

OPEN

# Rapid Brain Responses to Familiar vs. Unfamiliar Music – an EEG and Pupillometry study

Robert Jagiello<sup>1,2</sup>, Ulrich Pomper<sup>1,3\*</sup>, Makoto Yoneya<sup>4</sup>, Sijia Zhao<sup>1</sup> & Maria Chait<sup>1\*</sup>

Human listeners exhibit marked sensitivity to familiar music, perhaps most readily revealed by popular “name that tune” games, in which listeners often succeed in recognizing a familiar song based on extremely brief presentation. In this work, we used electroencephalography (EEG) and pupillometry to reveal the temporal signatures of the brain processes that allow differentiation between a familiar, well liked, and unfamiliar piece of music. In contrast to previous work, which has quantified gradual changes in pupil diameter (the so-called “pupil dilation response”), here we focus on the occurrence of pupil dilation events. This approach is substantially more sensitive in the temporal domain and allowed us to tap early activity with the putative salience network. Participants (N = 10) passively listened to snippets (750 ms) of a familiar, personally relevant and, an acoustically matched, unfamiliar song, presented in random order. A group of control participants (N = 12), who were unfamiliar with all of the songs, was also tested. We reveal a rapid differentiation between snippets from familiar and unfamiliar songs: Pupil responses showed greater dilation rate to familiar music from 100–300 ms post-stimulus-onset, consistent with a faster activation of the autonomic salience network. Brain responses measured with EEG showed a later differentiation between familiar and unfamiliar music from 350 ms post onset. Remarkably, the cluster pattern identified in the EEG response is very similar to that commonly found in the classic old/new memory retrieval paradigms, suggesting that the recognition of brief, randomly presented, music snippets, draws on similar processes.

The human auditory system exhibits a marked sensitivity to familiar music<sup>1–6</sup>. The concept of music familiarity heavily relies on long term memory traces<sup>7</sup>, auditory mental imagery<sup>8–10</sup> and is also linked to autobiographical memories, especially for emotionally relevant music<sup>11</sup>. Our prowess towards recognizing familiar musical tracks is anecdotally exemplified by “Name That Tune” games, in which listeners of a radio station are asked to name the title of a song on the basis of a very short excerpt. Even more fleeting recognition scenarios may occur when switching from one station to another while deciding which one to listen to—beloved songs often showing the ability to swiftly catch our attention, causing us to settle for a certain channel.

Here we seek to quantify, in a laboratory setting, one aspect of such recognition. Specifically, we aim to understand how quickly listeners’ brains can identify snippets from a familiar and personally relevant piece of music from among unfamiliar snippets, and pinpoint the neural signatures of this recognition. Beyond basic science, understanding the brain correlates of music recognition is useful for various music-based therapeutic interventions<sup>12</sup>. For instance, there is a growing interest in exploiting music to break through to dementia patients for whom memory for music appears well preserved despite an otherwise systemic failure of memory systems<sup>13,14</sup>. Pinpointing the neural signatures of the processes which support music identification may provide a clue to understanding the basis of the above phenomena, and how it may be objectively quantified.

Previous work using behavioral gating paradigms demonstrates that the latency with which listeners can identify a familiar piece of music (pop or instrumental) amongst unfamiliar excerpts ranges from 100 ms<sup>5</sup> to 500 ms<sup>3,15–17</sup>. It is likely that such fast recognition is driven by our memory of the timbre and other spectral distinctivenesses of the familiar piece<sup>15,18–20</sup>.

According to bottom-up theories of recognition memory, an incoming stimulus is compared to stored information, and upon reaching a sufficient congruence is then classified as familiar<sup>15,21</sup>. A particular marker

<sup>1</sup>Ear Institute, University College London, London, UK. <sup>2</sup>Present address: Institute of Cognitive and Evolutionary Anthropology, University of Oxford, Oxford, UK. <sup>3</sup>Present address: Faculty of Psychology, University of Vienna, Vienna, Austria. <sup>4</sup>NTT Communication Science Laboratories, NTT Corporation, Atsugi, 243-0198, Japan. \*email: [ulrich.pomper@univie.ac.at](mailto:ulrich.pomper@univie.ac.at); [m.chait@ucl.ac.uk](mailto:m.chait@ucl.ac.uk)

in the EEG literature that is tied to such recognition processes is the late positive potential (LPP)<sup>22–24</sup>. The correct identification of a familiar stimulus typically results in a sustained positivity ranging from 500 to 800 ms post-stimulation in left central-parietal regions, which is absent for unfamiliar stimuli<sup>25</sup>. This parietal “old versus new effect” has consistently been found across various domains, such as facial<sup>26</sup> and voice recognition<sup>27</sup> as well as paradigms that employed visually presented<sup>28</sup> and spoken<sup>29</sup> words as stimuli. In an fMRI study, Klostermann *et al.*<sup>30</sup> used 2-second long excerpts of newly composed music and familiarized their participants with one half of the snippet sample while leaving the other half unknown. Subsequently, participants were exposed to randomized trials of old and new snippets and were asked to make confidence estimates regarding their familiarity. Correct identification of previously experienced music was linked to increased activity in the posterior parietal cortex (PPC). However, due to the typically low temporal resolution of fMRI, the precise time course of the recognition process remains unknown.

Pupillometry is also increasingly used as a measure of music recognition and, more generally, of the effect of music on arousal. This is part of a broader understanding that cognitive states associated with vigilance, surprise and processing effort can be gleaned from measuring task-evoked changes in pupil size<sup>31–38</sup>. Pupil dilation also reliably co-occurs with musical chills<sup>39</sup>—a physiological phenomenon evoked by exposure to emotionally relevant and familiar pieces of music<sup>40</sup> and hypothesized to reflect autonomic arousal. Underlying these effects is the increasingly well understood link between non-luminance-mediated change in pupil size and the brain’s neuro-transmitter mediated salience and arousal network (specifically Acetylcholine and Norepinephrine)<sup>41–46</sup>.

In particular, abrupt changes in pupil size are commonly observed in response to salient<sup>47,48</sup> or surprising<sup>36,49,50</sup> events, including those in the auditory modality. Work in animal models has established a link between such phasic pupil responses and spiking activity within norepinephrine (NE, alternatively noradrenaline) generating cells in the brainstem nucleus locus coeruleus (LC). The LC projects widely across the brain and spinal cord<sup>51,52</sup> and is hypothesized to play a key role in regulating arousal. Phasic pupil responses are therefore a good measure of the extent to which a stimulus is associated with increased arousal or attentional engagement<sup>32,53,54</sup>. Here we aim to understand whether and how music familiarity drives pupil responses.

Pupil dilations have received much attention in recognition paradigms, analogue to the previously elaborated designs in which participants are first exposed to a list of stimuli during a learning phase and subsequently asked to identify old items during the recognition stage<sup>55</sup>. When identifying old words, participants’ pupils tend to dilate more than when confronted with novel words, a phenomenon which is referred to as the *pupil old/new effect*<sup>56–58</sup>. Otero *et al.*<sup>59</sup> replicated this finding using spoken words, thus extending this effect onto the auditory domain. The specific timing of these effects is not routinely reported. The bulk of previous work used analyses limited to measuring peak pupil diameter (e.g.<sup>56,60–62</sup>) or average pupil diameter change over the trial interval<sup>57,59</sup>. Weiss *et al.*<sup>63</sup> played a mix of familiar and unfamiliar folk melodies to participants and demonstrated greater pupil dilations in response to the known as opposed to the novel stimuli. In this study, the effect emerged late, around 6 seconds after stimulus onset. However, this latency may be driven by the characteristics of their stimuli (excerpts ranged in length from 12 to 20 seconds), as well as the fact that melody recognition may take longer than timbre-based recognition.

Here we combine EEG and pupillometry to investigate the temporal dynamics of the physiological processes that underlie the differentiation of a familiar and personally relevant from an unfamiliar piece of music. In contrast to previous work, which has quantified changes in pupil diameter (the so-called “pupil dilation response”), we focus on pupil dilation events (see Methods). This approach is substantially more sensitive in the temporal domain and allowed us to tap early activity with the putative salience network.

Our experimental paradigm consisted of exposing passively listening participants to randomly presented short snippets from a familiar and unfamiliar song—a design that constitutes a simplified version of the above-mentioned real-world scenarios such as radio channel switching, though we note that we have reduced the memory demands considerably by using only one familiar and one unfamiliar song per subject. A control group, unfamiliar with all songs, was also used. We sought to pinpoint the latency at which brain and pupil dynamics dissociate randomly presented familiar from unfamiliar snippets and understand the relationship between brain and pupil measures of this process.

## Methods

**Participants.** The participant pool encompassed two independent groups: A main group ( $N_{\text{main}} = 10$ ; 5 females;  $M_{\text{age}} = 23.56$ ;  $SD = 3.71$ ) and a control group ( $N_{\text{control}} = 12$ ; 9 females;  $M_{\text{age}} = 23.08$ ;  $SD = 4.79$ ). All reported no known history of hearing or neurological impairment. Two participants from the main group were excluded from the EEG analysis, as their respective pairs of familiar and unfamiliar songs elicited significant differences in EEG responses in the control group (see Results for details). This left 8 analyzed data sets (4 females;  $M_{\text{age}} = 23.62$ ;  $SD = 3.96$ ).

Experimental procedures were approved by the research ethics committee of University College London and were performed in accordance with the relevant guidelines and regulations. Written informed consent was obtained from each participant. Participants were paid for their participation.

**Stimuli and procedure.** *Preparatory stage.* Members of the main group filled out a music questionnaire, requiring them to list five songs that they have frequently listened to, bear personal meaning and are evocative of positive affect. One song per subject was then selected and matched with a control song, which was unknown to the participant, yet highly similar in terms of various musical aspects, such as tempo, melody, harmony, vocals and instrumentation. Since we are not aware of an algorithm that matches songs systematically, this process largely relied on the authors’ personal judgments, as well as the use of websites that generate song suggestions (e.g. <http://www.spotalike.com>, <http://www.youtube.com>). We provided each participant with the name as well as a 1500 ms snippet of the matched song to confirm that they are indeed unfamiliar with it. Upon completion,

	Familiar Song	Control Song
1.	M83 - Midnight City <a href="https://www.youtube.com/watch?v=dX3k_QDnzHE">https://www.youtube.com/watch?v=dX3k_QDnzHE</a>	Postiljonen - Atlantis <a href="https://www.youtube.com/watch?v=MnkzAPUHRog">https://www.youtube.com/watch?v=MnkzAPUHRog</a>
2.	Chuck Berry - You never can tell <a href="https://www.youtube.com/watch?v=uuM2FTq5f1o">https://www.youtube.com/watch?v=uuM2FTq5f1o</a>	Jerry Lee Lewis - Great Balls of Fire <a href="https://www.youtube.com/watch?v=Jt0mg8Z09SY">https://www.youtube.com/watch?v=Jt0mg8Z09SY</a>
3.	Barbra Streisand - The way we were <a href="https://www.youtube.com/watch?v=GNecQS4tXgQ">https://www.youtube.com/watch?v=GNecQS4tXgQ</a>	Carole King - So far away <a href="https://www.youtube.com/watch?v=UofY13dataU">https://www.youtube.com/watch?v=UofY13dataU</a>
4.	Kiss - Detroit Rock City <a href="https://www.youtube.com/watch?v=iZq3i94mSsQ">https://www.youtube.com/watch?v=iZq3i94mSsQ</a>	Black Sabbath - After Forever <a href="https://www.youtube.com/watch?v=WEsfkCX84s">https://www.youtube.com/watch?v=WEsfkCX84s</a>
5.	The Lumineers - Cleopatra <a href="https://www.youtube.com/watch?v=aN5s9N_pTUs">https://www.youtube.com/watch?v=aN5s9N_pTUs</a>	The Strumbellas - Shovels & Dirt <a href="https://www.youtube.com/watch?v=EJYtrvMxYVw">https://www.youtube.com/watch?v=EJYtrvMxYVw</a>
6.	Muse - Dead Inside <a href="https://www.youtube.com/watch?v=I5sJhSNUkwQ">https://www.youtube.com/watch?v=I5sJhSNUkwQ</a>	The Vaccines - Dream Lover <a href="https://www.youtube.com/watch?v=X60NXmTbunk">https://www.youtube.com/watch?v=X60NXmTbunk</a>
7.	Nightwish - Élan <a href="https://www.youtube.com/watch?v=8cGLKgT8S8">https://www.youtube.com/watch?v=8cGLKgT8S8</a>	After Forever - Energize Me <a href="https://www.youtube.com/watch?v=ml8rNN2WAec">https://www.youtube.com/watch?v=ml8rNN2WAec</a>
8.	Paolo Nutini - Candy <a href="https://www.youtube.com/watch?v=x3xYXGMRRYk">https://www.youtube.com/watch?v=x3xYXGMRRYk</a>	Josh Ritter - Good Man <a href="https://www.youtube.com/watch?v=C81SyunWMAQ">https://www.youtube.com/watch?v=C81SyunWMAQ</a>
9.	Iron Maiden - The Number Of The Beast <a href="https://www.youtube.com/watch?v=WxnN05vOuSM">https://www.youtube.com/watch?v=WxnN05vOuSM</a>	Riot - Riot <a href="https://www.youtube.com/watch?v=lvbFXo-M8ec">https://www.youtube.com/watch?v=lvbFXo-M8ec</a>
10	CHVRCHES - Recover <a href="https://www.youtube.com/watch?v=Jyqemlbcfg">https://www.youtube.com/watch?v=Jyqemlbcfg</a>	Grimes - Oblivion <a href="https://www.youtube.com/watch?v=m5H-YlcMSbc">https://www.youtube.com/watch?v=m5H-YlcMSbc</a>

**Table 1.** List of song dyads (“Familiar” and “Unfamiliar”) used in this study. Song were matched for style and timbre quality as described in the methods section. The songs were selected based on input from the “main” group. The “control” group were unfamiliar with all 20 songs.

this procedure resulted in ten dyads (one per participant), each containing one familiar and one unfamiliar song (see Table 1 for song details). All of the selected songs contained vocals, and pairs were matched according to the gender of the lead singer. Importantly, matching was also verified with the control group (see below).

The main role of the control group is to rule out any acoustic differences between song-pairs that might contribute to differences in brain responses. Participants were selected for the control group based on non-familiarity with any of the ten dyads, therefore the distinction between familiar and unfamiliar conditions does not apply to them. To check for their eligibility before entering the study, they had the opportunity to inspect the song list as well as to listen to corresponding excerpts from the chorus (1500 ms). Due to the fact that it was exceedingly difficult to identify participants unfamiliar with all 20 songs, those recruited to the control group comprised of international students enrolled at UCL who were largely inexperienced with western popular music. While English was not their native language, all were proficient in understanding spoken English (as per UCL’s admission requirements). Note that 6 of the participants of the main group were also non-native English speakers.

*Stimulus generation and testing paradigms.* The beginning and end of each song, which typically constitute silent or only gradually rising or fading parts of the instrumentation, were removed. Both songs from each pair were then divided into snippets of 750 ms. Out of these, 100 snippets were randomly chosen for each song. These song snippets were then used in two paradigms: (1) a passive listening task and (2) an active categorization task. In the passive listening task, participants listened to the snippets from each dyad in random order whilst their brain activity was recorded with EEG and their pupil diameters with an infrared eye-tracking camera. Each particular snippet was presented once only. Participants were instructed to attentively listen to the music. We chose a passive stimulation paradigm because it mimics an everyday listening situation, in which participants can focus on the music without distraction. Each block contained 200 trials, 100 of each song presented in random order with an inter-stimulus interval (ISI) randomized between 1 and 1.5 seconds. This resulted in a total duration of roughly 6.7 minutes per block. Participants from the main group were presented with only one block (pertaining to the dyad that contained their familiar song and the matched non-familiar song). Participants from the control group listened to all 10 dyads (each in a separate block) for a total of 10 blocks which were presented in random order.

The active categorization task was also divided into 1 block per dyad. In each block, participants were presented with 20 trials, each containing a random pairing of snippets, separated by 750 ms. In half of the trials, both snippets were drawn from the same song. In the other half, one snippet was taken from the familiar and the other from the unfamiliar song. Participants were instructed to indicate whether the two snippets were from the same song or from different songs, by pressing the corresponding buttons on the keyboard. There was no time limit imposed on the response, and trials were separated by 750 ms following a button press. Same as for the passive listening task, participants from the main group performed only one block, associated with their pairing of “familiar/unfamiliar” songs. Participants from the control group completed 10 blocks in random order.

*Procedure.* Participants were seated, with their heads fixed on a chinrest, in a dimly lit and acoustically shielded testing room (IAC Acoustics, Hampshire, UK). They were distanced 61 cm away from the monitor and 54 cm away from two loudspeakers, arranged at an angle of 30° to the left and right of the subject. Participants were instructed to continuously fixate on a white cross on a grey background, presented at the center of a 24-inch monitor (BENQ XL2420T) with a resolution of 1366 × 768 pixels and a refresh rate of 60 Hz. They first engaged in the passive listening task followed by the active categorization task. For both tasks, participants were instructed to

attentively listen to the snippets. EEG and pupil diameter were recorded during the first task, but not during the second, which was of purely behavioral nature.

**EEG acquisition, preprocessing and analysis.** EEG recordings were obtained using a Biosemi Active Two head-cap 10/20 system with 128 scalp channels. Eye movements and blinks were monitored using 2 additional electrodes, placed on the outer canthus and infraorbital margin of the right eye. The data were recorded reference-free with a passband of 0.016–250 Hz and a sampling rate of 2048 Hz. After acquisition, pre-processing was done in Matlab (The MathWorks Inc., Natick, MA, USA), with EEGLAB<sup>64</sup> (<http://www.sccn.ucsd.edu/eeglab/>) and FieldTrip software<sup>65</sup> (<http://www.ru.nl/fcdonders/fieldtrip/>). The data were downsampled to 128 Hz, low pass filtered at 40 Hz and re-referenced to the average across all electrodes. The data were not high pass filtered, to preserve low-frequency activity<sup>25</sup>, which is relevant when analyzing sustained responses. The data were segmented into stimulus time-locked epochs ranging from –500 ms to 1500 ms. Epochs containing artefacts were removed on the basis of summary statistics (variance, range, maximum absolute value, z-score, maximum z-score, kurtosis) using the visual artefact rejection tool implemented in Fieldtrip. On average, 2.1 epochs per song pair in the main group and 4.9 epochs in the control group were removed. The larger number for the control group is probably a consequence of the longer session duration (10 blocks instead of a single block for the main group) and associated fatigue. Artefacts related to eye movements, blinks and heartbeat were identified and removed using independent component analysis. Subsequently, the data were averaged over epochs of the same condition and baseline-corrected (200 ms preceding stimulus onset). In the control group, averaging was done separately for each dyad, resulting in 20 time series (2 (familiar/unfamiliar) × 10 dyads).

A cluster-based permutation analysis, which takes spatial and temporal adjacency into account<sup>65,66</sup> was used to investigate potential effects in the EEG responses. For both main and control participants, pairwise t-tests were calculated between ‘familiar’ and ‘unfamiliar’ snippets in all electrodes and over the entire epoch length. The significance threshold was chosen to control family-wise error-rate (FWER) at  $p = 0.05$ .

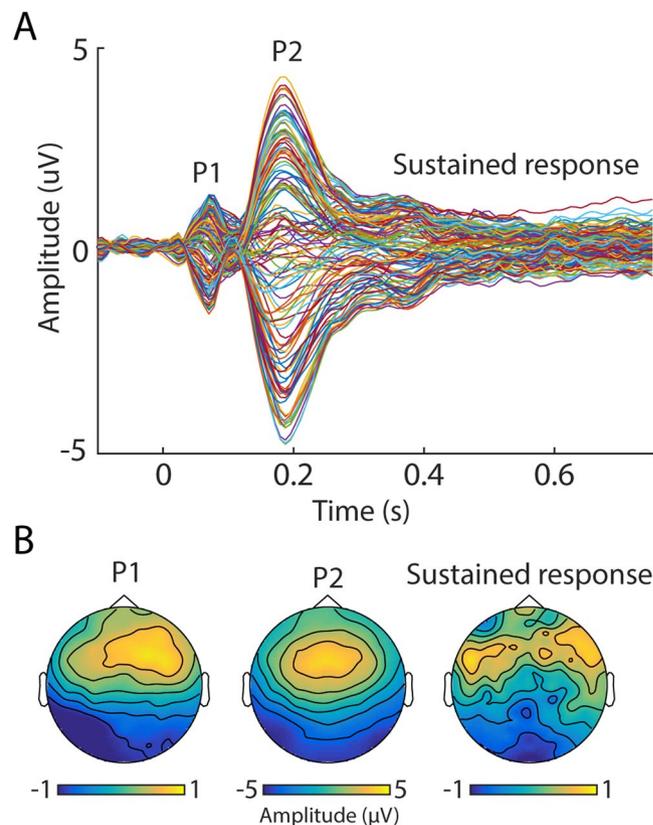
**Pupil measurement and analysis.** Gaze position and pupil diameter were continuously recorded by an infrared eye-tracking camera (Eyelink 1000 Desktop Mount, SR Research Ltd.), positioned just below the monitor and focusing binocularly at a sampling rate of 1000 Hz. The standard five-point calibration procedure for the Eyelink system was conducted prior to each experimental block. Due to a technical fault that caused missing data, the initial seven control participants were excluded from the pupillometry analysis, leaving five valid participants (5 females,  $Mage = 23.21$ ,  $SD = 4.37$ ) in the control group. Note that this still resulted in ample data since each control subject completed 10 blocks, one for each dyad. No participant was excluded from the main group.

To make sure that pupil data are of high quality and reflect a consistent gaze position, samples in which gaze position exceeded 1° away from fixation were removed from the analysis. Across participants, 6% of the data were rejected in this way.

The standard approach for analyzing pupillary responses involves across trial averaging of pupil diameter as a function of time. This is usually associated with relatively slow dynamics<sup>47,67–69</sup> which are not optimal for capturing potentially rapid effects within a fast-paced stimulus. Instead, the present analysis focused on examining pupil event rate. This analysis captures the incidence of pupil dilation events<sup>42</sup>, irrespective of their amplitude, and therefore provides a sensitive measure of subtle changes in pupil dynamics that may be evoked by the familiar vs. non-familiar stimuli. Pupil dilation events were extracted from the continuous data by identifying the instantaneous positive sign-change of the pupil diameter derivative (i.e. the time points where pupil diameter begins to positively increase). To compute the incidence rate of pupil dilation events, the extracted events were convolved with an impulse function (see also<sup>42,70</sup>), paralleling a similar technique for computing neural firing rates from neuronal spike trains<sup>71</sup>. For each condition, in each participant and trial, the event time series were summed and normalized by the number of trials and the sampling rate. Then, a causal smoothing kernel  $\omega(\tau) = \alpha^2 \times \tau \times e^{(-\alpha\tau)}$  was applied with a decay parameter of  $\alpha = 1/50 \text{ ms}$ <sup>70–72</sup>. The resulting time series was then baseline corrected over the pre-onset interval. For each condition, the pupil dilation rate averaged across participants is reported here.

To identify time intervals in which the pupil dilation rate was significantly different between the two conditions, a nonparametric bootstrap-based statistical analysis was used<sup>73</sup>. For the main group, the difference time series between the conditions was computed for each participant, and these time series were subjected to bootstrap re-sampling (with replacement; 1000 iterations). At each time point, differences were deemed significant if the proportion of bootstrap iterations that fell above or below zero was more than 99% (i.e.  $p < 0.01$ ). This analysis was conducted over the entire epoch duration. All significant intervals are reported.

Two control analyses were also conducted to verify the effects found in the main group. Firstly, permutation analysis on the data from the main group: in each iteration (1000 overall), 10 participants were selected with replacement. For each participant, all trials across conditions were randomly mixed and artificially assigned to the “familiar” or the “unfamiliar” condition (note that these labels are meaningless in this instance). This analysis yielded no significant difference between conditions. A second control analysis examined pupil dynamics in the control group. Data for each control participant consisted of 10 blocks (one per dyad), and these were considered as independent data sets for this analysis, resulting in 50 control datasets. On each iteration (1000 overall), 10 control datasets were selected with replacement from the pool of 50 and used to compute the mean difference between the two conditions. From here, the analysis was identical to the one described for the main group. This analysis also yielded no significant difference between conditions.



**Figure 1.** Grand-average event-related potential demonstrating the overall brain response to the music snippets. **(A)** Time-domain representation. Each line represents one EEG channel. The plotted data are averaged across the main and control groups, as well as familiar and unfamiliar songs. **(B)** Topographical representation of the P1 (71 ms), P2 (187 ms) peaks, as well as the sustained response (plotted is the average topography between 300–750 ms).

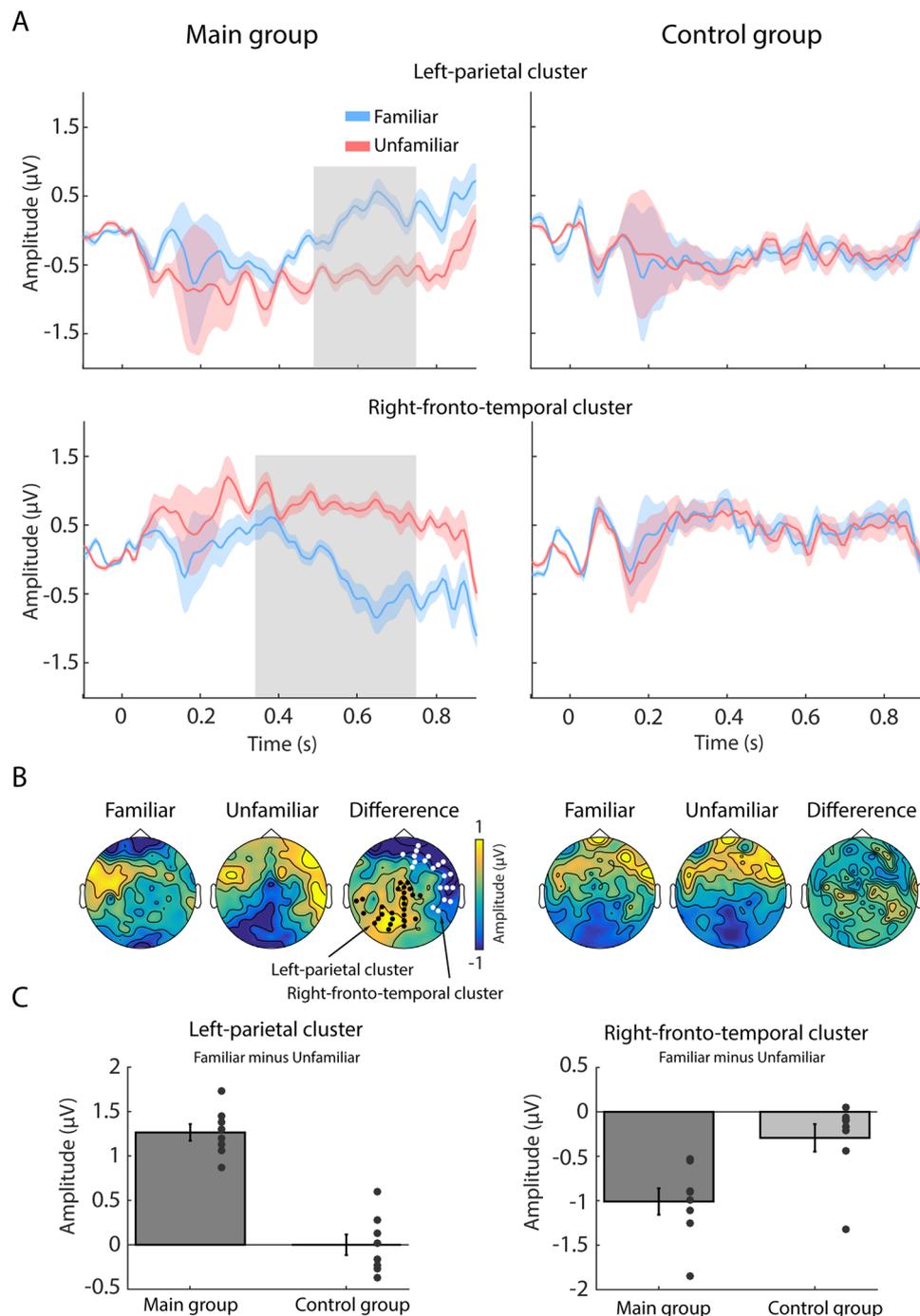
## Results

**EEG.** The overall EEG response to the sound snippets (collapsed across all participants and conditions) is shown in Fig. 1. The snippets evoked a characteristic onset response, followed by a sustained response. The onset response was dominated by P1 (at 71 ms) and P2 (at 187 ms) peaks, as is commonly observed for wide-band signals (e.g.<sup>74</sup>).

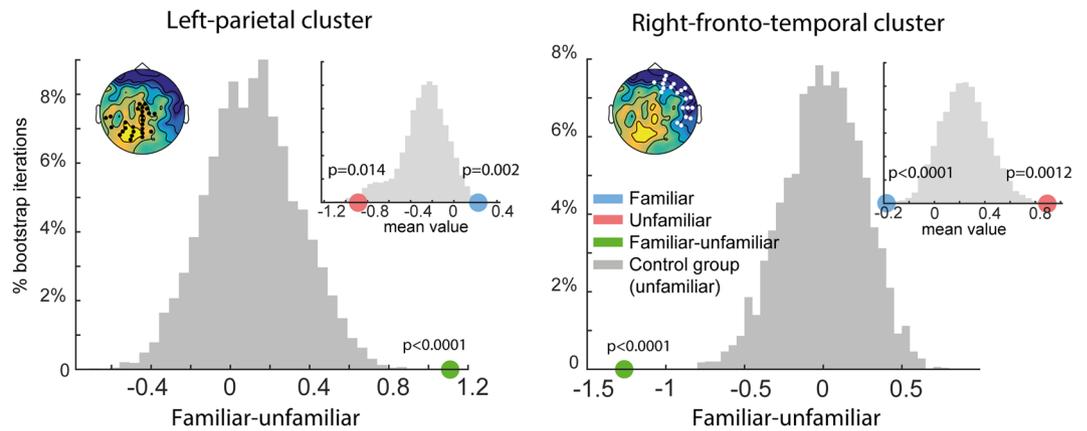
The main purpose of the control group was to verify that any significant differences which are potentially established for main participants were due to the manipulation of familiarity and not caused by any acoustic differences between the songs in each dyad. Because the control group participants were unfamiliar with the songs, we expected no differences in their brain responses to the songs in each dyad. However, the cluster-based permutation test revealed significant differences between conditions in dyad #2 (“familiar” more negative than “unfamiliar”) and #5 (“unfamiliar” more negative than “familiar”). This was taken as evidence that the songs were not matched properly, and the respective dyads (#2 and #5) were excluded from the subsequent main group analysis. Note that whilst this approach specifically targeted basic acoustic differences, it is possible, and indeed likely, that other, more abstract, differences between songs remain (e.g. differences in tonality or timbre). However, our use of different musical material for each dyad, spanning a wide range of genres (see Table 1), assures that any differences are not systematic across song material. This allows us to interpret the strongly consistent effects observed in the main group (see below) in terms of the symbolic difference between the “unfamiliar” and “familiar” songs, not linked to any physical stimulus parameters.

Comparing responses to “familiar” and “unfamiliar” snippets within the main group (Fig. 2), we identified two clusters of channels showing a significant difference between the two conditions (see Methods). A left-parietal cluster of 26 channels showing a significant difference between conditions from 540 to 750 ms ( $T_{\text{sum}} = -1238.71$ ), and a right frontotemporal cluster of 20 channels, showing a significant difference between 350 to 750 ms ( $T_{\text{sum}} = 1127.41$ ). These clusters are similar to those typically identified in old/new recognition memory studies<sup>22,75–78</sup>. Similar to what is observed here, the right frontotemporal cluster commonly exhibits earlier responses and is hypothesized to reflect familiarity, whereas the later responses in the left-parietal cortex are hypothesized to reflect retrieval<sup>22</sup>, consistent with dual-process theories of memory<sup>79</sup>.

To confirm that the observed differences are specific to the main group, we additionally performed a 2-factorial mixed ANOVA, with a within-subject factor of familiarity (familiar/unfamiliar) and a between-subjects factor of group (main/control). The average EEG amplitude across all channels and time points within the cluster served as the dependent variable. The interaction between the factors familiarity and group was significant for both the



**Figure 2.** Event-related potential results – differences between “familiar” and “unfamiliar” snippets in the main, but not control, group. **(A)** Time-domain ERPs for the left-parietal cluster (top row) and right frontotemporal cluster (bottom row), separately for the main (left column) and control (right column) group. Solid lines represent mean data (averaged across channels and dyads) for familiar (blue) and unfamiliar (red) songs (note that this labelling only applies to the ‘main’ group; both songs were unfamiliar to the control listeners). Significant differences between conditions, as obtained via cluster-based permutation tests, are indicated by grey boxes. Note that the shaded areas reflect standard error of the mean for each condition, whilst the reported statistics are conducted as a repeated measures analysis. **(B)** Topographical maps of the “familiar” and “unfamiliar” ERP responses (computed from 350 to 750 ms) as well as their difference, separately for the main (left column) and control (right column) group. Black and white dots indicate electrodes belonging to the left-parietal and right frontotemporal cluster, respectively. **(C)** Mean ERP amplitude differences. The main group, but not the control group, showed significantly larger responses to unfamiliar song snippets, at both the left-parietal and the right frontotemporal clusters. Error bars represent standard error of the mean, dots represent mean response differences (across participants) to each song (8 “familiar” and 8 matched “unfamiliar” songs). In the main group, each dot reflects data from a single subject. In the control group, each dot reflects the average across the 12 members of the control group.



**Figure 3.** Bootstrap analysis comparing mean differences between “familiar” and “unfamiliar” responses from the main group to baseline difference distributions sampled from the control group (grey). As expected, difference distributions in the control group are centered around 0. Left upper insets show the respective electrode cluster overlaid on the ‘familiar – unfamiliar’ difference of the main group. In both clusters, responses observed in the main group (green dots) are well outside the baseline distributions ( $p < 0.001$ ). Right upper inset figures show mean responses to “familiar” (blue dot) and “unfamiliar” (red dot) conditions from the main group, overlaid on distribution of responses from the control group (grey bars; collapsed across “familiar” and “unfamiliar”). Main group responses to both “familiar” and “unfamiliar” conditions are found on opposite sides of the control distributions, revealing a repulsive effect.

left-parietal cluster ( $F(1, 14) = 73.56$ ,  $p < 0.001$ ; partial  $\eta^2 = 0.84$ ) as well as the right frontotemporal cluster ( $F(1, 14) = 37.91$ ,  $p < 0.001$ ; partial  $\eta^2 = 0.73$ ). In both cases the interaction was driven by a significant difference between ‘familiar’ and ‘unfamiliar’ conditions in the main group, but a non significant difference among the control participants (left-parietal cluster:  $t_{\text{main}}(7) = -13.7$ ;  $p < 0.001$ ;  $t_{\text{control}}(7) = -0.03$   $p = 0.998$ ; frontotemporal cluster:  $t_{\text{main}}(7) = 8.14$ ;  $p < 0.001$ ;  $t_{\text{control}}(7) = 1.3$   $p = 0.21$ ). Note that the analysis presented so far was focused on dyads. Hence the control group data reflect data for each of 8 retained dyads, where responses to each song are averaged across the 12 members of the control group.

We also conducted a bootstrap resampling-based analysis to compare responses between the main group and matched subsets of the control group (Fig. 3). On each iteration (1000 overall) a single control subject was randomly assigned to each dyad (thus a subset of 8 control participants contributed 1 dyad each to the analysis). Data were then divided into “familiar” and “unfamiliar” conditions and averaged across participants (in the same way as for the main group; though the distinction between the two conditions is irrelevant for this group). Finally, the mean differences between “familiar” and “unfamiliar” conditions were computed over each of the ROI intervals. The grey histograms in Fig. 3 show the distribution of these values for each of the two ROIs ( $H_0$  distribution). The mean difference between conditions from the main group, indicated by the green dot, lies well beyond this distribution ( $p < 0.001$ ), further confirming that the effect observed for the main group was different from that in the control group.

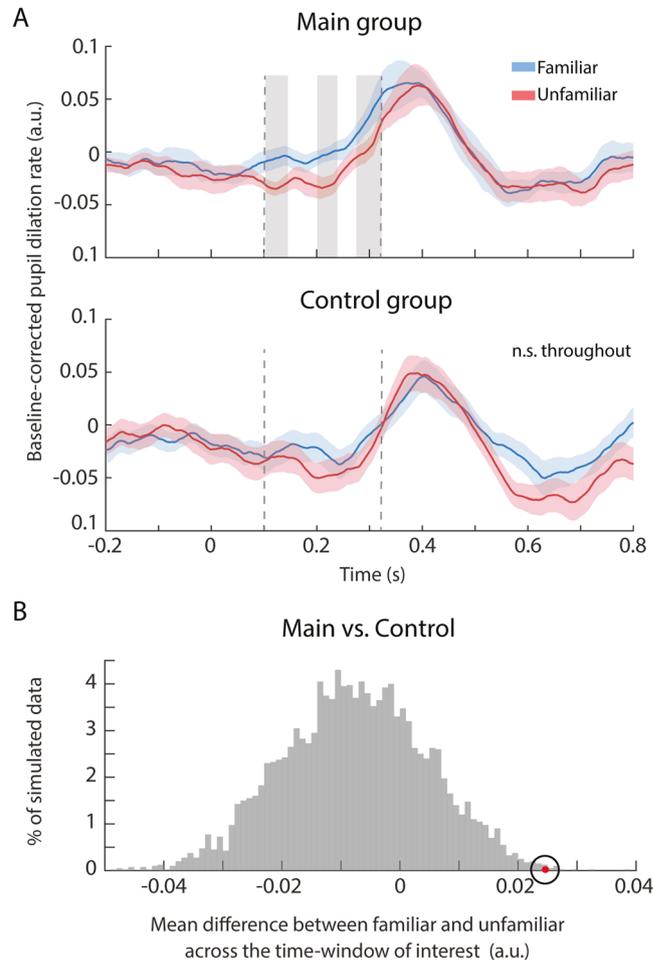
To understand how responses to “familiar” and “unfamiliar” music relate to those in the control group, we used a similar bootstrap routine, to obtain a distribution of mean control group responses in each ROI (collapsed across “familiar” and “unfamiliar” conditions) and compared to those in the main group. This analysis (Fig. 3; insets) demonstrates that responses to both “familiar” and “unfamiliar” snippets in the main group differed from those in the control group, such that, in both ROIs, “familiar” and “unfamiliar” responses lay on opposite edges of the control group distribution.

In terms of polarity and fieldmap distribution, main group responses to “unfamiliar” snippets were similar to the responses of the control group, though overall larger. It can be seen from the distributions plotted in Fig. 3 (insets) that the polarity of the “unfamiliar” response in the main group is consistent with the mean of the control group distribution. In contrast, the responses to familiar snippets in the main group lie on the opposite polarity and exhibit less deflection from 0. We will return to this point in the discussion.

**Pupil dilation.** Figure 4A (bottom) shows the pupil dilation rates for “familiar” and “unfamiliar” snippets in the control group. In response to the auditory stimuli, the pupil dilation rate increased shortly after the onset of a snippet, peaking at around 400 ms, before returning to baseline around 550 ms post-onset. No significant differences were observed between the two conditions throughout the entire epoch (see Methods), consistent with the fact that both were equally unfamiliar and hence equally salient to these participants.

In contrast, significant differences were observed in the main group. When compared with unfamiliar conditions, familiar snippets were associated with a higher pupil dilation rate between 108–319 ms post sound onset (Fig. 4A, top), i.e. during the rising slope of pupil dilation rate. This significant interval was absent in the shuffled data (see methods).

We also directly compared the difference between “familiar” and “unfamiliar” conditions between the two groups during the time interval (108–319 ms) identified as significant in the main group analysis. This was



**Figure 4.** Pupil dilation rate to familiar and unfamiliar snippets. **(A)** Top: Main group. The solid curves plot the average pupil dilation rate across participants for “familiar” (blue) and “unfamiliar” (red) conditions. The shaded area represents one standard deviation from the bootstrap mean. The grey boxes indicate time intervals where the two conditions were significantly different (108–135, 206–232, and 282–319 ms). The dashed lines indicate the time interval used for the resampling statistic in **(B)**. Bottom: Control group. The solid curves plot the average pupil dilation rate over 10 randomly selected control datasets. The shaded area represents one standard deviation of the bootstrap mean. No significant differences were observed throughout the entire epoch. **(B)** Results of the resampling analysis to compare the difference between “familiar” and “unfamiliar” conditions between the main and control groups (averages are computed over the time interval 108–319 ms, indicated via dashed lines in **(A)**). The grey histogram shows the distribution of differences between conditions for the control group ( $H_0$  distribution). The red dot indicates the observed difference in the main group.

achieved by computing a distribution of differences between conditions based on the control group data ( $H_0$  distribution). On each iteration (1000 overall) 10 datasets were randomly drawn from the control pool and used to compute the difference between conditions during the above interval. The grey histogram in Fig. 4B shows the distribution of these values. The mean difference from the main group, indicated via the red dot, lies well beyond this distribution ( $p = 0.0034$ ), confirming that the effect observed for the main group was different from that in the control group.

**Active categorization task.** This task, conducted after the EEG and pupillometry session, aimed to verify whether participants were able to differentiate between familiar and unfamiliar snippets and whether participants in the main group (who were highly familiar with one song in a pair) performed better than controls.

Main participants correctly identified whether or not the two presented snippets were from the same song in 92% of trials, whereas controls did so in 79% of trials. An independent samples  $t$ -test revealed that main participants scored significantly higher than controls  $t(18) = 6.19$ ,  $p < 0.00001$ . One-sample  $t$ -tests revealed that, in both groups, scores are at above-chance levels,  $t(9) = 13.61$ ,  $p < 0.00001$  for controls, and  $t(9) = 18.11$ ,  $p < 0.00001$  for the main group. Therefore, whilst there may have been enough information for control participants to consciously identify differences between snippets, this apparently did not affect the presently observed brain/pupil responses during passive listening.

## Discussion

We used EEG and pupillometry to identify brain responses which distinguish between a familiar, emotionally relevant and an unfamiliar piece of music. To tap rapid recognition processes, matched familiar and unfamiliar songs were divided into brief (750 ms) snippets which were presented in a mixed, random, order to passively listening participants. We demonstrate that despite the random presentation order, pupil and brain responses swiftly distinguished between snippets taken from familiar vs. unfamiliar songs, suggesting rapid underlying recognition. Specifically, we report two main observations: (1) pupil responses showed greater dilation rate to snippets taken from a familiar piece of music between ~100–300 ms post onset, and (2) brain activity measured with EEG showed differentiation between responses to familiar and unfamiliar music snippets from 350 ms post onset. The pattern of activation observed closely mirrored the ubiquitous old/new response patterns and thus suggest that similar mechanisms of recall and retrieval were recruited.

The implications of these results for our understanding of the neural correlates of music recognition (see also<sup>80</sup>) are discussed below. But to start with, we outline several important limitations which the reader must keep in mind: Firstly, “familiarity” is a multifaceted concept. In the present study, songs were explicitly selected to evoke positive feelings and memories. Therefore, for the main group, the familiar and unfamiliar songs did not just differ in terms of recognizability but also in terms of emotional engagement and affect. Whilst we continue to refer to the songs as familiar and unfamiliar, the effects we observed may also be linked to these parameters. Future studies could further investigate potential interactions between these factors, by independently varying the familiarity and emotional valence of the presented music. Furthermore, the present experiment used only one familiar and one control song. This (relative to a case where multiple songs from each category are used) significantly reduced the demands on memory processes and might have allowed the brain to achieve discrimination based on maintenance of a template of the “familiar” song in some form of working memory.

Relatedly, it was inevitable that participants in the main group were aware of the aim of the study, and might have listened with an intent that is different from that in the control group. This limitation is difficult to overcome given the present research question, and the results must be interpreted in this light.

Furthermore, consideration must be given to the control group. These participants were required to be unfamiliar with all of the 20 songs used in the present study. Finding participants to satisfy this constraint is exceedingly difficult. We therefore resorted to recruiting international students (predominantly from Asia). As a consequence, their native language and musical experience differed from that of the main group (which comprised of participants from a European background). Note that, while all of the presented songs contained vocals, it is unlikely that semantic processing of the lyrics contributed to the observed effects. Since snippets were cut at random time-points within each song, many snippets did not contain a vocal passage at all. Those that did contain vocals were extracted at random points within words or syllables, eliminating any possible reliance on semantic processing.

Lastly, though we took great care in the song matching process, ultimately this was done by hand due to lack of availability of appropriate technology. Advancements in automatic processing of music may improve matching in the future. The control group was used to make sure that the familiar and unfamiliar songs were sufficiently acoustically matched. Because those participants were unfamiliar with either song in a dyad, we expected no differences in brain activity. Such was the case for 8 out of the 10 song pairs, i.e. two participants from the main group were discarded from further analysis. This highlights a further limitation of the present study which is associated with limited group size. Despite the various limitations, we note that the effects we observed (Figs 2 and 3) are large – suggesting the presence of a substantial, and robust, effect in the population.

Analysis of pupil dilation rates, demonstrated a characteristic rapid increase in rate evoked by snippet onset, peaking at roughly 400 ms. In the main group, we observed differences between responses to familiar and unfamiliar snippets during the rising slope of this response, such that the snippets taken from the familiar song evoked a larger increase in dilation rate, possibly reflective of recognition-linked surge in arousal.

The timing of this effect – between ~100–300 ms after onset - is broadly consistent with previous behaviorally derived estimates, which place minimum identification time for music at 100–250 ms<sup>5,15</sup>. Though they were listening passively, it is possible that the main group maintained a template of their familiar song in working memory and compared incoming snippets to this representation. For differentiation to happen so quickly, the relevant features were likely related to the timbre of the familiar song. It is known that humans possess a remarkable sensitivity to, and long term memory of, the timbral properties of complex sounds<sup>19,20</sup> and these features are extracted early enough in the auditory processing hierarchy to be detectable within a short time of sound onset<sup>19,20,81,82</sup>.

Research in animal models has linked phasic pupil dilation events with increased firing in the LC<sup>42</sup>, hypothesized to reflect heightened arousal. Our approach of analyzing pupil dilation rates is particularly sensitive to capturing this activity. The present results can therefore be taken to indicate that the LC was differentially activated as early as ~100–300 ms after sound onset, possibly through projections from the inferior colliculus (where timbre cues may be processed<sup>81,82</sup>) to subcortical structures such as the hippocampus or amygdala, which are known to be linked to the LC<sup>83,84</sup>. Accumulating evidence<sup>42,44</sup> demonstrates very rapid connectivity within this network, with IC-spike triggered pupil dilation events peaking at a latency of ~200 ms.

Our paradigm is conceptually similar to the old/new paradigms commonly used to investigate recognition memory<sup>7,22,25</sup>, but with important differences: Typical recognition memory studies test brain responses to recently memorized “neutral” stimuli which are usually static (e.g. words, faces, pictures). Here we tested brain responses associated with positive affect-based representations of temporally dynamic stimuli. These representations likely consist of current auditory templates of the familiar song, fed by emotionally charged long-term memory traces. It is possible that recognition would affect subsequent responses to the snippets as they unfold. Therefore, the observed EEG responses likely reflect both the process of recognition and the effect of familiarity on the processing of the ensuing portion of the sound. Additionally, unlike the standard old/new experiments, participants here listened passively and were not required to make a response, though, as mentioned above, it is possible that

the main group were covertly making decisions about familiarity. For these reasons, comparison with previous results from the standard recognition memory paradigm may not be straightforward. It is noteworthy, however, that the cluster pattern identified in the EEG response is very similar to that commonly found in the classic old/new paradigms.

Specifically, in line with the ubiquitous findings in the memory literature, we blindly (i.e. using an unbiased whole scalp analysis) identified two clusters which distinguished familiar and unfamiliar responses: A right frontotemporal cluster, emerging from 350 ms after onset, and a left-parietal cluster emerging about 200 ms later - from 550 ms post onset. These responses are widely discussed in the memory literature as evidence for a two-stage memory process<sup>85,86</sup> - an initial processing stage associated with familiarity, and later activation associated with recollection. In agreement with that literature, we observed more negative responses to “familiar” relative to “unfamiliar” items in the frontotemporal cluster, and the opposite (more positive for “familiar”) in the left-parietal cluster. This suggests that even short, randomly mixed, sound snippets can draw upon memory retrieval processes similar to those observed in active old/new judgment paradigms.

Important insight was obtained from comparing “familiar” and “unfamiliar” responses to responses from the control group. In that group, all stimuli were unfamiliar, and hence unlikely to have evoked old/new processing. Instead, responses presumably reflect “baseline” processing of unfamiliar snippets. We observed that main group activation to both familiar and unfamiliar snippets lay at opposite edges of the distribution of responses in the control participants. This “repulsive” effect suggests that both “familiar” and “unfamiliar” representations changed relative to the baseline afforded by the control group. We return to this point further below.

Overall, “unfamiliar” responses were more similar to the responses in the control group in that they shared polarity and field distribution, consistent with the fact that in both cases snippets were unfamiliar. In contrast, the “familiar” response was of opposite polarity to that exhibited by the control group, hinting at a qualitatively different process. Since scalp EEG is a reference-based measure, it is tricky to make direct claims about response magnitude. However, taking the distance from the 0 baseline as a measure of response energy, may suggest that more energy is expended for processing the unfamiliar compared to the familiar items. In both the frontotemporal and parietal clusters, the magnitude of activation to the familiar snippets was smaller (closer to 0) than that to the unfamiliar snippets. One possible hypothesis, consistent with this pattern of results, is that, as discussed above, familiar snippets are recognized rapidly, mediated by fast-acting sub-cortical circuitry. This rapid dissociation between familiar and unfamiliar snippets may lead to later reduced cortical responses to the known stimulus and increased processing associated with the novel input e.g. as expected by predictive coding views of brain function<sup>87–89</sup> whereby surprising, unknown stimuli require more processing than familiar, expected, signals.

The present study does not have sufficient data for reliable source analysis, however from the overall field maps (Fig. 2B) it appears that the identified clusters encompass the right superior temporal gyrus (rSTG), right inferior and middle frontal gyri (rIFG/rMFG) and left posterior parietal cortex (LPPC). Interestingly, a recent meta-analysis of fMRI work seeking to identify the neural correlates of music familiarity<sup>80</sup>, has identified broadly consistent brain regions.

The rSTG and rIFG/MFG have been implicated in processes related to recognition, notably in the context of voices<sup>90,91</sup>. Zäske *et al.*<sup>92</sup> demonstrated that exposure to unfamiliar voices entailed increased activation in those areas. Similarly, old/new recognition paradigms have also demonstrated increased activation to unfamiliar (“new”) relative to familiar (“old”) words in these regions<sup>93</sup>. These increases in activation to unfamiliar items may be associated with explicit memory-driven novelty detection or else reflect a more general increase in activation related to attentional capture, or effort associated with processing of unfamiliar stimuli. Both the rIFG and rMFG have been implicated in a network that allocates processing resources to external stimuli of high salience/novelty<sup>94–96</sup>.

The left-parietal ROI is consistent with a large body of research which implicates left parietal regions (left posterior parietal cortex; LPPC) in episodic memory retrieval<sup>7,22,76,97–101</sup>. These areas are reciprocally connected to the para-hippocampal cortex and to the hippocampus, consistent with a role in the memory network, though the specific contribution of the PPC to episodic retrieval remain poorly understood.

Most reports, including EEG and BOLD-based investigations, found greater activation to familiar items in LPPC, which increases with the level of recollection of item details<sup>7,30,100,102</sup>. However, a recent human intracranial recording study<sup>99</sup> revealed the presence of two types of memory selective neurons in the PPC: those that increased their firing rates for familiar stimuli, and those that preferred novel stimuli. The presence of memory sensitive cells that respond to unfamiliar items suggests that the coding of “newness” is associated with processing that is different from a mere absence of a familiarity signal. Our finding that main group responses to both familiar and unfamiliar items were altered relative to those in the control group, is in-line with these results and indicates that, rather than familiarity per se, the parietal cortex may reflect broader aspects of memory search. The overall strength of activation to familiar vs. unfamiliar items may be related to the specific task demands: paradigms in which memory strength is relatively weak (as is the case for most experiments that involve recently memorized items) may result in reliance on the “familiar” sensitive cells whilst tasks which probe robust memories and in which novel items are therefore more salient (such as may have been the case here) may prompt overall stronger responses to the novel items.

Together, the eye tracking and EEG data reveal early effects of familiarity in the pupil dynamics measure and later effects in the EEG brain responses. The lack of earlier effects in EEG may result from various factors, including that early brain activity may not have been measurable with the current setup. The audio snippets were cut from random parts of the song. Therefore, the temporal dynamics at onset differed between snippets, which may have resulted in phase misalignment of single-trial activations, reducing the aggregate evoked response. Failure to capture early brain responses may also arise due to non-optimal source orientation, or if the early effects do not arise in cortex. Thus, we suggest the observed latency of EEG responses to be an upper limit, with the actual earliest difference likely to arise much closer in time to the effect in pupil responses. In particular, as discussed above, the rapid pupillometry effects are likely to arise from sub-cortical recognition pathways and are therefore not

measurable on the scalp. Future research combining sensitive pupil and brain imaging measurement is required to understand the underlying network.

In summary, the present results reveal that snippets from a familiar and unfamiliar song are differentiated rapidly in the brain. Even though only unique, randomly interleaved, snippets were presented, significant differences between responses were revealed from ~100–300 ms from sound onset. Initially, this was reflected by increased pupil dilation rate to the snippets from the familiar song. EEG responses differentiated from 350 ms after onset and recruited similar mechanisms to those previously identified in classic “old/new” memory paradigms. Both effects were remarkably stable across participants and song material (despite that different songs were used for each main subject). Due to the brief and random nature of snippet presentation, it is unlikely that these effects reflect recognition of melodic or semantic features but rather rely on a rapid match with a memory template of the familiar piece of music.

### Data availability

The data supporting the results in this manuscript are available at [<https://doi.org/10.5522/04/9975983>].

Received: 14 May 2019; Accepted: 7 October 2019;

Published online: 30 October 2019

### References

- Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R. & Pantev, C. Automatic Encoding of Polyphonic Melodies in Musicians and Nonmusicians. *J. Cogn. Neurosci.* **17**, 1578–1592 (2005).
- Trainor, L. J., Marie, C., Bruce, I. C. & Bidelman, G. M. Explaining the high voice superiority effect in polyphonic music: Evidence from cortical evoked potentials and peripheral auditory models. *Hear. Res.* **308**, 60–70 (2014).
- Filipic, S., Tillmann, B. & Bigand, E. Erratum to: Judging familiarity and emotion from very brief musical excerpts. *Psychon. Bull. Rev.* **17**, 601–601 (2010).
- Halpern, A. R. & Bartlett, J. C. Memory for Melodies. In *Music Perception* (eds Riess Jones, M., Fay, R. R. & Popper, A. N.) 233–258, [https://doi.org/10.1007/978-1-4419-6114-3\\_8](https://doi.org/10.1007/978-1-4419-6114-3_8) (Springer New York, 2010).
- Schellenberg, E. G., Iverson, P. & Mckinnon, M. C. Name that tune: Identifying popular recordings from brief excerpts. *Psychon. Bull. Rev.* **6**, 641–646 (1999).
- Koelsch, S. Investigating the Neural Encoding of Emotion with Music. *Neuron* **98**, 1075–1079 (2018).
- Wagner, A. D., Shannon, B. J., Kahn, I. & Buckner, R. L. Parietal lobe contributions to episodic memory retrieval. *Trends Cogn. Sci.* **9**, 445–453 (2005).
- Halpern, A. R. & Zatorre, R. J. When That Tune Runs Through Your Head: A PET Investigation of Auditory Imagery for Familiar Melodies. *Cereb. Cortex* **9**, 697–704 (1999).
- Kraemer, D. J. M., Macrae, C. N., Green, A. E. & Kelley, W. M. Sound of silence activates auditory cortex. *Nature* **434**, 158 (2005).
- Martarelli, C. S., Mayer, B. & Mast, F. W. Daydreams and trait affect: The role of the listener’s state of mind in the emotional response to music. *Conscious. Cogn.* **46**, 27–35 (2016).
- Janata, P., Tomic, S. T. & Rakowski, S. K. Characterisation of music-evoked autobiographical memories. *Memory* **15**, 845–860 (2007).
- Montinari, M. R., Giardina, S., Minelli, P. & Minelli, S. History of Music Therapy and Its Contemporary Applications in Cardiovascular Diseases. *South. Med. J.* **111**, 98–102 (2018).
- Cuddy, L. L. & Duffin, J. Music, memory, and Alzheimer’s disease: is music recognition spared in dementia, and how can it be assessed? *Med. Hypotheses* **64**, 229–235 (2005).
- Hailstone, J. C., Omar, R. & Warren, J. D. Relatively preserved knowledge of music in semantic dementia. *J. Neurol Neurosurg Psychiatry* **80**, 808–809 (2009).
- Bigand, E., Gérard, Y. & Molin, P. The Contribution of Local Features to Familiarity Judgments in Music. *Ann. N. Y. Acad. Sci.* **1169**, 234–244 (2009).
- Krumhansl, C. L. Plink: ‘Thin Slices’ of Music. *Music Percept. Interdiscip. J.* **27**, 337–354 (2010).
- Tillmann, B., Albouy, P., Caclin, A. & Bigand, E. Musical familiarity in congenital amusia: Evidence from a gating paradigm. *Cortex* **59**, 84–94 (2014).
- Bigand, E., Delbé, C., Gérard, Y. & Tillmann, B. Categorization of Extremely Brief Auditory Stimuli: Domain-Specific or Domain-General Processes? *PLoS One* **6**, e27024 (2011).
- Agus, T. R., Suied, C., Thorpe, S. J. & Pressnitzer, D. Fast recognition of musical sounds based on timbre. *J. Acoust. Soc. Am.* **131**, 4124–4133 (2012).
- Suied, C., Agus, T. R., Thorpe, S. J., Mesgarani, N. & Pressnitzer, D. Auditory gist: Recognition of very short sounds from timbre cues. *J. Acoust. Soc. Am.* **135**, 1380–1391 (2014).
- Wixted, J. T. Dual-process theory and signal-detection theory of recognition memory. *Psychol. Rev.* **114**, 152–176 (2007).
- Curran, T. Brain potentials of recollection and familiarity. *Mem. Cognit.* **28**, 923–938 (2000).
- Johnson, R. ERP Insights into the Neurobiology of Memory Systems. In *Handbook of neuropsychology* 135–163 (1995).
- Rugg, M. D. & Curran, T. Event-related potentials and recognition memory. *Trends Cogn. Sci.* **11**, 251–257 (2007).
- Kappenman, E. S. & Luck, S. J. ERP Components: The Ups and Downs of Brainwave Recordings. In *The Oxford handbook of event-related potential components* 3–30 (2012).
- Bobes, M. A., Martín, M., Olivares, E. & Valdés-Sosa, M. Different scalp topography of brain potentials related to expression and identity matching of faces. *Cogn. Brain Res.* **9**, 249–260 (2000).
- Záske, R., Volberg, G., Kovács, G. & Schweinberger, S. R. Electrophysiological Correlates of Voice Learning and Recognition. *J. Neurosci.* **34**, 10821–10831 (2014).
- Sanquist, T. F., Rohrbaugh, J. W., Syndulko, K. & Lindsley, D. B. Electrocutaneous Signs of Levels of Processing: Perceptual Analysis and Recognition Memory. *Psychophysiology* **17**, 568–576 (1980).
- Klostermann, E. C., Kane, A. J. M. & Shimamura, A. P. Parietal activation during retrieval of abstract and concrete auditory information. *NeuroImage* **40**, 896–901 (2008).
- Klostermann, E. C., Loui, P. & Shimamura, A. P. Activation of right parietal cortex during memory retrieval of nonlinguistic auditory stimuli. *Cogn. Affect. Behav. Neurosci.* **9**, 242–248 (2009).
- Beatty, J. Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychol. Bull.* **91**, 276–292 (1982).
- Bradley, M. M., Miccoli, L., Escrig, M. A. & Lang, P. J. The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology* **45**, 602–607 (2008).
- Hess, E. H. & Polt, J. M. Pupil Size in Relation to Mental Activity during Simple Problem-Solving. *Science* **143**, 1190–1192 (1964).
- Kahneman, D. & Beatty, J. Pupil Diameter and Load on Memory. *Science* **154**, 1583–1585 (1966).

35. Mathot, S. Pupillometry: Psychology, Physiology, and Function. *J. Cogn.* **1** (2018).
36. Preuschoff, K., Hart, B. M. & Einhäuser, W. Pupil Dilation Signals Surprise: Evidence for Noradrenaline's Role in Decision Making. *Front. Neurosci.* **5** (2011).
37. Privitera, C. M., Renninger, L. W., Carney, T., Klein, S. & Aguilar, M. Pupil dilation during visual target detection. *J. Vis.* **10**, 3–3 (2010).
38. Stelmack, R. M. & Siddle, D. A. T. Pupillary Dilation as an Index of the Orienting Reflex. *Psychophysiology* **19**, 706–708 (1982).
39. Laeng, B., Eidet, L. M., Sulutvedt, U. & Panksepp, J. Music chills: The eye pupil as a mirror to music's soul. *Conscious. Cogn.* **44**, 161–178 (2016).
40. Harrison, L. & Loui, P. Thrills, chills, frissons, and skin orgasms: toward an integrative model of transcendent psychophysiological experiences in music. *Front. Psychol.* **5** (2014).
41. Berridge, C. W. & Waterhouse, B. D. The locus coeruleus–noradrenergic system: modulation of behavioral state and state-dependent cognitive processes. *Brain Res. Rev.* **42**, 33–84 (2003).
42. Joshi, S., Li, Y., Kalwani, R. M. & Gold, J. I. Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. *Neuron* **89**, 221–234 (2016).
43. Reimer, J. *et al.* Pupil Fluctuations Track Fast Switching of Cortical States during Quiet Wakefulness. *Neuron* **84**, 355–362 (2014).
44. Reimer, J. *et al.* Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. *Nat. Commun.* **7**, 13289 (2016).
45. Sara, S. J. The locus coeruleus and noradrenergic modulation of cognition. *Nat. Rev. Neurosci.* **10**, 211–223 (2009).
46. Schneider, M. *et al.* Spontaneous pupil dilations during the resting state are associated with activation of the salience network. *NeuroImage* **139**, 189–201 (2016).
47. Liao, H.-I., Kidani, S., Yoneya, M., Kashino, M. & Furukawa, S. Correspondences among pupillary dilation response, subjective salience of sounds, and loudness. *Psychon. Bull. Rev.* **23**, 412–425 (2016).
48. Wang, C.-A. & Munoz, D. P. Modulation of stimulus contrast on the human pupil orienting response. *Eur. J. Neurosci.* **40**, 2822–2832 (2014).
49. Einhäuser, W., Koch, C. & Carter, O. L. Pupil Dilation Betrays the Timing of Decisions. *Front. Hum. Neurosci.* **4** (2010).
50. Lavin, C., San Martín, R. & Rosales Jubal, E. Pupil dilation signals uncertainty and surprise in a learning gambling task. *Front. Behav. Neurosci.* **7** (2014).
51. Samuels, E. R. & Szabadi, E. Functional Neuroanatomy of the Noradrenergic Locus Coeruleus: Its Roles in the Regulation of Arousal and Autonomic Function Part I: Principles of Functional Organisation. *Curr. Neuropharmacol.* **6**, 235–253 (2008).
52. Sara, S. J. & Bouret, S. Orienting and Reorienting: The Locus Coeruleus Mediates Cognition through Arousal. *Neuron* **76**, 130–141 (2012).
53. Partala, T., Jokiniemi, M. & Surakka, V. Pupillary Responses to Emotionally Provocative Stimuli. In *Proceedings of the 2000 Symposium on Eye Tracking Research & Applications* 123–129, <https://doi.org/10.1145/355017.355042> (ACM, 2000).
54. Wang, C.-A., Blohm, G., Huang, J., Boehnke, S. E. & Munoz, D. P. Multisensory integration in orienting behavior: Pupil size, microsaccades, and saccades. *Biol. Psychol.* **129**, 36–44 (2017).
55. Vö, M. L.-H. *et al.* The coupling of emotion and cognition in the eye: Introducing the pupil old/new effect. *Psychophysiology* **45**, 130–140 (2008).
56. Brocher, A. & Graf, T. Decision-related factors in pupil old/new effects: Attention, response execution, and false memory. *Neuropsychologia* **102**, 124–134 (2017).
57. Heaven, B. & Hutton, S. B. Keeping an eye on the truth? Pupil size changes associated with recognition memory. *Memory* **19**, 398–405 (2011).
58. Kafkas, A. & Montaldi, D. The pupillary response discriminates between subjective and objective familiarity and novelty. *Psychophysiology* **52**, 1305–1316 (2015).
59. Otero, S. C., Weekes, B. S. & Hutton, S. B. Pupil size changes during recognition memory. *Psychophysiology* **48**, 1346–1353 (2011).
60. Kafkas, A. & Montaldi, D. Familiarity and recollection produce distinct eye movement, pupil and medial temporal lobe responses when memory strength is matched. *Neuropsychologia* **50**, 3080–3093 (2012).
61. Kafkas, A. & Montaldi, D. Recognition memory strength is predicted by pupillary responses at encoding while fixation patterns distinguish recollection from familiarity. *Q. J. Exp. Psychol.* **64**, 1971–1989 (2011).
62. Papesch, M. H., Goldinger, S. D. & Hout, M. C. Memory strength and specificity revealed by pupillometry. *Int. J. Psychophysiol.* **83**, 56–64 (2012).
63. Weiss, M. W., Trehub, S. E., Glenn, E. & Habashi, P. Pupils dilate for vocal or familiar music. *J. Exp. Psychol. Hum. Percept. Perform.* **42**, 1061–1065 (2016).
64. Delorme, A. & Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. **134**, 9–21 (2004).
65. Oostenveld, R. *et al.* FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data, FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Comput. Intell. Neurosci. Comput. Intell. Neurosci.* **2011**, e156869 (2010).
66. Maris, E. & Oostenveld, R. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* **164**, 177–190 (2007).
67. Hoeks, B. & Levelt, W. J. M. Pupillary dilation as a measure of attention: a quantitative system analysis. *Behav. Res. Methods Instrum. Comput.* **25**, 16–26 (1993).
68. Murphy, P. R., Robertson, I. H., Balsters, J. H. & O'Connell, R. G. Pupillometry and P3 index the locus coeruleus-noradrenergic arousal function in humans. *Psychophysiology* **48**, 1532–1543 (2011).
69. Wang, C.-A. & Munoz, D. P. A circuit for pupil orienting responses: implications for cognitive modulation of pupil size. *Curr. Opin. Neurobiol.* **33**, 134–140 (2015).
70. Rolfs, M., Kliegl, R. & Engbert, R. Toward a model of microsaccade generation: The case of microsaccadic inhibition. *J. Vis.* **8**, 5–5 (2008).
71. Dayan, P. & Abbott, L. F. *Theoretical neuroscience* 806 (MIT Press, Cambridge, MA, 2001).
72. Widmann, A., Schröger, E. & Wetzl, N. Emotion lies in the eye of the listener: emotional arousal to novel sounds is reflected in the sympathetic contribution to the pupil dilation response and the P3. <https://doi.org/10.1101/250084> (2018).
73. Efron, B. & Tibshirani, R. J. *An Introduction to the Bootstrap*. (CRC Press, 1994).
74. Chait, M., Simon, J. & Poeppel, D. Auditory M50 and M100 responses to broadband noise: functional implications. *Neuroreport* **15**, 2455–2458 (2004).
75. Curran, T. & Cleary, A. M. Using ERPs to dissociate recollection from familiarity in picture recognition. *Cogn. Brain Res.* **15**, 191–205 (2003).
76. Woodruff, C. C., Hayama, H. R. & Rugg, M. D. Electrophysiological dissociation of the neural correlates of recollection and familiarity. *Brain Res.* **1100**, 125–135 (2006).
77. Evans, L. H. & Wilding, E. L. On the sensitivity of event-related fields to recollection and familiarity. *Brain Cogn.* **126**, 33–39 (2018).
78. Paller, K. A., Voss, J. L. & Boehm, S. G. Validating neural correlates of familiarity. *Trends Cogn. Sci.* **11**, 243–250 (2007).
79. Yonelinas, A. P. Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *J. Exp. Psychol. Learn. Mem. Cogn.* **20**, 1341–1354 (1994).
80. Freitas, C. *et al.* Neural Correlates of Familiarity in Music Listening: A Systematic Review and a Neuroimaging Meta-Analysis. *Front. Neurosci.* **12** (2018).

81. Zheng, Y. & Escabí, M. A. Proportional spike-timing precision and firing reliability underlie efficient temporal processing of periodicity and envelope shape cues. *J. Neurophysiol.* **110**, 587–606 (2013).
82. Zheng, Y. & Escabí, M. A. Distinct Roles for Onset and Sustained Activity in the Neuronal Code for Temporal Periodicity and Acoustic Envelope Shape. *J. Neurosci.* **28**, 14230–14244 (2008).
83. Bouret, S., Duvel, A., Onat, S. & Sara, S. J. Phasic Activation of Locus Ceruleus Neurons by the Central Nucleus of the Amygdala. *J. Neurosci.* **23**, 3491–3497 (2003).
84. Wagatsuma, A. *et al.* Locus coeruleus input to hippocampal CA3 drives single-trial learning of a novel context. *Proc. Natl. Acad. Sci.* **115**, E310–E316 (2018).
85. Wixted, J. T. & Mickes, L. A continuous dual-process model of remember/know judgments. *Psychol. Rev.* **117**, 1025–1054 (2010).
86. Yonelinas, A. P. The Nature of Recollection and Familiarity: A Review of 30 Years of Research. *J. Mem. Lang.* **46**, 441–517 (2002).
87. de Lange, F. P., Heilbron, M. & Kok, P. How Do Expectations Shape Perception? *Trends Cogn. Sci.* **22**, 764–779 (2018).
88. Friston, K. & Kiebel, S. Predictive coding under the free-energy principle. *Philos. Trans. R. Soc. B Biol. Sci.* **364**, 1211–1221 (2009).
89. Heilbron, M. & Chait, M. Great expectations: Is there evidence for predictive coding in auditory cortex? *Neuroscience*, <https://doi.org/10.1016/j.neuroscience.2017.07.061> (2018).
90. Blank, H., Wieland, N. & von Kriegstein, K. Person recognition and the brain: Merging evidence from patients and healthy individuals. *Neurosci. Biobehav. Rev.* **47**, 717–734 (2014).
91. von Kriegstein, K., Eger, E., Kleinschmidt, A. & Giraud, A. L. Modulation of neural responses to speech by directing attention to voices or verbal content. *Cogn. Brain Res.* **17**, 48–55 (2003).
92. Zäske, R., Awwad Shiekh Hasan, B. & Belin, P. It doesn't matter what you say: fMRI correlates of voice learning and recognition independent of speech content. *Cortex* **94**, 100–112 (2017).
93. Hoppstädter, M., Baeuchl, C., Diener, C., Flor, H. & Meyer, P. Simultaneous EEG–fMRI reveals brain networks underlying recognition memory ERP old/new effects. *NeuroImage* **116**, 112–122 (2015).
94. Corbetta, M., Patel, G. & Shulman, G. L. The Reorienting System of the Human Brain: From Environment to Theory of Mind. *Neuron* **58**, 306–324 (2008).
95. Corbetta, M. & Shulman, G. L. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* **3**, 201–215 (2002).
96. Shulman, G. L. *et al.* Interaction of Stimulus-Driven Reorienting and Expectation in Ventral and Dorsal Frontoparietal and Basal Ganglia-Cortical Networks. *J. Neurosci.* **29**, 4392–4407 (2009).
97. Curran, T. & Friedman, W. J. ERP old/new effects at different retention intervals in recency discrimination tasks. *Cogn. Brain Res.* **18**, 107–120 (2004).
98. Finnigan, S., Humphreys, M. S., Dennis, S. & Geffen, G. ERP 'old/new' effects: memory strength and decisional factor(s). *Neuropsychologia* **40**, 2288–2304 (2002).
99. Rutishauser, U., Aflalo, T., Rosario, E. R., Pouratian, N. & Andersen, R. A. Single-Neuron Representation of Memory Strength and Recognition Confidence in Left Human Posterior Parietal Cortex. *Neuron* **97**, 209–220.e3 (2018).
100. Sestieri, C., Shulman, G. L. & Corbetta, M. The contribution of the human posterior parietal cortex to episodic memory. *Nat. Rev. Neurosci.* **18**, 183–192 (2017).
101. Tsivilis, D. *et al.* Old-new ERP effects and remote memories: the late parietal effect is absent as recollection fails whereas the early mid-frontal effect persists as familiarity is retained. *Front. Hum. Neurosci.* **9** (2015).
102. Hutchinson, J. B., Uncapher, M. R. & Wagner, A. D. Posterior parietal cortex and episodic retrieval: Convergent and divergent effects of attention and memory. *Learn. Mem.* **16**, 343–356 (2009).

## Acknowledgements

This research was supported by a EC Horizon 2020 grant and a BBSRC international partnering award to M.C.

## Author contributions

R.J.: designed the study, conducted the research; analyzed the data; wrote the paper. U.P.: designed the study, conducted the research; analyzed the data; wrote the paper. M.Y.: conducted pilot experiments; commented on draft. S.Z.: analyzed the data; wrote the paper. M.C.: supervised the research; obtained funding; designed the study; wrote the paper.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to U.P. or M.C.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019