Pupil-linked arousal counteracts the positive effects of reward anticipation on incidental memory encoding

Lloyd, B.; Miletić, S.; Nieuwenhuis, S.T.

Citation

Lloyd, B., Miletić, S., & Nieuwenhuis, S. T. (2025). Pupil-linked arousal counteracts the positive effects of reward anticipation on incidental memory encoding. *Cognition*, *264*. doi:10.1016/j.cognition.2025.106237

Cognition 264 (2025) 106237

Contents lists available at ScienceDirect

Cognition

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

Pupil-linked arousal counteracts the positive effects of reward anticipation on incidental memory encoding

Beth Lloyd[a,*], Steven Miletić[a,b], Sander Nieuwenhuis[a]

[a] Institute of Psychology, Leiden University, Leiden, Netherlands
[b] Integrative Model-Based Cognitive Neuroscience Research Unit, University of Amsterdam, Amsterdam, Netherlands

ARTICLE INFO

Keywords:
Pupillometry
Arousal
Reward anticipation
Episodic memory

ABSTRACT

The anticipation of a potential reward has been shown to enhance episodic memory, but the underlying mechanisms remain unclear. While previous research has highlighted the role of attention in improving memory for reward-associated stimuli, the impact of arousal during encoding has been less explored. In this study, we used a monetary incentive delay task followed by a surprise memory test 24 h later to examine whether pupil-linked arousal mediates the beneficial effect of reward anticipation on memory. Contrary to our expectations, we found that higher arousal during encoding impaired subsequent memory performance, counteracting the dominant positive direct effect of reward anticipation on memory. This result suggests that, rather than facilitating memory, increased arousal during reward anticipation may disrupt encoding, possibly by heightening decision urgency. Our findings offer new insights into the complex interplay between reward anticipation, arousal, and memory encoding, highlighting the importance of considering time pressure as a potential influencing factor when studying the central arousal system.

1. Introduction

Research has shown that the anticipation of a potential reward can enhance the encoding of reward-associated stimuli in episodic memory (Miendlarzewska et al., 2016). The evidence is based on two lines of research. In one line of research, participants viewed a series of simple or complex neutral stimuli, and were explicitly cued before each stimulus that its successful retrieval in a later memory test would be rewarded or not (Adcock et al., 2006; Ariel & Castel, 2014; da Silva Castanheira et al., 2022). In the other line of research, participants incidentally encoded items, some of which signaled the opportunity to gain a reward in the next few seconds (Gieske & Sommer, 2023; Wittmann et al., 2005, 2011). A later surprise memory test then assessed recognition memory for reward-associated versus neutral stimuli; retrieval success itself was not rewarded. In both paradigms, one involving intentional encoding and one involving incidental encoding, the anticipation of a reward during the encoding stage enhanced the probability of successful retrieval.

The mechanisms by which reward-associated items enhance learning are not fully understood yet. One way in which reward promotes learning is by guiding attention to cues that signal reward (Miendlarzewska et al., 2016). Feature-based attention, as defined by the rate of information uptake, is enhanced for reward-predicting stimuli compared to neutral stimuli (Dix & Li, 2020; Spaniol et al., 2011). If the reward-predicting features are semantic in nature, associated with deeper encoding, then enhanced attention is known to result in stronger memory traces (Wittmann et al., 2011), possibly through interactions between the dopamine system and hippocampus (Shohamy & Adcock, 2010; Wittmann et al., 2005).

Here, we examined another mechanism by which reward anticipation might enhance memory: through an increase in arousal during memory encoding. Reward anticipation increases self-reported arousal (Samanez-Larkin et al., 2007), activates nuclei of the ascending arousal system (Gieske & Sommer, 2023; Lloyd et al., 2024), and is accompanied by an increase in pupil size (Dix & Li, 2020; Lloyd & Nieuwenhuis, 2024; Rudebeck et al., 2014; Schneider et al., 2018). Increased arousal during encoding in turn is often associated with improved subsequent memory. First, emotionally arousing experiences tend to be well remembered, a relationship that is mediated by noradrenergic influences and amygdala activation (McGaugh, 2004). Second, arousal induction enhances memory for salient and goal-relevant stimuli while impairing memory for less conspicuous stimuli (Mather & Sutherland, 2011). And third, a

---

* Corresponding author at: Pieter de la Court, Wassenaarseweg 52, 2333 AK Leiden, Netherlands.
E-mail address: b.lloyd@fsw.leidenuniv.nl (B. Lloyd).

https://doi.org/10.1016/j.cognition.2025.106237
Received 29 November 2024; Received in revised form 18 June 2025; Accepted 30 June 2025

number of studies have found a positive relationship between pupil size during the (intentional) encoding of neutral stimuli and subsequent memory of those stimuli (Miller & Unsworth, 2020; Papesh et al., 2012; but see Naber et al., 2013).

The research discussed above suggests that the positive relationship between reward association during encoding and subsequent memory may be mediated in part by arousal. However, to our knowledge, this possibility has never been examined directly. Here, we used a monetary incentive delay task, followed by a surprise memory test 24 h later, to examine if the boost in arousal associated with reward-associated items (partially) mediates the well-documented positive effect of reward anticipation on subsequent memory (Murayama & Kitagami, 2014; Wittmann et al., 2005, 2008, 2011). The study design was based on that of Wittmann et al. (2005), a pioneering study on reward-related and dopamine-dependent incidental memory formation in humans. The monetary incentive delay task is probably the most common task for inducing and measuring the (e.g., behavioural, pupillary, BOLD) effects of reward anticipation. In this task, participants first classify a cue that informs them whether later in the trial they can earn a reward (reward trial) or not (neutral trial) by making a speeded choice response. In our study, the same cues that signaled whether or not a reward could be obtained were also the objects of the incidental encoding task. This prevented dual-tasking, while the semantic nature of the cue (e.g., image of man-made item signaled a reward trial, natural item signaled a neutral trial) promoted deep rather than shallow memory encoding.

Research on reward-related memory enhancement suggests that its effects strengthen over time, with greater benefits emerging at longer retention intervals (e.g., Wittmann et al., 2005: 3 weeks vs. 20 min). This pattern aligns with evidence that dopamine plays a crucial role in post-encoding consolidation, as long-term potentiation over extended delays is dopamine-dependent (Chowdhury et al., 2012; Lisman et al., 2011). Consequently, immediate testing may yield different results, as memory traces for neutral and reward-predicting items might not yet diverge. In the present study, we chose a 24-hour delay to allow consolidation processes to unfold and to build on prior research demonstrating enhanced memory performance after longer retention intervals.

To foreshadow the results, we indeed found that pupil-linked arousal partially mediated the effect of reward anticipation on memory. But to our surprise this relationship was negative: higher arousal during encoding *impaired* subsequent memory, counteracting the dominant positive direct effect of reward anticipation on memory. We discuss a hypothesis that explains this negative relationship between phasic arousal and memory in terms of decision urgency, and contrast this with an alternative explanation in terms of mental effort (Kahneman, 1973; Mathôt et al., 2018).

2. Method

2.1. Participants

Forty healthy students from Leiden University took part in this study in return for 16.50 euros or course credits. Two participants did not complete the memory test and one participant's data were removed due to technical issues, leaving a final sample size of 37 (mean age = 25 years, range = 21–32 years, 28 women). None of the participants had a history of psychiatric or neurological disorders, learning disabilities or head traumas. Three participants were missing one block of pupil data due to technical issues with the eye-tracker. Participants were instructed to abstain from consuming alcohol and caffeine within 12 and three hours before each session, respectively. The study was approved by the Psychology Research Ethics Committee at Leiden University.

2.2. Stimuli

The stimuli consisted of 420 man-made and natural items obtained from an openly available stimulus set (Brady et al., 2008) and Google Image Search. Items were grey-scaled and luminance-matched with a grey background (RGB: 125, 125, 125 out of 255; Fig. 1A). All items were resized using in-house Matlab scripts. All text stimuli (i.e., fixation cross, number, and feedback) were matched in luminance with the background screen (RGB: 60, 60, 255).

2.3. Design and procedure

The study consisted of two tasks: a monetary incentive delay task and a surprise memory test, which were performed in separate sessions, spaced 24 h (± 2 h) apart. Both tasks were programmed in Python 3 using the Expyriment library (Krause & Lindemann, 2014). All task instructions are included in the Supplementary Materials. In the first session, participants familiarised themselves with the monetary incentive delay task by completing ten practice trials. The task consisted of 280 trials, divided into 4 blocks (Fig. 1A). Each trial began with a fixation cross (duration 2–11 s, pseudo-exponentially distributed with a mean of 5 s). Next, an image of a man-made or natural item was presented for a duration of 2 s. Participants were informed they could earn a reward on trials on which an item from one of the two categories was presented ("reward trials") and that they would earn no reward for the other category ("neutral trials"). Half of the participants expected a reward following man-made items (i.e., cup, bike, fork), while the other half of the participants expected a reward following natural items (i.e., shell, bird, tree). While the item was on the screen, participants were to answer the question 'Do you expect a reward on this trial?'. To respond 'yes', they pressed the left arrow key with their right index finger; to respond 'no' they pressed the right arrow key with their right ring finger. The item order was pseudo-randomised with a maximum of four items from the same category in a row.

Following item presentation, a fixation cross was presented (duration 2–11 s, pseudo-exponentially distributed with a mean of 5 s). Next, a target number (1, 4, 6, or 9 [randomised]) was presented for 100 ms. After the onset of the target number, participants were to respond to the question 'Is the number higher or lower than 5?', by pressing the up (higher) or down (lower) arrow key as quickly as possible using their right middle finger. The trial ended with a blank screen (1 s), followed by an outcome message for 0.5 s. On reward trials, participants could either win 10 eurocents (positive outcome, signaled with a euro sign) or lose 5 eurocents (negative outcome, signaled with a 'no signal' sign). On neutral trials no money could be won or lost (neutral outcome, signaled with a question mark). On reward trials, participants only received a positive outcome if they correctly classified the item and number and if their response to the number classification task met a response deadline. The corresponding task instructions were: "Your task is to indicate with a button press whether this number is higher or lower than 5. Whether or not you receive a reward on reward-predicting trials will depend on how quickly you make this response. If you respond correctly and quickly you will win €0.10, but if you respond incorrectly or too slow, you will lose €0.05, so try to respond as quickly as possible!" We used a staircase procedure to adjust the deadline individually based on the participant's reaction times in previous trials so as to yield a positive outcome on ~70 % of the reward trials. Reward trials with an incorrect response or a slow response on the number classification task always resulted in a negative outcome. At the end of each block, the total money earned on that block was presented on the screen. At the end of the first session, participants received the total reward money earned on all blocks and the practice task in cash.

Participants returned the next day to complete a surprise memory test (Fig. 3A); they were initially informed they would be completing a similar task as in session 1. The memory test was self-paced and consisted of 420 trials. On each trial, an item was presented in the centre of the screen with the words 'old' and 'new' on the bottom left and right corner, respectively. Participants had either seen the item one day earlier ('old', 280 trials) or the item was completely new ('new', 140 trials). First, participants indicated whether the item was old or new

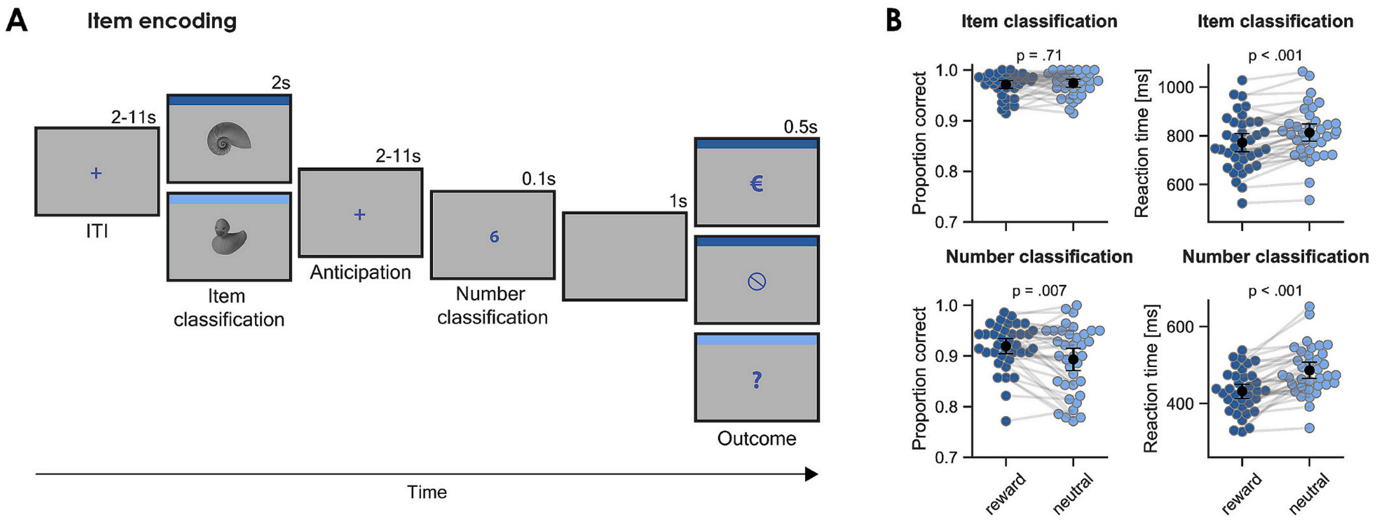

Fig. 1. Overview of monetary incentive delay task and task performance. A) Schematic overview of a trial in the monetary incentive delay task. Reward trials (neutral trials) are indicated by the dark blue (light blue) top border. Borders are used for illustrative purposes only. B) Accuracy and reaction time and of item classification (top) and number classification (bottom) responses. Data points and grey lines refer to individual participant scores. Error bars indicate ± SEM. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

using the left and right arrow keys. Next, participants were required to describe their memory for the item. If they responded 'old', they were asked to indicate whether they did 'remember' or 'know' the item (Düzel et al., 1997; Tulving, 1985; Wittmann et al., 2005). A 'remember' judgement meant they could recall vivid details about the item, whereas a 'know' judgement meant they had a sense of familiarity with the item. Following a 'new' response, participants were asked to indicate whether they were sure ('sure') or whether they were guessing ('guess'). The memory test lasted ~35 min. The experimental task including all the instructions is available on Github (https://github.com/bethlloyd/Reward-related_mem_LloydNieuwenhuis).

2.4. Behavioural analyses

The old and new responses in the memory test allowed us to classify trials according to signal detection theory: hits, misses, false alarms, and correct rejections. We then computed recognition memory (hits minus false alarms and d′), for all trials, and for remember and know judgements separately. We performed a condition-wise outlier detection analysis, whereby any individual who scored 3 standard deviations away from the condition mean was considered an outlier for that analysis. This procedure resulted in the following removals: one participant on the basis of the number classification reaction times, one participant on the basis of their proportion of hits ('know' judgement only), and one participant on the basis of recognition memory (hits – false alarms). Lastly, only trials with correctly classified items were used for the analyses involving memory performance.

2.5. Eye tracking

2.5.1. Data acquisition

During the monetary incentive delay task, we measured pupil size at a sampling rate of 40 Hz using a Tobii Pro X3-120 eye-tracker. A chin rest was used to ensure that the eye-tracker was positioned approximately 75 cm from the participant's eyes. The experiment was carried out in a cubicle without any distractions under constant dim lighting. Before the experiment started, we calibrated the eye-tracker using a default five-point calibration method from the eye-tracker manufacturer.

2.5.2. Pupil data preprocessing and analysis

Preprocessing of the pupil data was performed in PupCor (https://github.com/lindvoo/PupCor) and further preprocessing and analyses were carried out in Python 3 (scripts available here: https://github.com/bethlloyd/Reward-related_mem_LloydNieuwenhuis). Data points that were marked by the device manufacturer as blinks were removed by applying an automated interpolation procedure starting 100 ms before blink onset to 400 ms following blink offset. We analysed the pupil (left or right) for which more data points were available. The pupil time series were then manually checked and corrected if any artifacts remained. Lastly, to remove high-frequency noise from the time series, we applied a low-pass filter using a 10-Hz fourth-order Butterworth filter with zero-phase shift.

Next, the pupil time series were segmented into epochs ranging from 0.2 s before to 4 s after item onset. The mean pupil size in the 0.2 s before item onset served as the pre-item pupil baseline. Trials containing disproportionately high amounts of missing data points (>50 %) in either the pre-item pupil baseline (0.2 s) or the entire event epoch (4.2 s) were removed. After this, samples ranging more than 2.5 standard deviations above or below the mean of each trial were considered spurious and removed from the event epoch. After applying these criteria, an average of 73.9 ± 3.4 % trials remained. Finally, blocks containing >50 % invalid pupil trials were completely removed from all pupil analyses (23 blocks, including all four blocks from four participants), leaving a sample size of 33 participants for all analyses involving pupil data. Pupil diameter was then baseline-corrected on the trial level by subtracting the pre-item pupil size (0.2 s). The single-trial pupil dilation response was defined as the average pupil size from 1 s to 3 s after item onset. For follow-up tests and visualization, this event was averaged across trials separately per condition (reward, neutral), memory outcome (remembered, forgotten), and participant. Following the same outlier detection analysis as mentioned above (*Behavioural analysis*), the data from one participant were removed on the basis of the magnitude of the average pupil response. Only trials with correctly classified items were used for the pupil analyses.

2.6. Statistical analyses

Statistical analyses were carried out in Python 3 and Rstudio. Scripts are openly available here: https://github.com/bethlloyd/Reward-related_mem_LloydNieuwenhuis. Since some variables were not normally distributed, we used only non-parametric tests [$\alpha \leq 0.05$] (function: *scipy.stats.wilcoxon*; Virtanen et al., 2020). To estimate the potential effects of reward anticipation on the item and number classification responses, we conducted Wilcoxon signed-rank tests. To explore reward anticipation effects and subsequent memory effects on pupil size at the trial level, we used linear mixed-effects models (*lmerTest::lmer*; Kuznetsova et al., 2017) predicting item-related pupil size as a function of condition (reward, neutral), memory (remembered, forgotten) and their interaction. Intercepts were modeled as random effects. Similar models were used with pre-item pupil size and reaction time on the number classification task as the outcome variable. For comparisons between pupil waveforms, non-parametric tests were carried out across our event epoch (4.2 s), sample by sample, using the function *neurotools.stats.permtest_rel*. The *p*-values resulting from this procedure were corrected for false discovery rate (FDR).

We conducted a causal mediation analysis to explore the interaction between reward anticipation, pupil size, and subsequent memory. Unlike traditional mediation analysis, causal mediation analysis can adequately handle non-normally distributed variables, such as our binary treatment and outcome variables (Xu et al., 2023). To run the model, we used the *mediation* package (Tingley et al., 2014) in R and fit two regression models (*lmer* and *glmer* in the *lme4* package; Bates et al., 2015). The first model predicted item-related pupil size as a function of condition (reward vs. neutral), and the second predicted memory on the basis of condition and item-related pupil size. These models were then inserted into the 'mediate' function to estimate the average causal mediation effect (ACME) and the average direct effect (ADE). Throughout the *Results* section, data are expressed as the mean ± standard error of the mean (SEM).

3. Results

3.1. Manipulation successfully induced reward anticipation

In line with previous studies (Lloyd & Nieuwenhuis, 2024; Wittmann et al., 2005, 2011), behavioural responses on the item classification task ('Do you expect a reward on this trial?') and number classification task ('Is the number higher or lower than 5?') showed clear reward anticipation effects. Participants classified man-made and natural items faster on reward trials (772 ± 18 ms) than on neutral trials (813 ± 17 ms; $p < .001$; Fig. 1B). Item classification accuracy was good in both conditions (reward: 97.1 ± 0.4 %; neutral: 97.4 ± 0.4 %; $p = .71$). So, items with a reward association invigorated participants to respond more quickly, but not at the cost of more errors.

On the number classification task, participants gained 10 cents on 74.2 % of the reward trials (mean RT = 409 ± 51 ms), approximating the 70 % targeted with the staircase procedure. Of the 25.8 % reward trials on which participants lost 5 cents, 8.1 ± 4.5 % were incorrect and 17.8 ± 4.3 % had a correct response that exceeded the deadline (mean RT = 515 ± 88 ms). To allow a comparison with performance on neutral trials, we pooled all reward trials, regardless of RTs. On average, participants responded faster on reward trials (434 ± 9 ms) than on neutral trials (494 ± 12 ms; $p < .001$). Furthermore, number classification accuracy was higher on reward trials (92.0 ± 0.7 %) than on neutral trials (89.3 ± 1.2 %; $p = .007$). Altogether, these results show that our task induced a state of reward anticipation in our participants, observed through faster reaction times, and improved accuracy on the number classification task. For the remaining analyses, only trials with correctly classified natural and man-made images were included.

3.2. Reward anticipation increased arousal and strengthened memory encoding

A linear mixed-effects model predicting item-related pupil dilation at the trial level on the basis of condition (reward, neutral) and memory (remembered, forgotten) revealed that, as expected, the average pupil dilation response was larger for reward-associated items (0.15 ± 0.02 mm) than for neutral items (0.10 ± 0.02 mm; $b = 0.05$, $t_{32} = 3.75$, $p < .001$; Fig. 2A). Sample-by-sample pairwise comparisons showed that this effect occurred between 0.70 s and 3.95 s after item onset ($p < .05$; FDR-corrected for all timepoints). The pupil dilation response to the outcome was also stronger on reward trials than on neutral trials (Fig. 2B).

The results of the surprise memory test (Fig. 3A) showed better memory for items with a reward association than for neutral items (proportion of hits: $p < .001$; recognition memory [hits-false alarms]: $p = .002$; p = .002; Fig. 3B). The effect of reward remained when we included only trials with "remember" judgements (proportion of hits: $p = .02$; recognition memory [hits – false alarms]: $p = .03$). Similar analyses including only trials with "know" judgements only showed an effect of reward on the proportion of hits ($p = .04$; recognition memory: $p = .55$). These results were complemented by a signal detection theory analysis. Participants had a more liberal criterion (c) on reward trials (0.18 ± 0.45) than on neutral trials (0.40 ± 0.31, $W = 125$, $p < .001$), meaning that they were more inclined to say "old", even to new test items from the reward-associated image category. More importantly, the participants' sensitivity (d'), their ability to discriminate between old and new test items, was larger for reward-associated items (1.38 ± 0.37) than for neutral items (1.26 ± 0.39, $W = 206$, $p = .028$).

Together, these results confirm the expected effects of reward anticipation on pupil-linked arousal and memory strength, in particular "remember" judgements.

3.3. Reward anticipation effect on memory is partially mediated by pupil-linked arousal

Next, we asked whether the positive effect of reward anticipation on subsequent memory was mediated by pupil-linked arousal. To answer this question, we performed a model-based causal mediation analysis (Fig. 4). We found that the average causal mediation effect was highly significant, but surprisingly, this relationship was negative ($ACME = -0.003$, $p < .001$)—in the opposite direction of the average direct effect of reward on memory ($ADE = 0.099$, $p < .001$), which dominates the relationship. These results suggest that the effect of reward anticipation on pupil-linked arousal and ensuing effect on memory *counteract* the positive direct effect of reward anticipation on memory.

To understand the negative ACME of reward anticipation on memory, we took a closer look at the relationship between pupil dilation during encoding and memory on the next day. Our linear mixed-effects model revealed that, indeed, item-related pupil size was significantly associated with later memory ($b = -0.03$, $t_{31} = -2.84$, $p = .005$; Fig. 5A), and that this relationship was negative (remembered: 0.12 ± 0.02 mm, forgotten: 0.15 ± 0.02 mm). This subsequent-memory effect occurred between 1.7 s and 3.9 s after item onset ($p < .05$; FDR-corrected for all timepoints; Fig. 5A), and was preserved when trials with item classification errors were excluded. There was no interaction between condition and memory ($b = 0.002$, $t_{31} = 0.16$, $p = .87$). Indeed, follow-up tests showed that the negative relationship between pupil size and later memory was present on both reward trials ($b = -0.03$, 95 % CI [-0.05, -0.006], $t_{31} = -2.46$, $p = .02$) and neutral trials ($b = -0.03$, 95 % CI [-0.05, -0.009], $t_{31} = -2.82$, p = .005).

We then repeated our linear mixed-effects analysis, this time including pre-item baseline pupil size as an additional predictor. As is commonly found, pre-item baseline pupil size was negatively correlated with the magnitude of the pupil dilation response ($b = -0.29$, $t_{31} = -45.28$, $p < .001$). However, pre-item baseline pupil itself had no association with subsequent memory ($b = 0.009$, $t_{31} = 0.49$, $p = .62$; Fig. 5B); and the reward and subsequent memory effects on pupil dilation were still significant after statistically controlling for baseline pupil size ($ps < .002$).

Signal detection theory analyses revealed a significantly more conservative criterion for man-made (0.39 ± 0.31) than for natural test items (0.20 ± 0.46, $W = 154$, $p = .003$), meaning that participants were

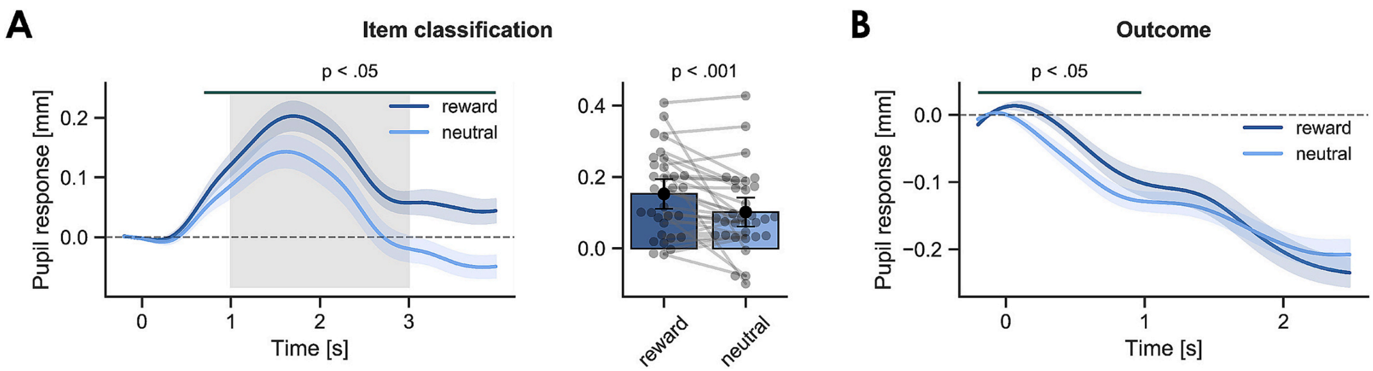

Fig. 2. Reward anticipation and reward outcome strengthened pupil response. A) Pupil waveforms in response to reward-associated and neutral items (left) and average pupil size in the 1–3 s after item onset (right). Data points and grey lines refer to individual participant scores. Grey-shaded rectangle indicates the window used to compute the average pupil response. B) Pupil waveforms in response to the outcomes. Error bars and shaded areas indicate ± SEM. Horizontal green lines indicate the time periods during which pupil size differed between reward and neutral trials ($p < .05$, FDR-corrected for all timepoints). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

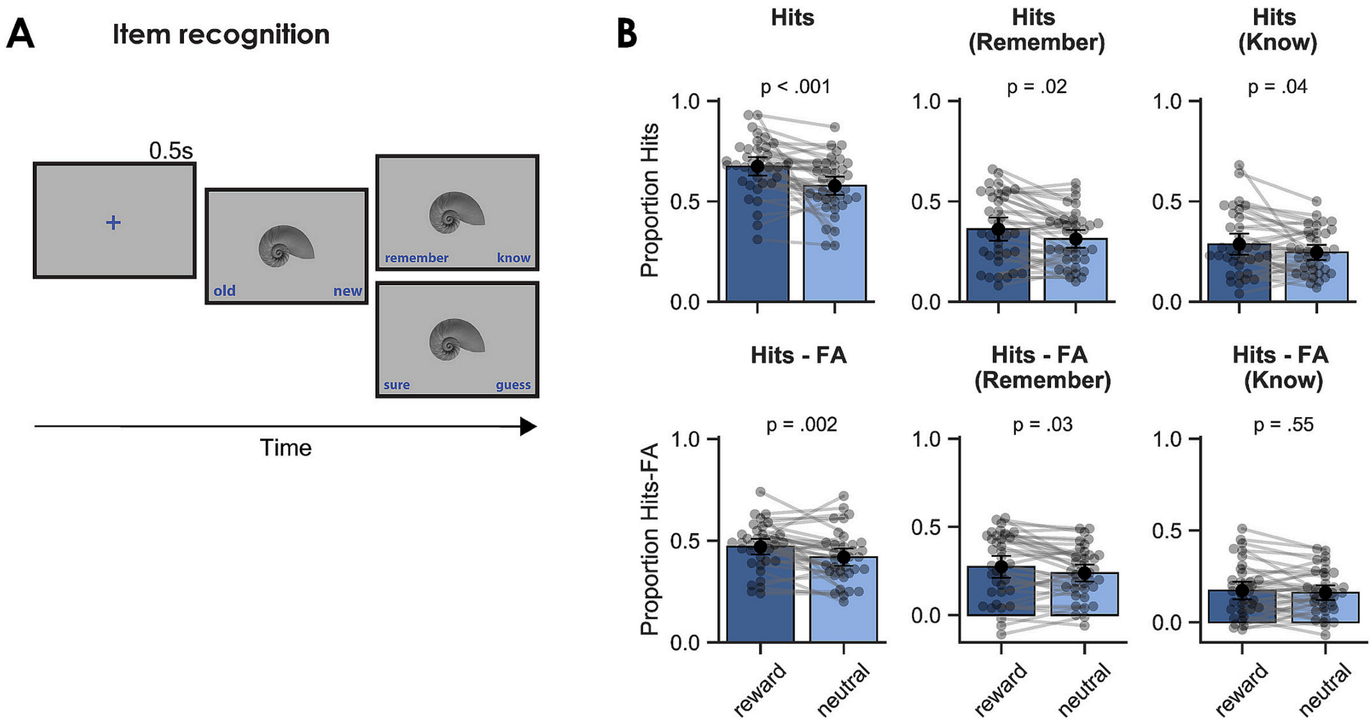

Fig. 3. Reward anticipation effects on subsequent memory. A) Schematic overview of a trial in the surprise memory test. B) Memory performance on reward and neutral trials. Uncorrected hits (top) and recognition memory (hits − false alarms, bottom), for all trials, and for trials with remember and know judgements separately. Data points and grey lines refer to individual participant scores. Error bars indicate ± SEM.

less inclined to evaluate old and new man-made test items as 'old' (Supplementary Fig. 1a). Because an increased criterion is associated with a smaller difference in average memory strength between remembered and forgotten items, we examined whether man-made items were associated with a smaller subsequent memory effect on pupil dilation. The negative effect of subsequent memory on pupil dilation was present for both the man-made ($W = 165$, $p = .039$) and natural image categories ($W = 109$, $p = .002$), and did not statistically differ between image categories (Supplementary Fig. 1b).

3.4. An urgency account of the negative relationship between pupil-linked arousal and subsequent memory

Recent studies have suggested a role for pupil-linked arousal (Gross & Dobbins, 2021; Lawlor et al., 2023; Murphy et al., 2016) and the ascending arousal system (Hauser et al., 2018; Lloyd et al., 2024) in generating decision urgency, an evidence-independent neural signal that expedites the evolving decision process by pushing it closer to a fixed decision threshold (Carland et al., 2019; Ditterich, 2006; Reddi & Carpenter, 2000; Standage et al., 2011). This growing urgency signal limits the time for deliberation when the strength of the sensory evidence is weak, enforcing a decision when the summed evidence and urgency reaches the threshold. Here, we consider a simple model that attempts to account for the negative relationship between pupil-linked arousal and subsequent memory in terms of decision urgency.

The model assumes that in the item classification task, a decision is made when a decision variable x(t) reaches a fixed decision threshold (Fig. 6). In line with typical sequential sampling models, we propose that the decision variable $x(t)$ is the sum of the accumulated evidence $v(t)$ and an evidence-independent, growing urgency signal $u(t) : x(t) =$

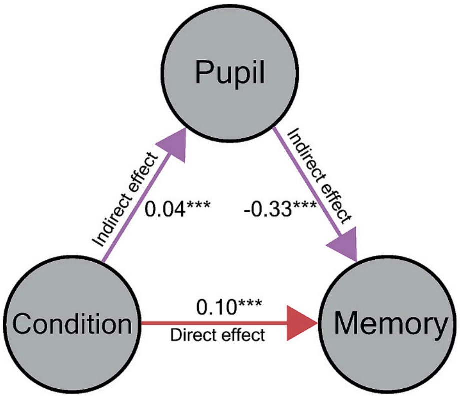

Fig. 4. Pupil size partially mediates the effect of reward anticipation on memory.

$v(t) + u(t)$. Because the value of the decision variable at the time of the decision crossing, $x(t_{dec})$, is the same on each trial (reflecting the fixed distance to the decision threshold), more accumulated evidence $v(t_{dec})$ means less accumulated urgency $u(t_{dec})$ and vice versa. For example, on trials with a lower evidence accumulation rate, the urgency signal has more time to build until the decision threshold is reached, while less evidence will have been accumulated by the time of the threshold crossing.

Importantly, we assume that the fidelity of encoding an item into memory is proportional to the total accumulated evidence $v(t_{dec})$, and maximum pupil dilation proportional to the level of urgency $u(t_{dec})$ at the moment the decision threshold is reached. If there is variability across trials in evidence accumulation rate (e.g., reflecting how much attention participants are paying or how difficult it is to identify the item), these assumptions impose a negative relationship between memory success and pupil dilation. This may explain why in our data (Fig. 5A) later forgotten items (lower evidence) tend to be characterized by a more protracted pupil dilation (urgency signal can mount longer), resulting in a higher peak dilation.

This urgency account makes several testable predictions. First, because the level of urgency rises until a decision is made (i.e., threshold crossing), maximum pupil dilation should be positively correlated with RT. To test this prediction, we added reaction time to the linear mixed-effects model predicting the pupil dilation response. In line with previous studies (Gross & Dobbins, 2021; Murphy et al., 2016; Tromp et al., 2024), we found that there was indeed a significant positive main effect of reaction time ($b = 9.721\text{e-}5$, $t_{31} = 6.09$, $p < .001$; Fig. 7A), while the effects of condition and memory remained significant ($ps < .002$). A second prediction is that item classification errors (caused by weak evidence for the correct choice) should have slower RTs than correct responses on the item classification task. We found this to be case (Fig. 7B; incorrect: 906 ± 39 ms; correct: 791 ± 17 ms; $p < .001$).

Some other predictions were not borne out by the data. Because item classification performance and the fidelity of encoding both depend on evidence accumulation rate, the model predicts that later forgotten items should be associated with more item classification errors and slower RTs than later remembered items. However, this was not the case in our data (Fig. 7C; remembered: 96.9 ± 0.3 %; forgotten: 97.6 ± 0.5 %; $p = .02$ [accuracy]; remembered: 799 ± 18 ms; forgotten: 781 ± 17 ms; $p = .002$ [RT]). These findings require additional assumptions for the urgency account to remain tenable. For example, the rare item classification errors may have generated a prediction error that may have strengthened the episodic memory trace (Rouhani et al., 2018), causing the later remembered items to have lower accuracy on the item classification task.

In any case, this urgency account cannot explain our finding that reward anticipation was positively related to both pupil-linked arousal and subsequent memory. As discussed above, the account predicts that any modulation of urgency should enforce a negative relationship

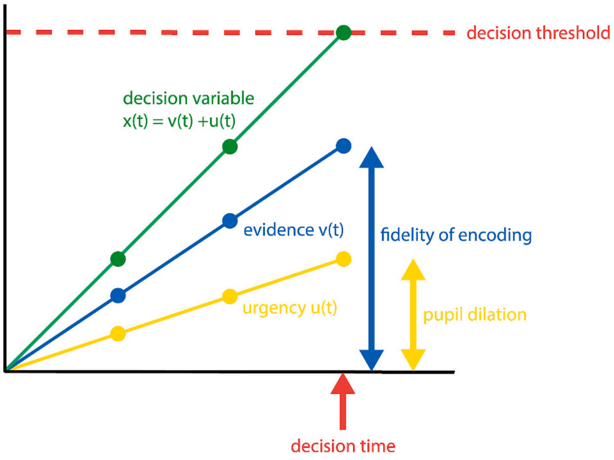

Fig. 6. Assumptions underlying urgency account. This diagram only illustrates the decision variable associated with the correct choice (e.g., natural item). An error is made when the decision variable associated with the incorrect choice (man-made item) reaches the threshold first.

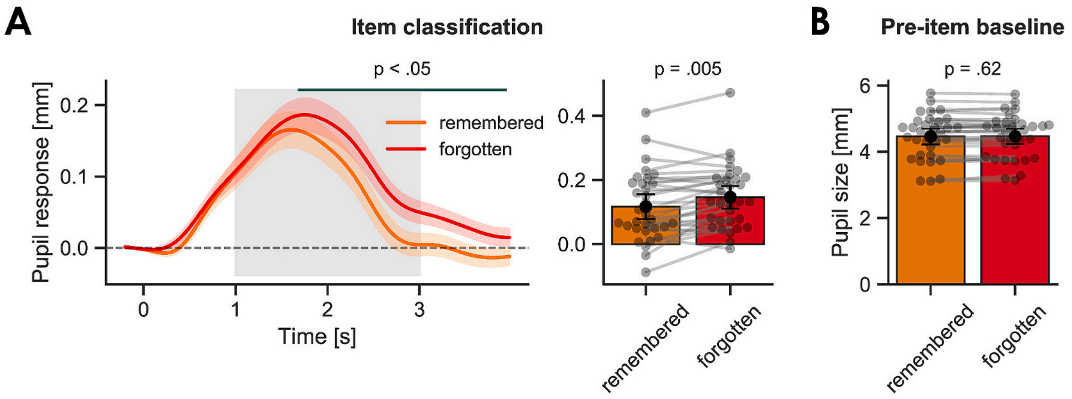

Fig. 5. Subsequent memory effect on pupil size. A) Item-related pupil dilation responses and B) pre-item baseline pupil size (0.2 s before item onset), separately for subsequently remembered and forgotten items. Grey-shaded rectangle indicates the window used to compute the average pupil dilation response. Horizontal green line indicates the time period during which pupil size differed between the remembered and forgotten items ($p < .05$, FDR-corrected for all timepoints). Data points and grey lines refer to individual participant scores. Error bars and shaded orange and red areas indicate ± SEM. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

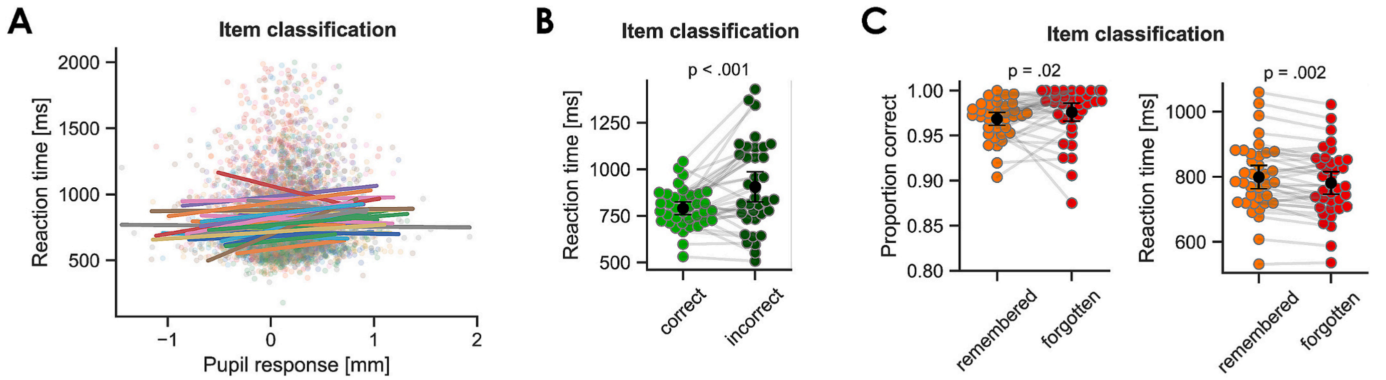

Fig. 7. Model prediction checks. A) Trial-by-trial reaction time on the item classification task was positively associated with pupil size response ($p < .001$). Regression lines are colored by participant. Points refer to individual trials colored by participant. B) Item classification errors were slower than correct responses. C) Accuracy on the item classification task was higher (left) and (correct) reaction times faster (right) for later forgotten vs. remembered items. Data points and grey lines refer to individual participant scores. Error bars indicate ± SEM.

between arousal and memory. Therefore, within the context of this account, the increased pupil response to reward-associated items cannot be linked to urgency. The difference in timing and shape of the pupil modulations associated with reward anticipation (Fig. 2A) and subsequent memory (Fig. 5A) provides a potential clue to this problem. While the memory modulation emerges late and reflects a more protracted increase in pupil size, the reward anticipation modulation starts relatively early and reflects a difference in steepness of the pupil response. Perhaps the transient central arousal response underlying the early, reward-related pupil modulation occurred before it could impact urgency. Or arousal does not influence urgency (Murphy et al., 2016) but urgency influences arousal (i.e., a reverse causal relationship), in which case reward anticipation does not necessarily increase urgency. We conclude that, whereas the urgency account is consistent with previous research linking pupil size to urgency, and can account for some of our key empirical findings, it cannot account for several other findings in our study.

Because researchers have often linked pupil size to mental effort (Kahneman, 1973; Mathôt et al., 2018), we also considered whether this link offers an alternative explanation for the negative across-trial relationship between pupil-linked arousal and subsequent memory. One may argue that some of the man-made and natural items in our task were harder to identify than others. Identification of these difficult items may have required more mental effort, and their encoding in memory may have been less successful. As a result, pupil dilation should have been stronger for items that were later forgotten (i.e., a negative relationship). Note that this account (cf. Kafkas & Montaldi, 2011), like the urgency account, predicts that forgotten items (that tended to be harder to identify) should be associated with slower item classification RTs–a prediction that is not in line with the findings shown in Fig. 7C: Forgotten items were characterized by *faster* item classification RTs, $W = 141$, $p = .002$. Alternatively, one may argue that for any level of identification difficulty a larger investment of mental effort should have increased the probability of successful encoding, resulting in a *positive* relationship between maximum pupil dilation and subsequent memory. These contrasting predictions reflect the fact that the concept of mental effort, at least in the current context, is computationally less well defined than the concept of urgency (e.g., does pupil size reflect *required* or *invested* effort?)

4. Discussion

The goal of our study was to determine if the previously established positive effect of reward anticipation on episodic memory is (partially) mediated by arousal. Participants were asked to categorize pictures as man-made or natural, one category of which invoked anticipation of a monetary reward, to be gained in a speeded reaction time task at the end of the trial. A surprise memory test carried out 24 h later assessed the participants' recognition memory for reward-associated versus neutral items. In line with previous studies, we found that reward-associated items elicited a phasic increase in pupil-linked arousal (Dix & Li, 2020; Lloyd & Nieuwenhuis, 2024; Schneider et al., 2018) and were more likely to be remembered the next day (Gieske & Sommer, 2023; Miendlarzewska et al., 2016; Murayama & Kitagami, 2014; Wittmann et al., 2005). Importantly, the phasic arousal response partially mediated the effect of reward anticipation on subsequent memory. However, to our surprise, the indirect effect of arousal *counteracted* rather than strengthened the positive direct effect of reward anticipation on memory.

The indirect negative effect of reward anticipation on episodic memory was driven by the finding that reward-predicting and neutral items that elicited a larger pupil dilation response during encoding tended to be remembered less well the following day. This finding supports recent work (Gross & Dobbins, 2021) suggesting that pupil dilation does not reflect the strength of memory encoding (cf. Papesh et al., 2012). We considered an alternative explanation in terms of mental effort, a construct that has long been linked to pupil-linked arousal (Kahneman, 1973; Mathôt et al., 2018). Items that are difficult to identify may require more effort during encoding and may be less likely to be remembered – an argument that predicts the negative relationship between pupil dilation and subsequent memory reported here (Kafkas & Montaldi, 2011). However, this explanation cannot explain why later forgotten items were associated with fewer item classification errors and faster RTs than later remembered items. More generally, as argued by others (cf. Gross & Dobbins, 2021), the mental effort account is not sufficiently constrained; had we found a positive correlation, proponents of this account might explain this by arguing that the more effort is invested in encoding an item, the larger the possibility that it will be later remembered.

Recent studies using pupillometry (Gross & Dobbins, 2021; Lawlor et al., 2023; Murphy et al., 2016), pharmacology (Hauser et al., 2018) and ultra-high-field fMRI (Lloyd et al., 2024) have suggested that central arousal modulates decision urgency, an evidence-independent neural signal that expedites the evolving decision process by driving it closer to a fixed decision threshold (Carland et al., 2019; Ditterich, 2006; Standage et al., 2011). As we argue above (see also Gross & Dobbins, 2021), this notion can explain several of our key results: (i) the negative correlation between maximum pupil dilation and subsequent memory; (ii) the positive correlation between maximum pupil dilation and RT; (iii) the observed relationship between speed and accuracy in the item classification task; and (iv) the finding that later forgotten items are associated with a more protracted but not steeper increase in pupil size.

As we explain above, other aspects of our results cannot be explained by this urgency account or require additional assumptions: (i) the relationship between item classification performance and subsequent memory; and (ii) the finding that reward anticipation was positively related to both pupil-linked arousal and subsequent memory. This indicates that more work is necessary to further uncover the interrelationships between reward anticipation, pupil-linked arousal and memory encoding.

How can we explain the negative relation between the magnitude of pupil dilations and subsequent memory, as reported here and elsewhere (e.g., Kafkas & Montaldi, 2011; Pilarczyk et al., 2022) in light of other studies that found no relationship (Võ et al., 2008) or a positive relationship? A potential clue may lie in the difference between incidental and intentional encoding. Although the amount of evidence is limited, incidental encoding studies including ours tend to find a negative relationship (Kafkas, 2021; Kafkas & Montaldi, 2011; Pilarczyk et al., 2022), whereas intentional encoding studies tend to find a positive relationship (Bergt, Urai, Donner, & Schwabe, 2018; Miller & Unsworth, 2020; Papesh et al., 2012; but see Naber et al., 2013). In intentional encoding studies, participants are asked to memorize the stimuli presented in the encoding phase, but do not have to explicitly respond to them (e.g., identify or categorize them using a button press). In these tasks, pupil dilations may be more directly linked to effortful memory encoding processes, as opposed to decision-making or response processes. On the other hand, the absence of choice and RT data during the encoding phase renders it hard to model information processing; for example, any urgency effects in the data would be harder to demonstrate. Kafkas (2021) manipulated the expectedness of pictures of man-made and natural objects that were presented during an incidental encoding task, and that the participants had not seen before. The relation between the magnitude of the pupil dilation response and subsequent memory performance was positive for the unexpected objects and negative for the expected objects. This striking dissociation could be a crucial clue, except that there seem to be no systematic differences in stimulus-induced surprise between previous studies that have reported positive and negative relations.

A peculiar aspect of our paradigm is that the monetary incentive delay task required processing of the category that each image belonged to (e.g., “natural item”), whereas the surprise memory test assessed whether participants recognized the specific image (e.g., “snail shell”), regardless of the category. That is, during encoding, the specific properties of each item were only goal-relevant to the extent that they were needed to determine the category. Interestingly, emotionally arousing events have been found to influence how surrounding non-emotional information is processed, enhancing memory for goal-relevant or highly salient (i.e., high-priority) stimuli, while impairing memory for low-priority stimuli (Knight & Mather, 2009; Sakaki et al., 2014). There are also some indications that the positive or negative impact of goal relevance and salience level on memory formation increases with pupil dilation (Clewett et al., 2018; Eldar et al., 2016). These findings, which formed the basis for the arousal-biased competition theory (Mather et al., 2016), seem consistent with the observed positive relationship between the magnitude of pupil dilation during encoding and subsequent memory for items that were either intentionally encoded (i.e., goal-relevant) or incidentally encoded but surprising (i.e., highly salient, because violating an expectation; Kafkas, 2021). The more complex question is whether incidentally encoded stimuli, such as the specific man-made and natural images in our study, can have such ‘low priority’ that pupil-linked arousal should impair their memory encoding according to the arousal-biased competition theory. In our study, that seems implausible, given the semantic categorical decisions required in the learning phase (promoting relatively deep encoding; Craik & Lockhart, 1972), and the recognition memory performance in the test phase. Therefore, there appears to be no single explanation yet that can account for the variety in subsequent memory effects on pupil size.

The benefits of an incidental encoding paradigm include its ecological validity and the low demand characteristics, which made it possible to simultaneously administer the monetary incentive delay task. However, a limitation of the incidental encoding paradigm is that experimenters have less control over participants’ encoding strategies and the compatibility of these strategies with the surprise memory test. Furthermore, the speeded response requirements of the monetary incentive delay task may have interfered with the memory encoding and consolidation processes. Future research should therefore explore the effects of reward anticipation and pupil-linked arousal on intentional encoding, in a paradigm that does not require simultaneous motor responses. For example, cues can signal the possibility of gaining a reward for memorizing an upcoming stimulus. It is known that such future incentives enhance memory encoding, as reflected in the probability of successful retrieval (Adcock et al., 2006; Ariel & Castel, 2014; da Silva Castanheira et al., 2022). An interesting question is whether the link between this form of reward anticipation and subsequent memory performance is also mediated by pupil-linked arousal.

If not arousal, then what cognitive process causes the direct, positive effect of reward anticipation on episodic memory? Drift diffusion model (DDM) analyses of performance on perceptual decision-making tasks have suggested that reward-predicting stimuli are processed with greater attention, as indexed by the drift rate parameter (Dix & Li, 2020; Spaniol et al., 2011). Enhanced attention can explain why in our experiment reward anticipation led participants to respond more quickly, without a concomitant increase in errors. Unfortunately, we could not perform DDM analyses to verify whether reward-associated items were processed with a higher drift rate; the number of incorrect responses was too small to allow accurate parameter estimation. DDM analyses have also been applied to responses in the recognition phase of a memory experiment (da Silva Castanheira et al., 2022). This work shows that when participants are explicitly cued before encoding a stimulus whether its successful recognition will be rewarded or not, the enhanced recognition of reward-associated items can be explained by an increased fidelity of the memory trace, again manifested as an increase in drift rate. Together, these findings suggest that reward-predicting items are processed with more attention, resulting in greater memory encoding fidelity.

In our version of the monetary incentive delay task, reward-associated cues were followed by a gain of 10 cents on ~70 % of the trials and a loss of 5 cents on ~30 % of the trials. This means that the effects that we have attributed to reward anticipation may reflect in part effects of punishment anticipation, especially in risk-averse participants. Previous research using a monetary incentive delay task found that punishment-predicting cues were remembered better than neutral cues, and activated similar areas in the ventral striatum and dopaminergic midbrain as reward-predicting cues (Wittmann et al., 2013). Furthermore, it is possible that the observed pupil response to reward-predicting cues does not signal expected reward per se, but instead signaled unsigned prediction error, a sign that the outcome of the number classification task is going to be better or worse than expected (Preuschoff et al., 2011). These findings suggest that the behavioural and physiological effects of reward anticipation and punishment anticipation show significant overlap. To overcome this limitation of our task design, future studies on motivation, pupil-linked arousal and subsequent memory will need to explicitly dissociate effects of reward and punishment anticipation (e.g., Wittmann et al., 2013).

Knowledge on the role of pupil-linked arousal in cognitive and brain function has been rapidly increasing in the last decade (Grujic et al., 2024; Strauch et al., 2022). We would like to highlight two of our findings that may help to further develop this field of research. First, our study, as well as other recent work (de Gee et al., 2024; O’Bryan et al., 2024), shows that single-trial measures of pupil-linked arousal can mediate the effects of task variables on cognitive task performance. We encourage other researchers to perform such mediation analyses, while keeping in mind that pupil size is a noisy measure and therefore any indirect mediation effect is going to be small in size. Second, we found a significant positive correlation between single-trial (within-condition) measures of reaction time and pupil dilation magnitude, replicating other studies (Gross & Dobbins, 2021; Lloyd et al., 2024; Murphy et al., 2016; Tromp et al., 2024). Adding reaction time as a predictor in mixed-effects models of pupil dilation magnitude will therefore increase explained variance and enhance the statistical power of such models, regardless of whether the positive correlations reflect urgency or another link between decision-making, responding and arousal.

Our findings suggest that reward-related incidental memory enhancements are not mediated by pupil-linked arousal. Indeed, rather than facilitating memory, increased arousal during reward anticipation seems to disrupt encoding, possibly by heightening decision urgency. These findings offer new insights into the complex interplay between reward anticipation, arousal, and memory encoding, and highlight the importance of considering time pressure as a potential influencing factor when studying the central arousal system.

CRediT authorship contribution statement

Beth Lloyd: Writing – original draft, Visualization, Formal analysis, Data curation, Conceptualization. **Steven Miletić:** Writing – review & editing, Formal analysis. **Sander Nieuwenhuis:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that no competing interests exist.

Acknowledgements

This work was supported by funding from the Vici grant awarded to Sander Nieuwenhuis, financed by the Dutch Research Council (NWO), with project number VI.C.181.032. We would like to thank Alexa Schrickel for help with data collection.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2025.106237.

Data availability

The raw and processed data as well as code to reproduce the results are publicly available without restriction: https://github.com/bethllloyd/Reward-related_mem_LloydNieuwenhuis.

References

Adcock, R. A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., & Gabrieli, J. D. E. (2006). Reward-motivated learning: Mesolimbic activation precedes memory formation. *Neuron, 50*(3), 507–517. https://doi.org/10.1016/j.neuron.2006.03.036

Ariel, R., & Castel, A. D. (2014). Eyes wide open: Enhanced pupil dilation when selectively studying important information. *Experimental Brain Research, 232*(1), 337–344. https://doi.org/10.1007/s00221-013-3744-5

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*. https://doi.org/10.18637/jss.v067.i01

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

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences of the United States of America, 105*(38), 14325–14329. https://doi.org/10.1073/pnas.0803390105

Carland, M. A., Thura, D., & Cisek, P. (2019). The urge to decide and act: Implications for brain function and dysfunction. *The Neuroscientist, 25*(5), 491–511. https://doi.org/10.1177/1073858419841553

Chowdhury, R., Guitart-Masip, M., Bunzeck, N., Dolan, R. J., & Düzel, E. (2012). Dopamine modulates episodic memory persistence in old age. *The Journal of Neuroscience, 32*(41), 14193–14204. https://doi.org/10.1523/JNEUROSCI.1278-12.2012

Clewett, D. V., Huang, R., Velasco, R., Lee, T. H., & Mather, M. (2018). Locus coeruleus activity strengthens prioritized memories under arousal. *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.2097-17.2017

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

da Silva Castanheira, K., Lalla, A., Ocampo, K., Otto, A. R., & Sheldon, S. (2022). Reward at encoding but not retrieval modulates memory for detailed events. *Cognition, 219*, Article 104957. https://doi.org/10.1016/j.cognition.2021.104957

de Gee, J. W., Mridha, Z., Hudson, M., Shi, Y., Ramsaywak, H., Smith, S., … McGinley, M. J. (2024). Strategic stabilization of arousal boosts sustained attention. *Current Biology: CB, 34*(18), 4114–4128.e6. https://doi.org/10.1016/j.cub.2024.07.070

Ditterich, J. (2006). Evidence for time-variant decision making. *The European Journal of Neuroscience, 24*(12), 3628–3641. https://doi.org/10.1111/j.1460-9568.2006.05221.x

Dix, A., & Li, S.-C. (2020). Incentive motivation improves numerosity discrimination: Insights from pupillometry combined with drift-diffusion modelling. *Scientific Reports, 10*(1), 2608. https://doi.org/10.1038/s41598-020-59415-3

Düzel, E., Yonelinas, A. P., Mangun, G. R., Heinze, H.-J., & Tulving, E. (1997). Event-related brain potential correlates of two states of conscious awareness in memory. *Proceedings of the National Academy of Sciences, 94*(11), 5973–5978. https://doi.org/10.1073/pnas.94.11.5973

Eldar, E., Niv, Y., & Cohen, J. D. (2016). Do You see the Forest or the tree? Neural gain and breadth versus focus in perceptual processing. *Psychological Science, 27*(12), 1632–1643. https://doi.org/10.1177/0956797616665578

Gieske, A., & Sommer, T. (2023). Independent effects of emotional arousal and reward anticipation on episodic memory formation. *Cerebral Cortex, 33*(8), 4527–4541. https://doi.org/10.1093/cercor/bhac359

Gross, M. P., & Dobbins, I. G. (2021). 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

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

Hauser, T. U., Moutoussis, M., Purg, N., Dayan, P., & Dolan, R. J. (2018). Beta-blocker propranolol modulates decision urgency during sequential information gathering. *Journal of Neuroscience, 38*(32), 7170–7178. https://doi.org/10.1523/JNEUROSCI.0192-18.2018

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*, Article 107412. https://doi.org/10.1016/j.nlm.2021.107412

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 (2006), 64*(10), 1971–1989. https://doi.org/10.1080/17470218.2011.588335

Kahneman, D. (1973). *Attention and effort*. Prentice-Hall inc.. https://www.semanticscholar.org/paper/Attention-and-Effort-Kahneman/a07ffad799cffee3ef6a2b33f4a56bffcc5b747d

Knight, M., & Mather, M. (2009). Reconciling findings of emotion-induced memory enhancement and impairment of preceding items. *Emotion (Washington, D.C.), 9*(6), 763–781. https://doi.org/10.1037/a0017281

Krause, F., & Lindemann, O. (2014). Expyriment: A Python library for cognitive and neuroscientific experiments. *Behavior Research Methods, 46*(2), 416–428. https://doi.org/10.3758/s13428-013-0390-6

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*, 1–26. https://doi.org/10.18637/jss.v082.i13

Lawlor, J., Zagala, A., Jamali, S., & Boubenec, Y. (2023). Pupillary dynamics reflect the impact of temporal expectation on detection strategy. *iScience, 26*(2), Article 106000. https://doi.org/10.1016/j.isci.2023.106000

Lisman, J., Grace, A. A., & Duzel, E. (2011). A neoHebbian framework for episodic memory; role of dopamine-dependent late LTP. *Trends in Neurosciences, 34*(10), 536–547. https://doi.org/10.1016/j.tins.2011.07.006

Lloyd, B., Miletic, S., Bazin, P.-L., Isherwood, S., Tse, D. H., Haberg, A., … Nieuwenhuis, S. (2024). Subcortical nuclei of the human ascending arousal system encode anticipated reward but do not predict subsequent memory. In *bioRxiv*, 2024–11.

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

Mather, M., Clewett, D., Sakaki, M., & Harley, C. W. (2016). Norepinephrine ignites local hotspots of neuronal excitation: How arousal amplifies selectivity in perception and memory. *The Behavioral and Brain Sciences, 39*, Article e200. https://doi.org/10.1017/S0140525X15000667

Mather, M., & Sutherland, M. R. (2011). Arousal-biased competition in perception and memory. *Perspectives on Psychological Science, 6*(2), 114–133. https://doi.org/10.1177/1745691611400234

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

McGaugh, J. L. (2004). The amygdala modulates the consolidation ofmemories of emotionally arousing experiences. *Annual Review of Neuroscience*. https://doi.org/10.1146/annurev.neuro.27.070203.144157

Miendlarzewska, E. A., Bavelier, D., & Schwartz, S. (2016). Influence of reward motivation on human declarative memory. *Neuroscience & Biobehavioral Reviews, 61*, 156–176. https://doi.org/10.1016/j.neubiorev.2015.11.015

Miller, A. L., & Unsworth, N. (2020). Variation in attention at encoding: Insights from pupillometry and eye gaze fixations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(12), 2277–2294. https://doi.org/10.1037/xlm0000797

Murayama, K., & Kitagami, S. (2014). Consolidation power of extrinsic rewards: Reward cues enhance long-term memory for irrelevant past events. *Journal of Experimental Psychology: General, 143*(1), 15–20. https://doi.org/10.1037/a0031992

Murphy, P. R., Boonstra, E., & Nieuwenhuis, S. (2016). Urgency during perceptual choice in humans. *Nature Communications, 7*(May), 1–14. https://doi.org/10.1038/ncomms13526

Naber, M., Frässle, S., Rutishauser, U., & Einhäuser, 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

O'Bryan, S. R., Price, M. M., Alquist, J. L., Davis, T., & Scolari, M. (2024). Changes in pupil size track self-control failure. *Experimental Brain Research, 242*(4), 829–841. https://doi.org/10.1007/s00221-024-06781-3

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

Pilarczyk, J., Sterna, R., Schwertner, E., Pacula, B., Bartoszek, M., & Kuniecki, M. (2022). Physiological reactions at encoding selectively predict recognition of emotional images. *Biological Psychology, 175*, Article 108429. https://doi.org/10.1016/j.biopsycho.2022.108429

Preuschoff, K., 't Hart, B. M., & Einhauser, W. (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

Reddi, B.a. J., & Carpenter, R. H. S. (2000). The influence of urgency on decision time. *Nature Neuroscience, 3*(8), 827–830. https://doi.org/10.1038/77739

Rouhani, N., Norman, K. A., & Niv, Y. (2018). Dissociable effects of surprising rewards on learning and memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 44*(9), 1430–1443. https://doi.org/10.1037/xlm0000518

Rudebeck, P. H., Putnam, P. T., Daniels, T. E., Yang, T., Mitz, A. R., Rhodes, S. E. V., & Murray, E. A. (2014). A role for primate subgenual cingulate cortex in sustaining autonomic arousal. *Proceedings of the National Academy of Sciences of the United States of America, 111*(14), 5391–5396. https://doi.org/10.1073/pnas.1317695111

Sakaki, M., Fryer, K., & Mather, M. (2014). Emotion strengthens high-priority memory traces but weakens low-priority memory traces. *Psychological Science, 25*(2), 387–395. https://doi.org/10.1177/0956797613504784

Samanez-Larkin, G. R., Gibbs, S. E. B., Khanna, K., Nielsen, L., Carstensen, L. L., & Knutson, B. (2007). Anticipation of monetary gain but not loss in healthy older adults. *Nature Neuroscience, 10*(6), 787–791. https://doi.org/10.1038/nn1894

Schneider, M., Leuchs, L., Czisch, M., Sämann, P. G., & Spoormaker, V. I. (2018). Disentangling reward anticipation with simultaneous pupillometry / fMRI. *NeuroImage, 178*(May), 11–22. https://doi.org/10.1016/j.neuroimage.2018.04.078

Shohamy, D., & Adcock, R. A. (2010). Dopamine and adaptive memory. *Trends in Cognitive Sciences, 14*(10), 464–472. https://doi.org/10.1016/j.tics.2010.08.002

Spaniol, J., Voss, A., Bowen, H. J., & Grady, C. L. (2011). Motivational incentives modulate age differences in visual perception. *Psychology and Aging, 26*(4), 932–939. https://doi.org/10.1037/a0023297

Standage, D., You, H., Wang, D.-H., & Dorris, M. C. (2011). Gain modulation by an urgency signal controls the speed-accuracy trade-off in a network model of a cortical decision circuit. *Frontiers in Computational Neuroscience, 5*, 7. https://doi.org/10.3389/fncom.2011.00007

Strauch, C., Wang, C.-A., Einhäuser, W., Van der Stigchel, S., & Naber, M. (2022). Pupillometry as an integrated readout of distinct attentional networks. *Trends in Neurosciences, 45*(8), 635–647. https://doi.org/10.1016/j.tins.2022.05.003

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). mediation: R package for causal mediation analysis. *Journal of Statistical Software, 59*, 1–38. https://doi.org/10.18637/jss.v059.i05

Tromp, J., Wurm, F., Lucchi, F., De Kleijn, R., & Nieuwenhuis, S. (2024). *Phasic alertness generates urgency and amplifies competition between evidence accumulators*. https://doi.org/10.1101/2024.06.18.599522

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie Canadienne, 26*(1), 1–12. https://doi.org/10.1037/h0080017

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., … van Mulbregt, P. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods, 17*(3), 261–272. https://doi.org/10.1038/s41592-019-0686-2

Võ, M. L.-H., 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

Wittmann, B. C., Dolan, R. J., & Düzel, E. (2011). Behavioral specifications of reward-associated long-term memory enhancement in humans. *Learning & Memory (Cold Spring Harbor, N.Y.), 18*(5), 296–300. https://doi.org/10.1101/lm.1996811

Wittmann, B. C., Schiltz, K., Boehler, C. N., & Düzel, E. (2008). Mesolimbic interaction of emotional valence and reward improves memory formation. *Neuropsychologia, 46*(4), 1000–1008. https://doi.org/10.1016/j.neuropsychologia.2007.11.020

Wittmann, B. C., Schott, B. H., Guderian, S., Frey, J. U., Heinze, H. J., & Düzel, E. (2005). Reward-related fMRI activation of dopaminergic midbrain is associated with enhanced hippocampus-dependent long-term memory formation. *Neuron, 45*(3), 459–467. https://doi.org/10.1016/j.neuron.2005.01.010

Wittmann, B. C., Tan, G. C., Lisman, J. E., Dolan, R. J., & Düzel, E. (2013). DAT genotype modulates striatal processing and long-term memory for items associated with reward and punishment. *Neuropsychologia, 51*(11), 2184–2193. https://doi.org/10.1016/j.neuropsychologia.2013.07.018

Xu, S., Coffman, D. L., Luta, G., & Niaura, R. S. (2023). Tutorial on causal mediation analysis with binary variables: An application to health psychology research. *Health Psychology : Official Journal of the Division of Health Psychology, American Psychological Association, 42*(11), 778–787. https://doi.org/10.1037/hea0001299