Neural correlates of emotional responses to music: An EEG study

Ian Daly a, Asad Malik a, Faustina Hwang b, Etienne Roesch a, James Weaver a, Alexis Kirke b, Duncan Williams b, Eduardo Miranda b, Slawomir J. Nasuto a

a Brain Embodiment Lab, School of Systems Engineering, University of Reading, Reading, RG6 6AY, UK
b Interdisciplinary Centre for Computer Music Research, University of Plymouth, Plymouth, PL4 8AA, UK

HIGHLIGHTS

• We record EEG while participants listen to a large set of music.
• Participants report their induced emotional responses.
• Asymmetry in beta and gamma frequency bands relate to induced emotion.
• Sparse long range networks are significantly modulated by music induced emotion.

ARTICLE INFO

Article history:
Received 6 March 2014
Received in revised form 14 April 2014
Accepted 1 May 2014
Available online 10 May 2014

Keywords:
EEG
Emotion
Music
Neural correlates
Functional connectivity

ABSTRACT

This paper presents an EEG study into the neural correlates of music-induced emotions. We presented participants with a large dataset containing musical pieces in different styles, and asked them to report on their induced emotional responses.

We found neural correlates of music-induced emotion in a number of frequencies over the pre-frontal cortex. Additionally, we found a set of patterns of functional connectivity, defined by inter-channel coherence measures, to be significantly different between groups of music-induced emotional responses.

© 2014 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Music is widely accepted to produce changes in affective (emotional) states in the listener [1]. However, the exact nature of the emotional response to music is an open question and it is not immediately clear that induced emotional responses to music would have the same neural correlates as those observed in response to emotions induced by other modalities. For example, in [2] individuals with post-antomedial temporal lobe excision reported being unable to recognise “scary” music, while retaining the ability to recognise “scary” faces, suggesting the existence of modality specific cerebral networks for emotions.

Neural correlates of emotional responses have been explored by a number of researchers [3]. However, although there is an emerging picture of the relationship between induced emotions and brain activity, there is a need for further refinement and exploration of neural correlates of emotional responses induced by music.

Differentiating between affective induction and affective recognition is well documented by music psychologists (for example, [4]). However, within music psychology, the terminology used to differentiate these responses varies: induced emotions, is synonymous with felt or experienced emotions, whilst perceived emotions are synonymous with conveyed or observed emotions. Great care has been taken throughout this work to ensure that induced emotional responses are targeted.

When considering the electroencephalogram (EEG), relationships have been found between the emotional content of music and specific frequency bands. Pleasantness of music has been reported to be positively correlated with power spectral density (PSD) in the theta band (4–7 Hz) over the prefrontal cortex [5]. The reported valence (pleasantness/unpleasantness) and arousal (intensity/energy) of musical stimuli have been reported to correlate with frontal alpha (8–13 Hz) asymmetry [3].

However, many previous studies investigating neural correlates of emotional responses to music have been limited by the range of
stimuli used [6]. Often music from just one style (e.g. classical), or only a small number of pieces, are used. For example, in [3] just four classic orchestral pieces were used. Thus, it is difficult to determine whether the reported neural correlates of emotional responses to music are in fact correlates of emotion, or instead correlates of syntactical or acoustic components of the stimuli. Additionally, there is a growing body of evidence highlighting the role of inter-regional connectivity in a range of cognitive phenomena. This includes evidence for connectivity changes related to motor control [7], emotional responses to audio-visual stimuli [8], and perception of music [9]. It is interesting to ask whether specific neural assemblies (connectivity maps) are involved in emotional responses to music.

For example, in [9] relationships between EEG electrodes are measured via a similarity index (SI). During music listening an increase in SI was noted over the pre-frontal cortex, suggesting music listening engages specific networks in this region. In [10] functional magnetic resonance imaging (fMRI) was used to identify increased activations in the left orbito and mid-dorsolateral frontal cortex during differences in emotional valence (pleasantness-unpleasantness) induced by changes in musical mode and tempo. This result may be contrasted with the laterality view of emotional processing, which suggests that different hemispheres process the valence of the emotion (for example as suggested in [3]). Finally, in [11] it is postulated that distinct cortical pathways may underlie emotions and different distributed networks may underpin emotions.

Together, these findings strongly suggest the involvement of networks of cortical and sub-cortical regions in emotion. However, it is not completely understood exactly which pathways are involved. For example, there is currently little evidence for how responses to music, such as sadness, are organised in the brain [11]. Therefore, we set out to explore neural correlates of emotional responses to music in order to answer the following questions:

1. What are the neural correlates of emotional responses to music when a larger and more varied set of stimuli, that is not specific to one style, is used?
2. What neural assemblies are involved in induced emotional responses to music?

We approach this problem by employing a large set of musical stimuli drawn from a range of styles, and that have a range of emotional content, as reported using a number of scales. Thus, our choice of stimuli offers both a large variety of stimuli and a level of comparative unfamiliarity for listeners (to minimise pre-selection bias), such that neural correlates of emotions may be better differentiated from syntactic and acoustic properties of the stimuli. This allows an investigation into the effect of complex musical pieces on emotions using a large set of systematically selected musical stimuli, a component noted to be absent in the majority of studies conducted into musically induced emotions [1].

2. Materials and methods

2.1. Measurements

2.1.1. Participants

Thirty-one individuals participated, comprising staff and students from the University and members of the public. All participants gave informed consent, and the study was approved per the procedures of the University of Reading research ethics committee. The median age was 35 (range 18–66, standard deviation 14.46), 18 were female, and 29 right handed. Participants were paid £10. Using a self-report questionnaire 9 participants indicated no musical training, 7 some basic training (e.g. trained as a child but not actively playing), 11 some moderate training (e.g. currently playing at an amateur level), 3 a high level of training (e.g. currently playing at a professional level), and 1 did not provide this information.

2.1.2. EEG recording

Electroencephalogram (EEG) was recorded from 19 channels positioned according to the international 10/20 system for electrode placement (Fp1, Fp2, F7, F3, Fz, F4, F8, T3, T5, C3, Cz, C4, T4, T7, T3, Pz, P4, T6, O1, and O2), referenced to electrode FCz and a ground at AFz.

EEG was sampled at a rate of 1000Hz via a Brain Products BrainAmp EEG amplifier (Brain Products, Germany). Impedances were kept below 5 kΩ for 24 of the participants and below 15 kΩ for all participants.

2.1.3. Musical stimuli

The musical stimuli were taken from [6] that presents a new set of stimuli for the study of the modulation of emotions via music. This stimulus set contains 110 excerpts from film scores spanning a range of styles were rated on induced emotion by 116 participants. The selected excerpts broadly spanned the emotional response space.

2.1.4. Paradigm

The experimental paradigm consisted of 6 runs of EEG recordings. For the first and last run, baseline recordings were made of the participants resting (sitting still while looking at the screen for 300 s).

The four intervening runs contained 10 trials each. For a single trial, participants were presented with a fixation cross from the beginning of the trial (t = −3) until 15 s had elapsed. Three seconds after the appearance of the fixation cross (second 0) a randomly selected musical clip was played for 12 s.

After musical stimuli presentation the participants were given a short (0.5 s) pause before 8 questions were asked in random order, requesting them to rate the music on a scale of 1–9 in terms of the induced pleasantness, energy, tension, anger, fear, happiness, sadness, and tenderness. Each question was phrased in the form of a Likert questionnaire [12], for example “The music made me feel X”, with answers ranging from “1 = Strongly disagree” through to “9 = Strongly agree”. Before the experiment, special care was taken to explain to report only induced emotions. This took the form of a written information sheet, slideshow, and practice trial.

Inter-trial intervals lasted 2–4 s from when the final question was answered by the participants until the next fixation cross.

2.2. EEG processing

2.2.1. Pre-processing

The EEG was first visually inspected for artefacts. Time segments and channels that contained blinks, electromyographic (EMG) activity, electrooculographic (EOG) activity, movement, 50 Hz noise, and other visually apparent artefacts were labelled by an experimenter who was blinded to the trial content.

A notch filter (infinite impulse response, 35th order) at 50 Hz was applied to remove line noise. A band-pass filter was then applied (3rd order Butterworth) between 0.1 and 45 Hz to remove direct current drift and high frequency noise components.

Independent component analysis (ICA) was applied to the EEG to attempt to separate components of interest from artefacts, which had not been removed by the preceding filtering steps. Infomax ICA [13] was chosen and the implementation in the EEGLAB toolbox was used [14].
Independent components (ICs) which were visually judged to contain artefacts were removed. Then the ICA de-mixing matrix was inverted and multiplied by the ICs to reconstruct clean EEG.

Finally, trials were marked for inclusion in the final analysis stage if they met the following criteria:

1. They had not been previously labelled as containing EMG, movement, or failing electrode artefacts. These artefacts often exhibit a broad spatial and spectral distribution, while also exhibiting a short duration. Thus, they are not easily removed via ICA.
2. The peak frequency of the signal was estimated within ±100 µV (the accepted amplitude range of clean EEG [15]).

The EEG was then re-referenced to a common average reference scheme and, finally, down-sampled to 100 Hz.

2.3. Asymmetry measures

Measures of EEG asymmetry over the pre-frontal cortex were estimated within a particular frequency range \( f_{\text{min}}:f_{\text{max}} \) in the following way.

First, Laplacian source derivations were calculated over channels F3 and F4, as described in [16]. Power spectral densities (PSDs) were then calculated for each Laplacian derivation and asymmetry measured as

\[
A = \left| f_{\text{min}}:f_{\text{max}} \right| - \left| R_{\text{max}}:f_{\text{max}} \right|,
\]

where \( f_{\text{min}}:f_{\text{max}} \) denotes the PSD on the left Laplacian derivation between frequencies \( f_{\text{min}} \) and \( f_{\text{max}} \), \( R_{\text{max}}:f_{\text{max}} \) denotes the same derived from the right prefrontal cortex, and \( \ldots \) denotes the mean.

Asymmetry was calculated for resting state sessions for each participant and for each trial from 0 to 10 s.

2.4. Connectivity measures

Connectivity between EEG channel pairs was measured via the magnitude squared coherence [17]. This is defined as

\[
C_{x,y} = \frac{\left| G_{x,y} \right|^2}{G_{x,x}G_{y,y}},
\]

where \( G_{x,y} \) denotes the cross spectral density between signals \( x \) and \( y \), and \( G_{x,x} \) and \( G_{y,y} \) the auto-spectral density.

Coherence was measured between all channels for each trial between \( f_{\text{min}} \) and \( f_{\text{max}} \).

2.5. Analysis

Participant ratings of emotion induced by each stimulus were given on 8 axes: pleasantness, energy, sadness, anger, tenderness, happiness, fear, and tension. A number of these axes may correlate highly, for example, happiness and pleasantness are likely to be positively correlated. Therefore, to avoid over-fitting, the dimensionality of the emotional responses was reduced via a principal component analysis (PCA) approach [18]. Principal components (PCs) were listed in order of decreasing variance and the first \( n \) PCs selected such that they contained a minimum of 75% of the variance.

It was first desirable to identify any correlations between participant demographic details and emotions reported by the participants and projected along the PCs.

The demographic details were age, gender, handedness, and level of musical training. Additionally, the impedance of the EEG was entered as an independent variable to determine if there was an effect of impedance.

Asymmetry measures were investigated as a potential correlate of induced emotion. Asymmetry was calculated at frequencies of 5–45 Hz in a sliding window of width 2 Hz (step size 1 Hz). Correlations were calculated between the asymmetry measure and each PC of induced emotion. Peak frequencies were identified at which a maximum correlation was observed between asymmetry and each PC.

Stepwise linear regression was then used to determine if resting state asymmetry at these peak frequencies significantly related to the PCs when subject demographic details were also considered as factors. Thus, at each peak frequency, stepwise linear regression was attempted with the dependent variable resting state asymmetry at the peak frequency (±2 Hz) and independent variables PCs and subject details.

Subsequently, asymmetry was investigated as a neural correlate of induced emotional responses during music listening. Asymmetry was calculated at the peak frequency for each PC during each trial and stepwise linear regression applied to identify which PCs and/or participant information significantly related to asymmetry.

Finally, coherence measures of connectivity between channel pairs were compared between groups of high vs. low PC value trials using paired \( t \)-tests with Bonferroni correction for multiple comparisons. This was done for the peak frequency bands for each PC.

3. Results

3.1. Signal quality

During artefact removal a mean of 1.49 ± 0.75 ICs were marked for removal from each run. After application of the inclusion/exclusion criteria to the cleaned trials, a total of 360 trials were discounted from subsequent analysis (31.03% of the total). This left a total of 880 artefact-free trials.

3.2. Principal components of emotion

The first three principal components are observed to contain 75.8% of the total variance and each correlate with specific emotional responses. Specifically, principal component 1 (PC1) most strongly correlates with the participants’ response to the music in terms of pleasantness \( r = -0.837, p < 0.001 \), anger \( r = 0.608, p < 0.001 \), fear \( r = 0.850, p < 0.001 \), and happiness \( r = -0.819, p < 0.001 \). PC2 most strongly correlates with energy \( r = -0.640, p < 0.001 \), sadness \( r = 0.775, p < 0.001 \), and tenderness \( r = 0.579, p < 0.001 \). Finally, PC3 most strongly correlates with tension \( r = 0.535, p < 0.001 \).

Thus, the three principal components may be described as consistent with the Schimmack and Grob three dimensional model of emotional responses [19]. In their model emotional responses are projected onto three axes “Valence” (PC1; pleasant–unpleasant), “Energy arousal” (PC2; wakefulness–tiredness), and “Tension arousal” (PC3; relaxation–tension).

3.3. Participant variability

A significant relationship is identified between PC1 (valence) and age \( p < 0.001 \), gender \( p < 0.001 \), and musical training \( p < 0.001 \) \( R^2_{\text{adj}} = 0.031 \), with post-hoc correlation identifying a negative relationship with age \( r = -0.078, p = 0.020 \), greater valence reported by male participants \( (r = 0.011, p = 0.001) \), and a negligible correlation with training \( r = -0.151, p = 0.001 \).

PC2 (energy) significantly relates to age \( p < 0.001 \) and gender \( p < 0.001 \) \( R^2_{\text{adj}} = 0.031 \), with a non-significant relationship
between age and PC2 \( (r=0.023, \text{n.s.}) \) and male participants reporting higher PC2 (energy) values \( (p<0.001) \).

PC3 (tension) significantly relates to handedness \( (p<0.001) \) and musical training \( (p<0.001) \) \( (R_{adj}^2 = 0.034) \). Post-hoc tests show left handed participants to have lower PC3 values (less tension) \( (p<0.001) \) and a negative correlation is observed between training and PC3 \( (r = -0.139, p < 0.001) \).

3.4. Frequency bands of interest

Each principal component (PC) is observed to correlate with asymmetry in a particular frequency band. PC1 (valence) correlates most with frequencies in the range 18–22 Hz \( (r=0.121, p < 0.001) \). PC2 (energy) is non-significantly correlated to asymmetry in the range of 14–18 Hz \( (r=0.061, p < 0.071) \). Finally, PC3 (tension) significantly correlates with asymmetry in the gamma band (35–39 Hz) \( (r=0.081, p < 0.016) \). Therefore, these frequency bands are investigated further.

3.5. Resting state neural correlates of emotion

A significant relationship is observed between resting state asymmetry and PC3 (tension) in the range 18–22 Hz \( (r=-0.116, p<0.001) \), 14–18 Hz \( (r=-0.102, p<0.002) \), and 35–39 Hz \( (r=-0.084, p<0.012) \). No significant relationships are observed between resting state asymmetry and any participant details.

3.6. Task related neural correlates of emotion

We consider each frequency band of interest in turn. In the band 18–22 Hz asymmetry significantly relates to PC1 (valence) \( (p<0.001) \), age \( (p<0.001) \), and training \( (p<0.001) \) \( (R_{adj}^2 = 0.062) \). Post-hoc testing reveals a negative correlation with PC1 \( (r=-0.121) \), age \( (r=-0.087) \), and training \( (r=-0.203) \). Asymmetry in the range 14–18 Hz reveals no significant effects.

Asymmetry in the range 35–39 Hz relates to PC1 \( (p<0.021) \) and training \( (p<0.001) \) \( (R_{adj}^2 = 0.177) \). Post-hoc tests show a negative relationship between asymmetry and PC1 \( (r=-0.038) \) and a negative correlation with training \( (r=-0.245, p<0.001) \).

Additionally, each PC is coarse grained into high vs. low value groups. Comparisons are made between these groups in the frequency bands of interest. Significant differences are found between high vs. low PC1 (valence) groups in the 18–22 Hz band \( (p<0.001) \) with lower (right hemisphere oriented) asymmetry for higher PC1 (valence) music clips. This is illustrated in Fig. 1.

3.7. Neural assemblies

Coherence maps are used to attempt to differentiate classes of musical clips (e.g. high vs. low PC1 groups). In the 18–22 Hz frequency band a small set of connections significantly differentiate between high vs. low PC1 groups. Specifically, coherence between channel pairs F7-Pz, F3-F4, Fz-F8, and P7-T4 differentiate high vs. low PC1 groups \( (p<0.05, \text{Bonferroni corrected}) \). This is illustrated in Fig. 2 (created with eConnectome [20]).

In the gamma band (35–39 Hz) significant differences in coherence are found between high vs. low PC3 (tension) groups over right frontal and central channel pairs. Specifically, channel pairs FP1-T6, Fz-C4, F4-Cz, and P4-T6 differentiate between high vs. low PC3 groups. This is illustrated in Fig. 3.

4. Discussion

The principal components (PCs) identified to contain 75.8% of the variance may be described as corresponding to the three
dimensions of the Schimmack and Grob model of emotional responses [19]. However, this is only one of several models proposed to capture emotions [21] and it is likely other models are also amenable to fitting with our observed responses.

The cortical mapping of emotion to music may be described by two theories [11]. The first holds that the right hemisphere is responsible for processing all emotions with specialised "emotional modules". The second is the valence hypothesis and holds that the right hemisphere is involved more during negative and the left during positive emotions. Support is available for both views [11]. However, in recent years the valence hypothesis has gained ground, see for example, [3].

Our results also support the valence hypothesis. For example, we find asymmetry in the pre-frontal cortex relating to a number of induced emotions. In the beta frequency band, at frequencies of 18–22 Hz, asymmetry relates to the first principal component (PC) (valence), which is similar to findings reported in [3].

However, some differences may be observed between our findings and other evidence supporting the valence hypothesis. For example, in [3] alpha band (8–13 Hz) is reported as relating to valence, while by contrast we found no effects in the alpha band and instead found a relationship in the beta band. This may be due to a number of reasons, such as different stimuli or inter-participant differences. However, a likely explanation is simply that researchers in [3] did not look at the beta band.

Additionally, asymmetry in the gamma band also correlates with valence. However, while arousal (energy) has been suggested elsewhere to be related to beta band activity [22] [23] we did not observe this.

Crucially we identify similarities between our results and the common view of the valence hypothesis relating to hemispheric orientation. Inspection of the differences in asymmetry between high vs. low valence music identifies a smaller (right hemisphere oriented) asymmetry during high valence music. This matches results reported in [24] and the common view of the valence hypothesis, namely that the right hemisphere is more involved in negative emotion and, therefore, exhibits lower band powers, as reported in [3].

Additionally, our results are analogous to results reported in [25] in which pleasant brain states were reported to be associated with an increase in power spectral densities (PSDs) in the pre-frontal cortex, and [26] in which gamma PSD was modulated by the valence of images.

As noted in the introduction, the cortical activations of a number of emotions are not yet fully understood, but are thought to involve long range networks [11]. Our observed neural networks of induced emotions, show sparse long range connections involving both pre-frontal and occipital cortices and left-right hemispheres. These are modulated by the valence and tension induced by the music.

This further highlights findings reported elsewhere [8] that induced emotions involve a broad network of cortical activation. Additionally, it is interesting to note that there are a number of similarities in the network topology to those reported elsewhere (e.g. left-right hemisphere connections [8]) for emotions induced by other modalities. This provides some support for the idea that EEG correlates of emotions induced by music are somewhat analogous to emotions induced by other modalities.

Finally, it is interesting that these long range networks all involve activations of the prefrontal cortex. This adds weight to the view that emotional processing, while largely involving prefrontal cortex areas, also impacts a range of other regions such as the right inferior parietal cortex, reported to be involved in emotional recognition of facial expressions [27].

5. Conclusions

We have observed that many neural correlates of musical stimuli-induced emotion, such as pre-frontal cortex asymmetry, are highly analogous to other modalities (e.g. visual modalities [26]). Additionally, specific neural assemblies are modulated by music-induced emotions, suggesting the involvement of long range cortical networks in emotional processing.

Acknowledgements

EPSRC grants (EP/J003077/1 and EP/J002135/1). We thank Sajeel Ahmed, Isil Poyraz Bilgin, Katherine Heseltine Flynn, and Maryam Karimidomal for help with EEG recording.

References


Fig. 3. Connectivity networks differentiating high vs. low PC3 (tension) groups of trials in the frequency range 35–39 Hz.
...