

Atkinson, R., & Shiffrin, R. (1968). Human Memory: A Proposed System and its Control Processes. The Psychology of learning and motivation (pp. 89–195). https://doi.org/10.1016/s0079-7421(08)60422-3

Bach, D. R., Hulme, O., Penny, W. D., & Dolan, R. J. (2015). The known unknowns: Neural representation of second-order uncertainty, and ambiguity. *Cerebral Cortex*, *25*(6), 1138–1147. https://doi.org/10.1093/cercor/bht317

Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering event structure in continuous narrative perception and memory. *Neuron*, *95*(3), 709-721.e5. https://doi.org/10.1016/j.neuron.2017.06.041

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., Hämäläinen, M. S., Marinkovic, K., Schacter, D. L., Rosen, B. R., & Halgren, E. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences*, *103*(2), 449–454. https://doi.org/10.1073/pnas.0507062103

Bergt, A., Urai, A. E., Donner, T. H., & Schwabe, L. (2018). Reading memory formation from the eyes. *European Journal of Neuroscience*, *47*(12), 1525–1533. https://doi.org/10.1111/ejn.13984

Biggs, J. B. (1987). *Student approaches to learning and studying*. https://ci.nii.ac.jp/ncid/BA67137366

Bone, M. B., Ahmad, F., & Buchsbaum, B. R. (2020). Feature-specific neural reactivation during episodic memory. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-15763-2

Bloem, B., Poorthuis, R. B., & Mansvelder, H. D. (2014). Cholinergic modulation of the medial prefrontal cortex: the role of nicotinic receptors in attention and regulation of neuronal activity. *Frontiers in Neural Circuits*, *8*. https://doi.org/10.3389/fncir.2014.00017

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences*, *8*(12), 539–546. https://doi.org/10.1016/j.tics.2004.10.003

Buckner, R. L., Koutstaal, W., Schacter, D. L., Wagner, A. D., & Rosen, B. R. (1998). Functional–Anatomic study of episodic retrieval using FMRI. I. Retrieval effort versus retrieval success. *NeuroImage*, *7*(3), 151–162. https://doi.org/10.1006/nimg.1998.0327

Campbell, R. A. A. (2024). shadedErrorBar (Version 1.65.0.4) [MATLAB code]. GitHub. https://github.com/raacampbell/shadedErrorBar

Choi, Y., Kim, K., Moon, S., Jung, G., Cha, J., & Yim, H. (2025). Context-dependent memory in the real world: the role of frequency and context dwell time. *Frontiers in Psychology*, *15*. https://doi.org/10.3389/fpsyg.2024.1489039

Cole, A. E., & Nicoll, R. A. (1984). The pharmacology of cholinergic excitatory responses in hippocampal pyramidal cells. *Brain Research*, *305*(2), 283–290. https://doi.org/10.1016/0006-8993(84)90434-7

Collins, L., Francis, J., Emanuel, B., & McCormick, D. A. (2023). Cholinergic and noradrenergic axonal activity contains a behavioral-state signal that is coordinated across the dorsal cortex. *eLife*, *12*. https://doi.org/10.7554/elife.81826

Cosquer, B., De Vasconcelos, A. P., Fröhlich, J., & Cassel, J. (2005). Blood–brain barrier and electromagnetic fields: Effects of scopolamine methylbromide on working memory after whole-body exposure to 2.45GHz microwaves in rats. *Behavioural Brain Research*, *161*(2), 229–237. https://doi.org/10.1016/j.bbr.2005.02.025

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology General*, *104*(3), 268–294. https://doi.org/10.1037/0096-3445.104.3.268

Culemann, W., Neuber, L., & Heine, A. (2023). Pupil vs. Eyelid: Evaluating the Accuracy of Blink Detection in Pupil-based Eye Tracking Devices. *Procedia Computer Science*, *225*, 2008–2017. https://doi.org/10.1016/j.procs.2023.10.191

Daumas, S., Halley, H., Francés, B., & Lassalle, J. M. (2005). Encoding, consolidation, and retrieval of contextual memory: Differential involvement of dorsal CA3 and CA1 hippocampal subregions. *Learning &Amp; Memory*, *12*(4), 375–382. https://doi.org/10.1101/lm.81905

Dercksen, T. T., Widmann, A., & Wetzel, N. (2023). Salient omissions—pupil dilation in response to unexpected omissions of sound and touch. *Frontiers in Psychiatry*, *14*. https://doi.org/10.3389/fpsyt.2023.1143931

De Rosa, E., & Hasselmo, M. E. (2000). Muscarinic cholinergic neuromodulation reduces proactive interference between stored odor memories during associative learning in rats. *Behavioral Neuroscience*, *114*(1), 32–41. https://doi.org/10.1037/0735-7044.114.1.32

DiNuzzo, M., Mascali, D., Moraschi, M., Bussu, G., Maugeri, L., Mangini, F., Fratini, M., & Giove, F. (2019). Brain networks underlying eye’s pupil dynamics. *Frontiers in Neuroscience*, *13*. https://doi.org/10.3389/fnins.2019.00965

Dolan, R. J., & Fletcher, P. F. (1999). Encoding and retrieval in human medial temporal lobes: An empirical investigation using functional magnetic resonance imaging (fMRI). Hippocampus, 9(1), 25–34. https://doi.org/10.1002/(SICI)1098-1063(1999)9:1<25::AID-HIPO3>3.0.CO;2-4

Ebbinghaus, H. (1885). Memory: A Contribution to Experimental Psychology. H. A. Ruger & C. E. Bussenius, Trans., Teachers College. https://psycnet.apa.org/fulltext/2004-16275-000-FRM.pdf

Ferreira, C., Charest, I., & Wimber, M. (2019). Retrieval aids the creation of a generalised memory trace and strengthens episode-unique information. *NeuroImage*, *201*, 115996. https://doi.org/10.1016/j.neuroimage.2019.07.009

Fu, X., Franchak, J. M., MacNeill, L. A., Gunther, K. E., Borjon, J. I., Yurkovic-Harding, J., Harding, S., Bradshaw, J., & Pérez-Edgar, K. E. (2024). Implementing mobile eye tracking in psychological research: A practical guide. *Behavior Research Methods*, *56*(8), 8269–8288. https://doi.org/10.3758/s13428-024-02473-6

Gimbel, S. I., & Brewer, J. B. (2010). Reaction time, memory strength, and fMRI activity during memory retrieval: Hippocampus and default network are differentially responsive during recollection and familiarity judgments. *Cognitive Neuroscience*, *2*(1), 19–26. https://doi.org/10.1080/17588928.2010.513770

Glanzer, M., Kim, K., Hilford, A., & Adams, J. K. (1999). Slope of the receiver-operating characteristic in recognition memory. *Journal of Experimental Psychology Learning Memory and Cognition*, *25*(2), 500–513. https://doi.org/10.1037/0278-7393.25.2.500

Ghoneim, M. M., & Mewaldt, S. P. (1975). Effects of diazepam and scopolamine on storage, retrieval and organizational processes in memory. *Psychopharmacology*, *44*(3), 257–262. https://doi.org/10.1007/bf00428903

Ghoneim, M. M., & Mewaldt, S. P. (1977). Studies on human memory: The interactions of diazepam, scopolamine, and physostigmine. *Psychopharmacology*, *52*(1), 1–6. https://doi.org/10.1007/bf00426592

Grujic, N., Polania, R., & Burdakov, D. (2024). Neurobehavioral meaning of pupil size. *Neuron*, *112*(20), 3381–3395. https://doi.org/10.1016/j.neuron.2024.05.029

Guskjolen, A., & Cembrowski, M. S. (2023). Engram neurons: Encoding, consolidation, retrieval, and forgetting of memory. *Molecular Psychiatry*, *28*(8), 3207–3219. https://doi.org/10.1038/s41380-023-02137-5

Gross, M. P., & Dobbins, I. G. (2020). Pupil dilation during memory encoding reflects time pressure rather than depth of processing. *Journal of Experimental Psychology Learning Memory and Cognition*, *47*(2), 264–281. https://doi.org/10.1037/xlm0000818

Papesh, M. H., Goldinger, S. D., & Hout, M. C. (2011). Memory strength and specificity revealed by pupillometry. *International Journal of Psychophysiology*, *83*(1), 56–64. https://doi.org/10.1016/j.ijpsycho.2011.10.002

Goldinger, S. D., & Papesh, M. H. (2012). Pupil dilation reflects the creation and retrieval of memories. *Current Directions in Psychological Science*, *21*(2), 90–95. https://doi.org/10.1177/0963721412436811

Gotardi, G. C., Rodrigues, S. T., Barbieri, F. A., Brito, M. B., Bonfim, J. V., & Polastri, P. F. (2020). Wearing a head-mounted eye tracker may reduce body sway. *Neuroscience Letters*, *722*, 134799. https://doi.org/10.1016/j.neulet.2020.134799

Hämmerer, D., Hopkins, A., Betts, M. J., Maaß, A., Dolan, R. J., & Düzel, E. (2017). Emotional arousal and recognition memory are differentially reflected in pupil diameter responses during emotional memory for negative events in younger and older adults. *Neurobiology of Aging*, *58*, 129–139. https://doi.org/10.1016/j.neurobiolaging.2017.06.021

Hansen, N. (2017). The longevity of Hippocampus-Dependent memory is orchestrated by the locus Coeruleus-Noradrenergic system. *Neural Plasticity*, *2017*, 1–9. https://doi.org/10.1155/2017/2727602

Hansen, D. W., & Hammoud, R. I. (2006). An improved likelihood model for eye tracking. *Computer Vision and Image Understanding*, *106*(2–3), 220–230. https://doi.org/10.1016/j.cviu.2006.06.012

Hayat, H., Regev, N., Matosevich, N., Sales, A., Paredes-Rodriguez, E., Krom, A. J., Bergman, L., Li, Y., Lavigne, M., Kremer, E. J., Yizhar, O., Pickering, A. E., & Nir, Y. (2020). Locus coeruleus norepinephrine activity mediates sensory-evoked awakenings from sleep. *Science Advances*, *6*(15). https://doi.org/10.1126/sciadv.aaz4232

Hasselmo, M. E. (1995). Neuromodulation and cortical function: modeling the physiological basis of behavior. *Behavioural Brain Research*, *67*(1), 1–27. https://doi.org/10.1016/0166-4328(94)00113-t

Hasselmo, M. E., & Bower, J. M. (1993). Acetylcholine and memory. *Trends in Neurosciences*, *16*(6), 218–222. https://doi.org/10.1016/0166-2236(93)90159-j

Hasselmo, M. E., & McGaughy, J. (2004). High acetylcholine levels set circuit dynamics for attention and encoding and low acetylcholine levels set dynamics for consolidation. *Progress in Brain Research*, 207–231. https://doi.org/10.1016/s0079-6123(03)45015-2

Hasselmo, M., & Schnell, E. (1994). Laminar selectivity of the cholinergic suppression of synaptic transmission in rat hippocampal region CA1: computational modeling and brain slice physiology. *Journal of Neuroscience*, *14*(6), 3898–3914. https://doi.org/10.1523/jneurosci.14-06-03898.1994

Headley, D. B. (1981). Pupillometric assessment of retrieval operations in factual long-term memory. *Acta Psychologica*, *49*(2), 109–126. https://doi.org/10.1016/0001-6918(81)90023-8

Hockley, W. E. (1982). Retrieval processes in continuous recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*(6), 497–512. https://doi.org/10.1037/0278-7393.8.6.497

Hopstaken, J. F., Van Der Linden, D., Bakker, A. B., & Kompier, M. A. (2015). The window of my eyes: Task disengagement and mental fatigue covary with pupil dynamics. *Biological Psychology*, *110*, 100–106. https://doi.org/10.1016/j.biopsycho.2015.06.013

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology*, *46*(3), 269–299. https://doi.org/10.1006/jmps.2001.1388

Isabella, S. L., Cheyne, J. A., & Cheyne, D. (2021). Inhibitory control in the absence of awareness: interactions between frontal and motor cortex oscillations mediate implicitly learned responses. *Frontiers in Human Neuroscience, 15*. https://doi.org/10.3389/fnhum.2021.786035

Ishibashi, K., Umeda, S., & Koike, T. (2020). Enhancement of ambiguous object recognition by anodal tDCS over the left prefrontal cortex. *Frontiers in Human Neuroscience, 14*, 36. https://doi.org/10.3389/fnhum.2020.00036

Jainta, S. (2011). The pupil reflects motor preparation for saccades – even before the eye starts to move. *Frontiers in Human Neuroscience*, *5*. https://doi.org/10.3389/fnhum.2011.00097

Jones, R. S. (1993). Entorhinal-hippocampal connections: a speculative view of their function. *Trends in Neurosciences, 16*(2), 58–64. https://doi.org/10.1016/0166-2236(93)90018-h

Johansson, R., Pärnamets, P., Bjernestedt, A., & Johansson, M. (2018). Pupil dilation tracks the dynamics of mnemonic interference resolution. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-23297-3

Lapteva, A., & Martarelli, C. S. (2024). Pupil old/new effect as an objective measure of recognition memory: a meta-analysis of 17 eye-tracking experiments. *Memory*, *32*(9), 1211–1226. https://doi.org/10.1080/09658211.2024.2395367

Lohnas, L. J., Polyn, S. M., & Kahana, M. J. (2015). Expanding the scope of memory search: Modeling intralist and interlist effects in free recall. *Psychological Review*, *122*(2), 337–363. https://doi.org/10.1037/a0039036

Kafkas, A. (2021). Encoding-linked pupil response is modulated by expected and unexpected novelty: Implications for memory formation and neurotransmission. *Neurobiology of Learning and Memory*, *180*, 107412. https://doi.org/10.1016/j.nlm.2021.107412

Kafkas, A., Brown, T., Olusola, N., & Guo, C. (2023). Pupil response patterns distinguish true from false memories. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-44362-6

Kafkas, A., & Montaldi, D. (2015). The pupillary response discriminates between subjective and objective familiarity and novelty. *Psychophysiology*, *52*(10), 1305–1316. https://doi.org/10.1111/psyp.12471

Kafkas, A., & Montaldi, D. (2012). Familiarity and recollection produce distinct eye movement, pupil and medial temporal lobe responses when memory strength is matched. *Neuropsychologia*, *50*(13), 3080–3093. https://doi.org/10.1016/j.neuropsychologia.2012.08.001

Kafkas, A., & Montaldi, D. (2011). Recognition Memory Strength is Predicted by Pupillary Responses at Encoding While Fixation Patterns Distinguish Recollection from Familiarity. *Quarterly Journal of Experimental Psychology*, *64*(10), 1971–1989. https://doi.org/10.1080/17470218.2011.588335

Kesner, R. P., & Rolls, E. T. (2014). A computational theory of hippocampal function, and tests of the theory: New developments. *Neuroscience & Biobehavioral Reviews*, *48*, 92–147. https://doi.org/10.1016/j.neubiorev.2014.11.009

Kirschner, P.A., Kirschner, F. (2012). Mental Effort. In: Seel, N.M. (eds) Encyclopedia of the Sciences of Learning. Springer, Boston, MA. https://doi.org/10.1007/978-1-4419-1428-6_226

Kloosterman, N. A., Meindertsma, T., Van Loon, A. M., Lamme, V. a. F., Bonneh, Y. S., & Donner, T. H. (2015). Pupil size tracks perceptual content and surprise. *European Journal of Neuroscience*, *41*(8), 1068–1078. https://doi.org/10.1111/ejn.12859

Knapen, T., de Gee, J. W., Brascamp, J., Nuiten, S., Hoppenbrouwers, S., & Theeuwes, J. (2016). Cognitive and ocular factors jointly determine pupil responses under equiluminance. *PloS one*, *11*(5), e0155574.

Kolnes, M., Uusberg, A., & Nieuwenhuis, S. (2023). Broadening of attention dilates the pupil. *Attention Perception & Psychophysics*, *86*(1), 146–158. https://doi.org/10.3758/s13414-023-02793-3

Kozak, R., Bruno, J. P., & Sarter, M. (2005). Augmented Prefrontal Acetylcholine Release during Challenged Attentional Performance. *Cerebral Cortex*, *16*(1), 9–17. https://doi.org/10.1093/cercor/bhi079

Krnjević, K., Pumain, R., & Renaud, L. (1971). The mechanism of excitation by acetylcholine in the cerebral cortex. *The Journal of Physiology*, *215*(1), 247–268. https://doi.org/10.1113/jphysiol.1971.sp009467

Kuhl, B. A., Dudukovic, N. M., Kahn, I., & Wagner, A. D. (2007). Decreased demands on cognitive control reveal the neural processing benefits of forgetting. *Nature Neuroscience*, *10*(7), 908–914. https://doi.org/10.1038/nn1918

Kucewicz, M. T., Dolezal, J., Kremen, V., Berry, B. M., Miller, L. R., Magee, A. L., Fabian, V., & Worrell, G. A. (2018). Pupil size reflects successful encoding and recall of memory in humans. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-23197-6

Larsen, R. S., & Waters, J. (2018). Neuromodulatory correlates of pupil dilation. *Frontiers in Neural Circuits*, *12*. https://doi.org/10.3389/fncir.2018.00021

Liang, J. C., & Preston, A. R. (2016b). Medial temporal lobe reinstatement of content-specific details predicts source memory. *Cortex*, *91*, 67–78. https://doi.org/10.1016/j.cortex.2016.09.011

Lloyd, B., & Nieuwenhuis, S. (2024). The effect of reward-induced arousal on the success and precision of episodic memory retrieval. *Scientific Reports*, *14*(1). https://doi.org/10.1038/s41598-024-52486-6

Loewenfeld, I. E. (1958). Mechanisms of reflex dilatation of the pupil; historical review and experimental analysis. Doc. Ophthalmol. 12, 185–448. doi: 10.1007/BF00913471

Lutz, S. T., & Huitt, W. G. (2003). Information Processing and Memory: Theory and applications. In Valdosta State University, *Educational Psychology Interactive*. Valdosta State University. https://www.edpsycinteractive.org/papers/infoproc.pdf

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Martin, J. T., Whittaker, A. H., & Johnston, S. J. (2022). Pupillometry and the vigilance decrement: Task-evoked but not baseline pupil measures reflect declining performance in visual vigilance tasks. *European Journal of Neuroscience*, *55*(3), 778–799. https://doi.org/10.1111/ejn.15585

Mathôt, S. (2018). Pupillometry: psychology, physiology, and function. *Journal of Cognition*, *1*(1). https://doi.org/10.5334/joc.18

Mathôt, S., Dalmaijer, E., Grainger, J., & Van Der Stigchel, S. (2014). The pupillary light response reflects exogenous attention and inhibition of return. *Journal of Vision*, *14*(14), 7. https://doi.org/10.1167/14.14.7

Mathôt, S., Fabius, J., Van Heusden, E., & Van Der Stigchel, S. (2018). Safe and sensible preprocessing and baseline correction of pupil-size data. *Behavior Research Methods*, *50*(1), 94–106. https://doi.org/10.3758/s13428-017-1007-2

Mathôt, S., & Vilotijević, A. (2023). Methods in cognitive pupillometry: Design, preprocessing, and statistical analysis. *Behavior Research Methods*, *55*(6), 3055–3077. https://doi.org/10.3758/s13428-022-01957-7

May, P. J., Reiner, A., & Gamlin, P. D. (2019). Autonomic regulation of the eye. *Oxford Research Encyclopedia of Neuroscience*. https://doi.org/10.1093/acrefore/9780190264086.013.276

Miles, W. R. (1930). Ocular dominance in human adults. *The Journal of General Psychology*, *3*(3), 412–430. https://doi.org/10.1080/00221309.1930.9918218

Mill, R. D., O’Connor, A. R., & Dobbins, I. G. (2016). Pupil dilation during recognition memory: Isolating unexpected recognition from judgment uncertainty. *Cognition*, *154*, 81–94. https://doi.org/10.1016/j.cognition.2016.05.018

Miller, G. A. (2003). The cognitive revolution: a historical perspective. Trends in Cognitive Sciences, 7(3), 141–144. https://doi.org/10.1016/s1364-6613(03)00029-9

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63(2), 81–97. https://doi.org/10.1037/h0043158

Murphy, P. R., O'Connell, R. G., O'Sullivan, M., Robertson, I. H., & Balsters, J. H. (2014). Pupil diameter covaries with BOLD activity in human locus coeruleus. *Human Brain Mapping*, *35*(8), 4140–4154. https://doi.org/10.1002/hbm.22466

Naber, M., Frassle, S., Rutishauser, U., & Einhauser, W. (2013). Pupil size signals novelty and predicts later retrieval success for declarative memories of natural scenes. *Journal of Vision*, *13*(2), 11. https://doi.org/10.1167/13.2.11

Naicker, P., Anoopkumar-Dukie, S., Grant, G. D., Neumann, D. L., & Kavanagh, J. J. (2016). Central cholinergic pathway involvement in the regulation of pupil diameter, blink rate and cognitive function. *Neuroscience*, *334*, 180–190. https://doi.org/10.1016/j.neuroscience.2016.08.009

Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasly, B., & Gold, J. I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature Neuroscience*, *15*(7), 1040–1046. https://doi.org/10.1038/nn.3130

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology General*, *106*(3), 226–254. https://doi.org/10.1037/0096-3445.106.3.226

Niehorster, D. C., Nyström, M., Hessels, R. S., Andersson, R., Benjamins, J. S., Hansen, D. W., & Hooge, I. T. C. (2025). The fundamentals of eye tracking part 4: Tools for conducting an eye tracking study. *Behavior Research Methods*, *57*(1). https://doi.org/10.3758/s13428-024-02529-7

Nyström, M., Andersson, R., Niehorster, D. C., Hessels, R. S., & Hooge, I. T. C. (2024). What is a blink? Classifying and characterizing blinks in eye openness signals. *Behavior Research Methods*, *56*(4), 3280–3299. https://doi.org/10.3758/s13428-023-02333-9

Oliveira, J., Fernandes, M., Rosa, P. J., & Gamito, P. (2021). Is pupil activity associated with the strength of memory signal for words in a continuous recognition memory paradigm? *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.686183

O'Reilly, R. C., & McClelland, J. L. (1994). Hippocampal conjunctive encoding, storage, and recall: Avoiding a trade-off. *Hippocampus*, *4*(6), 661–682. https://doi.org/10.1002/hipo.450040605

Oster, J., Huang, J., White, B. J., Radach, R., Itti, L., Munoz, D. P., & Wang, C. (2022). Pupillary responses to differences in luminance, color and set size. *Experimental Brain Research*, *240*(6), 1873–1885. https://doi.org/10.1007/s00221-022-06367-x

Pajkossy, P., Gesztesi, G., & Racsmány, M. (2023). How uncertain are you? Disentangling expected and unexpected uncertainty in pupil-linked brain arousal during reversal learning. *Cognitive Affective & Behavioral Neuroscience*, *23*(3), 578–599. https://doi.org/10.3758/s13415-023-01072-w

Pajkossy, P., & Racsmány, M. (2019). How the size of the to-be-learned material influences the encoding and later retrieval of associative memories: A pupillometric assessment. *PLoS ONE*, *14*(12), e0226684. https://doi.org/10.1371/journal.pone.0226684

Palacios-Filardo, J., Udakis, M., Brown, G. A., Tehan, B. G., Congreve, M. S., Nathan, P. J., Brown, A. J. H., & Mellor, J. R. (2021). Acetylcholine prioritises direct synaptic inputs from entorhinal cortex to CA1 by differential modulation of feedforward inhibitory circuits. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-25280-5

Papesh, M. H., Goldinger, S. D., & Hout, M. C. (2012). Memory strength and specificity revealed by pupillometry. *International Journal of Psychophysiology*, *83*(1), 56–64. https://doi.org/10.1016/j.ijpsycho.2011.10.002

Park, S., Kang, T., & Koo, D. (2024). Investigating eye movement and Postural stability relationships using mobile Eye-Tracking and Posturography: a Cross-Sectional study. *Bioengineering*, *11*(8), 742. https://doi.org/10.3390/bioengineering11080742

Parmentier, F. B. R., English, M., & Maybery, M. T. (2024). Auditory context-dependent distraction by unexpected visual stimuli. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-024-02527-y

Patil, M. M., & Hasselmo, M. E. (1999). Modulation of inhibitory synaptic potentials in the piriform cortex. *Journal of Neurophysiology*, *81*(5), 2103–2118. https://doi.org/10.1152/jn.1999.81.5.2103

Petersen, R. C. (1977). Scopolamine induced learning failures in man. *Psychopharmacology*, *52*(3), 283–289. https://doi.org/10.1007/bf00426713

Piquado, T., Isaacowitz, D., & Wingfield, A. (2010). Pupillometry as a measure of cognitive effort in younger and older adults. *Psychophysiology*, *47*(3), 560–569. https://doi.org/10.1111/j.1469-8986.2009.00947.x

Poe, G. R., Foote, S., Eschenko, O., Johansen, J. P., Bouret, S., Aston-Jones, G., Harley, C. W., Manahan-Vaughan, D., Weinshenker, D., Valentino, R., Berridge, C., Chandler, D. J., Waterhouse, B., & Sara, S. J. (2020). Locus coeruleus: a new look at the blue spot. *Nature Reviews. Neuroscience*, *21*(11), 644–659. https://doi.org/10.1038/s41583-020-0360-9

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review*, *116*(1), 129–156. https://doi.org/10.1037/a0014420

Porter, G., Troscianko, T., & Gilchrist, I. D. (2006). Effort during visual search and counting: Insights from pupillometry. *Quarterly Journal of Experimental Psychology*, *60*(2), 211–229. https://doi.org/10.1080/17470210600673818

Preuschoff, K. (2011). Pupil dilation signals surprise: evidence for noradrenaline’s role in decision making. *Frontiers in Neuroscience*, *5*. https://doi.org/10.3389/fnins.2011.00115

Ranganath, C., Johnson, M. K., & D'Esposito, M. (2000). Left anterior prefrontal activation increases with demands to recall specific perceptual information. J Neurosci, 20(22), RC108.

Ritchey, M., Wing, E. A., LaBar, K. S., & Cabeza, R. (2012b). Neural Similarity Between Encoding and Retrieval is Related to Memory Via Hippocampal Interactions. *Cerebral Cortex*, *23*(12), 2818–2828. https://doi.org/10.1093/cercor/bhs258

Robb, W. G. K., & Rugg, M. D. (2002). Electrophysiological dissociation of retrieval orientation and retrieval effort. *Psychonomic Bulletin & Review*, *9*(3), 583–589. https://doi.org/10.3758/bf03196316

Rotello, C. M. (2017). Signal detection theories of recognition memory. In *Elsevier eBooks* (pp. 201–225). https://doi.org/10.1016/b978-0-12-809324-5.21044-4

Rotello, C. M., & Zeng, M. (2008). Analysis of RT distributions in the remember—know paradigm. *Psychonomic Bulletin & Review*, *15*(4), 825–832. https://doi.org/10.3758/pbr.15.4.825

Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, *11*(6), 251–257. https://doi.org/10.1016/j.tics.2007.04.004

Rugg, M. D., Johnson, J. D., & Uncapher, M. R. (2015). Encoding and retrieval in episodic memory: Insights from fMRI. In D. R. Addis, M. Barense, & A. Duarte (Eds.), *The Wiley handbook on the cognitive neuroscience of memory* (pp. 84–107). Wiley Blackwell. https://doi.org/10.1002/9781118332634.ch5

Ryals, A. J., Kelly, M. E., & Cleary, A. M. (2021). Increased pupil dilation during tip-of-the-tongue states. *Consciousness and Cognition*, *92*, 103152. https://doi.org/10.1016/j.concog.2021.103152

Sara, S. J. (2009). The locus coeruleus and noradrenergic modulation of cognition. *Nature Reviews. Neuroscience*, *10*(3), 211–223. https://doi.org/10.1038/nrn2573

Sarter, M., Hasselmo, M. E., Bruno, J. P., & Givens, B. (2004). Unraveling the attentional functions of cortical cholinergic inputs: interactions between signal-driven and cognitive modulation of signal detection. *Brain Research Reviews*, *48*(1), 98–111. https://doi.org/10.1016/j.brainresrev.2004.08.006

Sarter, M., Givens, B., & Bruno, J. P. (2001). The cognitive neuroscience of sustained attention: where top-down meets bottom-up. *Brain Research Reviews*, *35*(2), 146–160. https://doi.org/10.1016/s0165-0173(01)00044-3

Schacter, D. L. (1999). The seven sins of memory: Insights from psychology and cognitive neuroscience. *American Psychologist*, 54(3), 182–203. https://doi.org/10.1037/0003-066X.54.3.1824

Schacter, D. L., Alpert, N. M., Savage, C. R., Rauch, S. L., & Albert, M. S. (1996). Conscious recollection and the human hippocampal formation: evidence from positron emission tomography. Proc Natl Acad Sci U S A, 93(1), 321-325.

Schneider, M., Hathway, P., Leuchs, L., Sämann, P. G., Czisch, M., & Spoormaker, V. I. (2016). Spontaneous pupil dilations during the resting state are associated with activation of the salience network. *NeuroImage*, *139*, 189–201. https://doi.org/10.1016/j.neuroimage.2016.06.011

Schwetlick, L., Graupner, H., Dimigen, O., & Engbert, R. (2025). Distinctive pupil and microsaccade-rate signatures in self-recognition. *Journal of Vision*, *25*(4), 16. https://doi.org/10.1167/jov.25.4.16

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. Journal of Neurology Neurosurgery & Psychiatry, 20(1), 11–21. https://doi.org/10.1136/jnnp.20.1.11

Sibley, C., Foroughi, C. K., Brown, N. L., Phillips, H., Drollinger, S., Eagle, M., & Coyne, J. T. (2020). More than Means: Characterizing Individual Differences in Pupillary Dilations. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *64*(1), 57–61. https://doi.org/10.1177/1071181320641017

Squire, L. R. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology of Learning and Memory*, *82*(3), 171–177. https://doi.org/10.1016/j.nlm.2004.06.005

Staresina, B. P., Michelmann, S., Bonnefond, M., Jensen, O., Axmacher, N., & Fell, J. (2016). Hippocampal pattern completion is linked to gamma power increases and alpha power decreases during recollection. *ELife*, *5*. https://doi.org/10.7554/elife.17397

Staudigl, T., & Hanslmayr, S. (2019). Reactivation of neural patterns during memory reinstatement supports encoding specificity. *Cognitive Neuroscience*, *10*(4), 175–185. https://doi.org/10.1080/17588928.2019.1621825

Stigliani, A., Weiner, K. S., & Grill-Spector, K. (2015). Temporal processing capacity in high-level visual cortex is domain specific. *Journal of Neuroscience*, *35*(36), 12412-12424.

St Peters, M., Demeter, E., Lustig, C., Bruno, J. P., & Sarter, M. (2011). Enhanced control of attention by stimulating Mesolimbic-Corticopetal cholinergic circuitry. *Journal of Neuroscience*, *31*(26), 9760–9771. https://doi.org/10.1523/jneurosci.1902-11.2011

Straube, B. (2012). An overview of the neuro-cognitive processes involved in the encoding, consolidation, and retrieval of true and false memories. *Behavioral and Brain Functions*, *8*(1), 35. https://doi.org/10.1186/1744-9081-8-35

Taikh, A., & Bodner, G. E. (2022). Pupil dilation during recognition reflects the subjective recollection/familiarity experience at test rather than the level of processing at encoding. *Canadian Journal of Experimental Psychology/Revue Canadienne De Psychologie Expérimentale*, *76*(3), 186–192. https://doi.org/10.1037/cep0000283

Thurley, K. (2022). Naturalistic neuroscience and virtual reality. *Frontiers in Systems Neuroscience*, *16*. https://doi.org/10.3389/fnsys.2022.896251

Rolls, E. T., Treves, A., Foster, D., & Perez-Vicente, C. (1996). Simulation studies of the CA3 hippocampal subfield modelled as an attractor neural network. *Neural Networks*, *10*(9), 1559–1569. https://doi.org/10.1016/s0893-6080(97)00092-0

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson, Organization of memory. Academic Press.

Tulving, E., & Thomson, D. M. (1973). *Encoding specificity and retrieval processes in episodic memory*. *Psychological Review*, 80(5), 352–373.

Tyukhova, Y., & Waters, C. E. (2018). Discomfort Glare from Small, High-Luminance Light Sources When Viewed against a Dark Surround. *LEUKOS the Journal of the Illuminating Engineering Society of North America*, *14*(4), 215–230. https://doi.org/10.1080/15502724.2018.1434415

Wetzel, N., Einhäuser, W., & Widmann, A. (2020). Picture-evoked changes in pupil size predict learning success in children. *Journal of Experimental Child Psychology*, *192*, 104787. https://doi.org/10.1016/j.jecp.2019.104787

Wierda, S. M., Van Rijn, H., Taatgen, N. A., & Martens, S. (2012). Pupil dilation deconvolution reveals the dynamics of attention at high temporal resolution. *Proceedings of the*

National Academy of Sciences, *109*(22), 8456–8460. https://doi.org/10.1073/pnas.1201858109

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*, *114*(1), 152–176. https://doi.org/10.1037/0033-295x.114.1.152

Wixted, J. T., & Mickes, L. (2010). A continuous dual-process model of remember/know judgments. *Psychological Review*, *117*(4), 1025–1054. https://doi.org/10.1037/a0020874

Wixted, J. T., & Rohrer, D. (1993). Proactive interference and the dynamics of free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(5), 1024–1039. https://doi.org/10.1037/0278-7393.19.5.1024

Wheeler, M. E., Petersen, S. E., & Buckner, R. L. (2000). Memory's echo: Vivid remembering reactivates sensory-specific cortex. *Proceedings of the National Academy of Sciences*, *97*(20), 11125–11129. https://doi.org/10.1073/pnas.97.20.11125

Unsworth, N., & Miller, A. L. (2021). Encoding dynamics in free recall: Examining attention allocation with pupillometry. *Memory & Cognition*, *49*(1), 90–111. https://doi.org/10.3758/s13421-020-01077-7

Urai, A. E., Braun, A., & Donner, T. H. (2016). Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. *Nature Communications*, *8*(1). https://doi.org/10.1038/ncomms14637

Van Der Wel, P., & Van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychonomic Bulletin & Review*, *25*(6), 2005–2015. https://doi.org/10.3758/s13423-018-1432-y

Van Rijn, H., Dalenberg, J. R., Borst, J. P., & Sprenger, S. A. (2012). Pupil Dilation Co-Varies with Memory Strength of Individual Traces in a Delayed Response Paired-Associate Task. *PLoS ONE*, *7*(12), e51134. https://doi.org/10.1371/journal.pone.0051134

Viglione, A., Mazziotti, R., & Pizzorusso, T. (2023). From pupil to the brain: New insights for studying cortical plasticity through pupillometry. *Frontiers in Neural Circuits*, *17*. https://doi.org/10.3389/fncir.2023.1151847

Võ, M. L., Jacobs, A. M., Kuchinke, L., Hofmann, M., Conrad, M., Schacht, A., & Hutzler, F. (2008). The coupling of emotion and cognition in the eye: Introducing the pupil old/new effect. *Psychophysiology*, *45*(1), 130–140. https://doi.org/10.1111/j.1469-8986.2007.00606.x

Wang, C., Boehnke, S. E., Itti, L., & Munoz, D. P. (2014). Transient pupil response is modulated by Contrast-Based saliency. *Journal of Neuroscience*, *34*(2), 408–417. https://doi.org/10.1523/jneurosci.3550-13.2014

Wierda, S. M., Van Rijn, H., Taatgen, N. A., & Martens, S. (2012). Pupil dilation deconvolution reveals the dynamics of attention at high temporal resolution. *Proceedings of the National Academy of Sciences*, *109*(22), 8456–8460. https://doi.org/10.1073/pnas.1201858109

Yellin, D., Berkovich-Ohana, A., & Malach, R. (2014). Coupling between pupil fluctuations and resting-state fMRI uncovers a slow build-up of antagonistic responses in the human cortex. *NeuroImage*, *106*, 414–427. https://doi.org/10.1016/j.neuroimage.2014.11.034

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

Yoo, K., Ahn, J., & Lee, S. (2021). The confounding effects of eye blinking on pupillometry, and their remedy. *PLoS ONE*, *16*(12), e0261463. https://doi.org/10.1371/journal.pone.0261463