

This is the author's final version of an article that has been accepted for publication in *Journal of Cognitive Neuroscience*. Published by MIT Press. Kuipers JR & Phillips WA (2020) Variations in pupil size related to memory for recently presented words and event related potentials. *Journal of Cognitive Neuroscience*. [https://doi.org/10.1162/jocn\\_a\\_01638](https://doi.org/10.1162/jocn_a_01638)

# Variations in pupil size related to memory for recently presented words and event related potentials

JR Kuipers<sup>1\*</sup>

WA Phillips<sup>1</sup>

1. University of Stirling

School of Natural sciences

Psychology

\*Corresponding author  
[jk28@stir.ac.uk](mailto:jk28@stir.ac.uk)

## Abstract

Pupillometry has been found to be correlated with activity of Cholinergic (ACh) and Noradrenergic (NE) neuromodulator systems. These systems regulate the level of cortical arousal and therefore perception, attention, and memory. Here we tested how different types of pupil size variance (prestimulus baseline and prestimulus hippus power) may correlate with behavioural and electrophysiological brain responses (ERPs). We recorded pupil size and ERPs whilst participants were presented with a series of words and then asked whether they had been in the initial list when they were later presented intermixed with unrepresented words. We found that a smaller prestimulus baseline pupil size during the study phase was associated with better memory performance. Study items also evoked a larger P3 response at presentation and a greater old/new memory ERP effect at test when prestimulus pupil size was small rather than large. Prestimulus hippus power was found to be a between-subjects factor affecting the robustness of memory encoding with less power being associated with a greater old/new memory ERP effect. These results provide evidence relating memory and ERPs to variables defined on pupil size that are thought to reflect varying states of parasympathetic and sympathetic arousal.

## Introduction

Memory depends on the ability to prioritize signals that are relevant to the current situation whilst attenuating irrelevant signals. However, signal transmission within the brain is inherently noisy (Averbeck, Latham, & Pouget, 2006; Schmitz & Duncan, 2018) and dependent on continuously varying levels of arousal of the individual (Cohen & Maunsell, 2011; McGinley et al., 2015). The level of arousal is predominantly determined by the activity of noradrenergic (NE) and cholinergic (ACh) neuromodulator systems (Berridge & Waterhouse, 2003; McCormick, Pape, & Williamson, 1991). These neuromodulator systems influence whether a signal is amplified or attenuated depending on its relevance or prioritisation (Mather, Clewett, Sakaki, & Harley, 2016; Phillips, Larkum, Harley, & Silverstein, 2016). The effects of prioritisation increase with arousal, and a highly influential theory proposes that higher levels of arousal imply higher neural gain (Aston-Jones & Cohen, 2005). It has been hypothesized that with high neural gain, the net amount of neural activity related to an attended stimulus is large compared to other neural activity, which leads to better stimulus detection (Corbetta, Patel, & Shulman, 2008) and memory encoding (Mather et al., 2016).

Several studies suggest that neural gain is modulated by activity of the Locus Coeruleus (LC; Berridge & Waterhouse, 2003; Gilzenrat, Nieuwenhuis, Jepma, & Cohen, 2010; Jepma & Nieuwenhuis, 2011), which is the sole source of NE for the forebrain (Samuels & Szabadi, 2008). How the LC-NE system may regulate neural gain is addressed in the GANE model by Mather et al. (Mather et al., 2016). In GANE, high arousal leads to a surge of NE release (Counts & Mufson, 2012), which amplifies already active representations and inhibits low activity representations. Spillover of synaptic glutamate release, which is a characteristic of high neural activity, stimulates the release of more NE from nearby LC varicosities, which in turn increases neural firing and spillover of local glutamate. This creates local “hotspots” of very high NE concentrations which contrast with low activity neural

regions where there is a low level of glutamate and, thus low levels of NE. Importantly, low levels of NE have an inhibiting effect, which makes the highly active representations stand out most. GANE therefore explains on a neural level how salient details are remembered whereas other details may be forgotten (Cahill, Gorski, & Le, 2003).

Activity of the LC-NE neuromodulator system is correlated with pupil size (Murphy, O'Connell, O'Sullivan, Robertson, & Balsters, 2014; Reimer et al., 2016), which makes it possible to study how LC-NE activity influences human cognitive performance without the need for intracranial recordings (Aston-Jones & Cohen, 2005; Gilzenrat et al., 2010; McGinley et al., 2015). However, different proposals have been put forward regarding how LC-NE activity may relate to pupil size variance. For example, greater neural gain has been assumed to be reflected in (1) a large pupil dilation response to a stimulus (Aston-Jones & Cohen, 2005; Mather et al., 2016; Reimer et al., 2014; Shine, 2019), or (2) a small pupil size response (Eldar, Niv, & Cohen, 2016; Hauser, Eldar, Purg, Moutoussis, & Dolan, 2019), or (3) the maximal derivative of pupil size fluctuations (McGinley et al., 2015; Reimer et al., 2016). These different views may reflect the complex interactions between ACh, NE, and pupil size, which are not yet adequately understood (Larsen & Waters, 2018). Moreover, in some studies arousal levels are typically low (e.g. in a signal detection task; Eldar et al., 2016), whereas in other studies participants' arousal may have implicitly been increased by including anticipation of reward (as in Jepma & Nieuwenhuis, 2011). Critically, arousal changes the way NE influences neural excitability. Under low arousal, NE binds to low-threshold alpha-2 receptors which have an inhibiting effect on neural firing (Berridge & Waterhouse, 2003). However, high arousal leads to surge of NE release leading to saturation of alpha-2 receptors. The surplus NE will bind to high-threshold alpha-1 and beta receptors, which increase firing rate of cortical layer 5 pyramidal cells when two conditions are met. The first of these is that the cell was already activated by bottom-up signals arriving at the soma (Larkum, 2013). The second is that the cortical layer 1 apical tuft, where alpha-1 and beta

receptors can be activated by NE, is connected to the soma by cholinergic and metabotropic glutamate receptors in the apical trunk (Suzuki & Larkum, 2020). The link between pupil size variance and neural gain is therefore far from straightforward and probably not the same under different levels of arousal.

For example, in the rat, baseline and evoked pupil responses are not linked to neuromodulatory activity in the same way. Baseline pupil size correlates more strongly with ACh activity than with NE activity and fast phasic responses correlate more strongly with NE activity (Reimer et al., 2016). Further, pupil size is never stationary and it usually fluctuates with a frequency around 0.3Hz, called Hippus (Bouma & Baghuis, 1971). Hippus has been associated with cholinergic activity (Turnbull, Irani, Lim, & Phillips, 2017), but further study of this association is needed.

In the current study we investigated how different types of human pupil size variation correlates with behavioural measures and event-related brain potentials (ERPs) in a test of memory for recently presented words. ERPs are sensitive indices of cognitive processes and cortical arousal. They may therefore provide insight into how temporary fluctuations in net arousal, as indexed by pupil size, are linked to both electrophysiological measures and performance in the memory task. Importantly, we maintained low to medium arousing testing conditions by having a slow pace of stimulus presentation, no requirement of a response on the part of the participant (in the learning phase), and they were seated alone in a quiet testing room performing a simple, unrewarded task.

We used a memory task in which participants were presented with a series of words during an encoding phase, which were intermixed with new words during a recall phase. We measured amplitude of the P3 (here P3b), which reflects the strength of the attention-orienting response (Crowley & Colrain, 2004; Donchin & Coles, 1988; Polich, 2007). Previous studies have already associated the P3 ERP with the pupil dilation response (Friedman, Hakerem, Sutton, & Fleiss, 1973), and indirectly with LC-NE activity and neural gain (Jepma &

Nieuwenhuis, 2011; Murphy, Robertson, Balsters, & O'Connell, 2011; Nieuwenhuis, Aston-Jones, & Cohen, 2005). We also recorded the left-parietal old/new memory effect which reflects encoding strength of learned items (Rugg & Curran, 2007). We reasoned that the cortical arousal during study would be reflected in the attention-orienting response to words during study, whilst the strength of encoding would be reflected in the amplitude of the old/new ERP effect and recollection accuracy measured in the testing phase of the experiment. Optimal levels of arousal during study may therefore be reflected in a greater P3 amplitude during study, better recall, and a greater old/new memory effect.

We linked the P3, the old/new ERP effect and recollection accuracy markers of cortical arousal to two different types of pupil size variance by splitting the trials of each participant based on prestimulus baseline pupil size and prestimulus hippus power. In a related study, Murphy et al. (2011) reported that pupil size has an inverted U-relationship with task performance and P3 amplitude in an auditory oddball task. Notably, they only found this for trials sorted in five bins by prestimulus pupil size and not for the evoked pupil dilation responses. However, the inverted-U effects of arousal on performance are suggested to be lowest at the extremes of the arousal spectrum and optimal at an intermediate level (Aston-Jones & Cohen, 2005). The oddball task Murphy et al. (2011) used is rather monotonous and participants only need to respond in 20% of the trials. Therefore, one can assume that arousal levels are low to medium in this task, which makes it unclear why they found the full inverted-U pattern (which spans from almost asleep to a state of panic) in their data. In the study phase of our experiment we expected arousal levels to be low to intermediate, which is why we sorted trials into two bins for each measure of pupil size variance.

## Methods

### Participants

55 psychology undergraduates were recruited to take part in this study. However, 15 participants were excluded from analysis due to excessive artefacts in the EEG and/or pupillometry data, too many errors in the task or not doing the task as instructed, or self-reported memory problems. The remaining 40 participants had a mean age of 20.9 years, 27 were female, and 34 were right-handed (Oldfield, 1971).

### Materials

#### *Stimuli*

Of 200 highly familiar English words (Table 1), four pseudo-randomised lists were created in which 100 were presented as to-be-learned stimuli, randomly intermixed with the other 100 presented during the recall phase. Participants were presented with one of the four lists. Probe words were presented on top of a picture of an indoor (furnished room) or outdoor (coastline) scene to assess source memory (Rugg & Curran, 2007). Response mapping and list number were counterbalanced across participants.

*Table 1. Values for means and standard deviations derived from the stimulus list.*

	Concreteness	Familiarity	Kucera-Francis freq.	Number of letters
M	499.53	500.21	12.46	5.49
SD	105.67	50.99	5.76	1.08
N	101	112	200	200

Note. The n-row depicts how many words were included in each measure since the MRC database does not hold information on all words used,.

### *Apparatus*

Stimuli were presented on a Tobii 1750 eye tracker using E-Prime (www.pstnet.com) while electrophysiological data was recorded using a Synamps II amplifier (www.neuroscan.com) connected to a 64-channel Ag/Ag/Cl Quick-Cap (NeuroMedical Supplies). Electrodes on the cap were arranged according to the extended version of the 10-20 system. Vertical and horizontal eye movements were monitored using pairs of electrodes placed near the outer canthi of both eyes and one above and one below the left eye. An additional electrode was placed over each mastoid bone.

### *Data pre-processing*

EEG data was recorded at a rate of 1 KHz while band-pass filtered between 0.1 Hz and 200 Hz, using a reference electrode located between CZ and CPZ. Impedance of all electrodes was kept below 5 KOhm. Offline EEG data was filtered using a low-pass filter at 30 Hz (48 db/Oct). The continuous data was re-referenced to the average of the mastoid electrodes, mathematically corrected for eye blinks (Gratton, Coles, & Donchin, 1983), visually inspected, and epoched from -100 to 900 ms relative to stimulus onset. Epochs were baseline corrected over the pre-stimulus interval and rejected when any cap electrode exceeded +/- 75 $\mu$ V.

Pupil size data was pre-processed using in-house R scripts in which we averaged the size of the two eyes, applied a linear interpolation to blinks (maximal gap 240 ms), and applied a low-pass filter of 10Hz.

### **Procedure**

Each trial started with presentation of 5 hash symbols at fixation on which participants were asked to keep focussed. If more than 1s fixation was achieved the hash marks were replaced by a word printed in black on top of a picture of a coastline (outdoors context) or a



furnished room (indoors context) presented for 1s. The trial finished with presentation of hash marks at fixation for 5 s creating an inter trial interval of 6 s, or more when fixation had not been achieved (Figure 1). Visual angle of the word was maximally 11 degrees and the picture's was 15 degrees.

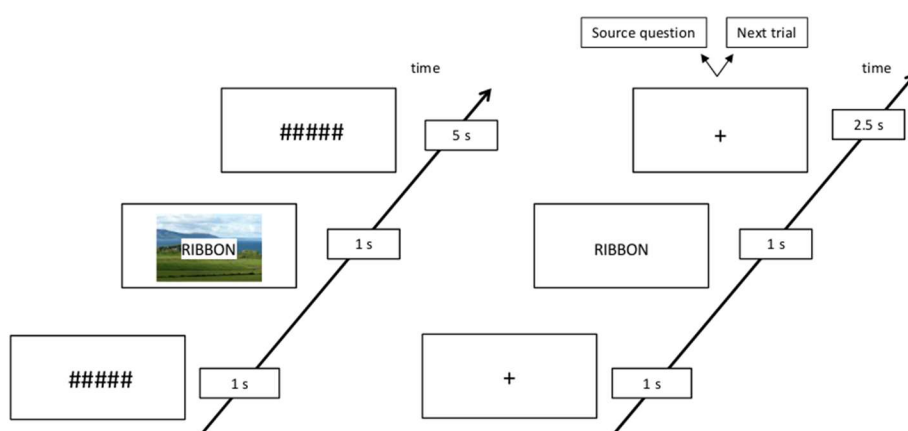


Figure 1. Trial procedures.

In the study phase, left, each trial began with presentation of 5 hash marks at fixation for 1 s to check the participant's fixation. Next, the target stimulus was presented for 1s followed by a fixation cross for 5 s. In the test phase, right, each trial started with presentation of a fixation cross for 1 s, followed by the target word for 1s, and a fixation cross for 2.5 s during which participants could make their response. The relative sizes of the stimuli and screen are not to scale. In case a word was correctly remembered a question appeared about which picture the word was presented with during study. Pupil size and ERPs were recorded during the study phase, whilst and the participant's response time and accuracy are recorded.

In the test phase of the experiment, participants were asked to make an old/new judgement on all 200 words using a button press. Each trial began with presentation of a fixation cross for 500 ms, followed by the word for 1s, which was replaced by a fixation cross presented for 2.5 s. After each correctly remembered "old" item, they were also asked which

context picture the word had been presented with. However, these source memory responses were not included in the analyses due to a too low number of usable trials after all conditions for inclusion were satisfied (noise-free EEG and pupil data, correctly remembered item, and correctly remembered context picture).

## Analysis

Prestimulus baseline pupil size was calculated using a 1s time-window before each target-stimulus onset. For each participant, we sorted trials based on the average prestimulus pupil size and made two conditions (large vs. small) by assigning each half to one of these conditions. This resulted in an average difference of .53 mm prestimulus size between the two conditions ( $t_{(39)} = -17, p < .001$ ). Grand average plots of these conditions (Figure 2) showed a small baseline was followed by an increase in pupil size whereas a large one was followed by a decrease in pupil size. Therefore, it appears the pupil made an accommodation response upon presentation of the word and picture, causing pupil size to converge to approximately the same size after target onset. Therefore, the evoked pupil size response in this case may reflect processes of accommodation rather than reflecting the state of neuromodulator systems or arousal.

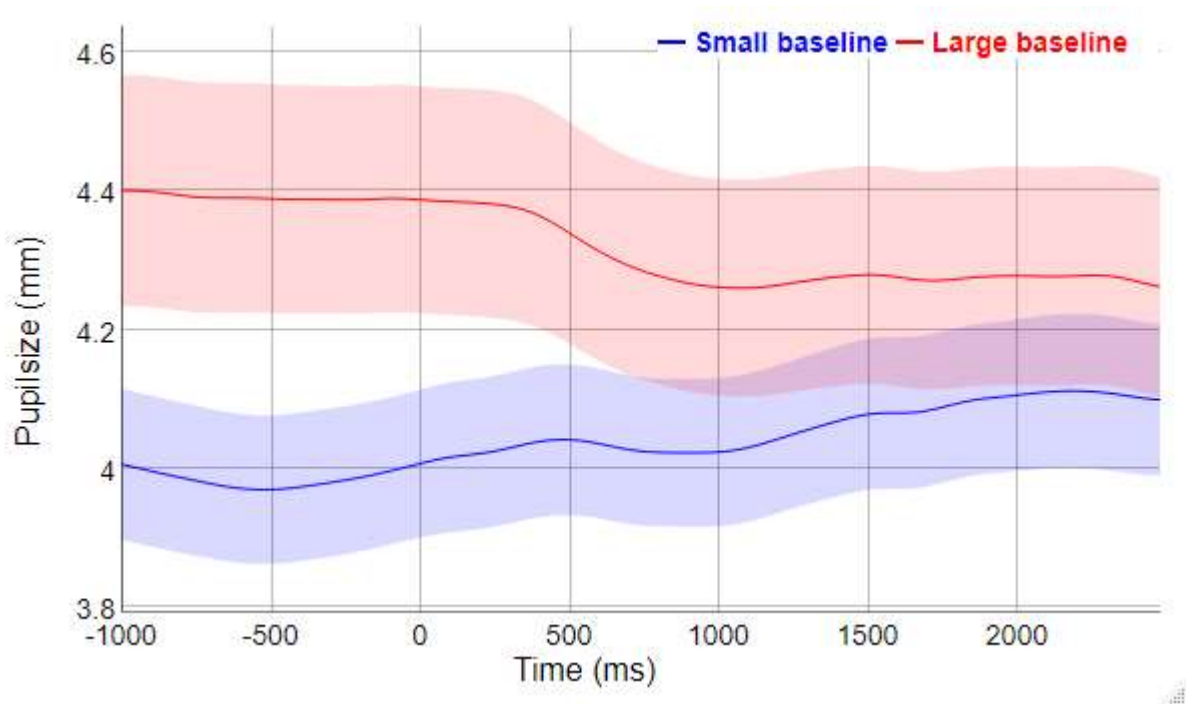


Figure 2. Prestimulus baseline pupil size plots

Grand average pupil size plot for trials sorted by prestimulus pupil size. Stimulus onset is at 0 ms and shaded areas depict the SE.

Given that the maximum power of hippus is at approximately .3 Hz (Turnbull et al., 2017), we used a bandwidth of .2 to .4 Hz determine hippus power from 5s prestimulus epochs. The amount of hippus power differed significantly between conditions (much power: 42.8 little power: 3.9;  $t_{(40)} = -7.5$ ,  $p < .001$ ; Figure 3).

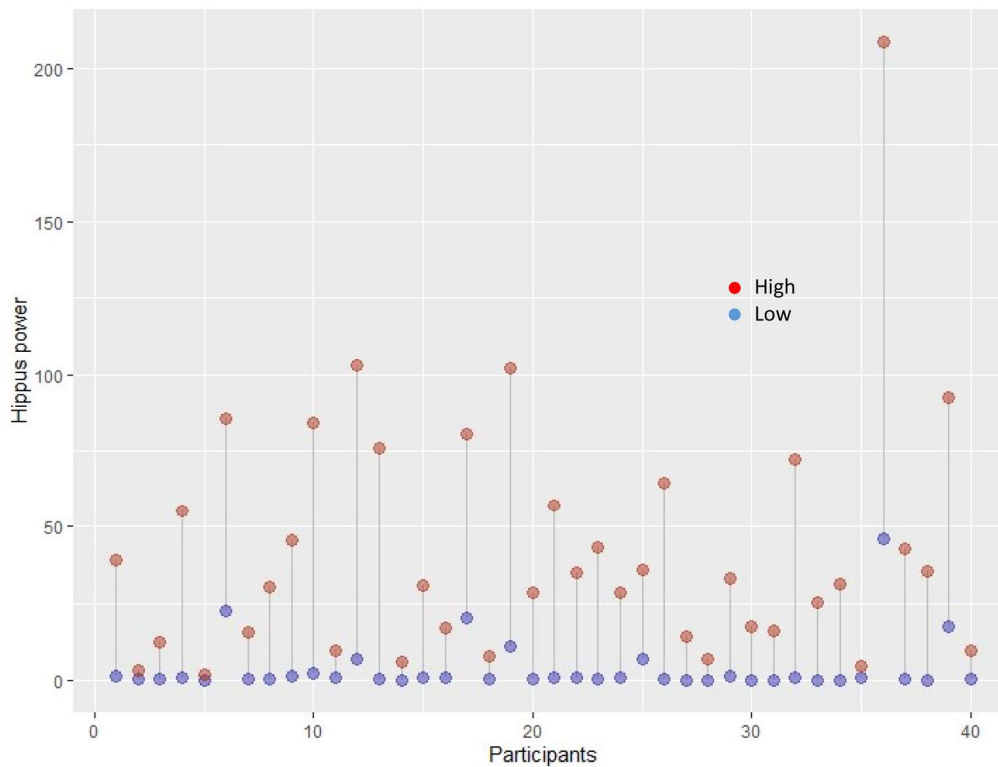


Figure 3. Hippus difference for each participant  
Lollipop plots of the sum of hippus power (.2 to .4 Hz) during the 5 s. prestimulus interval for each participant. Participant 36 was an outlier with very high hippus power which is why this participant was excluded from the correlation analysis. They did not influence the rest of the results and was therefore included in the other analyses.

As mentioned above, ERPs of interest were the (parietal) P3 during study, and, for the test phase of the experiment, the Left-Parietal old/new memory effect (Rugg & Curran, 2007). Visual inspection of grand average ERP plots of words presented at the different conditions showed the expected modulations of these ERPs, albeit not in the same way for all conditions. The P3 was analysed over centroparietal electrodes (CP1, CP2, CPz; Polich, 2007) from 310 ms to 390 ms relative to word onset. The Left-Parietal Old-New effect was analysed over electrodes CP1, CP3, CP5 from 600 ms to 800 ms relative to target word onset (Rugg & Curran, 2007). Finally, we correlated individual prestimulus and poststimulus pupil size, and hippus power with the average amplitude of the ERPs and behavioural measures to investigate how between-subjects differences in pupil size may be associated with the ERP

components we recorded. Greenhouse-Geisser corrections are reported, where applicable, in the analyses below.

## Results

### *Section 1. prestimulus baseline pupil size during study*

Analysis of response accuracy revealed that response accuracy was greater for small prestimulus pupil size than large prestimulus pupil size (Table 2). The difference in number of hits (correctly identified “old” items) between large and small prestimulus baseline pupil size approached significance (large: 21.8, SE =1.3; small: 22.9, SE =1.2;  $t_{(39)} = -1.9$ ,  $p = .06$ ), whilst the difference in the number of misses (failures to identify an item as “old”) was significant (large: 16.0, SE =1.2; small: 14.6, SE =1.2;  $t_{(39)} = -2.5$ ,  $p = .016$ ). RTs did not differ between conditions (large: 916 ms SE = 51, small: 904 ms, SE = 47;  $p = .5$ ).

Table 2. Average signal detection values in the different conditions. The \* characters indicate significant differences ( $p < .05$ , 2-tailed).

		Hits	Miss	CR	FA	D'	Crit
Prestimulus	large	21.8	16.1*	26.7	82.3	0.88*	0.25*
pupil size	small	22.9	14.6	26.7	82.3	0.97	0.20
Prestimulus	High	22.3	15.7	26.7	82.3	0.91	0.24
Hippus Power	Low	22.6	14.8	26.7	82.3	0.95	0.21

Next, we analysed P3 amplitude elicited by the to-be-memorised words and the Old-new effect these words elicited in the test phase of the experiment. Visual inspection of the grand average waveforms revealed a large modulation of the P3 (Fig. 4). Analysis of mean amplitude of these components showed the P3 modulation was significant (a difference of  $1.8\mu\text{V}$ ;  $t_{(39)} = -4.1$ ,  $p < .001$ ), showing that words presented at small prestimulus pupil size during study elicited a greater P3 response than words presented when pupil size was relatively large. The Anova on the old/new memory effect of the test phase revealed that the

three conditions (new,  $2.3\mu\text{V}$  SE = .4; old-large,  $3.1\mu\text{V}$  SE = .6; old-small,  $3.57\text{V}$  SE = .6; Fig. 4) differed significantly ( $F(1.6,65) = 3.5$ ,  $p = .04$ ). Planned comparisons showed the old/new effect was significant for small prestimulus pupil sizes ( $p = .01$ ), but surprisingly, not for large prestimulus pupil sizes ( $p = .26$ ), whilst the two “old” conditions did not differ ( $p = .09$ ).

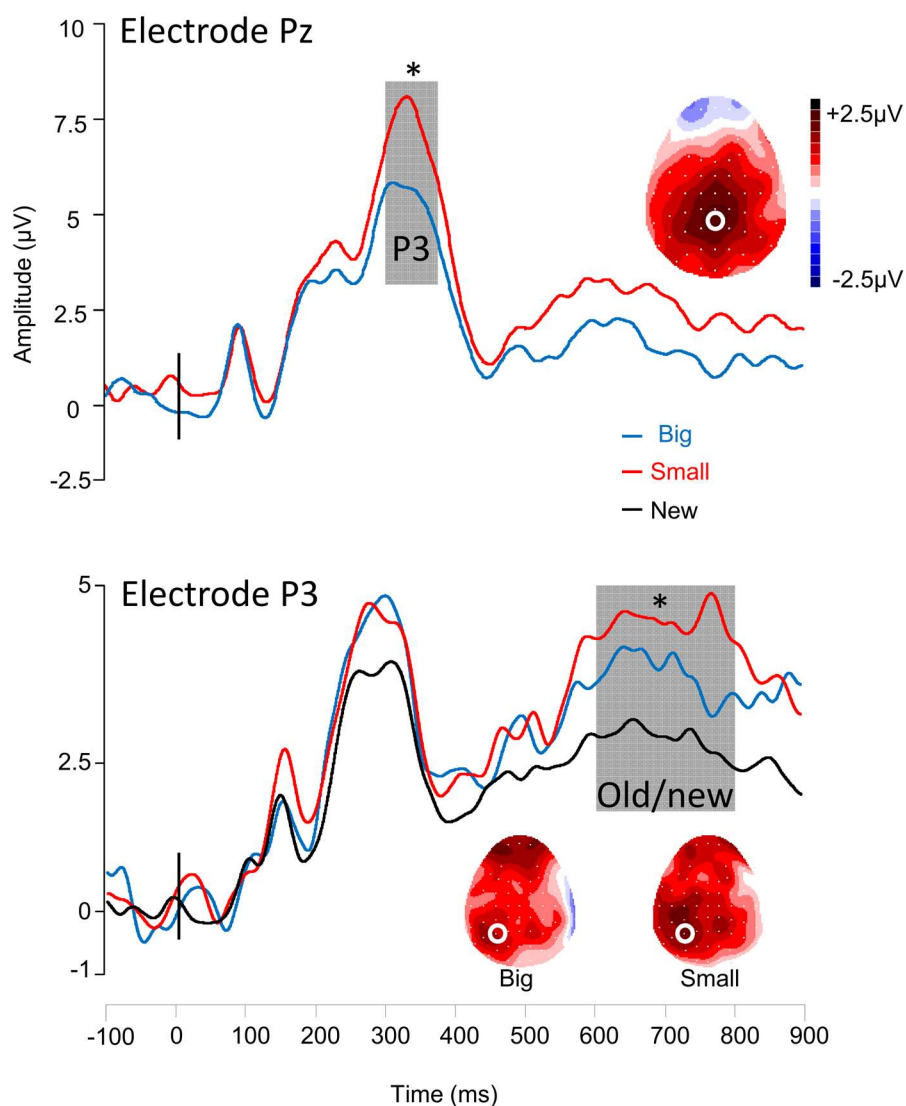


Figure 4. Prestimulus baseline ERP plots

Grand average plot of 1 s. prestimulus intervals for each of the small (blue lines) and large (red lines) prestimulus pupil size. Stimulus onset is at 0 ms and shaded areas show the SE.

*Section 2. Prestimulus Hippus power during the study phase*

Prestimulus hippus power did not influence the number of hits (high: 22.2 SE = 1.3, low: 22.6 SE = 1.2), the number of misses (high: 15.6 SE = 1.2, low: 14.8 SE = 1.1) or RTs (high: 917 ms SE = 50; low: 908 ms SE = 49; all  $p > .1$ ). Grand average ERP plots of the encoding phase revealed a P3 modulation by the amount of Hippus power, albeit not as strongly as in the prestimulus pupil size conditions (Fig. 5). Accordingly, the P3 effect over the same electrodes as used in the other analyses revealed that the difference failed to reach significance ( $t_{(39)} = -1.6$ ,  $p = .1$ ), and was only significant at the electrode of maximal modulation, Pz (high hippus power:  $5.8\mu\text{V}$  SE = .6; low hippus power:  $6.6\mu\text{V}$  SE = .6;  $t_{(39)} = -2.3$ ,  $p = .029$ ). The ANOVA on the old/new memory effect revealed the overall effect of condition was significant (new:  $2.3\mu\text{V}$  SE = .4; high  $4.0\mu\text{V}$  SE = .6; low:  $4.3\mu\text{V}$  SE = .7;  $F_{(2,78)} = 8.1$ ,  $p = .001$ ). Planned comparisons showed the old/new effect was significant for both hippus power conditions ( $p = .002$ ), whilst these conditions did not differ from each other ( $p = .6$ ).

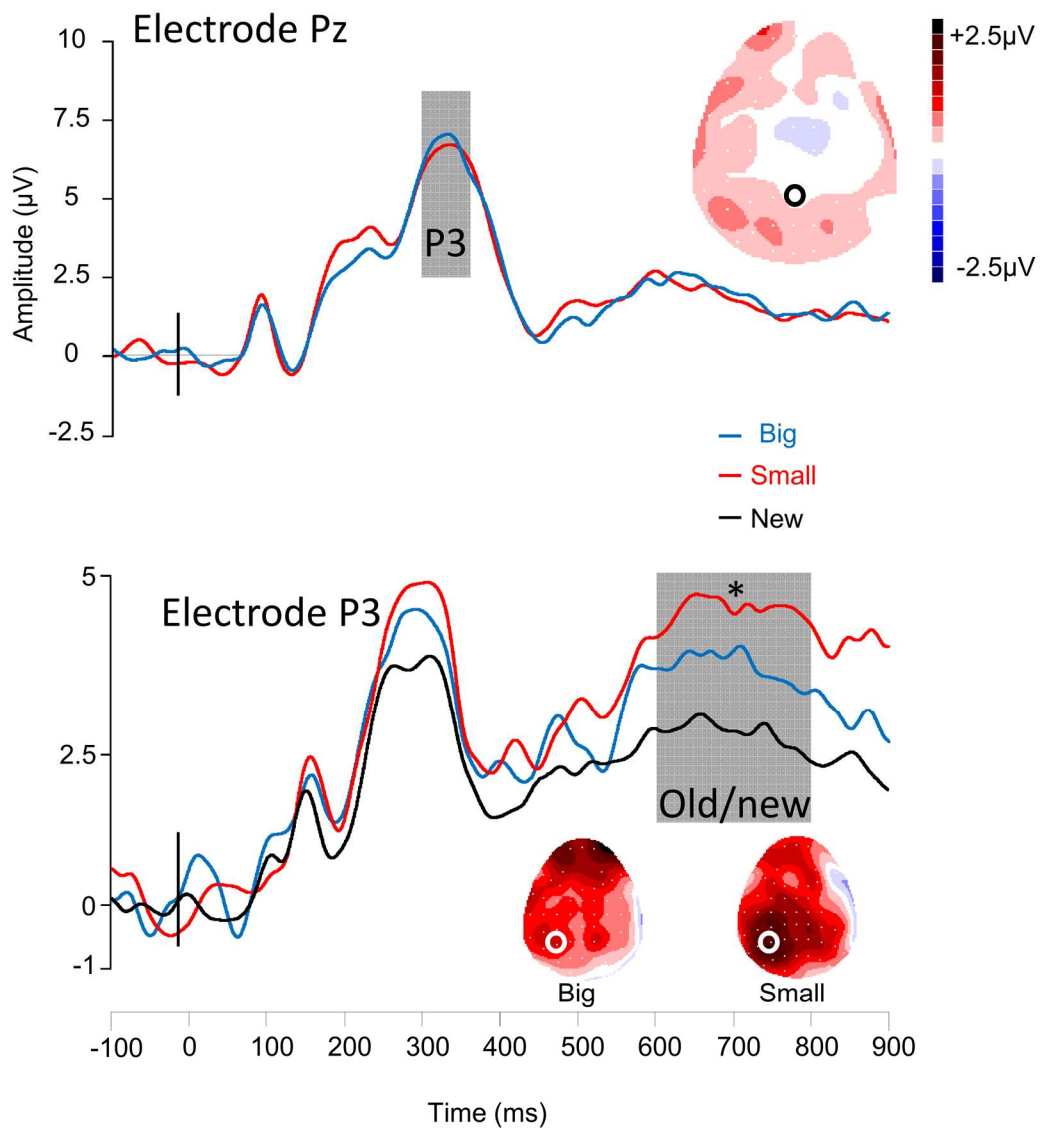


Figure 5. Hippus power ERP plots.

Grand average plot of trials that were sorted by their maximum during the 4 s. post-stimulus interval of the small (blue lines) and large (red lines) prestimulus conditions. Plots were baseline corrected over the 1 s prestimulus interval. Stimulus onset is at 0 ms, depicted by the vertical line. Blue lines represent the small evoked responses and red lines the large ones.



#### *Section 4. Cross-measure correlations*

Finally, we averaged the values of each measure across conditions for a correlation analysis. The results revealed that no other measure was correlated with prestimulus pupil size, which shows the within-participants variation in these pupil size measures is driving the effects reported above. However, Hippus power was significantly anticorrelated with the Old/New effect ( $R = -.33$ ,  $p = .04$ ; excluding participant 36 who would increase this correlation to  $R = -.6$ ), showing that participants with more prestimulus Hippus during study display a smaller old/new memory effect. Further, significant correlations were observed between P3 amplitude and the number of hits ( $R = .35$ ,  $p = .028$ ) and between P3 amplitude and RTs ( $R = -.33$ ,  $p = .041$ ), which shows that participants who display a greater P3 to words in the learning phase have a greater response accuracy and shorter response times. Finally, the number of hits and RTs were significantly correlated ( $R = -.46$ ,  $p = .003$ ), showing that participants who have greater recollection accuracy displayed shorter RTs on average.

#### **Discussion**

Although there is sufficient evidence for a link between LC-NE activity and pupil size, different proposals have been put forward regarding how pupil variance may be associated with LC-NE activity and cognitive functioning. Here we investigated how two types of pupil variation are related to ERPs and behavioural measures. The main goals of the current experiment were to investigate which, if any, ERP component would correlate with these types of pupil size variance, and to relate those types of variance to memory performance.

We found better memory performance for stimuli presented at small than large prestimulus pupil size at study. This finding is in line with previous reports such that a small baseline pupil size is associated with better target stimulus detection (Gilzenrat et al., 2010; Olivia, 2019), and here we show that stimuli are more robustly encoded into memory when

prestimulus pupil baseline is relatively small. We also found that P3 amplitude in response to to-be-memorised words was larger for small than large prestimulus pupil size and that that greater P3 amplitude was associated with shorter RTs in the recall stage of the experiment. The parietal P3 is well known to reflect target detection/categorisation and response selection, but it has also been associated with the amount of available neural resources engaged in the attention orienting response (Donchin & Coles, 1988; Polich, 2007). The words in the study phase did not require a manual response which implies that the P3 we observed likely reflects a phasic noradrenergic arousal response while words were encoded into memory. Finally, we observed that words presented when pupil size was relatively small during encoding elicited a greater old/new memory effect than those presented at a relatively large prestimulus pupil size. This left-parietal old/new effect reflects the quality of stimulus encoding into memory (Rugg & Curran, 2007). Together, these results firmly support the interpretation that small prestimulus pupil size under low arousing conditions is associated with a greater attentional response and a more robust stimulus encoding. This pattern of results fits a scenario in which low tonic LC-NE activity, and thus little engagement of alpha-1 and beta adrenergic receptors, facilitates a phasic LC-NE response and activation of their receptors on the neurons involved in encoding the target stimulus.

By contrast, prestimulus hippus power did not influence performance in our memory task. Instead, low hippus power was associated with greater P3 amplitude and participants with less Hippus on average displayed a larger old/new memory effect. Hippus has been associated with parasympathetic activity (Turnbull et al., 2017) and our finding that high hippus power is associated with a low attentional ERP response and less robust memory encoding supports this notion. However, given that the within-participants effect was small and the between subjects modulation strong, the variation in hippus power does not appear to change much on a relatively short timescale (e.g., 30 minutes) under sustained low arousal

conditions. Instead, some participants may have been more tired than others leading to the between-subjects differences in hippus power activity.

A limitation of the current study is the fact we used visual (word) stimuli which complicates analysis of the evoked pupil response that has been proposed to reflect phasic LC activity (e.g., Aston-Jones & Cohen, 2005). Here, we found that pupil size converged after stimulus onset, which suggests that visual (word) stimuli may not be the most suitable for studying evoked pupil size responses. In addition, emotional stimuli, surprising stimuli, or the requirement of a response (in the learning phase) all induce a large phasic pupil dilation response (Beatty, 1982; Bradley, Miccoli, Escrig, & Lang, 2008; Hess & Polt, 1964), which needs to be taken into account when studying the relation between evoked pupil size responses, cognitive processes and performance.

To conclude, we investigated potential links between pupillometry measures, ERPs, encoding, and retrieval when arousal levels were low to medium. As mentioned in the introduction, baseline pupil size under low arousal is associated with both cholinergic and noradrenergic activity whilst large phasic pupil size changes reflect bursts of noradrenergic activity (Reimer 2016). Given that we observed that trials with smaller prestimulus pupil size were associated with greater electrophysiological and behavioural measures of attention and memory encoding, low baseline levels of ACh and NE are probably associated with more effective stimulus processing than high baseline levels. Periods of relatively large baseline pupil size (under low arousal conditions), appear sub-optimal for processing a new stimulus, similar to mind wandering which has also been associated with relatively large baseline pupil size (Smallwood et al., 2011).

Little is known about Hippus so our study systematically investigated the relation between Hippus and brain state during tests of memory. Our findings are in line with the suggestion that hippus is associated with parasympathetic activity such that high hippus power reflects high parasympathetic activity. Our results suggest different types of pupil size

variation reflect different aspects of cortical arousal and may therefore provide insight into both parasympathetic and sympathetic neuromodulatory activity. However, it is unknown how the net level of arousal would influence the relationships we report here, since increases in arousal are associated with increases in baseline pupil size, phasic pupil responses, and LC activity (McGinley et al., 2015; Reimer et al., 2016).

## References

- Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*, 403-450. doi:10.1146/annurev.neuro.28.061604.135709
- Averbeck, B. B., Latham, P. E., & Pouget, A. (2006). Neural correlations, population coding and computation. *Nature Reviews Neuroscience*, *7*(5), 358-366. doi:10.1038/nrn1888
- Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, *91*(2), 276-292. doi:10.1037/0033-2909.91.2.276
- Berridge, C. W., & Waterhouse, B. D. (2003). The locus coeruleus-noradrenergic system: modulation of behavioral state and state-dependent cognitive processes. *Brain Research Reviews*, *42*(1), 33-84. doi:10.1016/s0165-0173(03)00143-7
- Bouma, H., & Baghuis, L. C. J. (1971). Hippus of the pupil - periods of slow oscillations of unknown origin. *Vision Research*, *11*(11), 1345-1351. doi:10.1016/0042-6989(71)90016-2
- Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, *45*(4), 602-607. doi:10.1111/j.1469-8986.2008.00654.x
- Cahill, L., Gorski, L., & Le, K. (2003). Enhanced human memory consolidation with post-learning stress: Interaction with the degree of arousal at encoding. *Learning & Memory*, *10*(4), 270-274. doi:10.1101/lm.62403
- Cohen, M. R., & Maunsell, J. H. R. (2011). When Attention Wanders: How Uncontrolled Fluctuations in Attention Affect Performance. *Journal of Neuroscience*, *31*(44), 15802-15806. doi:10.1523/jneurosci.3063-11.2011
- Corbetta, M., Patel, G., & Shulman, G. L. (2008). The reorienting system of the human brain: From environment to theory of mind. *Neuron*, *58*(3), 306-324. doi:10.1016/j.neuron.2008.04.017
- Counts, S. E., & Mufson, E. J. (2012). Locus Coeruleus. In J. K. Mai & G. Paxinos (Eds.), *The Human Nervous System* (pp. 425-438): Academic Press.
- Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: age, sleep and modality. *Clinical Neurophysiology*, *115*(4), 732-744. doi:10.1016/j.clinph.2003.11.021
- Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, *11*(3), 357-374. doi:10.1017/s0140525x00058027
- 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. doi:10.1177/0956797616665578
- Friedman, D., Hakerem, G., Sutton, S., & Fleiss, J. L. (1973). Effect of stimulus uncertainty on pupillary dilation response and vertex evoked-potential. *Electroencephalography and Clinical Neurophysiology*, *34*(5), 475-484. doi:10.1016/0013-4694(73)90065-5
- Gilzenrat, M. S., Nieuwenhuis, S., Jepma, M., & Cohen, J. D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive Affective & Behavioral Neuroscience*, *10*(2), 252-269. doi:10.3758/cabn.10.2.252
- Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*(4), 468-484. doi:10.1016/0013-4694(83)90135-9
- Hauser, T. U., Eldar, E., Purg, N., Moutoussis, M., & Dolan, R. J. (2019). Distinct Roles of Dopamine and Noradrenaline in Incidental Memory. *Journal of Neuroscience*, *39*(39), 7715-7721. doi:10.1523/jneurosci.0401-19.2019

- Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science*, *143*(361), 1190-1191. doi:10.1126/science.143.3611.1190
- Jepma, M., & Nieuwenhuis, S. (2011). Pupil Diameter Predicts Changes in the Exploration-Exploitation Trade-off: Evidence for the Adaptive Gain Theory. *Journal of Cognitive Neuroscience*, *23*(7), 1587-1596. doi:10.1162/jocn.2010.21548
- Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. *Neuron*, *89*(1), 221-234. doi:10.1016/j.neuron.2015.11.028
- Larkum, M. (2013). A cellular mechanism for cortical associations: an organizing principle for the cerebral cortex. *Trends in Neurosciences*, *36*(3), 141-151. doi:10.1016/j.tins.2012.11.006
- Larsen, R. S., & Waters, J. (2018). Neuromodulatory Correlates of Pupil Dilation. *Frontiers in Neural Circuits*, *12*, 9. doi:10.3389/fncir.2018.00021
- 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. *Behavioral and Brain Sciences*, *39*. doi:10.1017/s0140525x15000667
- McCormick, D. A., Pape, H. C., & Williamson, A. (1991). Actions of norepinephrine in the cerebral-cortex and thalamus - implications for function of the central noradrenergic system. *Progress in Brain Research*, *88*, 293-305.
- McGinley, M. J., Vinck, M., Reimer, J., Batista-Brito, R., Zaghera, E., Cadwell, C. R., . . . McCormick, D. A. (2015). Waking State: Rapid Variations Modulate Neural and Behavioral Responses. *Neuron*, *87*(6), 1143-1161. doi:10.1016/j.neuron.2015.09.012
- 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. doi:10.1002/hbm.22466
- Murphy, P. R., Robertson, I. H., Balsters, J. H., & O'Connell, R. G. (2011). Pupillometry and P3 index the locus coeruleus-noradrenergic arousal function in humans. *Psychophysiology*, *48*(11), 1531-1542. doi:10.1111/j.1469-8986.2011.01226.x
- MRC Psycholinguistic Database (1987). Informatics Division Science and Engineering Research Council Rutherford Appleton Laboratory Chilton.
- Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the p3, and the locus coeruleus-norepinephrine system. *Psychological Bulletin*, *131*(4), 510-532. doi:10.1037/0033-2909.131.4.510
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97-113. doi:10.1016/0028-3932(71)90067-4
- Olivia, M. (2019). Pupil size and search performance in low and high perceptual load. *Cognitive, Affective & Behavioral Neuroscience*, *19*, 366-376. doi: 10.3758/s13415-018-00677-w
- Phillips, W. A. (2017). Cognitive functions of intracellular mechanisms for contextual amplification. *Brain and Cognition*, *112*, 39-53. doi:10.1016/j.bandc.2015.09.005
- Phillips, W. A., Larkum, M. E., Harley, C. W., & Silverstein, S. M. (2016). The effects of arousal on apical amplification and conscious state. *Neuroscience of Consciousness*(1), 13. doi:10.1093/nc/niv015
- Polich, J. (2007). Updating p300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128-2148. doi:10.1016/j.clinph.2007.04.019
- Reimer, J., Froudarakis, E., Cadwell, C. R., Yatsenko, D., Denfield, G. H., & Tolias, A. S. (2014). Pupil Fluctuations Track Fast Switching of Cortical States during Quiet Wakefulness. *Neuron*, *84*(2), 355-362. doi:10.1016/j.neuron.2014.09.033
- Reimer, J., McGinley, M. J., Liu, Y., Rodenkirch, C., Wang, Q., McCormick, D. A., & Tolias, A. S. (2016). Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. *Nature Communications*, *7*. doi:10.1038/ncomms13289

- Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, 11(6), 251-257. doi:10.1016/j.tics.2007.04.004
- Samuels, E. R., & Szabadi, E. (2008). Functional neuroanatomy of the noradrenergic locus coeruleus: Its roles in the regulation of arousal and autonomic function part I: Principles of functional organisation. *Current Neuropharmacology*, 6(3), 235-253. doi:10.2174/157015908785777229
- Sara, S. J. (2009). The locus coeruleus and noradrenergic modulation of cognition. *Nature Reviews Neuroscience*, 10(3), 211-223. doi:10.1038/nrn2573
- Schmitz, T. W., & Duncan, J. (2018). Normalization and the Cholinergic Microcircuit: A Unified Basis for Attention. *Trends in Cognitive Sciences*, 22(5), 422-437. doi:10.1016/j.tics.2018.02.011
- Shine, J. M. (2019). Neuromodulatory Influences on Integration and Segregation in the Brain. *Trends in Cognitive Sciences*, 23(7), 572-583. doi:10.1016/j.tics.2019.04.002
- Suzuki, M., & Larkum, M. E. (2020). General Anesthesia Decouples Cortical Pyramidal Neurons. *Cell*, 180(4), 666-+. doi:10.1016/j.cell.2020.01.024
- Turnbull, P. R. K., Irani, N., Lim, N., & Phillips, J. R. (2017). Origins of Pupillary Hippus in the Autonomic Nervous System. *Investigative Ophthalmology & Visual Science*, 58(1), 197-203. doi:10.1167/iovs.16-20785