Network connectivity differences between perceptual and emotional music listening

Author List: Faber, S.[1,2], & McIntosh, A.R.[2]

[1]University of Toronto

[2]Institute for Neuroscience and Neurotechnology, Simon Fraser University

Abstract

Music is an intricate stimulus that engages numerous brain networks at various temporal scales. Past work has successfully correlated brain regions and networks of brain regions with features in musical signals and behavioural ratings. Current dynamic systems frameworks allow for the incorporation of multiple types of high-dimensional stimuli into the same model. In the present study, we used networks of music features and continuous behavioural ratings to examine patterns of brain network connectivity captured via EEG in two music listening conditions. Participants listened to music while continuously rating either perceptual features in the musical signal or emotional states induced by the music. Networks of brain regions and networks of music features were estimated using hidden Markov modelling and were analyzed together with continuous behavioural ratings using partial least squares analysis. We found task-specific differences in between-network brain activity. In the perceptual task, network activity in medial frontoparietal, pericentral, anterior frontal, and bilateral temporal networks was most strongly positively correlated with acoustic features in the musical signal. In the emotional task, transitions between networks were more positively correlated with low-arousal low-valence emotional states in anterior cingulate, frontal, and occipital networks. We conclude that attention to specific aspects of a musical experience employs different network connectivity patterns in music listening and discuss the implications of these methods for future work on feature-rich naturalistic behaviours.

Keywords: Computational modelling, EEG, Music perception

Introduction

Music listening is a sophisticated behaviour that combines perceptual, cognitive, and emotional information into a singular experience. Parcelling this experience into its constituent parts is one way of trying to disambiguate and quantify these unique systems of perception and evocation. Perception includes parsing timbral, syntactic, and semantic information from the music, and evocation includes affective, reward, and memory responses to the music.

Perception can be evaluated by correlating brain region activity with features extracted from music over time, with prior work showing specific brain regions corresponding to specific

musical features (Alluri et al., 2012). Evocation is less straightforward and depends on subjective musical preference, familiarity, and memory. While reward networks of brain states have been shown to emerge during music listening (Fasano et al., 2020; Gold et al., 2019; Salimpoor et al., 2011), how a listener moves from acoustic perception to specific emotional states, and whether this process changes with different attentional foci, remains ambiguous.

Work by Carpentier and colleagues (2019) showed that how an individual listens to music affects brain signal complexity. Perceptual listening corresponds to relatively lower brain signal complexity, and emotional music listening (listening while attending to inner emotional states) corresponds to relatively higher brain signal complexity as measured using multi-scale entropy (Costa et al., 2002). The authors conclude that the interoceptive monitoring required during emotional listening adds another layer of brain activity on top of parsing the features provided by the stimulus.

In the present study, we used the dataset from Carpentier et al. (2019) to further explore the time-varying nature of music listening during perceptual and emotional music listening. We used music information retrieval (MIR; Downie, 2003) to extract acoustic and structural features from musical excerpts comprising high- and low-level information. we then employed hidden Markov modelling (HMM, Vidaurre et al., 2017) to extract a multi-network time series from source-localized EEG and from the music feature time series, and used partial least squares (PLS, McIntosh & Lobaugh, 2004) analysis to compare temporal network characteristics between music listening and rest directly, and between different music listening tasks and behavioural reports.

Methods

Participants

Eighteen healthy right-handed adults (age range 19-35, mean age 26, 8 males) with normal hearing as established via audiogram were recruited for a music listening task. All participants provided written informed consent as per the Baycrest Centre-University of Toronto Research Ethics Committee.

Stimuli and apparatus

Stimuli were 40 excerpts of Western art music divided into control (perceptual) and experimental (emotional) tasks. Excerpts were between 40 seconds and two minutes in length and were from the Renaissance to contemporary eras (Table 1). Excerpts were orchestrated in a variety of styles including solo instrumental, orchestral, chamber, choral, and operatic works. Vocal excerpts were in English, Italian, French, and German.

Emotional Songs	
• Adams "Nixon in China, 'Beginning'" • Adams "Disappointment Lake" • Bach "No. 3 Aria 'Es Ist Vollbracht'" • Barber "Adagio for Strings" • Brahms "Intermezzo No. 2 in a Major, Op. 118" • Delibes "Lakmé/Flower Duet" • Elgar "Variation IX (Adagio) 'Nimrod'" • Galvany "Oh My Son" • Gluck "Armide Act Iv Air Sicilien" • Goodall "Belief" • Ives "Three Places in New England Orchestral Set No.1" • Liszt "Totentanz" • Monteverdi "Zefiro Torna" • Mozart "Cosi fan tutte" • Mozetich "The Passion of Angels" • Part "Spiegel im Spiegel" • Penderecki "Threnody to the Victims of Hiroshima" • Puccini "O soave fanciulla"	• Rameau "Entrée de Polymnie" • Richter "Vivaldi's Summer" • Rossini "Barbiere di Siviglia: Largo Al Factotum" • Schroer "Field of Stars" • Schumann-Liszt "Liebeslied (Widmung)" • Staniland "Solstice Songs No. 2 Interlude" • Stravinsky "Glorification of the Chosen One" • Tarrega "Recuerdos De La Alhambra" • Verdi "Messa Da Requiem: Dies Irae-Tuba Mirum Part 1" • Verdi "Messa Da Requiem: Dies Irae-Tuba Mirum Part 2" • Wagner "Die Walkurie, Act 3: Ride of the Valkyries" • Wagner "Tristan Und Isolde/Act 2–Prelude"
Perceptual Songs	
• Beethoven "Sonata in A Major Op. 69" • Brahms "Violin Concerto in D, Op. 77-3" • Glass "Glassworks Opening • Haydn "Cello Concerto in D Major" • Mozart "Symphony No. 40 in G-minor, K.550, Finale"	• Praetorious "Praeambulum" • Strauss "Der Rosenkavalier Act III/Duet-Denouement and Grand Waltz–Coda" • Strauss "September" • Vivaldi "Concerto for Violin, Stings and Harpsichord in G" • Vivaldi "Stabat Mater

Table 1: *List of stimuli for each task. Table adapted from "Complexity matching: brain signals mirror environment information patterns during music listening and reward" by Carpentier et al., 2020, Journal of cognitive neuroscience, 32 (4), p. 736.*

During music listening, participants completed a continuous ratings task. A two-dimensional cross was projected onto a computer screen and participants were instructed to track a cursor around the cross. In the control task, participants continuously tracked the perceived pitch and tempo of each excerpt with the axes of the cross corresponding to pitch (low-high) and tempo (low-high). In the experimental task, participants continuously reported their own inner emotional states with the axes of the cross corresponding to valence (unpleasant-pleasant) and arousal (relaxing-stimulating). The cursor reset to the centre of the cross at the beginning of each excerpt, and the ratings data were divided into five quadrants (Figure 1).

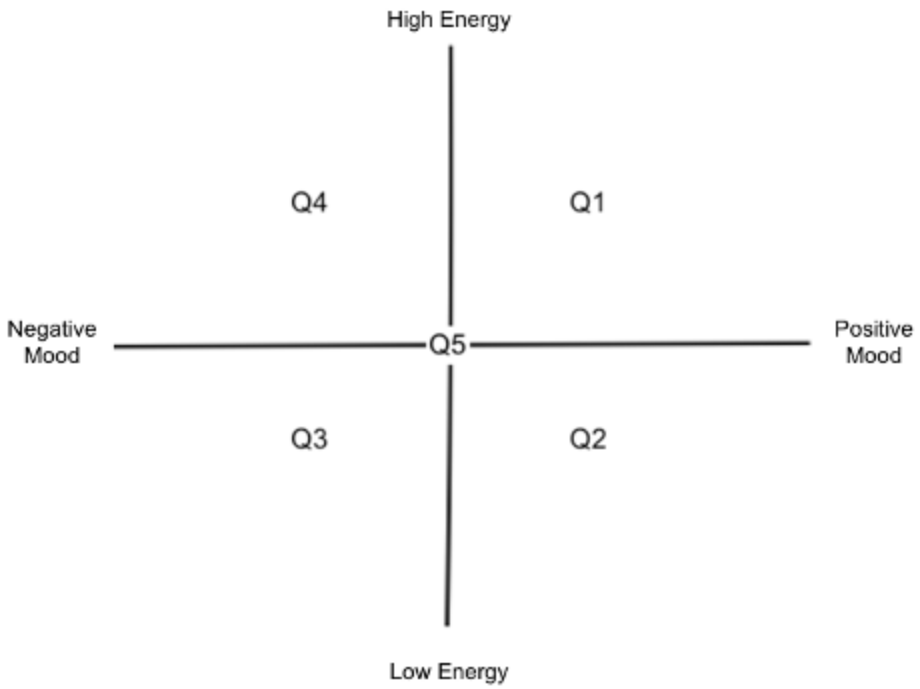

Figure 1: *Mouse tracker cross with quadrant labels 1-5. The fifth "quadrant" is the centre starting position (x- and y-coordinates = 0,0)*

Procedure

Participants were provided orientation to the behavioural task apparatus and completed six practice trials. Participants completed five minutes of pre-task, eyes open resting state, then began the music listening tasks. Five control pieces were presented at the beginning of the task, followed by 30 experimental pieces. Five additional control pieces were presented at the end followed by five minutes of eyes open post-task resting state. Pieces were presented in one of three counterbalance orders: in the first, stimuli were ordered to "flow" from one piece to the next without sudden transitions. The second order was the reverse of the first, and the third order was random. In preliminary analyses, counterbalance order did not have an effect on EEG or behavioural results and the data were pooled for the remainder of this analysis.

Data acquisition and pre-processing

Continuous EEG and behaviour data were collected simultaneously. EEG data were collected using a 64 + 10 channel Biosemi Active Two System with a sampling rate of 512 Hz. Data were imported to Brainstorm for pre-processing (Tadel et al., 2011). Raw EEG signals were bandpass filtered between 0.5 and 55 Hz, and re-referenced to the average of all electrodes. We used ICA for artifact detection, and manually selected artifactual components for each participant. Data were epoched to 40 seconds (the length of the shortest excerpt) by task and each epoch was visually inspected for noise. Data were source-localized using sLORETA and projected into

regions of interest using the Desikan-Killiany atlas (Desikan et al., 2006; see Supplementary materials for a list of regions), resulting in a 68-region*time matrix for each task for each participant. Resting state data were likewise divided into 40-second continuous epochs, inspected for noise, and source localized. Each clean resting state trial pre- and post-experiment was exported as a 68-region*time matrix for each participant.

We imported the source-localized time series data to Matlab R2016b, and completed hidden Markov modelling analysis (HMM) using the HMM-MAR toolbox (Vidaurre et al., 2017). HMM is a way of reducing dimensionality in space while preserving time series information. The user selects a value for K corresponding to the desired number of brain states, and the function calculates the K functional connectivity states that best describe the data. Each time point is allocated to one of the states, and the probability of transitioning to and from each state is also calculated. The model's output provides several metrics that describe the temporal and spatial properties of a time series.

We implemented HMM in a principal components analysis-reduced space preserving 90% of the variance of the raw data, and resampled the data to 256 Hz to maximize computational efficiency. To find an optimal value for K, we ran multiple HMM analyses with values for K between 5 and 20. The estimation with a K of 12 returned the lowest amount of free energy as calculated by the model (Vidaurre et al., 2017), and we proceeded with the output from this estimation for further analysis. We extracted the following metrics for each task block for each participant:

State Visits: how many times a state is visited during each task
State Lifetimes: how long, on average, a visit to a state lasts during a task
Fractional occupancy: the amount of time a participant spends in a state in total during each task
Task-wise transitional probability matrices: the probability of transitions between states

We generated task-wise probability matrices for each participant, and averaged them together by condition (rest, control task music listening, experiment task music listening). The resulting matrices were vectorized and used as input for partial least squares (PLS) analysis.

PLS is a multivariate analysis technique that describes the relationship(s) between matrices of measurements using singular value decomposition (SVD). Two common varieties are mean-centred PLS, which examines distributed activity differences between tasks, and PLS correlation, which examines the multivariate correlations between two input matrices, often brain and behaviour. Both varieties return latent variables (LVs) that explain the relationships between matrices, accounting for progressively smaller amounts of covariance between matrices. Each LV includes a singular vector of weights for items within each input matrix, and a singular value, which is the covariance between the matrices (McIntosh et al., 1996). Statistical significance is

established via permutation testing, and bootstrap estimation is used to determine the reliability of each LV (McIntosh & Lobaugh, 2004; Efron & Tibshirani, 1996).

The x- and y-position data from the mouse tracker were imported to Matlab R2016b and each timepoint was assigned to one of five mouse tracker states: quadrants one through four, and the starting neutral position. This allowed us to calculate state fractional occupancy, lifetimes, and visits consistent with the brain HMM metrics.

To better link together brain and behavioural data, we extracted features from the music data (Table 2) using the Music Information Retrieval Toolbox (MIRToolbox; Lartillot & Toiviainen, 2007). These features were extracted for each of the 40 pieces at 10 Hz and, using a K of 5, we estimated HMM from each excerpt. We calculated the fractional occupancy for each of the five states for each excerpt, and used these in PLS correlation analyses. Due to the low sampling rate, we did not calculate state visits and lifetimes from the MIR HMM state path.

Feature	*Description*
Zero Crossing	*Noisiness: how many times the signal crosses the 0 amplitude line*
Roughness	*Sensory dissonance from beating phenomenon*
Amplitude (root mean squared)	*Amplitude, a measure of loudness/energy*
Spectral Flux	*Distance between successive spectral frames*
Sub-band Flux 1-10	*Same as above, decomposed into 10 sub-bands (ascending Hz)*
Pulse Clarity	*Estimate of beat regularity*
Mode	*Estimate of major or minor modality*
Key Clarity	*Strength associated with the most likely key(s)*
Novelty	*Probability of transitions between successive frames*

Table 2: *MIR features*

Network estimation: Brain

The mean of each brain state extracted from HMM were calculated using the getMean function from the HMM-MAR toolbox (Vidaurre et al., 2017). State means were thresholded at the absolute 70th percentile and plotted onto a cortical surface using the ggseg package in R Studio (Mowinckel & Vidal-Piñeiro, 2020). The plots are displayed in Figure 2 and regions of interest and network labels are outlined in Table 3.

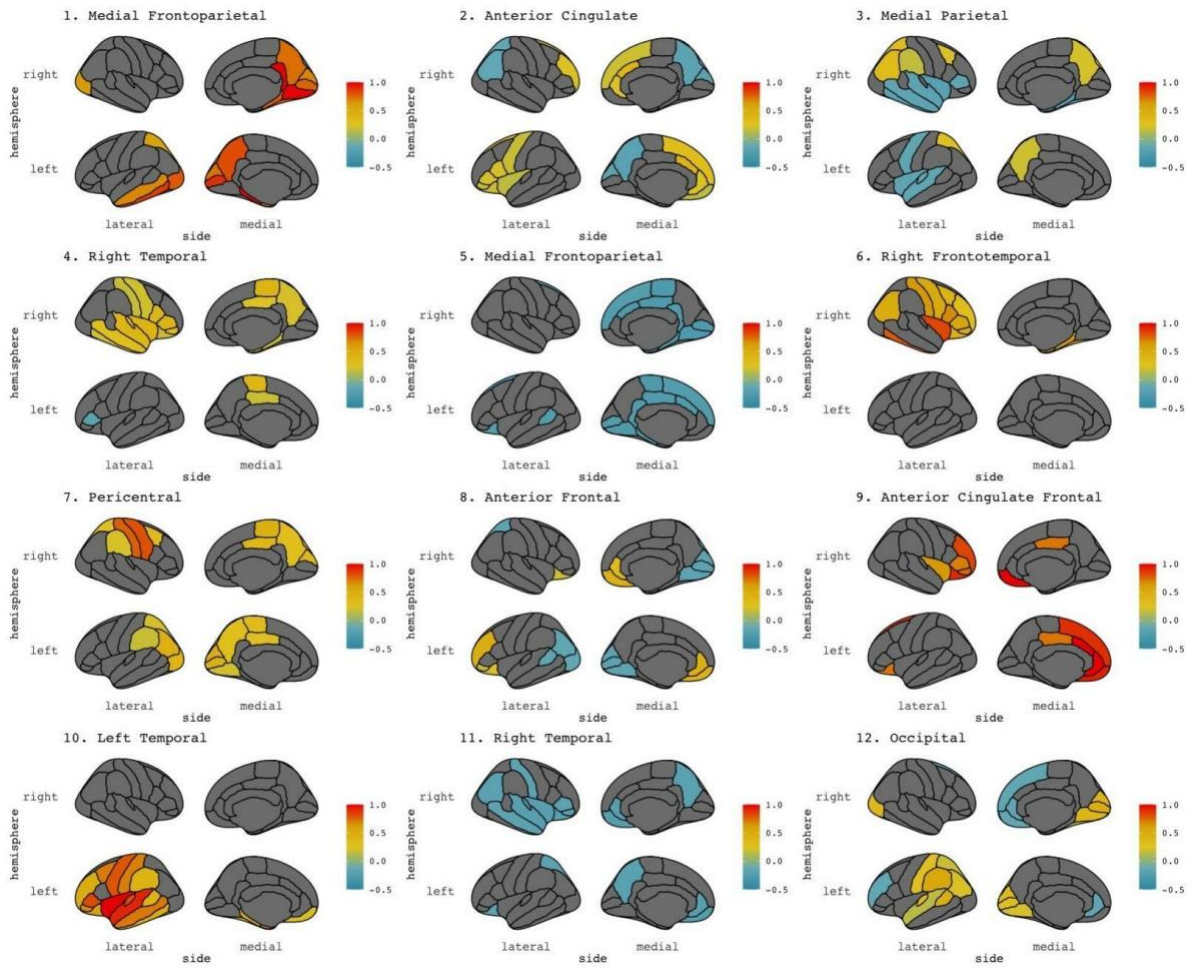

Figure 2: *Mean state activity maps projected on cortical surfaces. Colours represent relative activity of the states and all have been normalized within-state.*

State	*Main Regions*	*Network*
1	Activity in bilateral occipital cortices, bilateral precuneus, bilateral parahippocampal gyri, and left medial temporal lobe	Medial frontoparietal network
2	Activity in bilateral anterior cingulate cortex regions	Anterior cingulate network

3	Activity in bilateral precuneus and superior parietal, right precentral; lower relative activity in left precentral, bilateral temporal regions	Medial parietal network
4	Activity in left posterior cingulate and paracentral, right frontal, temporal, and parietal regions	Right temporal network
5	Activity in bilateral paracentral, cingulate, pericalcarine, and lingual regions	Medial frontoparietal network
6	Activity in right frontal, temporal, and parietal regions	Right frontotemporal network
7	Activity in bilateral parieto-occipital regions	Peri-central network
8	Activity in bilateral frontal regions, lower relative activity in bilateral parieto-occipital regions	Anterior frontal network
9	Activity in bilateral orbitofrontal regions, right inferior frontal gyrus regions	Anterior cingulate frontal network
10	Activity in left frontal, temporal, and parietal regions	Left temporal network
11	Activity in bilateral frontal, temporal, and insular regions	Right temporal network
12	Activity in left bank of superior temporal sulcus, left-side supramarginal gyrus, and occipital regions	Occipital network

Table 3: *Brain state regions of interest and network labels. Network labels are based on the work of Uddin et al. (2019).*

Network Estimation: Music

The MIR feature HMM returned five states, which we interpret as follows: State 1 corresponds to the acoustic texture of the sound, State 2 to prevalence in the mid-high frequency range, State 3 to prevalence in the low frequency range, State 4 to suddenness, and State 5 to predictability (Figure 3).

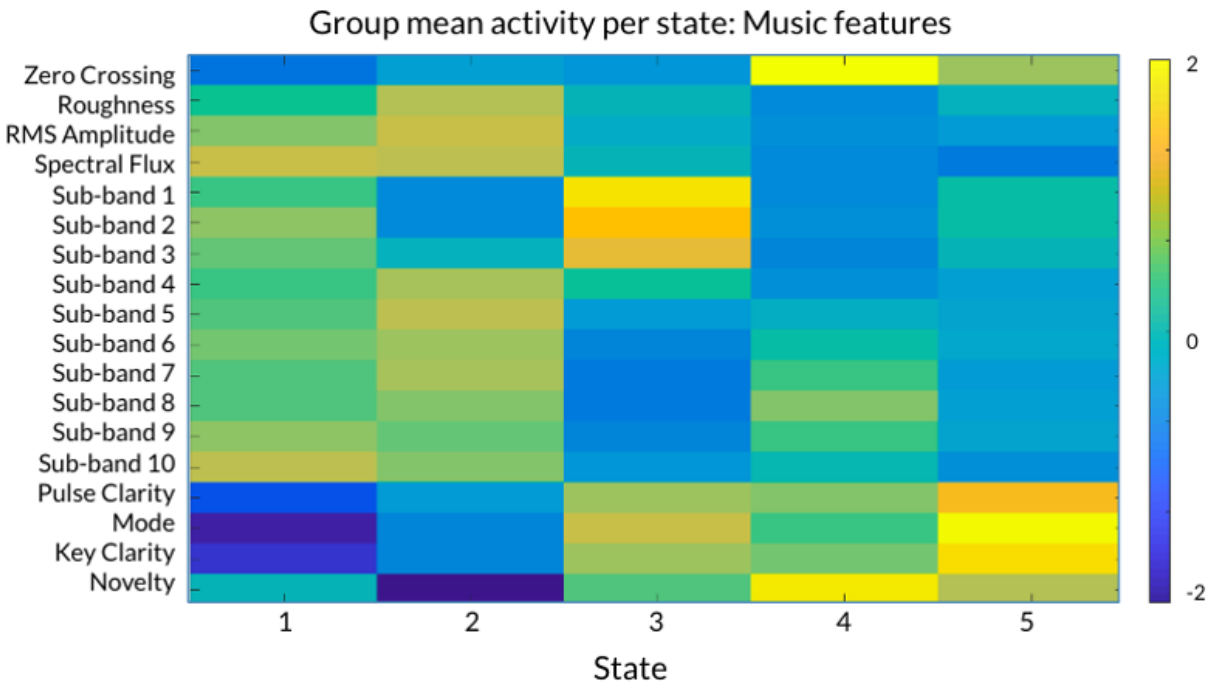

Figure 3: *MIR feature states*

States 1 through 3 describe acoustic properties of the sound. The musical information in these states is conveyed at rapid timescales and not dependent on prior information, making states 1-3 context-free. States 4 and 5 describe context-dependent structural properties of the sound. The musical information in these states is conveyed more gradually and is contingent on prior information. For clarity, we have grouped these states into two categories: *context-free features and context-dependent features*.

We extracted state fractional occupancy from the music HMM, but not lifetimes or visits due to the low sampling rate of the music data.

Results

Brain and behaviour: between-task analyses

We extracted state fractional occupancy, visits, and lifetimes from the EEG HMM estimation and analyzed the data for task differences between rest, control, and experimental conditions using mean-centred PLS. One significant LV was returned from the fractional occupancy and lifetime analyses, showing higher fractional occupancy in brain networks 3, 4, and 8 during rest and higher fractional occupancy in network 5 during music listening in the experimental condition (p = .034). Higher lifetimes were observed in networks 1 and 5 during music listening in the experimental condition (p = .026). No significant LVs were returned from the visit analysis. Follow-up analyses were performed between the music listening tasks, returning no significant results (p = 0.14).

We next extracted rest- and task-wise transitional probability matrices and completed a mean-

centred PLS analysis. No significant LVs were returned (p = 0.33). We ran a second mean-centred PLS on the transitional probability matrices between the two music listening conditions, which returned no significant LVs (p = 0.97).

Brain and behaviour: within-task analyses

We next modelled the brain visit, lifetime, and fractional occupancy data together with the mouse tracker data for each music listening condition using behavioural PLS. No significant effects were returned at the individual task level (p = 0.10 for the control task, p = 0.38 for the experiment task). Combined behavioural PLS analyses were not performed as the mouse tracker quadrants correspond to features that are unique to each task and cannot be equated.

We completed behavioural PLS between transitional probability data and mouse tracker values within-task, which returned one significant LV (p = 0.01) showing an interaction between lifetimes in the low valence/low arousal quadrant and likelihood of transitioning to other networks in the experiment condition (Figure 4).

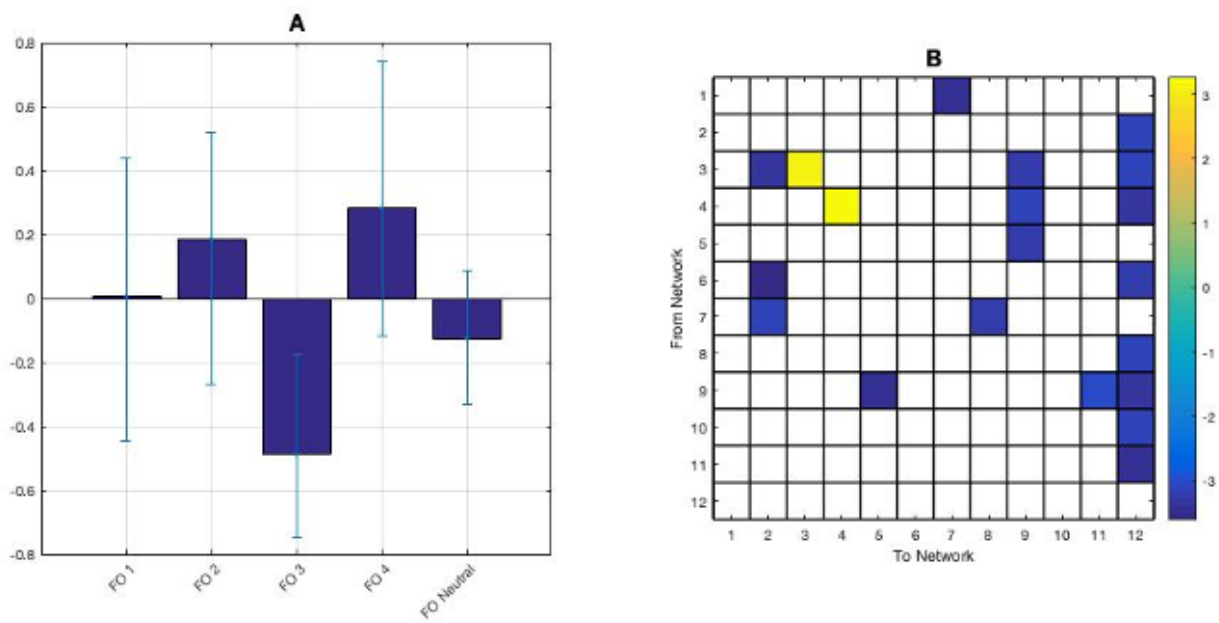

Figure 4: *(A) PLS contrasts between lifetimes in mouse tracker quadrants 1-5. Error bars were calculated using bootstrap resampling and reflect the 95% confidence interval. The contrasts show an effect of behaviour on transitional probabilities (B) with a higher likelihood of transitioning to other networks and a lower likelihood of staying in networks 3 and 4 while reporting a low arousal/low valence emotional state.*

Brain and stimulus: between-task analyses

We introduced stimulus data into the analysis and correlated the EEG visits, lifetimes, and fractional occupancy with the fractional occupancy from the music HMM estimation. We ran

mean-centred PLS on these correlation matrices, which returned one significant LV ($p < .01$) showing a task effect. In the control task, brain network visits and lifetimes were more positively correlated with MIR state fractional occupancy, particularly in brain networks 4, 5, 7, 8, and 10. In the experiment task, MIR fractional occupancy was more positively correlated with brain fractional occupancy, particularly in brain networks 3, 8, and 9 (Figure 5).

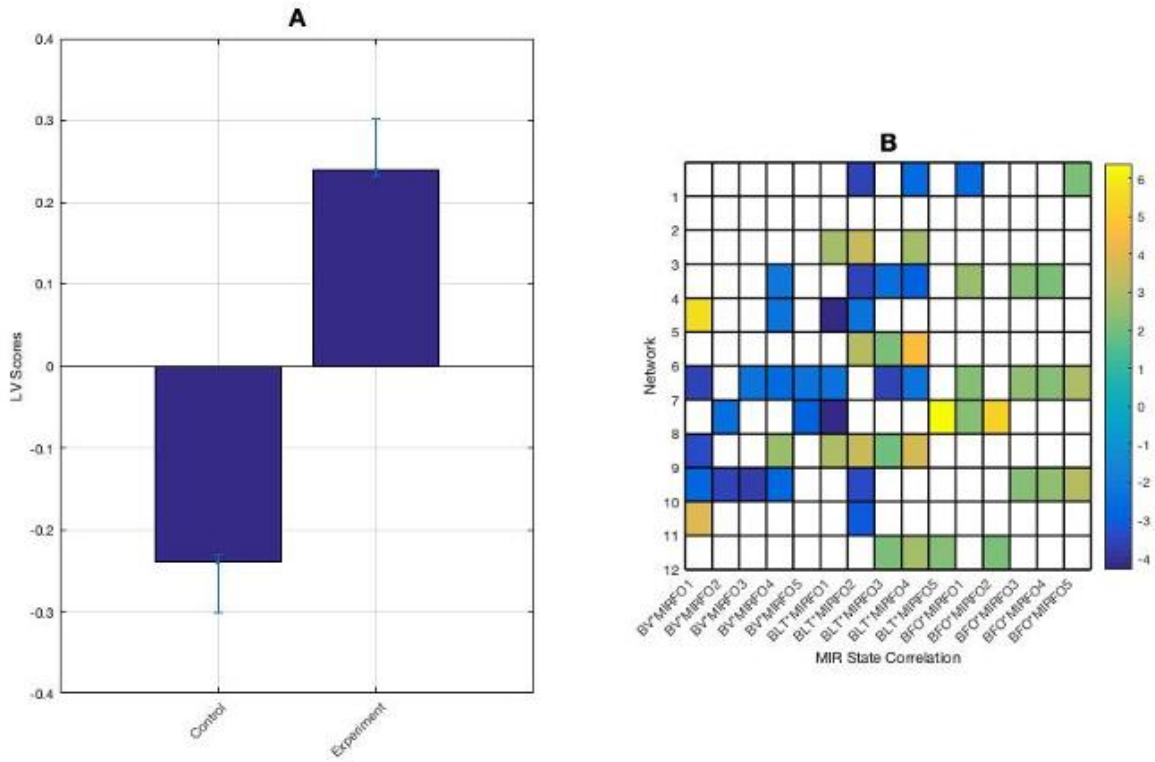

Figure 5: *(A) PLS contrasts between tasks and correlations between brain and stimulus measures. Error bars were calculated using bootstrap resampling and reflect the 95% confidence interval. The contrasts show an effect of task on correlations between brain and stimulus measures (B).*

We next correlated the MIR state fractional occupancy data with the brain transitional probability matrices. To do this, we vectorized the task-averaged transitional probability matrices and correlated them with the task-averaged MIR FO. An initial between-task mean-centered PLS returned no significant LVs ($p = 0.89$). We subsequently correlated the piece-wise transitional probability matrices with the piece-wise MIR fractional occupancy and again found no significant differences between the tasks via mean-centred PLS ($p = 0.78$).

We then averaged together the correlation matrices by MIR fractional occupancy circuit (context free, and context-dependent) and ran a mean-centred PLS. This analysis returned one significant LV ($p = .001$) showing the contrast between transitional probabilities correlated with context-free and context-dependent MIR states in the control task only. Context-free MIR states were positively correlated with transitions involving brain networks 3, 5, 6, 7, 9, and 12; while

context-dependent MIR states were positively correlated with transitions involving brain networks 2, 5, 10, and 11 (Figure 6).

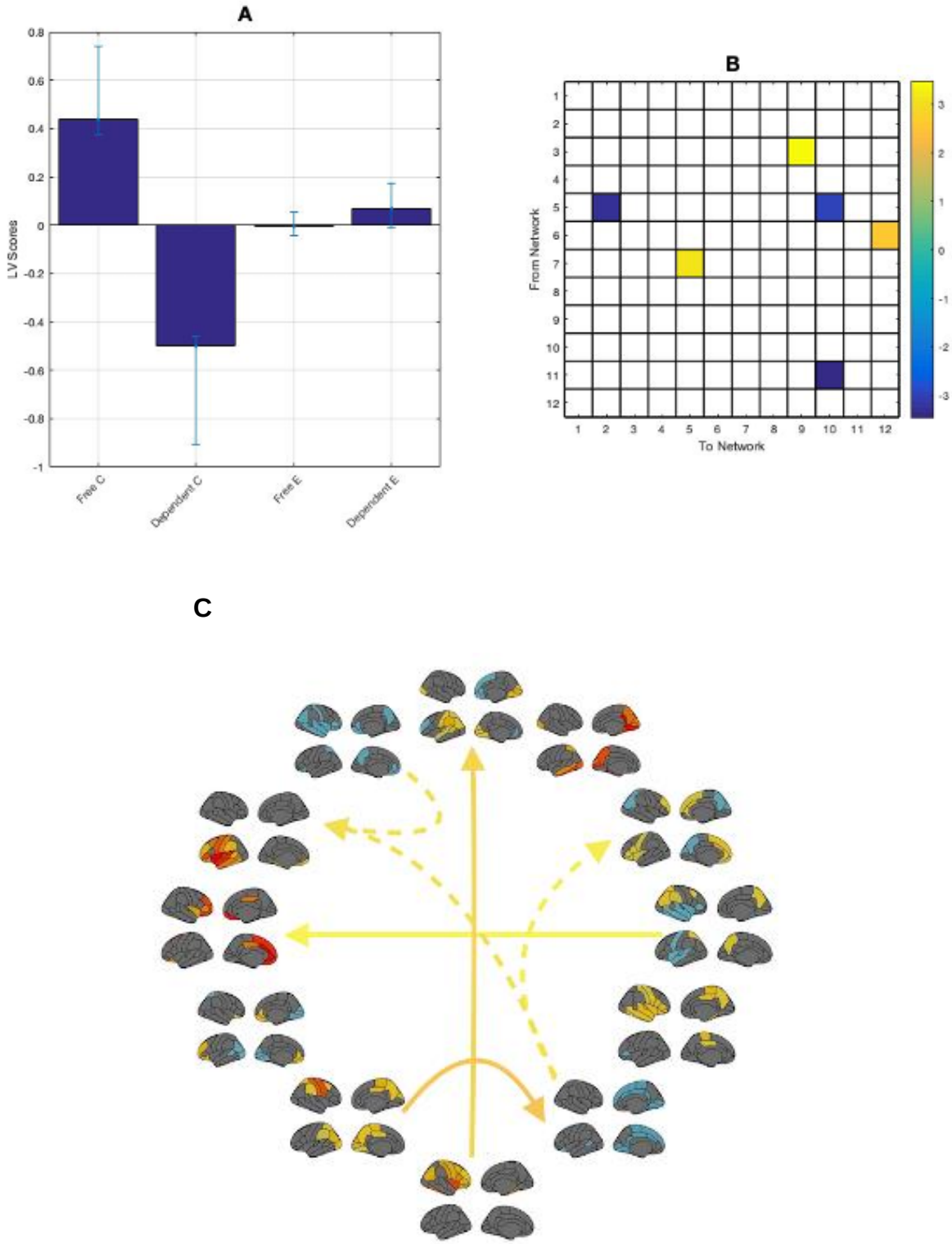

Figure 6: *(A) PLS contrasts between tasks and correlations between brain measures averaged*

by music feature categories and transitional probabilities. Error bars were calculated using bootstrap resampling and reflect the 95% confidence interval. The contrasts show an effect of task and feature category in correlations between transitional probabilities and MIR features (B). Panel C shows the directed transitional probability correlations with the r-value magnitude. Solid lines represent correlations with context-free MIR features, dashed lines represent correlations with context-dependent MIR features.

Behaviour and stimulus

To study the interactions between the stimulus and behaviour data, we correlated the mouse tracker data with the MIR state fractional occupancy data and ran a between-task mean-centred PLS. The analysis returned one significant LV (p <.01) showing a task effect in the correlations between behavioural ratings and MIR state features. In the control task, correlations were more strongly positive between fractional occupancy in MIR state 5 and fractional occupancy and lifetimes in mouse quadrant 1; between fractional occupancy in MIR states 1 and 3 and fractional occupancy and lifetimes in quadrant 2, and between fractional occupancy in MIR state 3 and fractional occupancy and lifetimes in quadrant 3. Control task correlations were more strongly negative compared to the experiment task between fractional occupancy in MIR states 1 and 3 and fractional occupancy and lifetimes in quadrant 1, fractional occupancy in MIR state 5 and fractional occupancy and lifetimes in quadrant 2, and fractional occupancy in MIR state 3 and fractional occupancy and lifetimes in quadrant 4 (Figure 7).

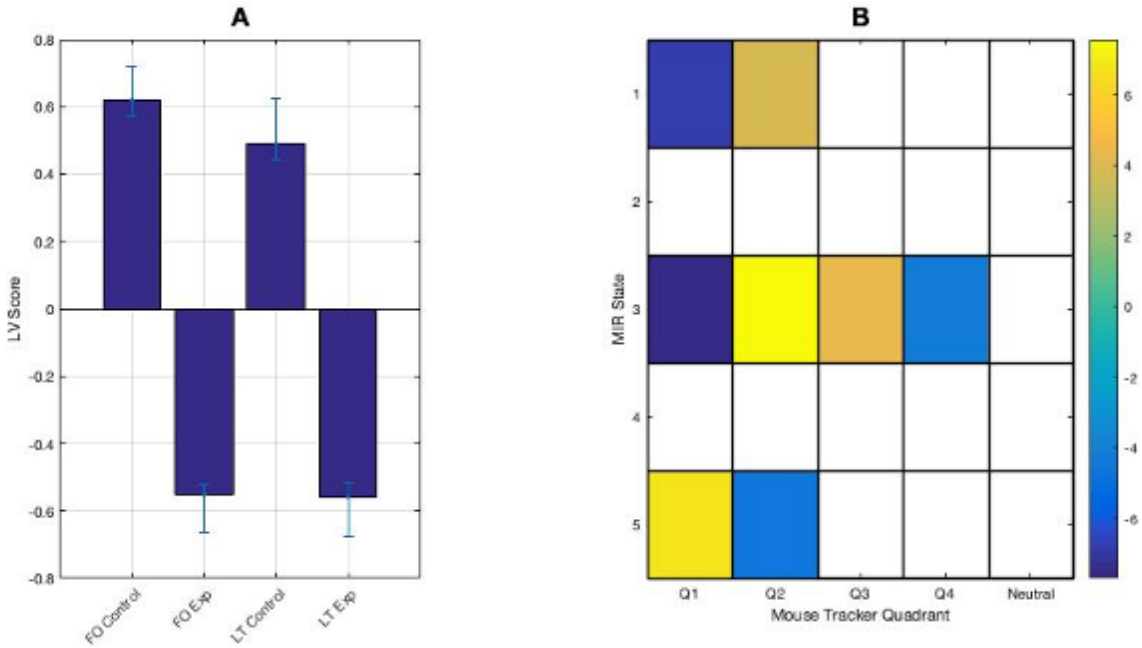

Figure 7: *(A) PLS contrasts between tasks and correlations between behaviour and stimulus measures. Error bars were calculated using bootstrap resampling and reflect the 95% confidence interval. The contrasts show an effect of task on correlations between behaviour and stimulus measures (B).*

Follow-up mean-centred PLS analyses on averaged correlation data revealed task differences in one significant LV ($p < 0.001$). Correlations were more strongly positive between fractional occupancy in MIR state 2 and mouse tracker fractional occupancy and lifetimes in the control task. Correlations were more strongly negative between fractional occupancy in MIR state 1 and fractional occupancy and lifetimes in the control task, and more strongly positive between fractional occupancy in MIR state 1 and fractional occupancy and lifetimes in the experiment

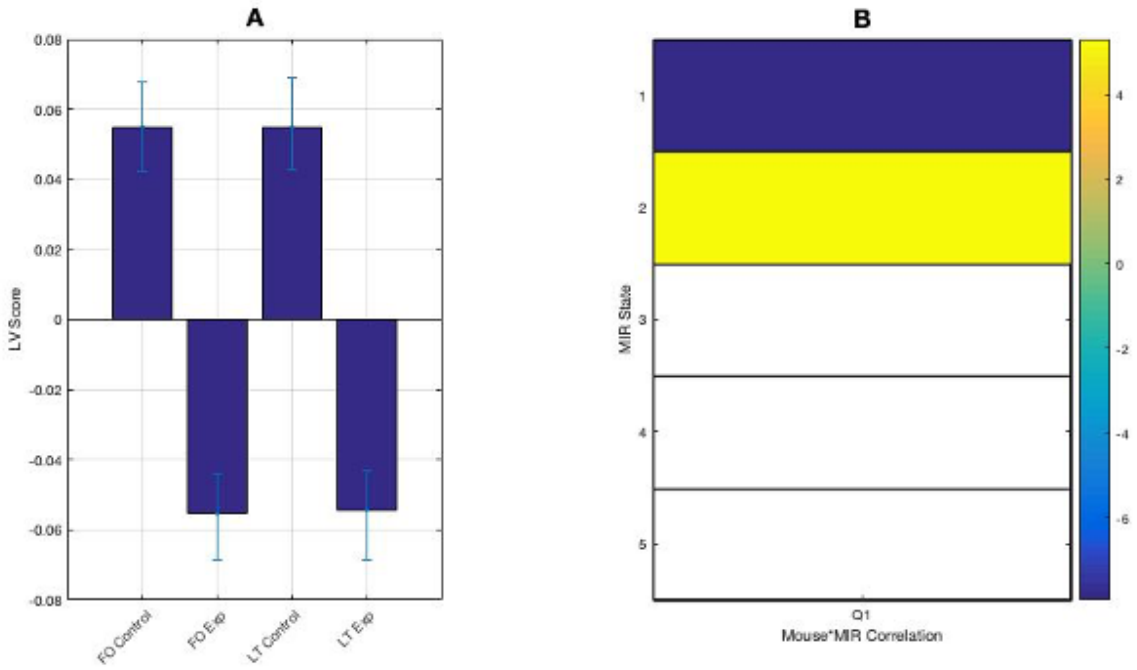

Figure 8: *(A) PLS contrasts between tasks and correlations between behaviour and stimulus measures. Error bars were calculated using bootstrap resampling and reflect the 95% confidence interval. The contrasts show an effect of task on correlations between behaviour and stimulus measures (B).*

Brain, behaviour, and stimulus

Finally, we ran a behavioural PLS between mouse tracker fractional occupancy and lifetimes in the starting neutral quadrant only (this quadrant was uniquely comparable between music listening tasks) and the correlations between MIR fractional occupancy and the brain transitional probability matrices. The analysis did not return any significant LVs ($p = 0.26$).

Discussion

We showed that network activity varies between perceptual and emotional music listening. The initial PLS analyses showed modality-specific fractional occupancy patterns between rest and the emotional music listening experiment task. At rest, there was more overall occupancy in the medial parietal, right temporal, and anterior frontal networks; while in emotional music listening,

there was more overall occupancy in the medial frontoparietal network. Analysis of the lifetimes, or average visit length, returned different results, with participants spending longer on average per visit in both medial frontoparietal networks while listening to music and attending to their emotions. These analyses demonstrate that HMM measures extracted from finely temporally-resolved timeseries' can be used to describe task-level differences.

Regions in the default mode network (DMN), consisting of many medial frontoparietal regions (Uddin et al., 2019) have shown higher levels of activity while listening to music that has been rated as liked (Wilkins et al., 2014; Pereira et al., 2011) and rated as more sad-sounding (Taruffi et al., 2017). These findings support DMN involvement in music listening with the added layer of complexity introduced by the continuous behavioural task.

Testing between the control and experiment music listening tasks returned no significant differences for either the HMM-extracted metrics or transitional probability matrices. This is not entirely surprising - both tasks, at the surface level, are quite similar: participants are listening to music and providing continuous ratings throughout each excerpt. The main difference is in the ratings themselves. In the perceptual control task, participants were asked to rate perceptual features in the musical signal, while in the emotional experiment task, they were asked to rate music-induced changes to their own emotional state. Both tasks were inherently subjective. Participants were not given training on perceptual feature rating, and as a result, perceptual ratings varied between participants (Carpentier et al., 2017).

The emotional data likewise varied between participants, and we did not initially expect to see a group-level effect between rating data and brain metrics. One effect emerged in the experiment task, however. Overall time spent in the low arousal/low valence quadrant (analogous to feeling sad) was significantly positively correlated with the likelihood of transitioning between states, especially into networks involving the anterior cingulate, frontal, and occipital regions. These regions are linked to emotion processing and attention (Uddin et al., 2019), and visual processing, indicating a community of networks involved in attending to and reporting the target emotion during music listening (Figure 4). Time spent in this state was negatively correlated with the likelihood of staying in the medial-parietal and right temporal networks. Regions in the medial parietal cortex have been shown to be active as part of the default mode network, while right-lateral temporal regions have been implicated in music processing, especially melody (Alluri et al., 2012). During reported low-valence, low-arousal mood, these networks are less likely to persist, indicating that default mode and auditory processing networks are less continuously occupied during interoceptive monitoring in this reported emotional state.

The emotional condition of sadness, represented in the low-arousal, low-valence quadrant is interesting here. Past music work has shown sadness is more consistent between individuals in a music-based emotion identification task (Faber & Fiveash, 2014), and that sad-sounding music causes participants to engage in more introspective behaviours (Taruffi et al., 2017). Sad-

sounding music does not necessarily predict the induction of a sad mood (Garrido et al., 2011), but the consistency in identifying sad music and the accompanying introspection may feed into the experience of sadness, offering a possible explanation as to why this quadrant is the lone group-level effect present.

When we introduced the stimulus data into the analysis, group level effects emerged in the control task (Figure 5). More correlations were observed between brain visits and lifetimes between music features and the right temporal, medial frontoparietal, pericentral, anterior frontal, and left temporal networks. These networks are associated with auditory processing of melodic and rhythmic information (Brown et al., 2006; Alluri et al., 2012), as well as executive function. In the control task, participants were asked to listen to music and make continuous judgements on perceptual features in the music. That this task employs bilateral temporal and anterior networks is pleasingly intuitive.

In the emotional listening experiment task, more correlations were present between music features and overall occupancy in the medial parietal, anterior frontal, and anterior cingulate networks. These networks are more implicated in executive function and emotion processing than auditory processing (Uddin et al., 2019). These results suggest that, in the emotional listening task, fine-grained auditory processing may be of secondary importance to activity in interoceptive monitoring networks. These networks also track musical features, perhaps to continually validate the reported emotional state for continued accuracy.

When we correlated brain network measures with the MIR tracker states, the perceptual task data showed two distinct between-network connectivity patterns: one correlated with context-free music features comprising numerous medial parietal, frontal, and cingulate networks; and the other correlated with context-dependent music features comprising medial, anterior cingulate, and bilateral temporal networks (Figure 6). Interestingly, involvement of the left temporal network is unique to the context-dependent circuit. Tempo is heavily context-dependent, and regions in the left temporal cortex have been previously linked to rhythmic perception specifically (Limb et al., 2006).

The emotional listening experiment task correlations did not return any significant results in this analysis, possibly due to the more internally-driven nature of the task. With perceptual listening, participants are, hopefully, listening carefully to the music, which should be reflected in correlations between the timecourse of the musical signal and the timecourse of the brain data. In emotional listening, participants are listening to the music, but attending to their internal states which may be influenced by the music, but do not follow the timecourse of the music in the same way perception will. As well, the induction of emotional states by music operates on a longer timescale which may only be correlated with the musical signal at coarser windows (Eich & Metcalfe, 1989; Västfjäll, 2002). When modelling the music excerpt and behaviour data together, we again found more correlations in the perceptual task data, particularly in context-free music

states, supporting the idea that participants are more in sync with the stimulus in the perceptual task (Figure 7).

This study is not without limitations. The ratings task was designed to give participants considerable interpretive freedom which resulted in high between-participant variability. This combined with the low sample size made individual-level analyses mathematically unfeasible, but provides a roadmap for future work, particularly the models incorporating brain, behaviour, and stimulus data together. To bridge resting state and behavioural tasks, a free listening task could have been included as a way to identify networks that may have been related to the physical operation of the mouse tracker. Further complicating matters, we did not collect data on participant familiarity with the stimuli, possibly confounding emotional response to both control and experimental pieces. The stimuli used were taken exclusively from Western art music which, though wide-ranging, describes only a sliver of extant music. Future studies should consider incorporating excerpts from outside the Western art canon, collecting data on familiarity, as well as inviting participants to select their own emotionally-evocative music.

The methods presented in this study have wide-reaching implications for the multi-modal study of naturalistic behaviour. Of particular note is the extraction of music feature networks using hidden Markov modelling. Where past work has quantified correlates and networks of brain regions that respond to isolated features in musical signals, our approach attempts to incorporate music's complexity into an analysis framework using feature networks. In the same way that functional brain networks allow us to understand how the whole brain acts and interacts, music features networks can help us understand, in a more holistic way, how music unfolds from an acoustic signal into a rich, human experience.

References

Alluri, V., Toiviainen, P., Jääskeläinen, I. P., Glerean, E., Sams, M., & Brattico, E. (2012). Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. *Neuroimage*, *59*(4), 3677-3689.

Brown, S., Martinez, M. J., & Parsons, L. M. (2006). Music and language side by side in the brain: a PET study of the generation of melodies and sentences. *European journal of neuroscience*, *23*(10), 2791-2803.

Carpentier, S. M., McCulloch, A. R., Brown, T. M., Faber, S. E., Ritter, P., Wang, Z., ... & McIntosh, A. R. (2020). Complexity matching: brain signals mirror environment information patterns during music listening and reward. *Journal of cognitive neuroscience*, *32*(4), 734-745.

Costa, M., Goldberger, A. L., & Peng, C. K. (2002). Multiscale entropy analysis of complex physiologic time series. *Physical review letters*, *89*(6), 068102.

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., ... & Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*, *31*(3), 968-980.

Downie, J. S. (2003). Music information retrieval. *Annual review of information science and technology*, *37*(1), 295-340.

Efron, B., & Tibshirani, R. (1996). Using specially designed exponential families for density estimation. *The Annals of Statistics*, *24*(6), 2431-2461.

Eich, E., & Metcalfe, J. (1989). Mood dependent memory for internal versus external events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(3), 443–455.

Faber, S., & Fiveash, A. (2014). Emotion Without Words: A Comparison Study of Music and Speech Prosody. *Canadian Journal of Music Therapy*, *20*(2).

Fasano, M. C., Cabral, J., Stevner, A., Vuust, P., Cantou, P., Brattico, E., & Kringelbach, M. L. (2020). The early adolescent brain on music: analysis of functional dynamics reveals engagement of orbitofrontal cortex reward system. BioRxiv.

Garrido, S., & Schubert, E. (2011). Individual differences in the enjoyment of negative emotion in music: A literature review and experiment. *Music Perception*, *28*(3), 279-296.

Gold, B. P., Mas-Herrero, E., Zeighami, Y., Benovoy, M., Dagher, A., & Zatorre, R. J. (2019). Musical reward prediction errors engage the nucleus accumbens and motivate learning. *Proceedings of the National Academy of Sciences*, *116*(8), 3310-3315.

Lartillot, O., & Toiviainen, P. (2007). A Matlab toolbox for musical feature extraction from audio. In *International conference on digital audio effects* (Vol. 237, p. 244).

Limb, C. J., Kemeny, S., Ortigoza, E. B., Rouhani, S., & Braun, A. R. (2006). Left hemispheric lateralization of brain activity during passive rhythm perception in musicians. *The Anatomical Record Part A: Discoveries in Molecular, Cellular, and Evolutionary Biology: An Official Publication of the American Association of Anatomists*, *288*(4), 382-389.

McIntosh, A. R., Bookstein, F. L., Haxby, J. V., & Grady, C. L. (1996). Spatial pattern analysis of functional brain images using partial least squares. *Neuroimage*, *3*(3), 143-157.

McIntosh, A. R., & Lobaugh, N. J. (2004). Partial least squares analysis of neuroimaging data: applications and advances. *Neuroimage*, *23*, S250-S263.

Mowinckel, A. M., & Vidal-Piñeiro, D. (2020). Visualization of brain statistics with R packages ggseg and ggseg3d. *Advances in Methods and Practices in Psychological Science*, *3*(4), 466-483.

Pereira, C. S., Teixeira, J., Figueiredo, P., Xavier, J., Castro, S. L., & Brattico, E. (2011). Music and emotions in the brain: familiarity matters. *PloS one*, *6*(11), e27241.

RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/

Salimpoor, V. N., Benovoy, M., Larcher, K., Dagher, A., & Zatorre, R. J. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nature neuroscience*, *14*(2), 257-262.

Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., & Leahy, R. M. (2011). Brainstorm: a user-friendly application for MEG/EEG analysis. *Computational intelligence and neuroscience*, *2011*, 1-13.

Taruffi, L., Pehrs, C., Skouras, S., & Koelsch, S. (2017). Effects of sad and happy music on mind-wandering and the default mode network. *Scientific reports*, *7*(1), 14396.

The MathWorks Inc. (2016). MATLAB version: 9.1.0 (R2016b), Natick, Massachusetts: The MathWorks Inc. https://www.mathworks.com

Uddin, L. Q., Yeo, B. T., & Spreng, R. N. (2019). Towards a universal taxonomy of macro-scale functional human brain networks. *Brain topography, 32*(6), 926-942.

Vidaurre, D., Smith, S. M., & Woolrich, M. W. (2017). Brain network dynamics are hierarchically organized in time. *Proceedings of the National Academy of Sciences*, *114*(48), 12827-12832.

Västfjäll, D. (2001). Emotion induction through music: A review of the musical mood induction procedure. *Musicae Scientiae, 5*(1), 173-211.

Wilkins, R. W., Hodges, D. A., Laurienti, P. J., Steen, M., & Burdette, J. H. (2014). Network science and the effects of music preference on functional brain connectivity: From Beethoven to Eminem. *Scientific reports*, *4*(1), 1-8.

Supplementary Material: Regions from the EEG data parcellation using the Desikan-Killiany atlas (Desikan et al., 2006)

Region	
caudal middle frontal L	caudal middle frontal R
entorhinal L	entorhinal R
frontal pole L	frontal pole R
lateral orbitofrontal L	lateral orbitofrontal R
medial orbitofrontal L	medial orbitofrontal R
pars opercularis L	pars opercularis R
pars orbitalis L	pars orbitalis R
pars triangularis L	pars triangularis R
rostral middle frontal L	rostral middle frontal R
superior frontal L	superior frontal R
insula L	insula R
paracentral L	paracentral R
banks superior temporal sulcus L	banks superior temporal sulcus R
inferior temporal L	inferior temporal R
middle temporal L	middle temporal R
parahippocampal L	parahippocampal R
superior temporal L	superior temporal R
temporal pole L	temporal pole R
transverse temporal L	transverse temporal R
caudal anterior cingulate L	caudal anterior cingulate R
isthmus cingulate L	isthmus cingulate R
posterior cingulate L	posterior cingulate R
rostral anterior cingulate L	rostral anterior cingulate R
fusiform L	fusiform R
inferior parietal L	inferior parietal R
postcentral L	postcentral R
precentral L	precentral R
precuneus L	precuneus R
superior parietal L	superior parietal R
supramarginal L	supramarginal R
cuneus L	cuneus R
lateral occipital L	lateral occipital R
lingual L	lingual R
pericalcarine L	pericalcarine R