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

Robert Jagiello[1,2], Ulrich Pomper[1,3], Makoto Yoneya[4], Sijia Zhao[1], Maria Chait[1]

[1] Ear Institute, University College London, London, UK
[2]Currently at Institute of Cognitive and Evolutionary Anthropology, University of Oxford, Oxford, UK
[3] Currently at Faculty of Psychology, University of Vienna, Austria
[4]NTT Communication Science Laboratories, NTT Corporation, Atsugi, 243-0198 Japan

Corresponding Author:

Ulrich Pomper
ulrich.pomper@univie.ac.at
Faculty of Psychology
University of Vienna
Liebiggasse 5
1010 Vienna, Austria

Keywords: Music; memory; familiarity; Electroencephalography; Pupillometry; rapid recognition;

Acknowledgements: This research was supported by a EC Horizon 2020 grant and a BBSRC international partnering award to MC.

When listening to music, we are able to identify well-known songs as familiar within fractions of a second. However, as of now it is unclear how fast the brain can differentiate between familiar and unfamiliar music, and what the underlying neural correlates are. Here, we used electro-encephalography (EEG) and pupillometry to reveal the temporal signatures of neural processes that allow differentiation between familiar and unfamiliar music. Participants passively listened to 750 ms snippets of familiar songs as well as of acoustically matched control songs, presented in random order. We observed greater pupil dilation rates to familiar music from 100-300 ms post stimulus onset. Brain responses measured with EEG showed a differentiation between familiar and unfamiliar music from 350 ms post onset, but in the opposite direction to that seen with pupillometry: Unfamiliar snippets were associated with greater responses than familiar snippets. A group of control participants, which was presented with the same song snippets but was unfamiliar with all of the songs, showed no such differences. Possible underlying mechanisms are discussed.

1. Introduction

The human auditory system exhibits a marked sensitivity when it comes to the processing of music and its various characteristics (Fujioka, Trainor, Ross, Kakigi, & Pantev, 2005; Trainor, Marie, Bruce, & Bidelman, 2014), potentially allowing for a rapid recognition of previously heard songs (Filipic, Tillmann, & Bigand, 2010; Halpern & Bartlett, 2010; Schellenberg, Iverson, & McKinnon, 1999). Our musical prowess towards recognizing familiar 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.

Despite the pleasurable and ubiquitous nature of music, reflected by its cultural omnipresence (Zatorre & Halpern, 2015) and likely evolutionary roots (Andrade & Bhattacharya, 2003; Cross, 2001), there are only few past research endeavors looking into the temporal course of the mechanisms that allow us to recognize a familiar song among unfamiliar ones. The current literature in the area of auditory recognition has predominantly focused on rapid categorization of less complex sounds (e.g. single instruments and voices; Agus, Suied, Thorpe, & Pressnitzer, 2012). In gating paradigms, the length of sound segments played to subjects is incremented step-wise, hence allowing to approximate the minimal stimulus duration that is necessary for an above-chance recognition (e.g. Bigand, Delbé, Gérard, & Tillmann, 2011). Remarkably, Suied et al.(2014) demonstrated that the minimal required time window for such instrument and voice

differentiation could be as short as 2 to 4 ms. An analogue procedure was applied to investigate minimal durations for the recognition of entire pieces of music (Bella, Peretz, & Aronoff, 2003). Bigand et al.(2009) employed a design with popular and unknown classical music, in which subjects required 250 ms for an above chance recognition. For pop music, minimal time windows for song identification range from 100 ms (Schellenberg et al., 1999) to up to 500 ms (Filipic et al., 2010; Krumhansl, 2010). However, purely behavioral investigations lack a fine temporal resolution due to their coarsely defined gating durations and also leave the question as to the perceptual and cognitive bases that are implicated in music recognition unanswered, hence necessitating the use of additional measures.

1.1. Electrophysiology of music recognition

Electroencephalography (EEG) is sensitive to functionally dissociable neural processes and their temporal signatures (Allan, Wilding, & Rugg, 1998; Cohen, 2017), and hence constitutes a suitable tool to shed light onto the temporal dynamics of music recognition. The concept of music familiarity heavily relies on long term memory traces (Wagner, Shannon, Kahn, & Buckner, 2005) in the form of auditory mental imagery (Halpern & Zatorre, 1999; Kraemer, Macrae, Green, & Kelley, 2005; Martarelli, Mayer, & Mast, 2016) and is also linked to autobiographical memories (Janata, Tomic, & Rakowski, 2007). 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 (Bigand et al., 2009; Wixted, 2007). A particular marker in the EEG literature that is tied to such recognition processes is the late positive potential (LPP; Curran, 2000; Johnson, 1995; Rugg & Curran, 2007): 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 (Kappenman & Luck, 2012). This parietal old versus new effect has consistently been found across various domains, such as facial (Bobes, Martín, Olivares, & Valdés-Sosa, 2000) and voice recognition (Zäske, Volberg, Kovács, & Schweinberger, 2014) as well as paradigms that employed visually presented- (Sanquist, Rohrbaugh, Syndulko, & Lindsley, 1980) and spoken words as stimuli (Klostermann, Kane, & Shimamura, 2008). In an fMRI study, Klostermann et al.(2009) 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, subjects 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.

1.2. Pupil responses to music

A growing body of work demonstrates that cognitive states associated with working memory load, vigilance, surprise or processing effort can be gleaned from measuring task-evoked changes in pupil diameter (Beatty, 1982; Bradley, Miccoli, Escrig, & Lang, 2008; Hess & Polt, 1964; Kahneman & Beatty, 1966; Mathôt, 2018; Preuschoff, Hart, & Einhäuser, 2011; Privitera, Renninger, Carney, Klein, & Aguilar, 2010; Stelmack & Siddle, 1982). Pupil dilation also reliably co-occurs with musical chills (Laeng, Eidet, Sulutvedt, & Panksepp, 2016) – a physiological phenomenon evoked by exposure to emotionally relevant and

familiar pieces of music (Harrison & Loui, 2014) 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 (specifically ACh and NE; Berridge & Waterhouse, 2003; Joshi, Li, Kalwani, & Gold, 2016; Reimer et al., 2014, 2016; Sara, 2009; Schneider et al., 2016) mediated salience and arousal network.

In particular, abrupt changes in pupil size are commonly observed in response to salient (Liao, Kidani, Yoneya, Kashino, & Furukawa, 2016; Wang & Munoz, 2014) or surprising (Einhäuser, Koch, & Carter, 2010; Lavín, San Martín, & Rosales Jubal, 2014; Preuschoff et al., 2011) events, including in the auditory modality. Work in animal models has established a link between such phasic pupil responses and spiking activity within Norepinephrine (alternatively Noradrenaline, NE) generating cells in the brainstem nucleus Locus Coeruleus (LC). The LC projects widely across the brain and spinal cord (Samuels & Szabadi, 2008; Sara & Bouret, 2012) 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 (Bradley et al., 2008; Partala, Jokiniemi, & Surakka, 2000; Wang, Blohm, Huang, Boehnke, & Munoz, 2017). Here we explore whether familiarity drives pupil responses.

Pupil dilations have received much attention in recognition paradigms, analogue to the previously elaborated designs in which subjects are first exposed to a list of stimuli during a learning phase and subsequently asked to identify old items during the recognition stage (Võ et al., 2008). When identifying old words, participants' pupils tend to dilate more than when confronted with novel ones, a phenomenon which is referred to as the *pupil old/new* effect (Brocher & Graf, 2017; Heaver & Hutton, 2011; Kafkas & Montaldi, 2015). In

one of their experiments, Otero et al.(2011) replicated this finding via the use of recorded words, 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. Brocher & Graf, 2017; Kafkas & Montaldi, 2011, 2012; Papesh, Goldinger, & Hout, 2012) or average pupil diameter change over the trial interval (Heaver & Hutton, 2011; Otero et al., 2011). Weiss et al. (2016) played a mix of familiar and unfamiliar folk melodies (either sung or piano-based) to subjects and demonstrated greater pupil dilations in response to the known- as opposed to the novel stimuli. In that 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.

1.3. Aims

We aimed to investigate the temporal dynamics of the neural processes that underlie the differentiation of familiar and unfamiliar music. Our approach consisted of exposing participants to randomly presented short snippets from a familiar and unfamiliar song—a design that is reminiscent of the above-mentioned real-world scenarios such as radio channel switching. Electroencephalography and pupillometry in passively listening participants served as a means to elucidate the processing stage at which familiar and unfamiliar music snippets first become dissociable, and to relate EEG and pupil-linked measures. In contrast to previous work which has focused on quantifying changes in pupil diameter (the so called 'pupil dilation response'), here 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.

2. Methods

2.1. Participants

The participant pool encompassed two independent groups: A **main group** (N_{main} = 10; 5 females; Mage= 23.56; SD = 3.71) and a **control group** ($N_{control}$= 12; 9 females; Mage= 23.08; SD = 4.79). All reported no known history of hearing or neurological impairment. Experimental procedures were approved by the research ethics committee of University College London, and written informed consent was obtained from each participant.

2.2. Stimuli and procedure

2.2.1. 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 (ratings were expressed on a 5-point scale ranging from '1 - Not at all' to '5 - Strong'). 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). Importantly, matching was also verified with the control group (see below). Further, we provided each participant with the name as well as a 1500 ms snipped of the matched song. Only if participants were unfamiliar with both, these songs were used as matched 'unfamiliar' songs. Upon completion, this procedure resulted in ten dyads (one per participant), each containing one 'familiar' and one 'unfamiliar' song.

Participants were selected for the **control group** based on non-familiarity with any of the ten dyads. 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 (1500 ms).The subjects in the control group (comprised of international students at UCL and hence relatively inexperienced with Western popular music) were unfamiliar with all songs used in this experiment, therefore the distinction between familiar and unfamiliar conditions does not apply to them.

2.2.2. Stimulus generation and testing paradigms

The beginning and end of each song, which typically constitute silent or only gradually rising parts of the instrumentation, were cut out. Both songs from each pair were 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, as well as (2) an active categorization task. In the passive listening task, subjects listened to snippets from each dyad (in separate blocks) in random order whilst their brain activity was recorded with EEG and their pupil diameters with an infrared eye-tracking camera. Each block contained 200 trials, 100 of each song presented in random order with an inter-

stimulus interval (ISI) randomized between 1000 ms and 1500 ms. This resulted in a total duration of 400 seconds 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, in random order, each in a separate block.

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 from each song, separated by 750 ms. They were instructed to indicate whether the two snippets were from the same song or from different ones, by pressing the corresponding buttons on the keyboard (with no time limit imposed). Same as for the passive listening task, subjects from the main group performed only one block, associated with their pairing of 'familiar/unfamiliar' songs. Subjects from the control group completed 10 blocks in random order.

2.2.3. 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. Subjects were instructed to continuously fixate on a white cross on grey background, presented at the center of a 24-inch monitor (BENQ XL2420T) with resolution of 1920x1080 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, subjects were instructed to attentively listen to the snippets. EEG and eye-tracking

were recorded during the first task, but not during the second, which was of purely behavioral nature.

2.3. 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 (http://www.sccn.ucsd.edu/eeglab/; Delorme & Makeig, 2004) and FieldTrip software (http://www.ru.nl/fcdonders/fieldtrip/; Oostenveld et al., 2010). 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 (Kappenman & Luck, 2012), 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. Artefacts related to eye movements, blinks and heartbeat were identified and removed using independent component analysis. Subsequently the data were re-referenced to the mean of all channels, averaged over epochs of the same condition and baseline-corrected (200 ms preceding stimulus onset).

A cluster-based permutation analysis which takes spatial and temporal adjacency into account (Maris & Oostenveld, 2007; Oostenveld et al., 2010) was used to investigate potential differences between the EEG responses to 'familiar' and 'unfamiliar' snippets within controls and main participants over the entire epoch length. The significance threshold was chosen to control family-wise error-rate (FWER) at 5%.

2.4. 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, 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. No participant was excluded in the main group.

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 (Hoeks & Levelt, 1993; Liao, Yoneya, Kidani, Kashino, & Furukawa, 2016; Murphy, Robertson, Balsters, & O'connell, 2011; Wang & Munoz, 2015) which are not optimal for capturing potentially rapid effects. Instead, the present analysis focused on examining pupil event rate. This analysis captures the incidence of pupil dilation events (Joshi et al, 2016), 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 Joshi, Li, Kalwani, & Gold, 2016; Rolfs, Kliegl, & Engbert, 2008), paralleling a similar technique for computing neural firing rates from neuronal spike trains (Dayan & Abbott, 2002). 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 $\alpha = 1/50$ ms (Dayan & Abbott, 2002; Rolfs et al., 2008; Widmann, Schröger, & Wetzel, 2018). 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 (Bradley & Tibshirani, 1994): For the main group, the difference time series between the conditions was computed for each participant, and these timeseries were subjected to bootstrap re-sampling (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. $p<0.01$).

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 theses 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. To calculate the differences between conditions, 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.

3. Results

3.1. Music questionnaire

Subjects in the **main group** were highly familiar with the selected songs (Familiarity Score; M = 3.91; SD = 0.85) and reported increased emotional ratings for items such as 'excited' (M = 3.50; SD = 1.08), 'happy' (M = 3.60; SD = 1.26), and 'memory inducing'. Conversely, control subjects gave ratings that were indicative of low familiarity for all the songs (M = 1.05; SD = 0.09), as well as their respective artists in general (M = 1.10; SD = 0.14).

3.2. EEG

The general EEG response to the sound snippets (collapsed across all participants and conditions) is shown in Figure 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. Chait et al, 2004).

3.3.1. Control group

The main purpose of the control group was to verify that any significant differences which are potentially established for main participants were due to our familiarity manipulation 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 brain activity. However, the cluster-based permutation test revealed significant clusters in dyad 2 (mean t = 3.15) and 5 (mean t = 2.84). Those dyads were excluded from the subsequent main group analysis.

3.3.2. Main group

Comparing responses to 'familiar' and 'unfamiliar' snippets within the main group, we identified two clusters of channels showing a significant difference between the two conditions. A centrally located cluster of 26 channels showing a significant difference between conditions from 540 to 750 ms (mean t = -4.52), and a right fronto-temporal cluster of 20 channels, showing significant difference between 350 to 750 ms (mean t =4.15). In both clusters, responses to unfamiliar snippets evoked stronger activity (more negative in the first cluster and more positive in the second cluster) than familiar snippets (Figure 2).

To confirm that this effect is specific to the main group, the identified clusters were used as ROIs for a factorial ANOVA across groups, with familiarity (familiar/unfamiliar) as

independent within-subjects factor and the average EEG response amplitudes within the cluster as dependent variable. The interaction was significant for both the central cluster ($F(1, 14) = 73.56$, $p < .001$) as well as the fronto-temporal cluster ($F(1, 14) = 37.91$, $p < .001$) and. More precisely, compared to controls, the main group showed increased negativity in the central regions and increased positivity in a right fronto-temporal area in response to unfamiliar snippets (Figure 2C).

3.3. Pupil dilation

3.3.1. Control group

Figure 3A (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 400ms, before returning to baseline around 550ms post-onset. No difference was observed between the two conditions throughout the entire epoch.

3.3.2. Main group

When compared with unfamiliar conditions, familiar snippets were associated with a higher pupil dilation rate from 108–315ms post sound onset (Figure 3A, top). 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- 315 ms) identified as significant in the main group analysis. This was achieved by computing a distribution of

differences between conditions based on the control group data (H0 distribution). On each iteration (5000 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 Figure 3B, shows the distribution of these values. The mean difference from the main group, indicated via the red dot, lies well beyond this distribution ($p=0.0022$), confirming that the effect observed for the main group was different from that in the control group.

3.4. Behavioral task

The behavioral task aimed to verify whether subjects 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 subjects correctly identified whether or not the two presented snippets were from the same song in 92% of trials, whereas controls in 79% of trials. One-sample t-tests revealed that these scores are at above-chance levels, $t(9) = 13.61$, $p< .00001$ for controls, and $t(9) = 18.11$, $p< .000001$ for the main group. An independent samples t-test revealed that main subjects scored significantly higher than controls $t(18) = 6.19$, $p< .00001$. Therefore, the behavioral results suggest that both groups were able to differentiate the songs in each dyad well above chance. That the control group participants achieved above chance performance suggests that there was sufficient acoustic information in the snippet to distinguish the songs.

4. Discussion

We used EEG and eye tracking to identify brain responses which distinguish between familiar and unfamiliar 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 subjects. 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 mechanisms. Specifically, we report two main observations: (1) pupil responses showed greater dilation rate to snippets of familiar music from ~100-300 ms post stimulus (2) brain activity measured with EEG showed a differentiation between responses to familiar and unfamiliar music snippets from 350 ms post onset, but, notably, in the opposite direction from that observed with pupillometry: Unfamiliar music snippets evoked stronger responses.

The implications of these results for our understanding of how music experiences are coded in the brain are discussed below (Koelsch, 2018). Beyond basic research, this knowledge is instrumental for various music-based therapeutic interventions (Montinari, Giardina, Minelli, & Minelli, 2018). 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 (Cuddy & Duffin, 2005; Hailstone, Omar, & Warren, 2009). Our results may provide a clue to understanding the basis of this phenomena, and how it may be objectively quantified.

4.1. Familiarity

'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 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 with these other factors.

4.2. Pupil responses

Pupil responses revealed a very early effect, dissociating the 'familiar' and 'unfamiliar' snippets from 108ms after onset. To our knowledge, this is the earliest pupil old/new effect documented. Previously, Kafkas and Montaldi (2015) found the duration of first fixation after stimulus onset discriminated between familiar and unfamiliar stimuli as early as 320ms post-onset, with the familiar condition associated with a longer first fixation. Whilst our subjects were instructed to listen passively, it is likely that the participants in the main group may have been listening with the intent to recognize the familiar songs. It is therefore possible that the rapid effects are driven by this perceptual state.

The timing at which 'familiar' snippets were distinguished from the unfamiliar ones is consistent with previous behaviorally derived estimates which place minimum identification time at 100-250 ms for music (Schellenberg et al. 1999; Bigand, 2009) and likely stems from remarkable sensitivity to, and memory of, the timbre of the familiar songs (Suied et al., 2014).

Research in animal models has linked phasic pupil dilation events with increased firing in the LC (Joshi et al., 2016). The present results can therefore be taken to indicate that the LC was differentially activated as early as 100 ms after sound onset, possibly through projections from the inferior colliculus (where timbre cues may be processed; Zheng & Escabi, 2008, 2013) or other subcortical structures such as the amygdala, which is known to activate the LC (Bouret, Duvel, Onat, & Sara, 2003). The Amygdala has been extensively implicated in music processing, especially for familiar pleasure-inducing music such as the one used here (Blood & Zatorre, 2001; Salimpoor et al., 2013).

It is notable that the pupil effects were restricted to the early portion of the trial. A possible explanation is that later in the trial, the relatively subtle effects of familiarity were masked by the more powerful effects associated with processing the snippet as a whole.

4.3. Brain responses

Contrary to previous demonstration of increased activation to familiar items -the so called old/ new effect (Kappenman & Luck, 2012; Wagner et al., 2005)- familiarity was associated with reduced responses between 350 to 750 ms in the right fronto-temporal regions. This might be due to the fact that here subjects were listening passively and not making active recognition judgments as it was the case in the majority of past paradigms (e.g. Sanquist et al., 1980). Further, it must be cautioned that the majority of observed LPPs in the literature occurred post stimulation, whereas in the present design the difference manifested during ongoing listening.

The present study does not have sufficient data for reliable source analysis, however from the overall field maps (Figure 2B) it appears that the pattern of increased activation

for 'unfamiliar' snippets encompasses areas such as the right superior temporal gyrus (rSTG),right inferior and middle frontal gyri (rIFG/rMFG), which have been implicated in the recognition of familiarity notably in the context of voices (Blank, Wieland, & von Kriegstein, 2014; von Kriegstein, Eger, Kleinschmidt, & Giraud, 2003) and are therefore potential neural generators of the present effect. Indeed, Zäske et al. (2017) demonstrated that exposure to unfamiliar voices entailed an increased activation in those areas. These increases in activation may be associated with explicit memory-driven novelty detection or else reflect more general increase in activation related to attentional capture, or listening effort associated with processing of unfamiliar sounds. Both the rIFG and rMFG have been implicated in a network that allocates processing resources to external stimuli of high salience (Corbetta, Patel, & Shulman, 2008; Corbetta & Shulman, 2002; Shulman et al., 2009).

Notably, no differences were observed in the control group, despite the behavioral evidence that indicated that they can differentiate the 'familiar' and 'unfamiliar' songs. This suggests that whilst there may have been enough information for active perceptual scanning to indicate acoustic or stylistic differences between songs which would lead to above chance recognition, these processes did not affect the presently observed brain responses during passive listening.

4.4. Relating pupil- and brain- responses

Together, the data reveal early effects of familiarity in the pupil dynamics measure and later (opposite direction) 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. This could happen if effects are small, or sources not optimally oriented for capture by EEG. 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. One possible hypothesis, consistent with the present pattern of results, is that familiar snippets are recognized rapidly, likely based on timbre cues, mediated by fast acting sub-cortical circuitry. This rapid dissociation between 'familiar' and 'unfamiliar' snippets leads to later reduced cortical processing of the known stimulus and increased processing associated with the novel input (e.g. as expected by predictive coding views of brain function; de Lange, Heilbron, & Kok, 2018; Friston & Kiebel, 2009; Heilbron & Chait, 2018).

4.5. Limitations

There are two important limitations for the present work: Firstly, 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. Alternatively, it may be possible to work with a composer to 'commission' a match to a familiar song. Another major limitation is the fact that 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 hard to overcome, and the results must be interpreted in this light.

5. References

Agus, T. R., Suied, C., Thorpe, S. J., & Pressnitzer, D. (2012). Fast recognition of musical sounds based on timbre. *The Journal of the Acoustical Society of America*, *131*(5), 4124. https://doi.org/10.1121/1.3701865

Allan, K., Wilding, E. L., & Rugg, M. D. (1998). Electrophysiological evidence for dissociable processes contributing to recollection. *Acta Psychologica*, *98*, 231–252. https://doi.org/Doi 10.1016/S0001-6918(97)00044-9

Andrade, P. E., & Bhattacharya, J. (2003). Brain tuned to music. *Journal of the Royal Society of Medicine*, *96*(6), 284–287. https://doi.org/10.1258/jrsm.96.6.284

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, *91*(2), 276–292. https://doi.org/10.1037/0033-2909.91.2.276

Bella, S. D., Peretz, I., & Aronoff, N. (2003). Time course of melody recognition: a gating paradigm study. *Percept Psychophys*, *65*(7), 1019–1028. https://doi.org/10.3758/BF03194831

Berridge, C. W., & Waterhouse, B. D. (2003). The locus coeruleus-noradrenergic system: Modulation of behavioral state and state-dependent cognitive processes. *Brain Research Reviews*. https://doi.org/10.1016/S0165-0173(03)00143-7

Bigand, E., Delbé, C., Gérard, Y., & Tillmann, B. (2011). Categorization of extremely brief auditory stimuli: Domain-specific or domain-general processes? *PLoS ONE*, *6*(10). https://doi.org/10.1371/journal.pone.0027024

Bigand, E., Gérard, Y., & Molin, P. (2009). The contribution of local features to familiarity judgments in music. *Annals of the New York Academy of Sciences*, *1169*, 234–244. https://doi.org/10.1111/j.1749-6632.2009.04552.x

Blank, H., Wieland, N., & von Kriegstein, K. (2014). Person recognition and the brain: Merging evidence from patients and healthy individuals. *Neuroscience and Biobehavioral Reviews*, *47*, 717–734. https://doi.org/10.1016/j.neubiorev.2014.10.022

Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences*, *98*(20), 11818–11823. https://doi.org/10.1073/pnas.191355898

Bobes, M. A., Martín, M., Olivares, E., & Valdés-Sosa, M. (2000). Different scalp topography of brain potentials related to expression and identity matching of faces. *Cognitive Brain Research*, *9*(3), 249–260. https://doi.org/10.1016/S0926-6410(00)00003-3

Bouret, S., Duvel, A., Onat, S., & Sara, S. J. (2003). Phasic activation of locus ceruleus neurons by the central nucleus of the amygdala. *J. Neurosci.*, *23*(8), 3491–3497. https://doi.org/23/8/3491

Bradley, E., & Tibshirani, R. J. (1994). *An Introduction to the Bootstrap*. CRC Press. https://doi.org/10.1214/ss/1063994971

Bradley, M. B., Miccoli, L. M., Escrig, M. a, & Lang, P. J. (2008). The pupil as a measure of emotional arousal and automatic activation. *Psychophysiology*, *45*(4), 602. https://doi.org/10.1111/j.1469-8986.2008.00654.

Chait, M., Simon, J. Z., & Poeppel, D. (2004). Auditory M50 and M100 responses to broadband noise: functional implications. *Neuroreport*, *15*(16), 2455-2458.

Cohen, M. X. (2017). Where Does EEG Come From and What Does It Mean? *Trends in Neurosciences*, *40*(4), 208–218. https://doi.org/10.1016/j.tins.2017.02.004

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The Reorienting System of the Human Brain: From Environment to Theory of Mind. *Neuron*. https://doi.org/10.1016/j.neuron.2008.04.017

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews. Neuroscience*, *3*(3), 201–15. https://doi.org/10.1038/nrn755

Cross, I. (2001). Music, cognition, culture, and evolution. *Annals of the New York Academy of sciences*, *930*(1), 28-42.

Cuddy, L. L., & Duffin, J. (2005). Music, memory, and Alzheimer's disease: Is music recognition spared in dementia, and how can it be assessed? *Medical Hypotheses*, *64*(2), 229–235. https://doi.org/10.1016/j.mehy.2004.09.005

Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition*, *28*(6), 923–938. https://doi.org/10.3758/BF03209340

Dayan, P., & Abbott, L. (2002). Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems (Computational Neuroscience). *Journal of Cognitive Neuroscience*, 480. https://doi.org/10.1016/j.neuron.2008.10.019

de Lange, F. P., Heilbron, M., & Kok, P. (2018). How Do Expectations Shape Perception? *Trends in Cognitive Sciences*, *22*(9), 764–779. https://doi.org/10.1016/j.tics.2018.06.002

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Einhäuser, W., Koch, C., & Carter, O. L. (2010). Pupil dilation betrays the timing of decisions. *Frontiers in Human Neuroscience*, *4*(February), 18. https://doi.org/10.3389/fnhum.2010.00018

Filipic, S., Tillmann, B., & Bigand, E. (2010). Erratum to: Judging familiarity and emotion from very brief musical excerpts. *Psychonomic Bulletin & Review*, *17*(4), 601–601. https://doi.org/10.3758/PBR.17.4.601

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1211–1221. https://doi.org/10.1098/rstb.2008.0300

Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2005). Automatic encoding of polyphonic melodies in musicians and nonmusicians. *Journal of Cognitive Neuroscience*, *17*(10), 1578–1592. https://doi.org/10.1162/089892905774597263

Hailstone, J. C., Omar, R., & Warren, J. D. (2009). Relatively preserved knowledge of music in semantic dementia. *Journal of Neurology, Neurosurgery and Psychiatry*, *80*(7), 808–809. https://doi.org/10.1136/jnnp.2008.153130

Halpern, A. R., & Bartlett, J. C. (2010). Memory for Melodies. In M. Jones, A. Popper, & R. Fay (Eds.), *Music Perception* (Vol. 36, pp. 4473–4482). New York: Springer-Verlag. https://doi.org/10.1007/978-1-4419-6114-3

Halpern, A. R., & Zatorre, R. J. (1999). When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cerebral Cortex*, *9*(7), 697–704. https://doi.org/10.1093/cercor/9.7.697

Harrison, L., & Loui, P. (2014). Thrills, chills, frissons, and skin orgasms: Toward an integrative model of transcendent psychophysiological experiences in music. *Frontiers in Psychology*, *5*(JUL). https://doi.org/10.3389/fpsyg.2014.00790

Heilbron, M., & Chait, M. (2018). Great Expectations: Is there Evidence for Predictive Coding in Auditory Cortex? *Neuroscience*, *389*, 54–73. https://doi.org/10.1016/j.neuroscience.2017.07.061

Hess, E. H., & Polt, J. M. (1964). Pupil Size in Relation to Mental Activity during Simple Problem-Solving. *Science*, *143*(3611), 1190–1192. https://doi.org/10.1126/science.143.3611.1190

Hoeks, B., & Levelt, W. J. M. (1993). Pupillary dilation as a measure of attention: a quantitative system analysis. *Behavior Research Methods, Instruments, & Computers*, *25*(1), 16–26. https://doi.org/10.3758/BF03204445

Janata, P., Tomic, S. T., & Rakowski, S. K. (2007). Characterisation of music-evoked autobiographical memories. *Memory*, *15*(8), 845–860. https://doi.org/10.1080/09658210701734593

Johnson, R. (1995). Event-related potential insights into the neurobiology of memory systems. *Handbook of neuropsychology*, *10*, 135-135.

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. https://doi.org/10.1016/j.neuron.2015.11.028

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science (New York, N.Y.)*, *154*(3756), 1583–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5924930

Luck, S. J., & Kappenman, E. S. (Eds.). (2011). *The Oxford handbook of event-related potential components*. Oxford university press. https://doi.org/10.1093/oxfordhb/9780195374148.001.0001

Klostermann, E. C., Kane, A. J. M., & Shimamura, A. P. (2008). Parietal activation during retrieval of abstract and concrete auditory information. *NeuroImage*, *40*(2), 896–901. https://doi.org/10.1016/j.neuroimage.2007.10.068

Klostermann, E. C., Loui, P., & Shimamura, A. P. (2009). Activation of right parietal cortex during memory retrieval of nonlinguistic auditory stimuli. *Cognitive, Affective & Behavioral Neuroscience*, *9*(3), 242–248. https://doi.org/10.3758/CABN.9.3.242

Koelsch, S. (2018). Investigating the Neural Encoding of Emotion with Music. *Neuron*, *98*(6), 1075–1079. https://doi.org/10.1016/j.neuron.2018.04.029

Kraemer, D. J. M., Macrae, C. N., Green, A. E., & Kelley, W. M. (2005). Musical imagery: sound of silence activates auditory cortex. *Nature*, *434*(7030), 158. https://doi.org/10.1038/434158a

Krumhansl, C. L. (2010). Plink: "Thin Slices" of Music. *Music Perception*, *27*(5), 337–354. https://doi.org/10.1525/mp.2010.27.5.337

Laeng, B., Eidet, L. M., Sulutvedt, U., & Panksepp, J. (2016). Music chills: The eye pupil as a mirror to music's soul. *Consciousness and Cognition*, *44*, 161–178. https://doi.org/10.1016/j.concog.2016.07.009

Lavín, C., San Martín, R., & Rosales Jubal, E. (2014). Pupil dilation signals uncertainty and surprise in a learning gambling task. *Frontiers in Behavioral Neuroscience*, *7*(January), 1–8. https://doi.org/10.3389/fnbeh.2013.00218

Liao, H.-I., Kidani, S., Yoneya, M., Kashino, M., & Furukawa, S. (2016). Correspondences among pupillary dilation response, subjective salience of sounds, and loudness. *Psychonomic Bulletin & Review*, *23*(2), 412–425. https://doi.org/10.3758/s13423-015-0898-0

Liao, H. I., Yoneya, M., Kidani, S., Kashino, M., & Furukawa, S. (2016). Human pupillary dilation response to deviant auditory stimuli: Effects of stimulus properties and voluntary attention. *Frontiers in Neuroscience*, *10*(FEB). https://doi.org/10.3389/fnins.2016.00043

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

Martarelli, C. S., Mayer, B., & Mast, F. W. (2016). Daydreams and trait affect: The role of the listener's state of mind in the emotional response to music. *Consciousness and Cognition*, *46*, 27–35. https://doi.org/10.1016/j.concog.2016.09.014

Mathôt, S. (2018). Pupillometry: Psychology, Physiology, and Function. *Journal of Cognition*. https://doi.org/10.5334/joc.18

Montinari, M. R., Giardina, S., Minelli, P., & Minelli, S. (2018). History of Music Therapy and Its Contemporary Applications in Cardiovascular Diseases. *Southern Medical Journal*, *111*(2), 98–102. https://doi.org/10.14423/SMJ.0000000000000765

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), 1532–1543. https://doi.org/10.1111/j.1469-8986.2011.01226.x

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.-M., Oostenveld, R., Fries, P., Schoffelen, J.-M. (2010). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience, Computational Intelligence and Neuroscience*, *2011*, *2011*, e156869. https://doi.org/10.1155/2011/156869, 10.1155/2011/156869

Partala, T., Jokiniemi, M., & Surakka, V. (2000). Pupillary responses to emotionally provocative stimuli. *Proceedings of the 2000 Symposium on Eye Tracking Research & Applications*, 123–129. https://doi.org/10.1145/355017.355042

Preuschoff, K., 't Hart, B. M., & Einhäuser, W. (2011). Pupil dilation signals surprise: Evidence for noradrenaline's role in decision making. *Frontiers in Neuroscience*, *5*(SEP), 1–12. https://doi.org/10.3389/fnins.2011.00115

Privitera, C. M., Renninger, L. W., Carney, T., Klein, S., & Aguilar, M. (2010). Pupil dilation during visual target detection. *Journal of Vision*, *10*(10), 3. https://doi.org/10.1167/10.10.3

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. https://doi.org/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*. https://doi.org/10.1038/ncomms13289

Rolfs, M., Kliegl, R., & Engbert, R. (2008). Toward a model of microsaccade generation: The case of microsaccadic inhibition. *Journal of Vision*, *8*(11), 5–5. https://doi.org/10.1167/8.11.5

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

Salimpoor, V. N., Van Den Bosch, I., Kovacevic, N., McIntosh, A. R., Dagher, A., & Zatorre, R. J. (2013). Interactions between the nucleus accumbens and auditory cortices predict music reward value. *Science*, *340*(6129), 216–219. https://doi.org/10.1126/science.1231059

Samuels, E., & 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. https://doi.org/10.2174/157015908785777229

Sanquist, T. F., Rohrbaugh, J. W., Syndulko, K., & Lindsley, D. B. (1980). Electrocortical Signs of Levels of Processing: Perceptual Analysis and Recognition Memory. *Psychophysiology*, *17*(6), 568–576. https://doi.org/10.1111/j.1469-8986.1980.tb02299.x

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

Sara, S. J., & Bouret, S. (2012). Orienting and Reorienting: The Locus Coeruleus Mediates Cognition through Arousal. *Neuron*, *76*(1), 130–141. https://doi.org/10.1016/j.neuron.2012.09.011

Schellenberg, E. G., Iverson, P., & McKinnon, M. C. (1999). Name that tune: Identifying popular recordings from brief excerpts. *Psychonomic Bulletin & Review*, *6*(4), 641–646. https://doi.org/10.3758/BF03212973

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

Shulman, G. L., Astafiev, S. V, Franke, D., Pope, D. L. W., Snyder, A. Z., McAvoy, M. P., & Corbetta, M. (2009). Interaction of stimulus-driven reorienting and expectation in ventral and dorsal frontoparietal and basal ganglia-cortical networks. *Journal of Neuroscience, 29*(14), 4392–4407. https://doi.org/10.1523/JNEUROSCI.5609-08.2009

Stelmack, R. M., & Siddle, D. A. T. (1982). Pupillary Dilation as an Index of the Orienting Reflex. *Psychophysiology*, *19*(6), 706–708. https://doi.org/10.1111/j.1469-8986.1982.tb02529.x

Suied, C., Agus, T. R., Thorpe, S. J., Mesgarani, N., & Pressnitzer, D. (2014). Auditory gist: recognition of very short sounds from timbre cues. *The Journal of the Acoustical Society of America*, *135*(3), 1380–91. https://doi.org/10.1121/1.4863659

Trainor, L. J., Marie, C., Bruce, I. C., & Bidelman, G. M. (2014). Explaining the high voice superiority effect in polyphonic music: Evidence from cortical evoked potentials and peripheral auditory models. *Hearing Research*, *308*, 60–70. https://doi.org/10.1016/j.heares.2013.07.014

von Kriegstein, K., Eger, E., Kleinschmidt, A., & Giraud, A. L. (2003). Modulation of neural responses to speech by directing attention to voices or verbal content. *Cognitive Brain Research*, *17*(1), 48–55. https://doi.org/10.1016/S0926-6410(03)00079-X

Wagner, A. D., Shannon, B. J., Kahn, I., & Buckner, R. L. (2005). Parietal lobe contributions to episodic memory retrieval. *Trends in Cognitive Sciences*, *9*(9), 445–453. https://doi.org/10.1016/j.tics.2005.07.001

Wang, C. A., Blohm, G., Huang, J., Boehnke, S. E., & Munoz, D. P. (2017). Multisensory integration in orienting behavior: Pupil size, microsaccades, and saccades. *Biological Psychology*. https://doi.org/10.1016/j.biopsycho.2017.07.024

Wang, C. A., & Munoz, D. P. (2014). Modulation of stimulus contrast on the human pupil orienting response. *European Journal of Neuroscience*, *40*(5), 2822–2832. https://doi.org/10.1111/ejn.12641

Wang, C. A., & Munoz, D. P. (2015). A circuit for pupil orienting responses: Implications for cognitive modulation of pupil size. *Current Opinion in Neurobiology*, *33*, 134–140. https://doi.org/10.1016/j.conb.2015.03.018

Widmann, A., Schröger, E., & Wetzel, N. (2018). 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. *Biological Psychology*, *133*, 10–17. https://doi.org/10.1016/j.biopsycho.2018.01.010

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

Zäske, R., Awwad Shiekh Hasan, B., & Belin, P. (2017). It doesn't matter what you say: FMRI correlates of voice learning and recognition independent of speech content. *Cortex*, *94*, 100–112. https://doi.org/10.1016/j.cortex.2017.06.005

Zäske, R., Volberg, G., Kovács, G., & Schweinberger, S. R. (2014). Electrophysiological correlates of voice learning and recognition. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *34*(33), 10821–31. https://doi.org/10.1523/JNEUROSCI.0581-14.2014

Zatorre, R. J., & Halpern, A. R. (2015). Music. *Brain Mapping*, 551–557. https://doi.org/10.1016/B978-0-12-397025-1.00283-9

Zheng, Y., & Escabi, M. A. (2008). Distinct Roles for Onset and Sustained Activity in the Neuronal Code for Temporal Periodicity and Acoustic Envelope Shape. *Journal of Neuroscience*, *28*(52), 14230–14244. https://doi.org/10.1523/JNEUROSCI.2882-08.2008

Zheng, Y., & Escabi, M. A. (2013). Proportional spike-timing precision and firing reliability underlie efficient temporal processing of periodicity and envelope shape cues. *Journal of Neurophysiology*, *110*(3), 587–606. https://doi.org/10.1152/jn.01080.2010

6. Figures

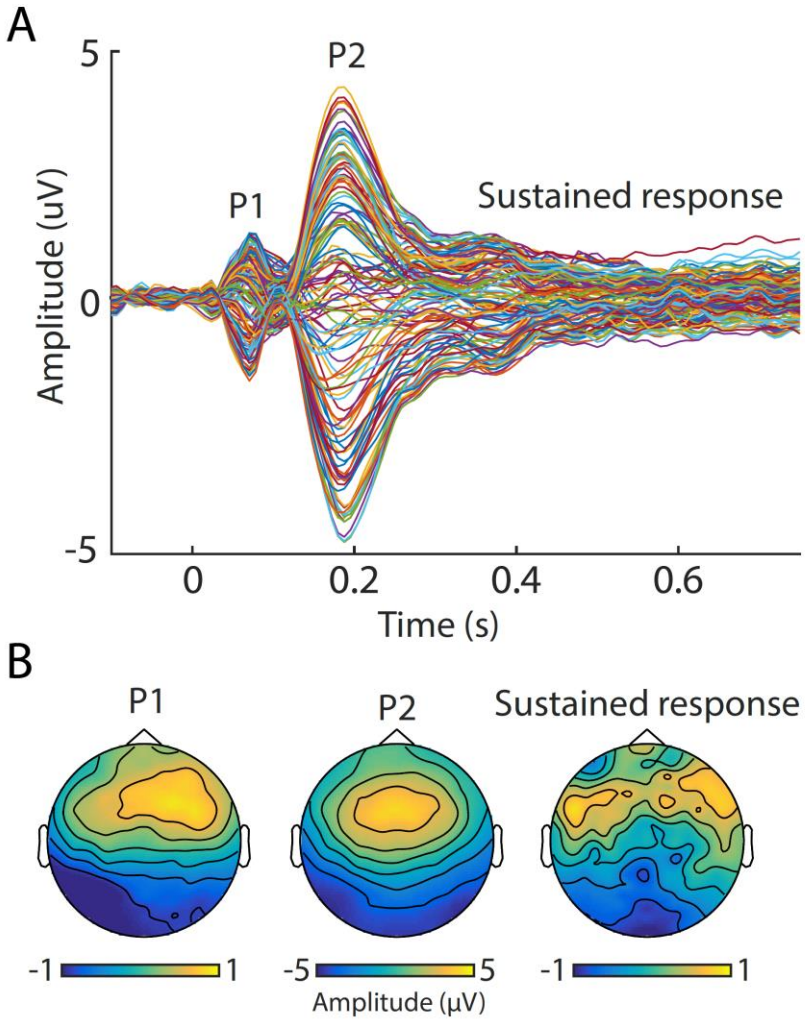

Figure 1: Grand-average event-related potential, averaged across the main and control group, as well as familiar and unfamiliar songs, demonstrating the brain response to the music snippets. (A)Time-domain representation. Each line represents one EEG channel. (B) Topographical representation of the P1(71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

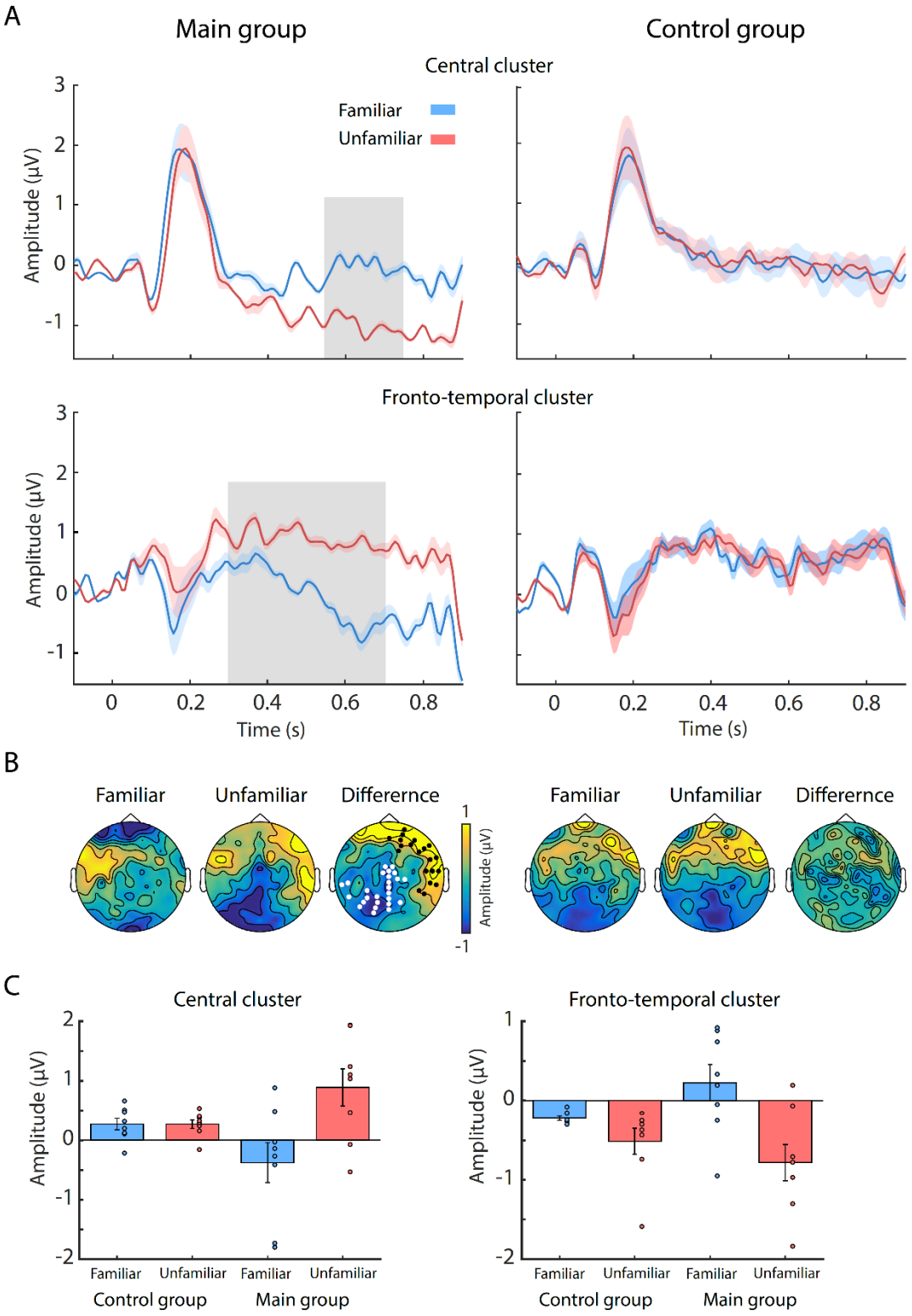
A
Main group
Control group
Central cluster
Familiar
Unfamiliar
Amplitude (µV)
Fronto-temporal cluster
Time (s)
B
Familiar
Unfamiliar
Differernce
Amplitude (µV)
C
Central cluster
Fronto-temporal cluster
Familiar
Unfamiliar
Control group
Main group

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 central cluster (top row) and fronto-central cluster (bottom row), separately for the main (left column) and control (right column) group. Solid lines represent mean data 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). Shaded areas represent standard error of the mean. Significant differences between conditions, as obtained via cluster-based permutation tests, are indicated by grey boxes. (B) Topographical maps of the familiar and unfamiliar ERP responses as well as their difference (from 350 to 750 ms), separately for the main (left column) and control (right column) group. White and black dots indicate electrodes belonging to the central and fronto-temporal cluster, respectively. (C) Bar plots of mean ERP amplitudes, as entered into the ANOVA. Main group, but not control group, showed significantly larger responses to unfamiliar song snippets, at both the central and the fronto-temporal cluster. Error bars represent standard error of the mean, coloured dots represent individual subjects' mean.

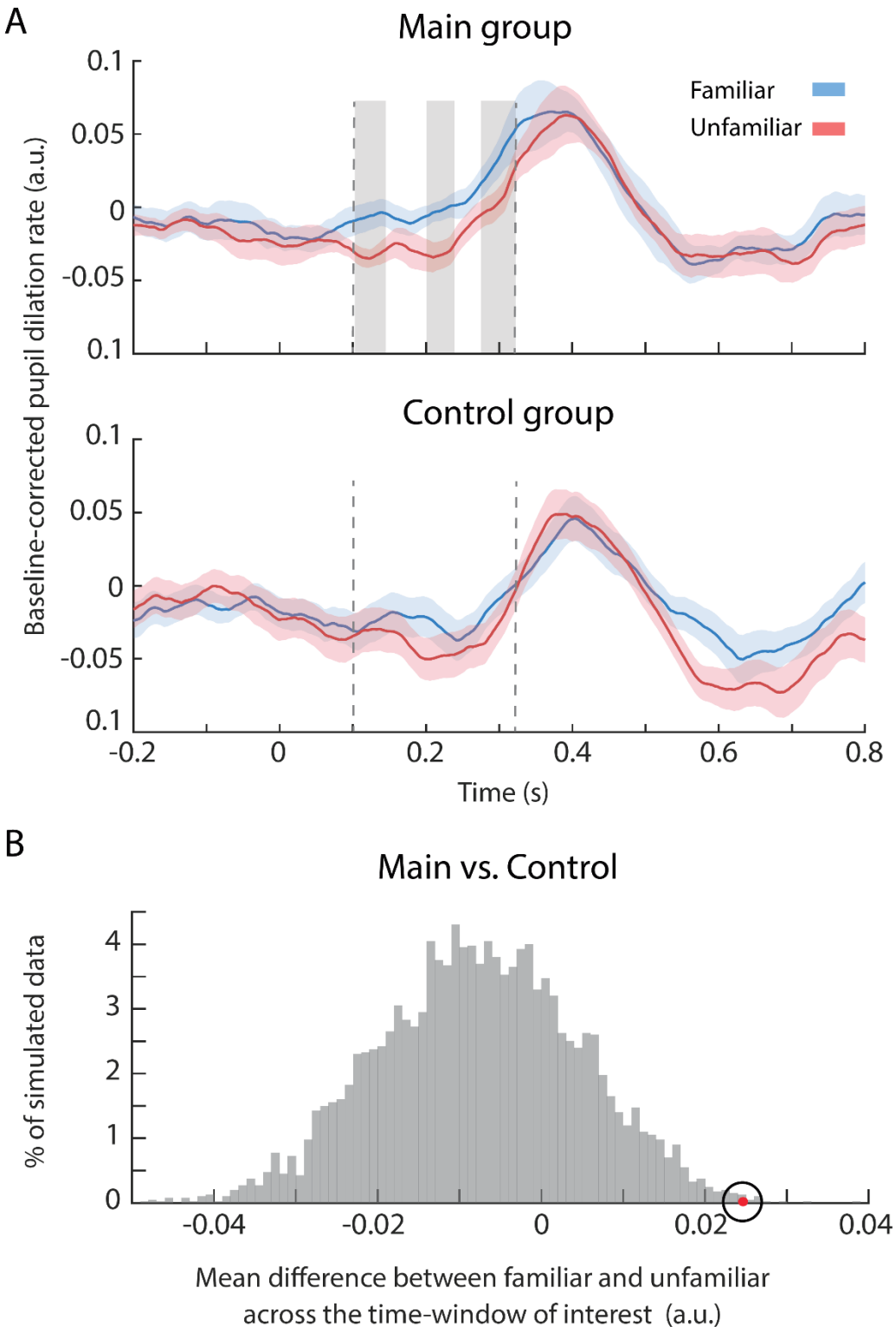
A
Main group
Familiar
Unfamiliar
Control group
Baseline-corrected pupil dilation rate (a.u.)
0.1
0.05
0
-0.05
0.1
-0.2
0
0.2
0.4
0.6
0.8
Time (s)
B
Main vs. Control
% of simulated data
4
3
2
1
0
-0.04
-0.02
0
0.02
0.04
Mean difference between familiar and unfamiliar
across the time-window of interest (a.u.)

Figure 3: Pupil dilation rate to familiar and unfamiliar snippets. (A)Top: Main group. The solid curves plot the average pupil dilation rate across subjects for familiar (blue) and unfamiliar (red) conditions. The shaded area represents one standard deviation of the bootstrap resampling. The grey boxes indicate time intervals where the two conditions were significantly different (114–137, 217–230, and 287–314ms). The dashed lines indicate the time interval use for the resampling statistic in B. Bottom: Control group. The solid curves plot the average PD rate over 10 randomly selected control datasets. The shaded area represents one standard deviation of the bootstrap resampling. No significant differences were observed (B) Results of the resampling test to compare the difference between familiar and unfamiliar conditions (during the time interval 108-315 ms, indicated via dashed lines in A) across the main and control groups. The grey histogram shows the distribution of differences between conditions for the control group. The red dot indicates the observed difference for the main group.