A Shared Neural Network for Highly Liked and Disliked Music

Pablo Ripollés[1,2,3§], Amy M. Belfi[4§], Anna Kasdan[1,5], Edward A. Vessel[6], Andrea R. Halpern[7], Jess Rowland[1], Robert Hopkins[8], G. Gabrielle Starr[9,10], David Poeppel[1,3]

[1] Department of Psychology, New York University, New York, NY, 10003, USA.

[2] Music and Audio Research Lab (MARL), New York University, New York, NY, 10003, USA.

[3] Center for Language, Music, and Emotion (CLaME), New York University, New York, NY, 10003, USA.

[4] Department of Psychological Science, Missouri University of Science and Technology, Rolla, MO, 65409, USA.

[5] AAAS Science & Technology Fellowship Program, Washington, DC.

[6] Department of Psychology, City College, The City University of New York, New York, NY, 10031, USA.

[7] Department of Psychology, Bucknell University, Lewisburg, Pennsylvania, 17837, USA.

[8] Department of Philosophy, New York University, New York, NY, 10003, USA.

[9] Department of English, Pomona College, Claremont, CA, 91711, USA

[10] Department of Neuroscience, Pomona College, Claremont, CA, 91711, USA

[§] Authors contributed equally to this work

***Corresponding authors:**

Pablo Ripollés
Dept. of Psychology, NYU
6 Washington Place
New York, NY 10003
pripolles@nyu.edu

Amy M. Belfi
Dept. of Psychological Science, Missouri S&T
500 W. 14th St.
Rolla, MO 65409
amybelfi@mst.edu

ABSTRACT

Music research has increasingly focused on neural responses to naturalistic stimuli. The neural correlates of music-induced peak moments of pleasure (e.g., chills) have been well-studied; data on other types of aesthetic responses, such as judgments of *beauty, goodness, or liking* are sparse. Among these, *liking* is of particular interest given its complex nature and its ubiquity in everyday—and musical—experiences. In this functional magnetic resonance imaging (fMRI) study, 26 participants listened to musical excerpts while continuously rating how much they *liked* the music. In addition, subjects provided an overall liking judgment at the end of each musical piece. Participants' musical preferences manifested remarkably early and were sustained over time, with the value of the continuous rating at the end of each piece showing the strongest correlation with the overall ratings. At the neural level, highly liked and disliked musical pieces activated a shared neural network previously associated with the processing of both positive and negative emotional stimuli. This network encompassed the anterior cingulate cortex, limbic system, basal ganglia, and precuneus. These activations reflected individual differences, such that participants with greater trait-level responsiveness to artistic stimuli showed stronger engagement of this network while listening to music. Finally, the default-mode network was associated with the slope of the continuous ratings, suggesting that the faster participants liked a musical excerpt, the more disengaged this network became. In linking behavior and brain function to better characterize the way humans make aesthetic judgments of music the data support a model in which preferences and aesthetic choices are modulated by brain circuits associated with emotion and reward.

Keywords: Music, aesthetics, liking, valence, default-mode network, ICA.

1. INTRODUCTION

There is growing interest in studying perception and its neural correlates in the context of complex naturalistic materials, including music (Aliko et al., 2020; Desai et al., 2016; Huth et al., 2016; Jääskeläinen et al., 2021; Kandylaki et al., 2016; Ki et al., 2016; Saarimäki, 2021; Sonkusare et al., 2019; Alluri et al., 2012, 2013, 2017; Burunat et al., 2014; Cong et al., 2014; Kaneshiro et al., 2020; Kern et al., 2022; Li et al., 2024; Sachs et al., 2020; Tervaniemi, 2023). Prior work has focused on unveiling the neural correlates of various components of music, such as acoustic features (Burunat et al., 2016; Haumann et al., 2021; Martin et al., 2018; Toiviainen et al., 2014), and components of subjective response, including uncertainty and surprise (Abrams et al., 2022; Cheung et al., 2019; Di Liberto et al., 2020; Robert et al., 2024), peak moments of pleasure (i.e., musical chills or frissons; Blood & Zatorre, 2001; Koelsch, 2014, 2020; Martínez-Molina et al., 2016; Mas-Herrero et al., 2021; Salimpoor et al., 2011, 2013), or specific emotional experiences (Daly et al., 2015; Putkinen et al., 2021; Singer et al., 2016), among others.

Beyond identifying the neural correlates of music's perceptual, cognitive, emotional, and reward-related features, there is a growing body of research examining more complex participants responses to music and visual arts, such as aesthetic judgments (Brattico et al., 2013; Chatterjee & Vartanian, 2014; Chuan-Peng et al., 2020; Pearce et al., 2016). These are subjective evaluations of, for example, beauty, goodness, or liking (Brielmann & Pelli, 2017; Bruns et al., 2023; Che et al., 2024; Egermann & Reuben, 2020; Jacobsen et al., 2004; Kant, 1764; Pelowski et al., 2017). Among these, *liking* is especially interesting given its ubiquity in our everyday lives as a description of subjective, personal experience, and its breadth as an aesthetic term. So, while the label "beautiful" or "good" is subject to arguments about individual taste, the claim of "liking" is not disputable by other individuals. In addition, "liking" covers a complex range of experiences: for example, although it is true that positive valence is highly correlated with liking (Altmann et al., 2012; Clemente et al., 2023; Cohrdes et al., 2017), that is not always the case. We can appreciate art that is unsettling, enjoy horror movies, delight in books or

films that end tragically, and, of course, find pleasure in listening to sad songs. Indeed, the emotional valence of music does not need to be in agreement with its aesthetic evaluation (i.e., *sad* music does not always induce sad feelings; note however, that the enjoyment of sad music does not necessarily correspond to a conscious reinterpretation of unpleasant sounds as positive experiences; Brattico et al., 2015; Eerola et al., 2018; Garrido & Schubert, 2011; Huron & Vuoskoski, 2020; Juslin, 2013; Menninghaus et al., 2017; Schubert, 2024; Taruffi & Koelsch, 2014; Vuoskoski & Eerola, 2017). Despite these common experiences, empirical research assessing the neural underpinnings of liking music is lacking.

Moreover, prior work using naturalistic musical excerpts has typically used (i) listener ratings obtained *post-hoc* or (ii) continuous ratings (Abrams et al., 2024; Bachorik et al., 2009; Belfi et al., 2018, 2021; Curzel et al., 2024; Isik & Vessel, 2019, 2021; Lucas et al., 2010; Mas-Herrero et al., 2014; Mori & Iwanaga, 2017; Upham & McAdams, 2018; Wen & Krumhansl, 2017). However, one open question which has received less focus is how summary judgments of liking rating relate to moment-to-moment judgments as a composition unfolds over time. Real-time measures capture the dynamic fluctuations in emotional or aesthetic responses as the music unfolds, whereas retrospective measures condense that experience into a summary judgment influenced by memory biases (e.g., the peak–end rule; Kahneman et al., 1993). Studying both provides a more comprehensive understanding of how immediate sensory, cognitive, and affective processes translate into final appraisals. In this vein, various metrics of the continuous response, such as the peak value, slope, or end value, have been used to predict an overall rating, with varying degrees of success (Rozin et al., 2004).

The aim of the present experiment is to investigate the relationship between continuous and overall aesthetic ratings of liking, and how this is translated into network-level brain responses. To do so, we collected both continuous (i.e., behavioral responses measured dynamically during stimulus presentation; Larsen et al., 2017) and overall (post-stimulus) liking judgments of music while concurrently recording fMRI. For the musical stimuli, we selected two historically distant and highly

distinct genres of music (classical and electronic) to allow for a better generalization across musical modes and to provide a range of liked and disliked musical materials. To characterize network-level responsiveness to liking, we used independent component analysis (ICA), a data-driven approach (Calhoun et al., 2008) that identifies temporally coherent networks of brain regions. ICA is well-suited to discern how multiple functional networks subserve different cognitive processes (Calhoun et al., 2001; Wu et al., 2009) and has several advantages over univariate analyses (e.g., disentangling different neural circuits occurring within the same brain area; Beldzik et al., 2013; McKeown et al., 2003; Xu, Potenza, et al., 2013; Xu, Zhang, et al., 2013).

At the behavioral level, we hypothesized that participants' preferences for the musical excerpts (i.e., whether the participant likes or dislikes the song) will be apparent very early in the trial, will be sustained over time (Belfi et al., 2018; Isik & Vessel, 2021), and will show strong agreement with the overall rating after the whole excerpt has been heard. Based on previous research in other aesthetic domains (Belfi, Vessel, et al., 2017; Clemente et al., 2024; Frame et al., 2024; Halpern et al., 2008; Isik & Vessel, 2019; Vessel & Rubin, 2010), we also hypothesized low agreement across raters for ratings of individual excerpts (i.e., *one person's trash is another person´s treasure;* Cohrdes et al., 2020; Gingras et al., 2015; Kuppens et al., 2013; Vessel et al., 2018; Yang et al., 2023).

At the neural level, we hypothesized that aesthetic judgments may show similar neural correlates across stimulus modalities (e.g., visual, auditory; Brown et al., 2011; Ishizu & Zeki, 2011; Taruffi et al., 2017; Wilkins et al., 2014). Recent studies, which have primarily focused on visual stimuli (Belfi et al., 2019; Chatterjee & Vartanian, 2014; Chuan-Peng et al., 2020; Sacheli et al., 2022; Skov & Nadal, 2023), suggest that aesthetic experiences consistently engage multiple brain networks, including those involved in sensory analysis, emotional responses, and self-referential processes (Alluri et al., 2013; Belfi et al., 2019; Brattico et al., 2013; Isik & Vessel, 2021; Luo et al., 2024; Mas-Herrero et al., 2021; Putkinen et al., 2021; Reybrouck et al., 2018; Singer et al., 2016; Vessel et al., 2012). Thus, we hypothesized that networks involved in the processing of music reward and emotion (Koelsch,

2014; Mas-Herrero et al., 2021), as well as self-referential processes (Sheline et al., 2009) would be modulated by continuous and/or overall liking judgments of music.

2. MATERIALS AND METHODS

2. 1. Participants

Thirty participants were recruited from the greater New York University community. The sample size was selected based on a previous study using a similar behavioral paradigm (i.e., participants providing continuous liking ratings while listening to music; Belfi et al., 2018). All participants were right-handed, had no previous history of neurological disorders and reported normal hearing. As prior work has indicated that musical experience as well as expertise influences neural responses to music (Criscuolo et al., 2022; Oechslin et al., 2012; Pinho et al., 2014), we sought to exclude participants with substantial musical training. Similarly, we sought to exclude participants who were "musically anhedonic," such that they show no emotional responses to music (Belfi, Evans, et al., 2017; Belfi & Loui, 2020; Mas-Herrero et al., 2013, 2014). To this end, all participants completed the Goldsmith's Musical Sophistication Index (Gold-MSI) prior to participating in the study (Müllensiefen et al., 2014). Participants were excluded if they scored more than 2 SD above the mean on the Musical Training subscale of the Gold-MSI, and more than 2 SD below the mean on the Emotions subscale. Participants were paid for their participation in the experiment. New York University's Institutional Review Board approved this study, and all participants gave written informed consent in accordance with the New York University Committee on Activities Involving Human Subjects. Three participants experienced technical problems while rating the musical excerpts and one participant did not complete all the fMRI runs after feeling claustrophobic. The reported results are thus based on data from 26 participants (mean age = 27.84 ± 6.21 years; 13 women; Gold-MSI Musical Training Scale = 10.96 ± 3.76, participants' mean falls within the 11th percentile of the norms; Gold-MSI Emotions Scale = 32.46 ± 4.36, participants' mean falls within the 31st percentile of the norms; Müllensiefen et al., 2014).

2.2. Musical stimulus selection

We selected two genres of music to better generalize and to allow for a range of liked and disliked musical materials. For classical music, all clips were taken from 19th-century small ensemble music from the Romantic era. The 19th-century Romantic Classical music was selected because it tends to have a wider range of dynamic and emotional intensity than earlier Classical-era music and would, we believed, therefore allow for a wider range of behavioral responses. No composer was included more than once, and no well-known composers were included; no pieces contained lyrics or the human voice. All the pieces were from the European tradition. Electronic music selections consisted of electronic dance music with a distinctive beat structure (i.e., no ambient music) that contained no lyrics or human vocalizations. All electronic pieces were contemporary and ranged in tempo from 60-150 beats per minute. All stimuli selected for this work have been previously validated and used in a study assessing the timing of music aesthetic judgements (Belfi et al., 2018). The complete set of stimuli consisted of 16 classical pieces and 16 electronic pieces. Each piece was clipped to a 60 s excerpt such that participants could hear enough of the piece to show temporal variability in their ratings, while still allowing for many trials. For electronic music, the 60 s excerpts were selected so that some dynamics occurred during that period (e.g., the excerpt contained a “drop” or a transition). Prior to the experiment, participants completed a brief (~10 minute) training session using musical excerpts from these genres not presented in the experiment to familiarize themselves with the task and response modalities.

2.3. Procedure

The experiment was programmed in PsychToolBox 3.0 (Brainard, 1997) using Matlab2017b. Behavioral and fMRI data were collected at the NYU Center for Brain Imaging using a 3T Siemens Allegra scanner. The experiment consisted of four experimental runs with eight musical excerpts each (32 excerpts: 16 classical and 16 electronic). Excerpts were blocked by genre, such that there were two blocks of classical excerpts and two blocks of electronic excerpts. Block and excerpt order were counterbalanced across participants.

In our fMRI experiment, we collected continuous behavioral ratings while participants listened to music, instead of allowing them to listen passively while scanning and collecting liking ratings in a second presentation of the stimuli outside the scanner (Sachs et al., 2020). We did so to capture participants' initial, unmodified responses to the auditory material. Repeated interactions with music can fundamentally alter the experience, including changes in participant engagement and in aesthetic judgments (Madison & Schiölde, 2017; Madsen et al., 2019; Margulis, 2013). Importantly, previous research has shown that providing continuous ratings during stimulus presentation does not have a significant impact on behavioral or brain responses (Isik & Vessel, 2019; Kühn & Gallinat, 2012). Participants were specifically instructed to use the slider to express how much they liked/disliked the song.

Each trial began with a 1 s blinking fixation cross, followed by the onset of the excerpt. During the excerpt, the participant saw a slider bar presented on the screen. At the onset of the excerpt, participants were instructed to continuously rate how much they *liked* the piece at the present moment. To make this continuous rating, participants used a roller ball mouse to move the visual slider bar. The mapping between roller ball and slider was counterbalanced across participants to de-confound movement direction and rating: half of the participants moved the ball up to move the cursor to the right (higher liking) and half the subjects moved the ball down to move the cursor to the right (higher liking). The endpoints of the slider were labeled with 'L' and 'H' to indicate "Low" and "High", respectively. Mouse movements were sampled at a rate of 10Hz. Following the offset of the stimulus was a 1s blank screen, which was then followed by an "overall rating" period. During this 4s period, the visual slider bar reappeared on the screen and participants were asked to rate, *overall,* how much they liked the piece. Participants used the slider ball followed by a button-press to make a single, discrete rating on the visual slider bar. For the overall rating, we did not explicitly ask participants to summarize their continuous ratings. Instead, we asked them to report the overall "gestalt" of their experience in terms of liking. Following this rating, the inter-trial-interval (i.e., rest) across all trials was 20s.

After the experiment had ended, participants completed the Aesthetic Responsiveness Assessment questionnaire (AReA; Schlotz et al., 2021). The AReA is a self-report questionnaire that assesses an individual's general aesthetic responsiveness to stimuli such as paintings, music, natural landscapes, architecture, dance, and poetry, among others. Examples of questions include "I experience joy, serenity, or other positive emotions when looking at art", "I am emotionally moved by music", "I appreciate the visual design of buildings", or "I notice beauty when I look at art". The English version of AReA has a reliability (ω) of 0.89 and has been validated using several independent datasets and in different languages (Schlotz et al., 2021).

Statistical analyses were carried out with the software JASP version 0.19.1.0 (JASP Team, 2025). All data were first assessed for normality using the Shapiro-Wilk test. Multiple data distributions deviated from normality and thus non-parametric statistics were used in all analyses (e.g., Wilcoxon signed rank), with the matched rank biserial correlation (r_{rb}=0.10, small effect size; r_{rb}=0.30, medium effect size; r_{rb}=0.50, large effect size; r_{rb}>=0.7 very large effect size; Maher et al., 2013) or Kendall's W (W=0.1, small effect size; W=0.3, medium effect size; W>=0.5, large effect size; In & Lee, 2024) as measures of effect size. When needed we computed Bayes factors using default priors (BF_{01}), which reflect how likely data are to support the null hypothesis (Morey & Rouder, 2015; Rouder & Morey, 2012; Wagenmakers et al., 2018).

2.4. Behavioral data

2.4.1. Inter-rater reliability

To assess the agreement across participants' liking ratings, we used the intraclass correlation coefficient (ICC). Using JASP, we calculated the "ICC(2,1)" (i.e., a random sample of raters rate all items; Shrout & Fleiss, 1979) with 95% confidence intervals using the overall ratings from all songs (e.g., 32 songs, 26 raters). An ICC smaller than 0.5 represents 'poor' agreement (Koo & Li, 2016).

2.4.2. Liking differences between genres

Although we anticipated low inter-rater reliability, it is possible that participants had a general preference for one genre over the other, even if their ratings for individual stimuli within each genre varied widely. To address this possibility, we averaged the overall ratings, and the four measures derived from the continuous ratings for each participant (*end, mean,* and *peak value*, and *slope;* see section 2.4.3 below), grouping the values by genre. These average values were analyzed using Wilcoxon signed-rank tests.

2.4.3. Relationship between overall and continuous behavioral ratings

First, we tested how the overall rating of liking provided at the end of each stimulus presentation related to the continuous ratings. For each trial, we computed the *slope*, *mean*, *peak* and *end value* as metrics of the continuous trace (Rozin et al., 2004; Schafer et al., 2014). The slope was calculated using MATLAB r2021b by fitting (in a least-squares sense) a simple linear regression between the liking ratings of each excerpt and the time (i.e., $Y = a + bX$, where Y represents the liking ratings, X represents time, a is the intercept, and b, the slope). The mean value was calculated as the mean of all ratings. For the peak and end value, we first divided the continuous ratings into 2 second bins. Then we averaged the ratings across the bins obtaining 30 values per musical excerpt. We did this to obtain more stable ratings over time. The peak value and the end value were defined as the maximum and the last of these values, respectively. Then, for each participant, we computed Spearman´s correlations between the different metrics of the continuous ratings and the overall responses. Each individual r coefficient was subsequently transformed to a normal distribution using Fischer's z transform. These z-transformed values were analyzed using a non-parametric Friedman test with one factor with four levels (Measure: End Value, Mean Value, Peak Value, Slope). For significant effects, post-hoc Conover tests were used with Holm correction for multiple comparisons. This analysis allowed us to compare the strength of the relationship between the different continuous metrics and the discrete, overall liking score.

For the remaining analyses, trials were designated as ‘high,’ ‘medium,’ and ‘low’ aesthetic appreciation based on the value of the overall rating given at the end of each trial. Specifically, for each participant, the 32 musical excerpts were ranked using these overall ratings and then divided into High (excerpts 1 to 11 in the ranking), Medium (excerpts 12 to 21) and Low (excerpts 22 to 32) trials. This choice of binning was based on previous work showing that neural responses to aesthetic stimuli can follow a nonlinear response based on discrete divisions of aesthetic ratings (Belfi et al., 2019; Vessel et al., 2012). Thus, binning the responses increased our sensitivity to such a nonlinear response compared to doing an analysis on the continuous response. The remaining behavioral and fMRI analyses were calculated according to this division between High, Medium, and Low.

2.4.4. Relationship between acoustic features and behavioral ratings

Previous research has shown that certain acoustic features can be predictive of emotional, aesthetic, and reward-related responses to music (Eerola et al., 2009; Gingras et al., 2014; Mori, 2022; Schubert, 2004). While the anticipated low agreement and high variability between raters (i.e., we expected people to like different songs and different genres) could potentially rule out a relationship between acoustic features and liking patterns, we nevertheless conducted an analysis to control for this possibility. First, we used the MIRToolbox 1.8.1 (Lartillot & Toiviainen, 2007) in MATLAB R2021b to extract acoustic features from each song, employing the default sampling rate and frame lengths. We chose a set of 14 features based on previous studies using music information retrieval techniques (Alluri et al., 2012; Carone & Ripollés, 2024; Gingras et al., 2014; Groves et al., 2023; Salakka et al., 2021). These features represent different aspects of music related to tonality, timbre, temporal properties, and other characteristics (Salakka et al., 2021). Regarding tonality, we extracted *mode* (using the *mirmode* function; the outcome value represents the difference between the strength of the best fitting major key and the best fitting minor key) and the *key clarity strength* (using the *mirkey* function; the outcome value represents the clarity with which the key of the song can be identified). For timbre, we extracted the *attack time* (using the function *mirattacktime:* the outcome is a vector with the time from the start of a

an event, identified with *mirevents*, to its peak amplitude; we took the average over all attack times per song), several measures extracted from the spectrum of frequencies (*spectral centroid, spread, flux, flatness,* and *entropy,* using the functions, *mircentroid, mirspread, mirflux, mirflatness, mirentropy,* respectively), and the *roughness* (using the function *mirroughness*; the outcome value represents the sensory dissonance stemming from frequencies with small perceptual distance between them). Regarding temporal features, we calculated the *fluctuation centroid and entropy* (using the *mirfluctuation* function; these measures describe rhythmic periodicities), and the *pulse clarity* (using the *mirpulseclarity* function; the outcome value represents the salience of the song beat). In addition, we also calculated measures related to *loudness* (using the function *mirrms*) and to *novelty* (using the function *mirnovelty,* this measure is related to musical expectations).

For each participant, we then calculated the average values for the 14 acoustic features for the songs in the High and Low liking conditions (i.e., for the top 11 and bottom 11 songs, ranked by the overall liking rate). For each participant and acoustic feature, we thus obtained one average value for songs with high and low ratings. These values were analyzed using Wilcoxon sign-ranked tests.

2.4.5. Behavioral Online Replication

To address concerns that scanner noise and physical discomfort might influence the accuracy of participants' music liking ratings, we conducted an online behavioral replication experiment. This replication followed the same experimental procedures, stimuli, and design as the main fMRI experiment, with the main differences being that: (1) the slider for the ratings went from 1 to 100 instead of 0 to 1 and participants provided responses using their mouse instead of using a roller ball, (2) the experiment was built using oTree (Chen et al., 2016), and (3) participants were recruited through Amazon Mechanical Turk (AMT). By replicating the experiment in a non-MRI environment, we aimed to validate the consistency and reliability of the behavioral data across contexts. Recruitment was restricted to participants living in the United States and, to ensure data quality, who had completed at least 100 tasks with 95% of approval rate on AMT. Results from online experiments with participants

recruited via AMT have consistently replicated well-established effects (Crump et al., 2013). For example, studies have demonstrated that AMT samples yield results comparable to those from traditional participant populations across various cognitive tasks, including, for example, flanker tasks (Crump et al., 2013; Orpella et al., 2024), experiments involving complex auditory stimuli (Lizcano-Cortés et al., 2022), and experimental designs that require participants to provide continuous ratings while listening to music (Abrams et al., 2024; Curzel et al., 2024). To address the higher attrition rates typically observed in online experiments (Lumsden et al., 2017), we recruited 50 participants. At the end of the experiment, participants completed the Musical Training Scale from the Gold-MSI (Müllensiefen et al., 2014). In this replication, we did not exclude participants with substantial musical training. To better control for participant characteristics in the online setting and exclude individuals with specific music anhedonia, we administered the Barcelona Music Reward Questionnaire (BMRQ) and excluded participants with scores below 63 (Mas-Herrero et al., 2013). The online experiment was approved by the Institutional Review Board at New York University. Participants provided informed consent by clicking an "accept" button after reading the consent form and were compensated for their participation. Eight participants were excluded from the analysis (1 had specific music anhedonia and 7 did not provide continuous ratings while listening to music). The reported results are thus based on data from 42 participants (mean age = 39.09 ± 12.13 years; 21 women; Gold-MSI Musical Training Scale = 24.76 ± 10.59, participants' mean falls within the 43rd percentile of the norms; BMRQ=81.28 ± 8.84; for BMRQ scores, a score in the 10th percentile is about 65 and a score in the 90th percentile is about 87; Mas-Herrero et al., 2013; Müllensiefen et al., 2014).

2.5. fMRI data analysis

2.5.1. Scanning parameters

All fMRI scans took place at the NYU Center for Brain Imaging using a 3T Siemens Allegra scanner with a Nova Medical head coil (NM011 head transmit coil). For the music fMRI task, four runs of 346 sequential whole-brain multi-echo (ME) echo-planar imaging (EPI) volumes were acquired (TR

= 2000 ms, right-to-left phase encoding, flip angle = 75º, voxel size = 3.0 × 3.0 × 3.0 mm^3, 34 axial slices, acquisition size = 80×64). The ME EPI sequence and a tilted slice prescription (15-20° tilt relative to the AC-PC line) were used to minimize dropout near the orbital sinuses. ME EPI images were reconstructed using a custom algorithm designed by the NYU Center for Brain Imaging to minimize dropout and distortion. In some of the participants, the sequence did not provide full coverage of the whole brain anatomy, and the most superior parts of the brain were left out (3 mm of the most superior part of the SMA and parietal cortex). To allow precise co-registration with the functional data, a high resolution T1 MPRAGE image was also acquired during this MRI session (TR = 2500 ms, TE = 3.93 ms, flip angle = 8º, voxel size = 1.0 × 1.0 × 1.0 mm^3, 176 sagittal slices, acquisition matrix = 256 × 256).

2.5.2. Preprocessing and ICA analysis

Data were preprocessed using Statistical Parameter Mapping software (SPM12, Wellcome Trust Centre for Neuroimaging, University College, London, UK, www.fil.ion.ucl.ac.uk/spm/) under MATLAB r2021b. For each participant, the four functional runs were first realigned to the mean image of all the EPIs. The high resolution T1 was then co-registered to this mean functional image and segmented using the Unified Segmentation algorithm (Ashburner & Friston, 2005). The deformation fields (i.e., normalization parameters) obtained during the segmentation process were then used to spatially normalize all functional images to the MNI template included with SPM12 (maintaining the voxel size of 3.0 × 3.0 × 3.0 mm^3). Finally, images were spatially smoothed with an 8 mm FWHM kernel.

Once fMRI data were preprocessed, Group Spatial ICA was applied using the Group ICA of fMRI Toolbox (GIFT v4.0b; http://mialab.mrn.org/software/gift/; Calhoun et al., 2001). The number of independent components (formed by a spatial map and a time-course) to be extracted was set to twenty, which has been shown to be an optimal dimension in previous research in both healthy and clinical populations (Lopez-Barroso et al., 2015; Orpella et al., 2022; Sihvonen et al., 2017; Smith et al., 2009). Data were intensity normalized, concatenated and reduced to 20 temporal dimensions (using

principal component analysis), and finally analyzed using the infomax algorithm (Bell & Sejnowski, 1995). No scaling was used, as with the intensity normalization step, the intensities of the spatial maps were in percentage of signal change. To obtain whole brain group-wise statistics, the spatial map of each of the ICA components (i.e., networks) retrieved for all participants was submitted to a second-level analysis using a one sample t-test under SPM12, which treats each subject's spatial map as a random effect (Calhoun et al., 2001). For the spatial maps of each ICA network, all statistics are reported at a $p < 0.05$ FWE-corrected threshold at the cluster level, with a $p < 0.001$ uncorrected threshold at the voxel level. Clusters with fewer than 50 voxels were not included in the analyses and are not depicted in the figures. Maxima and all coordinates are reported in MNI space. Anatomical and cytoarchitectonical areas were identified using the Automated Anatomical Labeling (Tzourio-Mazoyer et al., 2002) and the Talairach Daemon database atlases (Lancaster et al., 2000) included in the xjView toolbox (http://www.alivelearn.net/xjview/).

2.5.3. ICA Network modulation by music, and by overall and continuous liking ratings

To identify which of the networks were related to the task (i.e., music listening) and to the overall or continuous liking ratings of interest, several multiple regressions were conducted. GIFT allows for fitting each participant’s component time-course to a specific model. The models can be created using SPM12, by convolving a canonical hemodynamic response with the timing of specific conditions and their parametric modulators. First, we visually inspected each network, and those reflecting artifacts such as movements, ventricles, edges, or the presence of blood vessels were discarded. For the non-artifactual components, and to discard networks not modulated by music, we created a model including only one main condition of interest: music (including all musical excerpts regardless of their liking rating). A resting condition plus six movement regressors (obtained from the realignment step, added to take into account motion artifacts) were also included in the model. A multiple regression between this model and each network time-course was fitted and beta values—which represent network engagement—for the music condition were obtained (i.e., we obtained one beta value per network and

346 participant). These beta values were further analyzed using Wilcoxon signed-ranked tests to identify
347 networks which were significantly activated or deactivated by music listening. In order to control for the
348 number of networks being tested, we adjusted the p-values using FDR-correction as implemented in
349 the *mafdr* function on MATLAB r2021b, which follows the Benjamini and Hochberg method (Benjamini
350 & Hochberg, 1995).

351 Once the networks significantly engaged during music listening were identified, a series of finer-
352 grained analyses were performed using the “high”, “medium”, and “low” binning for each individual
353 participant, so that our analysis was more sensitive to a nonlinear response (see Materials and Methods
354 section 2.4.3). This binning was done separately for each individual participant and effectively controls
355 for any possible differences in acoustic features between the conditions, as participants tend to show
356 very low agreement in terms of which pieces are rated ‘high’ versus ‘low’ (i.e., one participant may have
357 rated a particular excerpt as a 0.9, while another participant may have rated the same excerpt as 0.1;
358 see Figure 1A and ICC results in section 3.1.1). Therefore, the same piece may appear in each of the
359 ‘high,’ ‘medium,’ and ‘low’ conditions depending on the participant’s preferences, eliminating any
360 potential confound of acoustic differences between stimuli. Note that, for completeness, we also directly
361 compared the acoustic properties of highly liked and disliked songs (see Results section 3.1.4; we
362 found no significant differences for any of the 14 acoustic features tested).

363 Once songs were classified into High, Medium, and Low liking conditions, two different models
364 were created and tested only for the networks significantly engaged by music. While both models
365 included the three conditions of interest (High, Medium and Low liked songs) plus the rest condition
366 and the six movement regressors, they differed in their parametric modulators: one model used the
367 *overall* judgments of liking and another the *slope* of the continuous ratings. The first model was used
368 to measure which networks were modulated by the overall liking judgments. A single parametric
369 modulator per musical condition (one for High, Medium and Low trials) was included in this model. This
370 modulator was formed by the ratings provided by participants after the end of each musical excerpt for

each condition (High, Medium, Low). The second model was used to identify which networks were modulated by the continuous liking ratings. For each trial, we fitted a linear model to the continuous ratings and obtained a slope that represented the speed with which participants reached asymptote on aesthetic ratings. We used the slope as a measure of the continuous ratings, as this metric was most different from the overall ratings (see Results section 3.1.3. and Figure 1C). Again, a single parametric modulator per condition was included in this model. In this case, the parametric modulator was formed by the slopes of the High, Medium and Low musical conditions.

A multiple regression between these two models and each network time-course was fitted and beta values for each parametric modulator were obtained. To normalize for individual differences, the beta values for the parametric modulator of the Music Medium condition were subtracted from the beta values of the parametric modulators of the Music High and the Music Low conditions for each participant. Thus, for each network, we obtained a series of corrected (from now on, $High_m$ and Low_m) beta values for a) the overall parametric modulator, and b) the continuous parametric modulator (i.e., slopes). The rationale behind this correction by the Medium condition is twofold, based on: (1) the fact that people show great variability in the way in which they process aesthetic stimuli (Belfi, Vessel, et al., 2017; Clemente et al., 2024; Halpern et al., 2008; Vessel & Rubin, 2010); and (2) the fact that variability in processing aesthetics has direct effects on brain activity within reward and emotion-related regions (Keller et al., 2013; Martínez-Molina et al., 2016; Matthews et al., 2024; Padrao et al., 2013). Thus, in this analysis, the Medium condition becomes an intrinsic baseline (specific for each participant), that allows us to take into account the individual differences in responses to aesthetic stimuli. We then focused on how high and low aesthetic stimuli *displace* this baseline in opposite or equal directions (Ferreri et al., 2019; Ripollés et al., 2018). Finally, for each network and condition of interest (continuous and overall parametric modulators) we used Wilcoxon signed-ranked tests with FDR correction to assess whether the network was significantly modulated by: (1) $High_m$ trials; 2) Low_m trials; or (3) significantly more modulated by $High_m$ than by Low_m trials.

3. RESULTS

3.1. Behavioral Results

3.1.1. Inter-rater variability

Replicating prior work in other aesthetic domains (poetry; visual arts; Belfi, 2019; Belfi, Vessel, et al., 2017; Halpern et al., 2008; Pombo et al., 2024; Vessel & Rubin, 2010), ICC was quite low, reflecting disagreement among raters (ICC value= 0.069, 95% CI: 0.033, 0.138). See Figure 1A and 1B for a visual depiction of the variability of the ratings. This low agreement supports the choice of our materials: we wanted pieces/genres on which people would disagree, to reduce reactions to low-level acoustic features.

3.1.2. Liking differences between genres

Even in the context of low inter-rater agreement, participants consistently liked classical music more than electronic music for all measures, including both overall [classical = 0.59 ± 0.12, electronic = 0.47 ± 0.10, W = 49, $p < 0.001$, $r_{rb} = 0.72$] and continuous ratings [*end value:* classical = 0.62 ± 0.13, electronic = 0.50 ± 0.12, W = 50, $p < 0.001$, $r_{rb} = 0.71$; *mean value:* classical = 0.59 ± 0.10, electronic = 0.49 ± 0.10, W = 54, $p = 0.001$, $r_{rb} = 0.69$; *peak value*: classical = 0.71 ± 0.10, electronic = 0.65 ± 0.09, W = 65, $p = 0.004$, $r_{rb} = 0.63$; *slope*: using a 1 to 100 liking scale for clarity, participants increased the liking score 0.4 points per second for classical and 0.006 for electronic on average, W = 50, $p < 0.001$, $r_{rb} = 0.71$].

3.1.3. Relationship between overall and continuous behavioral ratings

There was strong agreement between the overall ratings and various metrics extracted from the continuous responses (see Figure 1C). However, there was a significant effect [$X^2_F(3) = 38.35$, $p < 0.001$, Kendall's W = 0.49] with the *end value* of the continuous judgment (see Figure 1C) showing the strongest correlation with the overall ratings, significantly higher than those obtained using the *mean value* ($p_{holm} = 0.030$, $r_{rb} = 0.57$), the *peak value* ($p_{holm} < 0.001$, $r_{rb} = 0.81$), and the *slope* ($p_{holm} < 0.001$, $r_{rb} = 1$).

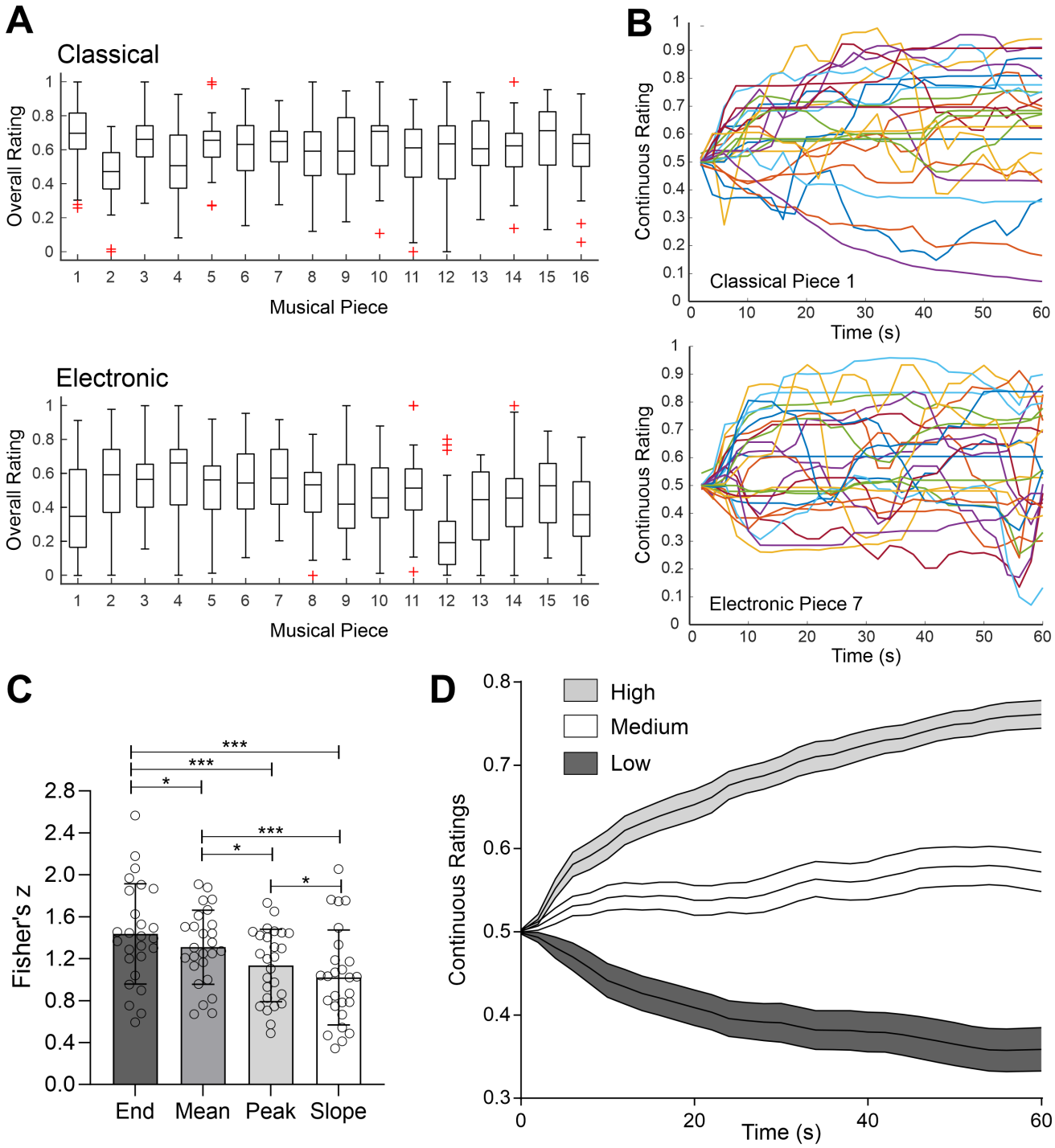


Figure 1. Continuous and overall behavioral aesthetic judgments. A. Boxplots illustrating participants' overall ratings for each musical piece. Center line indicates median, bottom and top edges of the box indicate 25th and 75th percentiles. Whiskers extend to the most extreme data points not considered outliers, and red + symbols indicate individual outliers. **B.** Individual continuous responses from all participants for two representative musical excerpts (top, classical piece number 1; bottom, electronic piece number 7). Each line represents the continuous rating provided by a specific participant. **C**. Barplots (mean ± SD) showing the strength (using Fisher's z transform) of the correlation between the *overall* rating and different measures extracted from the continuous ratings, including the *end, mean,* and *peak value*, and the *slope.* Each circle represents data from one participant. * $p < 0.05$, *** $p < 0.001$. **D.** Mean (solid lines) plus SEM (shaded error bands) of *continuous* ratings (between 0 and 1) across participants are shown for High (grey), Medium (white) and Low (black) musical excerpts classified using the overall ratings.

The *mean* value showed a significantly stronger correlation with the overall ratings than the *peak* value (p_{holm} = 0.013 r_{rb} = 0.67) and the *slope* (p_{holm} < 0.001, r_{rb} = 0.79). The *peak value* also showed a stronger relationship with the overall ratings than the *slope* (p_{holm} = 0.013, r_{rb} = 0.36). This suggests that there was a strong alignment between the overall and continuous ratings, with the *slope* being the metric showing the weakest relationship. As described in the methods, for the remaining analyses we divided the trials according to the overall rating bin (high, medium, low; see Materials and Methods; Belfi et al., 2018, 2019). As expected, the evolution of the behavioral continuous ratings revealed clear differences between high, medium and low-rated musical excerpts (Figure 1D). Strikingly, participants' average preferences for the musical excerpts were shown very early in the trial (within the first 3 seconds) and were sustained over time.

3.1.4. Relationship between acoustic features and behavioral ratings

For each participant in the fMRI experiment, acoustic features were averaged for high and low rated songs. For the 14 features extracted, no significant differences were found between the acoustic characteristics of these songs (see Table 1).

Feature	W	p-value	r_{rb}
Tonality: Mode	173	0.960	-0.014
Tonality: Key Clarity Strength	154	0.600	-0.123
Timbre: Attack Time	229	0.181	0.305
Timbre: Spectral Centroid	124	0.199	-0.293
Timbre: Spectral Spread	123	0.190	-0.299
Timbre: Spectral Flux	123	0.190	-0.299
Timbre: Spectral Flatness	119	0.157	-0.322
Timbre: Spectral Entropy	110	0.099	-0.373
Timbre: Roughness	139	0.367	-0.208
Temporal: Fluctuation Centroid	224	0.227	0.276
Temporal: Fluctuation Entropy	213	0.353	0.214
Temporal: Pulse Clarity	107	0.084	-0.390
Other: Loudness	132	0.280	-0.248
Other: Novelty	201	0.532	0.145

Table 1. Acoustic Features. Comparison of average acoustic features for highly vs. lowly liked songs for each participant.

3.1.5. Behavioral online replication

The online behavioral experiment replicated the behavioral results obtained for the main fMRI experiment: (1) we found very low agreement among raters (ICC value= 0.092, 95% CI: 0.054, 0.163);

(2) participants consistently liked classical music more than electronic [overall: classical = 60. ± 18, electronic = 49 ± 14, W = 251, p = 0.011, r_{rb} = 0.44; *end value:* classical = 62 ± 19, electronic = 0.50 ± 0.14, W = 250, p = 0.011, r_{rb} = 0.44; *mean value:* classical = 58 ± 14, electronic = 50 ± 10, W = 248, p = 0.010, r_{rb} = 0.45; *peak value*: classical = 73 ± 13, electronic = 66 ± 9, W = 229, p = 0.005, r_{rb} = 0.49; *slope*: using a 1 to 100 liking scale, participants increased the liking score 0.29 points per second for classical and 0.024 for electronic on average, W = 271, p = 0.023, r_{rb} = 0.40]; (3) regarding the relationship between overall and continuous ratings, there was a significant effect [$X^2_F(3) = 54.88$, $p < 0.001$, Kendall's W = 0.43] with the *end value* of the continuous judgment showing the strongest correlation with the overall ratings, significantly higher than with any other measure [*mean value*: $p_{holm} < 0.001$, r_{rb} = 0.60); *peak value:* $p_{holm} < 0.001$, r_{rb} = 0.81); *slope*: $p_{holm} < 0.001$, r_{rb} = 1]; and (4) when binning trials into high, medium, and low liking according to the overall ratings, participants' average preferences for the musical excerpts were shown very early in the trial (within the first 3 seconds) and were sustained over time. See Appendix Figure 1 for a summary of these results.

3.2. fMRI Results

3.2.1 *ICA decomposition* and *network modulation by music*

From the original 20 networks extracted using the Group ICA of fMRI Toolbox (GIFT; see Materials and Methods section 2.5.2), 6 components reflecting artifacts such as noise, movements, ventricles, edges, or the presence of blood vessels were discarded. Of the 14 remaining components, 11 were significantly and robustly modulated by music listening (by fMRI blocks in which music was presented). While 20 networks have been shown to be an optimal dimension for ICA analysis in previous research in both healthy and clinical populations (Lopez-Barroso et al., 2015; Orpella et al., 2022; Sihvonen et al., 2017; Smith et al., 2009), we repeated the analysis extracting 25 networks. The results were very similar, with the same networks being modulated by music listening. Note that most of these networks are typically found in task-related and resting state ICA analyses (Forn et al., 2013; Simó et al., 2018; Smith et al., 2009; Tie et al., 2008).

463 Of these networks, six were positively modulated by music (see Table 2): an *Auditory-Premotor*
464 *Network* (APMN) covering the bilateral superior and middle temporal gyrus, and premotor and
465 supplementary motor regions; a *Ventromedial Prefrontal Network* (VMPN), covering the ventromedial
466 prefrontal cortex and extending to the ventral striatum and caudate; a network covering the anterior
467 cingulate cortex and frontal regions, and extending to the limbic system, the basal ganglia, posterior
468 cingulate cortex and precuneus, which closely resembles a network previously known to respond to
469 stimuli with both positive and negative emotional valence (a *Valence Network*, VN; Lindquist et al.,
470 2016; Liu et al., 2011); an *Orbitofrontal Network* (OrbN), encompassing the lateral and medial
471 orbitofrontal cortex and regions in the occipital cortex; a *Superior Frontal Network* (SFN), extending
472 from superior frontal regions to angular and temporal gyri; and a *Medial Temporal Lobe Network*
473 (MTLN) encompassing the hippocampus, parahippocampus, amygdala and fusiform gyrus.

Network	**Activation Region**	***W***	p_{FDR}	r_{rb}
APMN	Bilateral sup/mid temporal gyrus; Bilateral Heschl gyrus; Bilateral insula; Right precentral gyrus; Bilateral postcentral gyrus; Supplementary motor area.	351	<0.001	1
VMPN	Ventromedial prefrontal cortex; Bilateral anterior cingulate; Bilateral ventral striatum; Bilateral caudate.	342	<0.001	0.949
VN	Bilateral ant/mid/superior cingulate cortex; Bilateral sup/inf frontal gyrus; Bilateral orbitofrontal cortex; Bilateral insula; Bilateral caudate; Bilateral putamen; Bilateral middle temporal gyrus; Bilateral hippocampus; Bilateral parahippocampus; Bilateral amygdala; Thalamus; Bilateral precuneus; Bilateral calcarine; Bilateral lingual gyrus; Right mid/inf occipital gyrus.	298	0.002	0.698
OrbN	Bilateral orbitofrontal cortex; Bilateral hippocampus; Bilateral sup/mid occipital gyrus.	295	0.003	0.681
SFN	Bilateral superior/middle frontal gyrus; Bilateral angular gyrus; Bilateral precuneus; Right middle temporal gyrus.	285	0.006	0.624
MTLN	Bilateral hippocampus; Bilateral parahippocampus; Bilateral fusiform gyrus; Bilateral amygdala; Bilateral temporal pole; Bilateral middle temporal gyrus.	270	0.019	0.538

Table 2. ICA networks *positively* modulated by music listening. The main areas of the network are reported using a $p < 0.05$ FWE-corrected threshold at the cluster level. The level of network engagement with the task is reported using Wilcoxon sign-rank tests and FDR-corrected p-values. APMN: Auditory-Premotor Network; VMPN: Ventromedial Prefrontal Network; VN: Valence Network; OrbN: Orbitofrontal Network; SFN, Superior Frontal Network; MTLN, Medial Temporal Lobe Network.

474 Five of the identified networks were significantly disengaged (see Table 3) while the participants
475 were listening to the musical excerpts: a *right Fronto-Parietal Network* (rFPN), covering right inferior
476 and middle frontal, angular and inferior parietal gyri; a S*ensory-Motor Network* (SMN) comprising
477 premotor and supplementary motor regions, the postcentral gyrus and basal ganglia; *Medial and*
478 *Lateral Visual Networks* (MVN and LVN), extending mainly through the occipital lobe; and finally the
479 *Default Mode Network* (DMN) comprised of its typical constituents, including the anterior and posterior
480 cingulate cortex, the precuneus, and parietal regions.

Network	Activation Region	W	p_{FDR}	r_{rb}
rFPN	Right sup/inf frontal gyrus; Bilateral middle frontal gyrus; Right precentral gyrus; Right postcentral gyrus; Right insula; Bilateral ant/mid cingulate cortex; Right angular gyrus; Right supramarginal gyrus; Bilateral inferior parietal gyrus; Right mid/inf temporal gyrus; Right precuneus; Right mid/sup occipital gyrus;	0	p<0.001	-1
SMN	Bilateral precentral gyri; Bilateral postcentral gyri; Supplementary motor area; Bilateral middle cingulate gyri; Bilateral thalamus; Bilateral caudate; Bilateral putamen; Bilateral sup/inf parietal gyrus; Bilateral supramarginal gyri; Bilateral cuneus	0	p<0.001	-1
MVN	Bilateral calcarine; Bilateral cuneus; Bilateral lingual gyrus; Bilateral precuneus; Bilateral middle/superior occipital gyrus; Bilateral precentral gyrus.	17	p<0.001	-0.903
LVN	Bilateral sup/mid/inf occipital gyrus; Bilateral lingual gyrus; Bilateral posterior superior/mid/inf temporal gyrus; Bilateral fusiform gyrus; Bilateral inferior parietal gyrus; Bilateral postcentral gyrus; Bilateral precentral gyrus.	26	p<0.001	-0.852
DMN	Bilateral ant/post/mid cingulate cortex; Bilateral precuneus; Bilateral cuneus; Bilateral sup/mid occipital gyrus; Bilateral calcarine; Bilateral lingual gyrus; Bilateral angular gyrus; Bilateral inferior parietal gyrus; Bilateral sup/middle frontal gyrus; Thalamus; Bilateral middle temporal; Bilateral hippocampus; Bilateral parahippocampus; Bilateral fusiform gyrus; Brainstem.	53	p=0.002	-0.698

Table 3. ICA networks *negatively* modulated by music listening. The main areas of the network are reported using a $p < 0.05$ FWE-corrected threshold at the cluster level. The level of network engagement with the task is reported using Wilcoxon sign-rank tests and FDR-corrected p-values. rFPN: right Fronto-Parietal Network; SMN: Sensorimotor Network; MVN: Medial Visual Network; LVN: Lateral Visual Network; DMN: Default Mode Network.

481 *3.2.2. ICA network modulation by overall and continuous liking ratings*

482 When assessing which networks *were modulated by the liking judgments* made by the
483 participants for each musical excerpt, we found significant results for both the overall and continuous
484 ratings. Note that parametric beta values for high and low rated songs were corrected using the Medium

trials as a baseline to account for individual differences in responses to aesthetic stimuli (corrected betas: $High_m$ and Low_m; see Materials and Methods section 2.5.3).

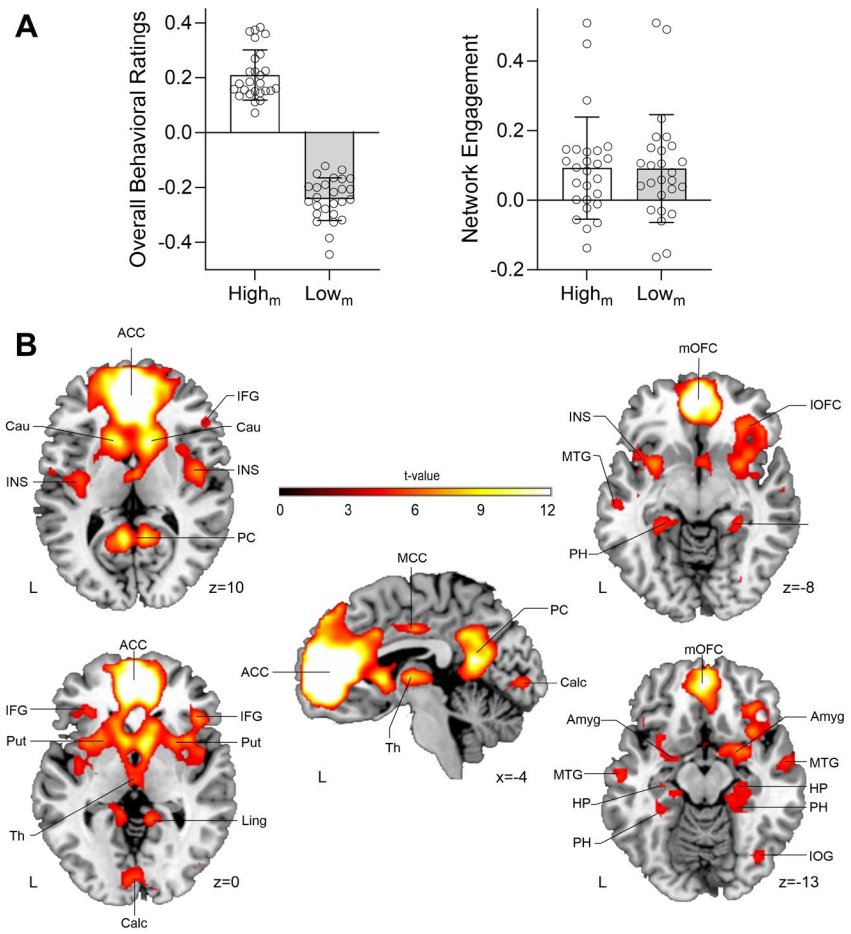


Figure 2. Network modulated by overall liking judgments. A. On the left, mean (plus SD) behavioral overall (i.e., discrete) liking ratings for $High_m$ (white) and Low_m (grey) trials (corrected using the Medium behavioral ratings). This is the expected distribution of corrected ratings, since we classified songs into high or low conditions based on the overall ratings (we did not compute statistics to compare between overall $High_m$ and Low_m overall ratings to avoid double-dipping). On the right panel, mean beta values (plus SD; corrected using Medium beta values), showing that the VN equally responds to $High_m$ (white) and Low_m (grey) trials. Each circle represents data from one participant. This suggests that the VN network is modulated by the overall aesthetic judgment of liking for both highly liked and disliked musical pieces. **B.** In red-yellow**,** spatial map corresponding to the VN ($p < 0.05$ FWE-corrected at the cluster level). Neurological convention is used with MNI coordinates at the bottom right of each slice. L, Left Hemisphere; ACC, Anterior Cingulate Cortex; MCC, Middle Cingulate Cortex Cau, Caudate; IFG, Inferior Frontal Gyrus; INS, Insula; PC, Precuneus; Put, Putamen; Ling, Lingual Gyrus; Th, Thalamus; Calc, Calcarine; mOFC, medial Orbitofrontal Cortex; lOFC, lateral Orbitofrontal Cortex; MTG, Middle Temporal Gyrus; PH, Parahippocampal Gyrus; Amyg, Amygdala; HP, Hippocampus; lOG, Inferior Occipital Gyrus.

The VN significantly responded to both $High_m$ ($W = 296$, $p_{FDR} = 0.021$, $r_{rb} = 0.68$) and Low_m ($W = 289$, $p_{FDR} = 0.031$, $r_{rb} = 0.64$) overall liking judgments. No differences were found for the $High_m > Low_m$

contrast for the VN (W = 178, p_{FDR} = 0.96, r_{rb} = 0.014; see Figure 2A), with this network being *equally modulated* by positive and negative overall judgments of liking (BF_{01} = 4.84). No other networks were significantly modulated by overall ratings or showed significant differences in engagement between the $High_m$ and Low_m parametric modulators after correction for multiple comparisons.

Since the VN resembles a network previously known to respond to stimuli with both positive and negative emotional valence, and in the presented data responds to both high and low overall liking ratings, we conducted an additional exploratory analysis. We compared the spatial component of the VN with several fMRI meta-analyses. First, we used NeuroSynth (a platform for large-scale, automated meta-analysis of fMRI data; www.neurosynth.org; Yarkoni et al., 2011) to calculate an association map based on the search term "*valence"* that resulted in 361 studies (search performed on December 19th, 2024). Second, we obtained the frequency maps from a valence meta-analysis of 397 fMRI and PET studies for regions consistently activated by both positive and negative affective experiences (Lindquist et al., 2016). As a control, we also included NeuroSynth association meta-analyses for *music* (163 studies; search performed on December 19th, 2024), *reward* (922 studies; search performed on December 19th, 2024), and a well-known meta-analysis of the DMN (Smith et al., 2009). We then used the spatial sorting function provided by GIFT, which uses multiple linear regression to compare the spatial pattern of a particular ICA network with a set of templates (Xu, Zhang, et al., 2013). We then spatially compared the VN with the meta-analyses of valence, reward, music, and the DMN. As a control, we repeated the same analysis, but using the DMN and APMN ICA networks obtained from our data (see Tables 2 and 3). We computed two multiple regressions, one per valence meta-analysis (i.e., one analysis included templates for the NeuroSynth meta-analysis of valence, reward, music, and the DMN from Smith and colleagues, 2009; the other included templates for the meta-analysis of valence from Lindquist and colleagues, 2016, the NeuroSynth meta-analyses of reward and music, and the DMN from Smith and colleagues, 2009). These analyses resulted in correlation coefficients (*r*) that

514 indicated the degree of spatial similarity between each ICA network being analyzed (VN, DMN, APMN)
515 and each template.

516 As expected, the highest *r* values for the APMN and DMN were those corresponding to the
517 music and DMN meta-analyses, respectively. For the VN, the highest *r* corresponded to both meta-
518 analyses related to *valence* (see Tables 4 and 5), and not to those corresponding to music, reward, or
519 the DMN. This further strengthens the finding that the liking judgments provided by the participants—
520 regardless of their *positive or negative value*—were supported by a series of brain regions that can
521 respond to stimuli with *positive and negative emotional valence* rather than by a network more
522 specifically related to reward (Koelsch, 2014; Mas-Herrero et al., 2021) or music processing in general
523 (Gordon et al., 2018; Kasdan et al., 2022; Rodriguez-Fornells et al., 2012).

ICA Network	**Valence** NeuroSynth Meta-Analysis	**Music** NeuroSynth Meta-Analysis	**Reward** NeuroSynth Meta-Analysis	**DMN** From Smith et al., 2009
VN	**0.201**	-0.016	0.050	0.061
DMN	-0.014	-0.008	-0.0002	**0.244**
APMN	-0.002	**0.663**	-0.010	0.004

Table 4. Spatial sorting of ICA networks. The results shown here correspond to the multiple regression that included the **NeuroSynth valence meta-analysis.** *r* values show correlations between ICA networks and templates, with the highest correlation value for each ICA network shown in bold.

524

ICA Network	**Valence** From Lindquist et al., 2016	**Music** NeuroSynth Meta-Analysis	**Reward** NeuroSynth Meta-Analysis	**DMN** From Smith et al., 2009
VN	**0.351**	-0.012	0.052	0.061
DMN	0.033	-0.007	-0.002	**0.244**
APMN	-0.017	**0.663**	-0.01	0.004

Table 5. Spatial sorting of ICA networks. The results shown here correspond to the multiple regression that included the **valence meta-analysis from Lindquist and colleagues (2016).** *r* values show correlations between ICA networks and templates, with the highest correlation value for each ICA network shown in bold.

525 Regarding continuous ratings, there was a significant difference between the engagement of the
526 DMN by the slopes for the $High_m > Low_m$ contrast ($W = 62$, $p_{FDR} = 0.033$, $r_{rb} = -0.64$; see Figure 3A and
527 B). Given that the DMN is consistently deactivated in fMRI task-blocked designs and that in our data
528 the DMN was negatively modulated by music listening (see Table 3), these results suggest that the

529 faster the participants became engaged by a musical excerpt (i.e., the faster they rated it), the faster
530 the DMN became disengaged. No other networks were significantly modulated by the slope of the
531 continuous ratings or showed significant differences in engagement between the $High_m$ and Low_m
532 parametric modulators after correction for multiple comparisons.

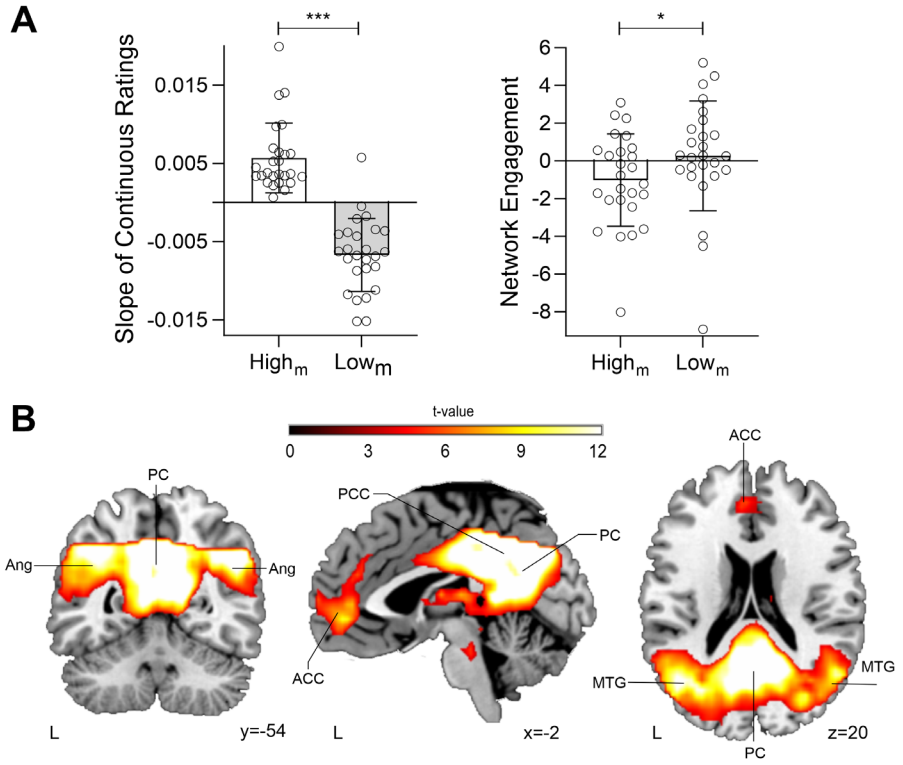


Figure 3. The Default Mode Network. A. On the left, mean (plus SD) behavioral slopes (representing the speed by which the participants provided positive or negative continuous ratings for a musical excerpt) for $High_m$ (white) and Low_m (grey) trials (corrected using the Medium behavioral slopes). In this case, since we classified songs into the highly or lowly liked condition using the overall ratings and not any measure extracted from the continuous ratings, we can compute statistics without double dipping. The corrected slopes were significantly higher for $High_m$ than for Low_m (Wilcoxon signed-rank test, $W = 351$, $p<0.001$, $r_{rb} = 1$). This suggests that participants were faster at rating songs they liked compared to ones they disliked. On the right panel, mean beta values (plus SD; also corrected using Medium beta values), showing that the DMN is more negatively modulated by $High_m$ (white) than by Low_m (grey) slopes. This suggests that the faster participants provided positive aesthetic judgments, the more disengaged the DMN became. Each circle represents data from one participant. **B.** In red-yellow**,** spatial map corresponding to the DMN ($p < 0.05$, FWE-corrected at the cluster level). Neurological convention is used with MNI coordinates at the bottom right of each slice. L, Left Hemisphere; ACC, Anterior Cingulate Cortex; PCC, Posterior Cingulate Cortex; PC, Precuneus; MTG, Middle Temporal Gyrus; Ang, Angular Gyrus. * $p < 0.05$,*** $p<0.001$.

533 To further clarify the possible functional roles of the VN and DMN in aesthetic appreciation, we
534 computed one final exploratory analysis focusing on individual differences in aesthetic responsivity.
535 Using Spearman's correlation, we assessed the relationship between the AReA (Schlotz et al., 2021)
536 and the general engagement of the VN and DMN with the musical excerpts (using the beta values from

537 the simplest model that only included one general musical condition). There was a significant correlation
538 between the engagement of the VN and the AReA: the higher a participant's general responsiveness
539 to artistic stimuli, the more the VN was engaged while listening to music (r = 0.496, p = 0.01; see Figure
540 4). There was no significant relationship between the AReA and DMN (r = 0.331, p = 0.098).

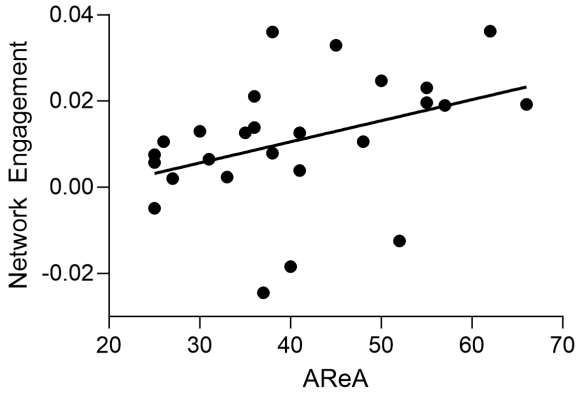


Figure 4. Individual differences in aesthetic responsiveness and VN network engagement. The scatter plot shows the significant association (r = 0.496, p = 0.01) between the AReA (which measures individual differences in general responsiveness to artistic stimuli) and the network engagement of the VN with music (each dot represents data from one participant). As a measure of general network engagement to music, we used the beta values extracted from the simplest model that only included one general musical condition.

541 **DISCUSSION**

542 This study assesses the neural underpinnings of both continuous and overall liking judgments
543 of music. Behaviorally, as expected, there was low agreement in liking ratings across participants.
544 Participants' preferences were revealed very early in the presentation of a piece and were sustained
545 over time (Belfi et al., 2018; Isik & Vessel, 2021), with the end value of the continuous liking judgement
546 being the measure most highly related to the overall liking rating (see Figure 1). These behavioral
547 results collected inside of the MRI scanner were replicated in a different sample of online participants
548 (see Appendix Figure 1). Regarding brain activity, the DMN was associated with the slope of the
549 continuous ratings, suggesting that the faster participants liked a musical excerpt, the more rapidly
550 disengaged this network became (Figure 3). Interestingly, a network comprised of regions known to
551 respond to stimuli with both positive and negative emotional *valence* (the VN, see Figure 2, Table 2;
552 Lindquist et al., 2016) was modulated by musical pieces rated both high and low on liking. Our results

553 regarding the VN extend to the individual differences domain, as participants with a greater general
554 responsiveness to artistic stimuli engaged the VN more while listening to music (Figure 4).

555 One aim of this work was to investigate the relationship between continuous and overall ratings
556 of liking during music listening. In agreement with previous research (Schafer et al., 2014), we found
557 that the final value (i.e., the value of the continuous trace at the end of an excerpt) showed the highest
558 correlation with the overall ratings. While other metrics such as the average or peak value of the
559 continuous ratings have also been related to overall ratings (Rozin et al., 2004; Schafer et al., 2014),
560 in both our fMRI and online experiments the end value showed a stronger statistical relationship with
561 the overall rating (see Figure 1C and Appendix Figure 1C). This difference in the relationship of the
562 average, peak, and final values with the overall rating could be due to the specific nature of the
563 measures: here, listeners rated *liking*, while prior work often focused on ratings of *emotional intensity*
564 of the music. It may also be the case that aesthetic judgments such as liking grow stronger over time,
565 reflecting temporal integration (Belfi et al., 2018; Brielmann & Pelli, 2017; Isik & Vessel, 2019, 2021).
566 As shown in Figure 1C and Appendix Figure 1C, on average, liked pieces tended to be rated as
567 increasingly more liked over the course of the trial, while ratings of disliked pieces grew more disliked.
568 This suggests that listeners form a quick initial impression that strengthens (or is reinforced) during the
569 course of the piece, which may explain why we found the strongest relationship between the end and
570 overall ratings (Jacobsen & Beudt, 2017b; Kahneman et al., 1993). The results might also align with
571 evidence accumulation models, which propose that humans make decisions by gradually gathering
572 evidence in favor of a particular choice until a threshold is reached, triggering the decision (Boag et al.,
573 2023; Parker and Ramsey, 2024). Conversely, emotional intensity (e.g., chills) may show more
574 variability over the course of the piece, which is why the peak or average may be a more representative
575 metric of that reaction (Martínez-Molina et al., 2016; Mas-Herrero et al., 2014; Salimpoor et al., 2011).

576 The fMRI analyses suggest an overlap between regions supporting liking judgments (the VN;
577 see Figure 2) and areas previously known to be associated with the processing of emotional stimuli

regardless of their positive or negative valence. Indeed, there is substantial overlap between our VN and modality-general (i.e., not specific to musical stimuli; Chikazoe et al., 2014; Kim et al., 2017) fMRI meta-analyses of valence (see Tables 4 and 5; Bartra et al., 2013; Lindquist et al., 2016; X. Liu et al., 2011). Several of the regions that form the VN—including the amygdala, hippocampus, thalamus, striatum, and the ventromedial prefrontal cortex—have also been related to emotional processing of music and other stimuli (Blood & Zatorre, 2001; Koelsch, 2014, 2020; Kühn & Gallinat, 2012; Martínez-Molina et al., 2016; Mas-Herrero et al., 2021; Salimpoor et al., 2011). While these studies usually focused on brain activity elicited during peak moments of high pleasure (e.g., chills), here we show that these regions are also modulated by the *liking* of the musical piece, both for high and low rated musical pieces.

It is noteworthy that individual differences in our participants' preferences for abstract aesthetic stimuli (e.g., music, paintings, natural landscapes, architecture, dance, poetry) were also related to the engagement of the VN (see Figure 4), which suggests that this network might be related to modality-general aesthetic judgments (Jacobsen & Beudt, 2017a). This is in agreement with previous findings showing that in order to assign emotional value to abstract stimuli, an interaction between perceptual (e.g., visual, auditory) and reward, valuation and emotion-related circuits (e.g., the VN) is essential (Isik & Vessel, 2021; C. Liu et al., 2017; Mori & Zatorre, 2024; Singer et al., 2016; Trost et al., 2012; Zatorre & Salimpoor, 2013), and that individual differences in the structure and function of the neural substrates that support this interaction can affect the way we perceive and judge artistic stimuli (Fasano et al., 2023; Gold et al., 2019; Loui et al., 2017; Martínez-Molina et al., 2016, 2019; Matthews et al., 2024; Sachs et al., 2016). Future research could examine whether the VN can be used as a neural proxy for measuring individual differences in aesthetic judgments of liking across different domains (music, poetry, visual art, etc.).

The data show that the DMN was modulated by the speed of the positive continuous liking ratings: the faster the participants liked a musical excerpt, the more disengaged the DMN became (see

Figure 3). This aligns with the DMN being consistently deactivated in task-based fMRI studies (Raichle et al., 2001; Smith et al., 2009). Previous work in the visual domain has reported greater activation for high versus low appeal artworks in several regions of the DMN, including the medial prefrontal cortex and anterior cingulate, orbitofrontal cortex, inferior frontal gyrus, insula and the striatum (Brown et al., 2011; Di et al., 2011; Jacobsen et al., 2006; Kawabata & Zeki, 2004; Satpute & Lindquist, 2019; Vessel et al., 2012). Engagement of the DMN has also been reported for highly aesthetically pleasing visual artwork (Belfi et al., 2019; Vessel et al., 2012, 2013, 2019). While it remains unclear why such a different (i.e., opposite) pattern was observed in many of these regions and networks in this study, one potential difference may be the role of prediction, expectations, and top-down processing for aesthetic experiences of visual art versus for extended musical pieces. Whereas a static image of a painting may invite active exploration, reflection, and top-down sense making, predictions and expectation in music are more directly determined by temporal, harmonic, and tonal structural properties of a composition that unfold over time. During a musical clip, it is possible that both top-down predictions and aesthetic pleasure vary in a manner that leads to a less direct relationship between the overall aesthetic judgment at the end of the piece and average network engagement. In addition, it is possible that some musical clips are aesthetically pleasing for different reasons than visual art, such as the urge to dance (e.g., "groove"; Janata et al., 2012; Matthews et al., 2020) or to hum along, that rely less on top-down reflective processes thought to be mediated by the DMN. A second possible interpretation consistent with the observed results is that regions within the DMN play a role in the initial evaluation of the song, which, when it happens quickly, engages the DMN for a shorter period of time than when that evaluation takes longer (Vessel et al., 2019).

The DMN has been shown to be also engaged during the aesthetic experience of music, with its activation influenced by stimuli of both positive and negative valence (Alluri et al., 2017; Koelsch et al., 2022; Y. Liu et al., 2021; Reybrouck et al., 2018; Sachs et al., 2020). For example, previous research has shown that liked music resulted in higher engagement of DMN brain regions (Wilkins et al., 2014).

Other work showed that sad music (as compared to happy music) increased functional connectivity among regions within the DMN (Taruffi et al., 2017). Moreover, one of the aforementioned studies showed that areas located in the DMN (e.g., precuneus) were activated for both liked and disliked music (Wilkins et al., 2014). The DMN, as identified in our ICA analysis, appeared more strongly disengaged in response to faster ratings of highly liked music (see Figure 3), a finding that initially seems to contradict previous research. However, the VN—which was influenced by both highly liked and disliked music—includes DMN regions such as the anterior cingulate cortex and precuneus (see Table 2). Moreover, as shown in Tables 4 and 5, the VN exhibits stronger spatial correlation with the DMN than with the broader meta-analysis of music, indicating substantial overlap between the VN and classical DMN areas. Because of this overlap and given that prior studies have reported increased DMN engagement during aesthetic experiences of music, our findings can be reconciled with those earlier observations. Another possible interpretation is that when regions of the DMN are engaged as parts of other networks (as happens here with the VN), overall DMN engagement may be suppressed; this can happen in some cases of music listening, especially as compared with engagement with visual art (Starr 2023a, 2023b). For example, listening to music or evaluating it may introduce processing demands that conflict with overall engagement of the DMN. Particular features of music, like tonal “brightness” (e.g., sounds like the high notes of trumpets or other brass instruments) recruit regions of the DMN (Alluri et al, 2012). By contrast, the DMN has been found to be engaged with listening to familiar music when individuals are not asked to make liking judgments (“free listening” paradigms; Wilkins et al. 2014), in line with task-specific suppression of the DMN. Finally, one possibility that could be explored in future studies is whether performing network decomposition on data collected during active music listening, vs. during independent rest scans, results in a different grouping of these overlapping regions.

One limitation of this study is that all participants had a predominantly Western musical background, which may limit the generalizability of our findings to other cultural contexts. In addition, we confined our study to two genres—classical and electronic—excluding “unconventional”

compositions within those genres (e.g., atonal classical works) and omitting less popular styles such as punk or free jazz. Consequently, our results may not extend to situations where listeners, despite unfamiliarity or initial dislike of a genre, ultimately find themselves enjoying the music.

In summary, we provide data linking behavior and brain function to better characterize a fundamental process: the way human listeners judge the aesthetic value of music. Participants determined whether they liked or disliked a particular piece (and maintained that decision) very early in time and their decisions—regardless of the positive or negative value of the ratings—were associated with a particular network composed of areas previously related to respond to modality-general valence. Our results extend to the individual differences domain, as we show that the engagement of the VN during music listening is predictive of general aesthetic experience. Taken together, our results advance the hypothesis that our individual preferences in aesthetic choices are modulated by a complex interaction of circuits associated with emotion and reward.

DATA AND CODE AVAILABILITY

Behavioral data and ICA networks are provided as supplementary materials. The code used followed standard behavioral and fMRI procedures.

AUTHOR CONTRIBUTIONS

Author contributions are listed following the CRediT (Contributor Roles Taxonomy) format. Pablo Ripollés: Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. Amy M. Belfi: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Writing – original draft, Writing – review & editing. Anna Kasdan: Data curation, Formal analysis, Investigation, Writing – review & editing. Edward A. Vessel: Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Software, Supervision, Writing – review & editing. Andrea R. Halpern: Conceptualization, Methodology, Resources, Writing – review & editing. Jess Rowland: Conceptualization, Resources,

Writing – review & editing. Rob Hopkins: Conceptualization, , Writing – review & editing. G. Gabrielle Starr: Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review & editing. David Poeppel: Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review & editing.

FUNDING

This work was supported by funding from the NYU Global Institute for Advanced Study.

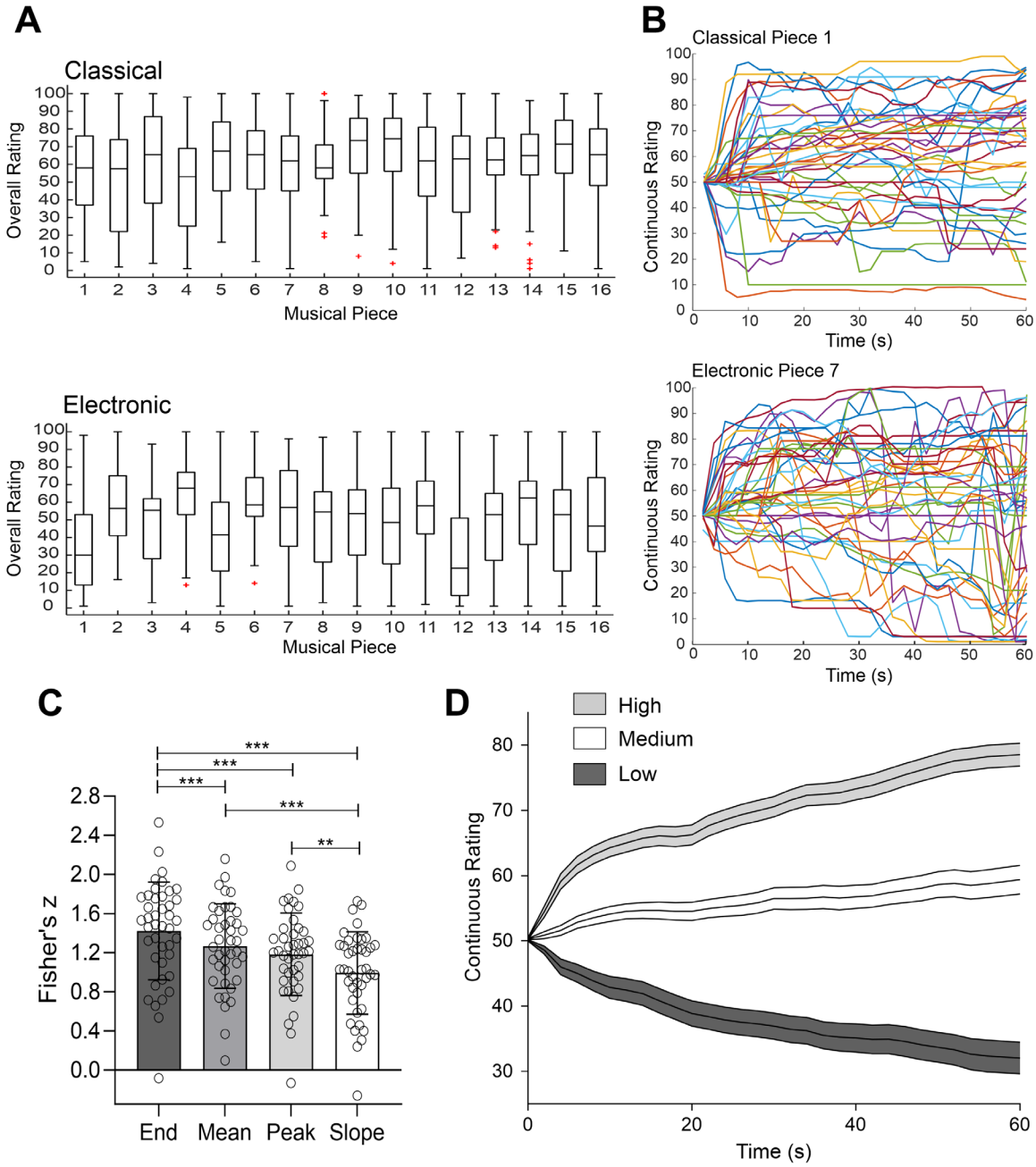
DECLARATION OF COMPETING INTEREST

The authors declare no potential conflicts of interest.

ACKNOWLEDGMENTS

We would like to thank Danielle Retcho for assistance with creating stimulus materials.

APPENDIX



Appendix Figure 1. Online behavioral replication experiment. A. Boxplots illustrating participants' overall ratings for each musical piece. Center line indicates median, bottom and top edges of the box indicate 25th and 75th percentiles. Whiskers extend to the most extreme data points not considered outliers, and red + symbols indicate individual outliers. **B.** Individual continuous responses from all participants for two representative musical excerpts (top, classical piece number 1; bottom, electronic piece number 7). Each line represents the continuous rating provided by a specific participant. **C**. Barplots (mean ± SD) showing the strength (using Fisher's z transform) of the correlation between the *overall* rating and different measures extracted from the continuous ratings, including the *end, mean,* and *peak value*, and the *slope.* Each circle represents data from one participant. ** $p < 0.005$, *** $p < 0.001$. **D.** Mean (solid lines) plus SEM (shaded error bands) of *continuous* ratings (between 1 and 100) across participants are shown for High (grey), Medium (white) and Low (black) musical excerpts classified using the overall ratings.

REFERENCES

Abrams, E., Namballa, R., He, R., Poeppel, D., & Ripollés, P. (2024). Elevator music as a tool for the quantitative characterization of reward. *Annals of the New York Academy of Sciences*, *1535*(1), Article 1. https://doi.org/10.1111/nyas.15131

Abrams, E., Vidal, E. M., Pelofi, C., & Ripollés, P. (2022). Retrieving musical information from neural data: How cognitive features enrich acoustic ones. *ISMIR*, 160–168. https://www.academia.edu/download/98780568/000018.pdf

Aliko, S., Huang, J., Gheorghiu, F., Meliss, S., & Skipper, J. I. (2020). A naturalistic neuroimaging database for understanding the brain using ecological stimuli. *Scientific Data*, *7*(1), 347. https://doi.org/10.1038/s41597-020-00680-2

Alluri, V., Toiviainen, P., Burunat, I., Kliuchko, M., Vuust, P., & Brattico, E. (2017). Connectivity patterns during music listening: Evidence for action-based processing in musicians. *Hum Brain Mapp*, *38*, 2955–2970.

Alluri, V., Toiviainen, P., Lund, T. E., Wallentin, M., Vuust, P., Nandi, A. K., Ristaniemi, T., & Brattico, E. (2013). From Vivaldi to Beatles and back: Predicting lateralized brain responses to music. *Neuroimage*, *83*, 627–636.

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

Altmann, U., Bohrn, I. C., Lubrich, O., Menninghaus, W., & Jacobs, A. M. (2012). The power of emotional valence—From cognitive to affective processes in reading. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00192

Ashburner, J., & Friston, K. J. (2005). Unified segmentation. *Neuroimage*, *26*, 839–851.

Bachorik, J. P., Bangert, J., Loui, P., Larke, K., Berger, J., Rowe, R., & Schlaug, G. (2009). *Emotion in Motion: Investigating the time-course of emotional judgements of musical stimuli*. Music Perception355-364.

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage*, *76*, 412–427.

Beldzik, E., Domagalik, A., Daselaar, S., Fafrowicz, M., Froncisz, W., Oginska, H., & Marek, T. (2013). Contributive sources analysis: A measure of neural networks' contribution to brain activations. *Neuroimage*, *76*, 304–312.

Belfi, A. M., Evans, E., Heskje, J., Bruss, J., & Tranel, D. (2017). Musical anhedonia after focal brain damage. *Neuropsychologia*, *97*, 29–37.

Belfi, A. M., Kasdan, A., Rowland, J., Vessel, E. A., Starr, G. G., & Poeppel, D. (2018). Rapid timing of musical aesthetic judgments. *J Exp Psychol Gen*.

Belfi, A. M., Samson, D. W., Crane, J., & Schmidt, N. L. (2021). Aesthetic Judgments of Live and Recorded Music: Effects of Congruence Between Musical Artist and Piece. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.618025

Belfi, A. M., Vessel, E. A., Brielmann, A., Isik, A. I., Chatterjee, A., Leder, H., Pelli, D. G., & Starr, G. G. (2019). Dynamics of aesthetic experience are reflected in the default-mode network. *NeuroImage*, *188*, 584–597. https://doi.org/10.1016/j.neuroimage.2018.12.017

Belfi, A. M., Vessel, E. A., & Starr, G. G. (2017). Individual ratings of vividness predict aesthetic appeal in poetry. *Psychology of Aesthetics, Creativity and the Arts*.

Belfi, A. M., & Loui, P. (2020). Musical anhedonia and rewards of music listening: Current advances and a proposed model. *Annals of the New York Academy of Sciences*, *1464*(1), 99–114. https://doi.org/10.1111/nyas.14241

Belfi, A. M. (2019). Emotional valence and vividness of imagery predict aesthetic appeal in music. *Psychomusicology: Music, Mind, and Brain*, *29*(2–3), 128–135. https://doi.org/10.1037/pmu0000232

Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Comput*, *7*, 1129–1159.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

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), Article 20.

Boag, R. J., Strickland, L., Heathcote, A., Neal, A., Palada, H., & Loft, S. (2023). Evidence accumulation modelling in the wild: Understanding safety-critical decisions. *Trends in Cognitive Sciences*, *27*(2), 175–188. https://doi.org/10.1016/j.tics.2022.11.009

Brainard D. H. (1997). The Psychophysics Toolbox. *Spat Vis*, *10*, 433–436.

Brattico, E., Bogert, B., Alluri, V., Tervaniemi, M., Eerola, T., & Jacobsen, T. (2015). It's Sad but I Like It: The Neural Dissociation Between Musical Emotions and Liking in Experts and Laypersons. *Front Hum Neurosci*, *9*(676).

Brattico, E., Bogert, B., & Jacobsen, T. (2013). Toward a neural chronometry for the aesthetic experience of music. *Front Psychol*, *4*(206).

Brielmann, A. A., & Pelli, D. G. (2017). Beauty Requires Thought. *Current Biology: CB*, *27*(10), 1506-1513.e3. https://doi.org/10.1016/j.cub.2017.04.018

Brown, S., Gao, X., Tisdelle, L., Eickhoff, S. B., & Liotti, M. (2011). Naturalizing aesthetics: Brain areas for aesthetic appraisal across sensory modalities. *Neuroimage*, *58*, 250–258.

Bruns, A., Pombo, M., Ripollés, P., & Pelli, D. G. (2023). Emotions of subject and object affect beauty differently for images and music. *Journal of Vision*, *23*(13), 6. https://doi.org/10.1167/jov.23.13.6

Burunat, I., Alluri, V., Toiviainen, P., Numminen, J., & Brattico, E. (2014). Dynamics of brain activity underlying working memory for music in a naturalistic condition. *Cortex*, *57*, 254–269.

Burunat, I., Toiviainen, P., Alluri, V., Bogert, B., Ristaniemi, T., Sams, M., & Brattico, E. (2016). The reliability of continuous brain responses during naturalistic listening to music. *Neuroimage*, *124*, 224–231.

Calhoun, V. D., Adali, T., McGinty, V. B., Pekar, J. J., Watson, T. D., & Pearlson, G. D. (2001). fMRI activation in a visual-perception task: Network of areas detected using the general linear model and independent components analysis. *Neuroimage*, *14*, 1080–1088.

Calhoun, V. D., Kiehl, K. A., & Pearlson, G. D. (2008). Modulation of temporally coherent brain networks estimated using ICA at rest and during cognitive tasks. *Hum Brain Mapp*, *29*, 828–838.

Carone, B. J., & Ripollés, P. (2024). SoundSignature: What Type of Music do you Like? *2024 IEEE 5th International Symposium on the Internet of Sounds (IS2)*, 1–10. https://doi.org/10.1109/IS262782.2024.10704174

Chatterjee, A., & Vartanian, O. (2014). Neuroaesthetics. *Trends Cogn Sci*, *18*, 370–375.

Che, J., Sun, X., Skov, M., & Nadal, M. (2024). Judging beauty and liking: The effects of personal intuitions and task instructions. *Psychology of Aesthetics, Creativity, and the Arts*, No Pagination Specified-No Pagination Specified. https://doi.org/10.1037/aca0000702

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97. https://doi.org/10.1016/j.jbef.2015.12.001

Cheung, V. K. M., Harrison, P. M. C., Meyer, L., Pearce, M. T., Haynes, J.-D., & Koelsch, S. (2019). Uncertainty and Surprise Jointly Predict Musical Pleasure and Amygdala, Hippocampus, and

Auditory Cortex Activity. *Current Biology: CB*, *29*(23), 4084-4092.e4. https://doi.org/10.1016/j.cub.2019.09.067

Chikazoe, J., Lee, D. H., Kriegeskorte, N., & Anderson, A. K. (2014). Population coding of affect across stimuli, modalities and individuals. *Nat Neurosci*, *17*, 1114–1122.

Chuan-Peng, H., Huang, Y., Eickhoff, S. B., Peng, K., & Sui, J. (2020). Seeking the "Beauty Center" in the Brain: A Meta-Analysis of fMRI Studies of Beautiful Human Faces and Visual Art. *Cognitive, Affective & Behavioral Neuroscience*, *20*(6), 1200–1215. https://doi.org/10.3758/s13415-020-00827-z

Clemente, A., Friberg, A., & Holzapfel, A. (2023). Relations between perceived affect and liking for melodies and visual designs. *Emotion*, *23*(6), 1584–1605. https://doi.org/10.1037/emo0001141

Clemente, A., Kaplan, T. M., & Pearce, M. T. (2024). Perceptual representations mediate effects of stimulus properties on liking for music. *Annals of the New York Academy of Sciences*, *1533*(1), 169–180. https://doi.org/10.1111/nyas.15106

Cohrdes, C., Wrzus, C., Frisch, S., & Riediger, M. (2017). Tune yourself in: Valence and arousal preferences in music-listening choices from adolescence to old age. *Developmental Psychology*, *53*(9), 1777–1794. https://doi.org/10.1037/dev0000362

Cohrdes, C., Wrzus, C., Wald-Fuhrmann, M., & Riediger, M. (2020). "The sound of affect": Age differences in perceiving valence and arousal in music and their relation to music characteristics and momentary mood. *Musicae Scientiae*, *24*(1), 21–43. https://doi.org/10.1177/1029864918765613

Cong, F., Puolivali, T., Alluri, V., Sipola, T., Burunat, I., Toiviainen, P., Nandi, A. K., Brattico, E., & Ristaniemi, T. (2014). Key issues in decomposing fMRI during naturalistic and continuous music experience with independent component analysis. *J Neurosci Methods*, *223*, 74–84.

Criscuolo, A., Pando-Naude, V., Bonetti, L., Vuust, P., & Brattico, E. (2022). An ALE meta-analytic review of musical expertise. *Scientific Reports*, *12*(1), 11726. https://doi.org/10.1038/s41598-022-14959-4

Crump, M. J., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PloS One*, *8*(3), Article 3.

Curzel, F., Osiurak, F., Trân, E., Tillmann, B., Ripollés, P., & Ferreri, L. (2024). Enhancing musical pleasure through shared musical experience. *iScience*, *27*(6). https://doi.org/10.1016/j.isci.2024.109964

Daly, I., Williams, D., Hallowell, J., Hwang, F., Kirke, A., Malik, A., Weaver, J., Miranda, E., & Nasuto, S. J. (2015). Music-induced emotions can be predicted from a combination of brain activity and acoustic features. *Brain and Cognition*, *101*, 1–11. https://doi.org/10.1016/j.bandc.2015.08.003

Desai, R. H., Choi, W., Lai, V. T., & Henderson, J. M. (2016). Toward Semantics in the Wild: Activation to Manipulable Nouns in Naturalistic Reading. *Journal of Neuroscience*, *36*(14), 4050–4055. https://doi.org/10.1523/JNEUROSCI.1480-15.2016

Di, D. C., Canessa, N., Cappa, S. F., & Rizzolatti, G. (2011). Specificity of esthetic experience for artworks: An FMRI study. *Front Hum Neurosci*, *5*(139).

Di Liberto, G. M., Pelofi, C., Bianco, R., Patel, P., Mehta, A. D., Herrero, J. L., de Cheveigné, A., Shamma, S., & Mesgarani, N. (2020). Cortical encoding of melodic expectations in human temporal cortex. *eLife*, *9*, e51784. https://doi.org/10.7554/eLife.51784

Eerola, T., Lartillot, O., & Toiviainen, P. (2009). Prediction of multidimensional emotional ratings in music from audio using multivariate regression models. *Proceedings of the 10th International Society for Music Information Retrieval Conference, ISMIR*, 621–626.

Eerola, T., Vuoskoski, J. K., Peltola, H.-R., Putkinen, V., & Schäfer, K. (2018). An integrative review of the enjoyment of sadness associated with music. *Physics of Life Reviews*, *25*, 100–121. https://doi.org/10.1016/j.plrev.2017.11.016

Egermann, H., & Reuben, F. (2020). “Beauty Is How You Feel Inside”: Aesthetic Judgments Are Related to Emotional Responses to Contemporary Music. *Frontiers in Psychology*, *11*, 510029. https://doi.org/10.3389/fpsyg.2020.510029

Fasano, M. C., Cabral, J., Stevner, A., Vuust, P., Cantou, P., Brattico, E., & Kringelbach, M. L. (2023). The early adolescent brain on music: Analysis of functional dynamics reveals engagement of orbitofrontal cortex reward system. *Human Brain Mapping*, *44*(2), 429–446. https://doi.org/10.1002/hbm.26060

Ferreri, L., Mas-Herrero, E., Zatorre, R. J., Ripollés, P., Gomez-Andres, A., Alicart, H., & Rodriguez-Fornells, A. (2019). Dopamine modulates the reward experiences elicited by music. *Proceedings of the National Academy of Sciences*, *116*(9), Article 9.

Forn, C., Ripolles, P., Cruz-Gomez, A. J., Belenguer, A., Gonzalez-Torre, J. A., & Avila, C. (2013). Task-load manipulation in the Symbol Digit Modalities Test: An alternative measure of information processing speed. *Brain Cogn*, *82*, 152–160.

Frame, J., Mehl, K., Head, K., & Belfi, A. M. (2024). The influence of sensory modality on aesthetic judgments of poetry. *Psychology of Aesthetics, Creativity, and the Arts*, No Pagination Specified-No Pagination Specified. https://doi.org/10.1037/aca0000673

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.

Gingras, B., Marin, M. M., & Fitch, W. T. (2014). Beyond intensity: Spectral features effectively predict music-induced subjective arousal. *Quarterly Journal of Experimental Psychology*, *67*, 1428–1446.

Gingras, B., Marin, M. M., Puig-Waldmüller, E., & Fitch, W. T. (2015). The Eye is Listening: Music-Induced Arousal and Individual Differences Predict Pupillary Responses. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00619

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 of the United States of America*, *116*(8), 3310–3315. https://doi.org/10.1073/pnas.1809855116

Gordon, C. L., Cobb, P. R., & Balasubramaniam, R. (2018). Recruitment of the motor system during music listening: An ALE meta-analysis of fMRI data. *PLOS ONE*, *13*(11), e0207213. https://doi.org/10.1371/journal.pone.0207213

Groves, K., Farbood, M. M., Ripolles, P., & Zuanazzi, A. (2023). *Through the lens of music: Instrumental soundtrack listening elicits extra-musical meaning*. OSF. https://doi.org/10.31234/osf.io/aj97z

Halpern, A. R., Ly, J., Elkin-Frankston, S., & O’Connor, M. G. (2008). I know what I like": Stability of aesthetic preference in Alzheimer’s patients. *Brain Cogn*, *66*, 65–72.

Haumann, N. T., Lumaca, M., Kliuchko, M., Santacruz, J. L., Vuust, P., & Brattico, E. (2021). Extracting human cortical responses to sound onsets and acoustic feature changes in real music, and their relation to event rate. *Brain Research*, *1754*, 147248. https://doi.org/10.1016/j.brainres.2020.147248

Huron, D., & Vuoskoski, J. K. (2020). On the Enjoyment of Sad Music: Pleasurable Compassion Theory and the Role of Trait Empathy. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.01060

Huth, A. G., Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, *532*, 453–458.

In, J., & Lee, D. K. (2024). Alternatives to the P value: Connotations of significance. *Korean Journal of Anesthesiology*, *77*(3), 316–325. https://doi.org/10.4097/kja.23630

Ishizu, T., & Zeki, S. (2011). Toward a brain-based theory of beauty. *PLoS One*, *6:e21852*.

Isik, A. I., & Vessel, E. A. (2019). Continuous ratings of movie watching reveal idiosyncratic dynamics of aesthetic enjoyment. *PLOS ONE*, *14*(10), e0223896. https://doi.org/10.1371/journal.pone.0223896

Isik, A. I., & Vessel, E. A. (2021). From Visual Perception to Aesthetic Appeal: Brain Responses to Aesthetically Appealing Natural Landscape Movies. *Frontiers in Human Neuroscience*, *15*. https://doi.org/10.3389/fnhum.2021.676032

Jääskeläinen, I. P., Sams, M., Glerean, E., & Ahveninen, J. (2021). Movies and narratives as naturalistic stimuli in neuroimaging. *NeuroImage*, *224*, 117445. https://doi.org/10.1016/j.neuroimage.2020.117445

Jacobsen, T., & Beudt, S. (2017a). Domain generality and domain specificity in aesthetic appreciation. *New Ideas in Psychology*, *47*, 97–102. https://doi.org/10.1016/j.newideapsych.2017.03.008

Jacobsen, T., & Beudt, S. (2017b). Stability and Variability in Aesthetic Experience: A Review. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00143

Jacobsen, T., Buchta, K., Köhler, M., & Schröger, E. (2004). The primacy of beauty in judging the aesthetics of objects. *Psychological Reports*, *94*(3 Pt 2), 1253–1260. https://doi.org/10.2466/pr0.94.3c.1253-1260

Jacobsen, T., Schubotz, R. I., Hofel, L., & Cramon, D. Y. (2006). Brain correlates of aesthetic judgment of beauty. *Neuroimage*, *29*, 276–285.

Janata, P., Tomic, S. T., & Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *J Exp Psychol Gen*, *141*, 54–75.

JASP Team. (2025). *JASP (Version 0.19.3)[Computer software]*. https://jasp-stats.org/

Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews*, *10*(3), 235–266. https://doi.org/10.1016/j.plrev.2013.05.008

Kahneman, D., Friedrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When mpre pain is preferred to less: Adding a better end. *Psychol Sci*, *4*, 401–405.

Kandylaki, K. D., Nagels, A., Tune, S., Kircher, T., Wiese, R., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2016). Predicting “When” in Discourse Engages the Human Dorsal Auditory Stream: An fMRI Study Using Naturalistic Stories. *J Neurosci*, *36*, 12180–12191.

Kaneshiro, B., Nguyen, D. T., Norcia, A. M., Dmochowski, J. P., & Berger, J. (2020). Natural music evokes correlated EEG responses reflecting temporal structure and beat. *NeuroImage*, *214*, 116559. https://doi.org/10.1016/j.neuroimage.2020.116559

Kant, I. (1764). *Observations on the Feeling of the Beautiful and Sublime*.

Kahneman, D., Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When More Pain Is Preferred to Less: Adding a Better End. *Psychological Science,* 4(6), 401-405. https://doi.org/10.1111/j.1467-9280.1993.tb00589.x

Kasdan, A. V., Burgess, A. N., Pizzagalli, F., Scartozzi, A., Chern, A., Kotz, S. A., Wilson, S. M., & Gordon, R. L. (2022). Identifying a brain network for musical rhythm: A functional neuroimaging meta-analysis and systematic review. *Neuroscience & Biobehavioral Reviews*, *136*, 104588. https://doi.org/10.1016/j.neubiorev.2022.104588

Kawabata H., & Zeki S. (2004). Neural correlates of beauty. *J Neurophysiol*, *91*, 1699–1705.

Keller, J., Young, C. B., Kelley, E., Prater, K., Levitin, D. J., & Menon, V. (2013). Trait anhedonia is associated with reduced reactivity and connectivity of mesolimbic and paralimbic reward pathways. *J Psychiatr Res*, *47*, 1319–1328.

Kern, P., Heilbron, M., de Lange, F. P., & Spaak, E. (2022). Cortical activity during naturalistic music listening reflects short-range predictions based on long-term experience. *eLife*, *11*, e80935. https://doi.org/10.7554/eLife.80935

Ki, J. J., Kelly, S. P., & Parra, L. C. (2016). Attention Strongly Modulates Reliability of Neural Responses to Naturalistic Narrative Stimuli. *J Neurosci*, *36*, 3092–3101.

Kim, J., Shinkareva, S. V., & Wedell, D. H. (2017). Representations of modality-general valence for videos and music derived from fMRI data. *Neuroimage*, *148*, 42–54.

Koelsch, S. (2014). Brain correlates of music-evoked emotions. *Nature Reviews Neuroscience*, *15*(3), Article 3.

Koelsch, S. (2020). A coordinate-based meta-analysis of music-evoked emotions. *NeuroImage*, *223*, 117350.

Koelsch, S., Andrews-Hanna, J. R., & Skouras, S. (2022). Tormenting thoughts: The posterior cingulate sulcus of the default mode network regulates valence of thoughts and activity in the brain's pain network during music listening. *Human Brain Mapping*, *43*(2), 773–786. https://doi.org/10.1002/hbm.25686

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kühn, S., & Gallinat, J. (2012). The neural correlates of subjective pleasantness. *NeuroImage*, *61*(1), 289–294. https://doi.org/10.1016/j.neuroimage.2012.02.065

Kuppens, P., Tuerlinckx, F., Russell, J. A., & Barrett, L. F. (2013). The relation between valence and arousal in subjective experience. *Psychological Bulletin*, *139*(4), 917–940. https://doi.org/10.1037/a0030811

Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., Kochunov, P. V., Nickerson, D., Mikiten, S. A., & Fox, P. T. (2000). Automated Talairach atlas labels for functional brain mapping. *Hum Brain Mapp*, *10*, 120–131.

Larsen, J. T., Hershfield, H. E., Stastny, B. J., & Hester, N. (2017). On the relationship between positive and negative affect: Their correlation and their co-occurrence. *Emotion*, *17*, 323–336.

Lartillot, O., & Toiviainen, P. (2007). A Matlab toolbox for musical feature extraction from audio. *Proceedings of the 9th International Conference on Digital Audio Effects*, *237*, 244.

Li, Q., Liu, G., Zhang, Y., Wu, J., & Huang, R. (2024). Neural correlates of musical familiarity: A functional magnetic resonance study. *Cerebral Cortex (New York, N.Y.: 1991)*, *34*(4), bhae177. https://doi.org/10.1093/cercor/bhae177

Lindquist, K. A., Satpute, A. B., Wager, T. D., Weber, J., & Barrett, L. F. (2016). The Brain Basis of Positive and Negative Affect: Evidence from a Meta-Analysis of the Human Neuroimaging Literature. *Cereb Cortex*, *26*, 1910–1922.

Liu, C., Brattico, E., Abu-Jamous, B., Pereira, C. S., Jacobsen, T., & Nandi, A. K. (2017). Effect of Explicit Evaluation on Neural Connectivity Related to Listening to Unfamiliar Music. *Frontiers in Human Neuroscience*, *11*, 611. https://doi.org/10.3389/fnhum.2017.00611

Liu, X., Hairston, J., Schrier, M., & Fan, J. (2011). Common and distinct networks underlying reward valence and processing stages: A meta-analysis of functional neuroimaging studies. *Neurosci Biobehav Rev*, *35*, 1219–1236.

Liu, Y., Lian, W., Zhao, X., Tang, Q., & Liu, G. (2021). Spatial Connectivity and Temporal Dynamic Functional Network Connectivity of Musical Emotions Evoked by Dynamically Changing Tempo. *Frontiers in Neuroscience*, *15*, 700154. https://doi.org/10.3389/fnins.2021.700154

Lizcano-Cortés, F., Gómez-Varela, I., Mares, C., Wallisch, P., Orpella, J., Poeppel, D., & Assaneo, M. F. (2022). Speech-to-Speech Synchronization protocol to classify human participants as high or low auditory-motor synchronizers. *STAR Protocols*, *3*(2), Article 2.

Lopez-Barroso, D., Ripolles, P., Marco-Pallares, J., Mohammadi, B., Munte, T. F., Bachoud-Levi, A. C., Rodriguez-Fornells, A., & Diego-Balaguer, R. (2015). Multiple brain networks underpinning word learning from fluent speech revealed by independent component analysis. *Neuroimage*, *110*, 182–193.

Loui, P., Patterson, S., Sachs, M. E., Leung, Y., Zeng, T., & Przysinda, E. (2017). White Matter Correlates of Musical Anhedonia: Implications for Evolution of Music. *Front Psychol*, *8*(1664).

Lucas, B. J., Schubert, E., & Halpern, A. R. (2010). Perception of Emotion in Sounded and Imagined Music. *Music Perception*, *27*, 399–412.

Lumsden, J., Skinner, A., Coyle, D., Lawrence, N., & Munafo, M. (2017). Attrition from Web-Based Cognitive Testing: A Repeated Measures Comparison of Gamification Techniques. *Journal of Medical Internet Research*, *19*(11), e395. https://doi.org/10.2196/jmir.8473

Luo, Y., Yu, Q., Wu, S., & Luo, Y. (2024). Distinct Neural Bases of Visual Art- and Music-Induced Aesthetic Experiences. *NeuroImage*, 120962. https://doi.org/10.1016/j.neuroimage.2024.120962

Madison, G., & Schiölde, G. (2017). Repeated Listening Increases the Liking for Music Regardless of Its Complexity: Implications for the Appreciation and Aesthetics of Music. *Frontiers in Neuroscience*, *11*. https://doi.org/10.3389/fnins.2017.00147

Madsen, J., Margulis, E. H., Simchy-Gross, R., & Parra, L. C. (2019). Music synchronizes brainwaves across listeners with strong effects of repetition, familiarity and training. *Scientific Reports*, *9*, 3576. https://doi.org/10.1038/s41598-019-40254-w

Maher, J. M., Markey, J. C., & Ebert-May, D. (2013). The Other Half of the Story: Effect Size Analysis in Quantitative Research. *CBE Life Sciences Education*, *12*(3), 345–351. https://doi.org/10.1187/cbe.13-04-0082

Margulis, E. H. (2013). Repetition and emotive communication in music versus speech. *Frontiers in Psychology*, *4*, 167. https://doi.org/10.3389/fpsyg.2013.00167

Martin, S., Mikutta, C., Leonard, M. K., Hungate, D., Koelsch, S., Shamma, S., Chang, E. F., Millán, J. del R., Knight, R. T., & Pasley, B. N. (2018). Neural Encoding of Auditory Features during Music Perception and Imagery. *Cerebral Cortex*, *28*(12), 4222–4233. https://doi.org/10.1093/cercor/bhx277

Martínez-Molina, N., Mas-Herrero, E., Rodríguez-Fornells, A., Zatorre, R. J., & Marco-Pallarés, J. (2016). Neural correlates of specific musical anhedonia. *Proceedings of the National Academy of Sciences*, *113*(46), Article 46.

Martínez-Molina, N., Mas-Herrero, E., Rodríguez-Fornells, A., Zatorre, R. J., & Marco-Pallarés, J. (2019). White Matter Microstructure Reflects Individual Differences in Music Reward Sensitivity. *Journal of Neuroscience*, *39*(25), 5018–5027. https://doi.org/10.1523/JNEUROSCI.2020-18.2019

Mas-Herrero, E., Maini, L., Sescousse, G., & Zatorre, R. J. (2021). Common and distinct neural correlates of music and food-induced pleasure: A coordinate-based meta-analysis of neuroimaging studies. *Neuroscience & Biobehavioral Reviews*, *123*, 61–71.

Mas-Herrero, E., Marco-Pallares, J., Lorenzo-Seva, U., Zatorre, R. J., & Rodriguez-Fornells, A. (2013). Individual differences in music reward experiences. *Music Perception: An Interdisciplinary Journal*, *31*(2), Article 2.

Mas-Herrero, E., Zatorre, R. J., Rodriguez-Fornells, A., & Marco-Pallarés, J. (2014). Dissociation between musical and monetary reward responses in specific musical anhedonia. *Current Biology*, *24*(6), Article 6.

Matthews, T. E., Lumaca, M., Witek, M. A. G., Penhune, V. B., & Vuust, P. (2024). Music reward sensitivity is associated with greater information transfer capacity within dorsal and motor white matter networks in musicians. *Brain Structure & Function*, *229*(9), 2299–2313. https://doi.org/10.1007/s00429-024-02836-x

Matthews, T. E., Witek, M. A. G., Lund, T., Vuust, P., & Penhune, V. B. (2020). The sensation of groove engages motor and reward networks. *NeuroImage*, *214*, 116768. https://doi.org/10.1016/j.neuroimage.2020.116768

McKeown, M. J., Hansen, L. K., & Sejnowski, T. J. (2003). Independent component analysis of functional MRI: What is signal and what is noise? *Current Opinion in Neurobiology*, *13*(5), 620. https://doi.org/10.1016/j.conb.2003.09.012

Menninghaus, W., Wagner, V., Hanich, J., Wassiliwizky, E., Jacobsen, T., & Koelsch, S. (2017). The Distancing-Embracing model of the enjoyment of negative emotions in art reception. *The Behavioral and Brain Sciences*, *40*, e347. https://doi.org/10.1017/S0140525X17000309

Morey, R. D., & Rouder, J. N. (2015). *BayesFactor (Version 0.9.10-2)[Computer software*.

Mori, K. (2022). Decoding peak emotional responses to music from computational acoustic and lyrical features. *Cognition*, *222*, 105010. https://doi.org/10.1016/j.cognition.2021.105010

Mori, K., & Iwanaga, M. (2017). Two types of peak emotional responses to music: The psychophysiology of chills and tears. *Sci Rep*, *7*(46063).

Mori, K., & Zatorre, R. (2024). State-dependent connectivity in auditory-reward networks predicts peak pleasure experiences to music. *PLoS Biology*, *22*(8), e3002732. https://doi.org/10.1371/journal.pbio.3002732

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The musicality of non-musicians: An index for assessing musical sophistication in the general population. *PLoS One*, *9*, 89642.

Oechslin, M. S., Läge, D., & Vitouch, O. (2012). Training of Tonal Similarity Ratings in Non-Musicians: A “Rapid Learning” Approach. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00142

Orpella, J., Assaneo, M. F., Ripollés, P., Noejovich, L., López-Barroso, D., Diego-Balaguer, R. de, & Poeppel, D. (2022). Differential activation of a frontoparietal network explains population-level differences in statistical learning from speech. *PLOS Biology*, *20*(7), e3001712. https://doi.org/10.1371/journal.pbio.3001712

Orpella, J., Bowling, D. L., Tomaino, C., & Ripollés, P. (2024). Effects of Music Advertised to Support Focus on Mood and Processing Speed. *PLOS ONE*.

Padrao, G., Mallorqui, A., Cucurell, D., Marco-Pallares, J., & Rodriguez-Fornells, A. (2013). Neurophysiological differences in reward processing in anhedonics. *Cogn Affect Behav Neurosci*, *13*, 102–115.

Parker, S., & Ramsey, R. (2024). What can evidence accumulation modelling tell us about human social cognition? *Quarterly Journal of Experimental Psychology*, *77*(3), 639–655. https://doi.org/10.1177/17470218231176950

Pearce, M. T., Zaidel, D. W., Vartanian, O., Skov, M., Leder, H., Chatterjee, A., & Nadal, M. (2016). Neuroaesthetics: The Cognitive Neuroscience of Aesthetic Experience. *Perspect Psychol Sci*, *11*, 265–279.

Pelowski, M., Markey, P. S., Forster, M., Gerger, G., & Leder, H. (2017). Move me, astonish me… delight my eyes and brain: The Vienna Integrated Model of top-down and bottom-up processes in Art Perception (VIMAP) and corresponding affective, evaluative, and neurophysiological correlates. *Physics of Life Reviews*, *21*, 80–125. https://doi.org/10.1016/j.plrev.2017.02.003

Pinho, A. L., MO, Fransson, P., Eriksson, H., & Ullen, F. (2014). Connecting to create: Expertise in musical improvisation is associated with increased functional connectivity between premotor and prefrontal areas. *J Neurosci*, *34*, 6156–6163.

Pombo, M., Igdalova, A., & Pelli, D. G. (2024). Consensus and contention in beauty judgment. *iScience*, *27*(7). https://doi.org/10.1016/j.isci.2024.110213

Putkinen, V., Nazari-Farsani, S., Seppälä, K., Karjalainen, T., Sun, L., Karlsson, H. K., Hudson, M., Heikkilä, T. T., Hirvonen, J., & Nummenmaa, L. (2021). Decoding Music-Evoked Emotions in the Auditory and Motor Cortex. *Cerebral Cortex*, *31*(5), 2549–2560. https://doi.org/10.1093/cercor/bhaa373

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proc Natl Acad Sci U S A*, *98*, 676–682.

Reybrouck, M., Vuust, P., & Brattico, E. (2018). Brain Connectivity Networks and the Aesthetic Experience of Music. *Brain Sciences*, *8*(6), 107. https://doi.org/10.3390/brainsci8060107

Ripollés, P., Ferreri, L., Mas-Herrero, E., Alicart, H., Gómez-Andrés, A., Marco-Pallares, J., Antonijoan, R. M., Noesselt, T., Valle, M., Riba, J., & Rodriguez-Fornells, A. (2018).

Intrinsically regulated learning is modulated by synaptic dopamine signaling. *eLife*, *7*, e38113. https://doi.org/10.7554/eLife.38113

Robert, P., Cang, M. P. V., Mercier, M., Trébuchon, A., Bartolomei, F., Arnal, L. H., Morillon, B., & Doelling, K. (2024). *Multi-stream predictions in human auditory cortex during natural music listening* (p. 2024.11.27.625704). bioRxiv. https://doi.org/10.1101/2024.11.27.625704

Rodriguez-Fornells, A., Rojo, N., Amengual, J. L., Ripollés, P., Altenmüller, E., & Münte, T. F. (2012). The involvement of audio-motor coupling in the music-supported therapy applied to stroke patients. *Annals of the New York Academy of Sciences*, *1252*, 282–293. https://doi.org/10.1111/j.1749-6632.2011.06425.x

Rouder, J. N., & Morey, R. D. (2012). Default Bayes factors for model selection in regression. *Multivariate Behavioral Research*, *47*(6), Article 6.

Rozin, A., Rozin, P., & Goldberg, E. (2004). The feeling of music past: How listeners remember musical affect. *Music Perception*, *4*, 15–39.

Saarimäki, H. (2021). Naturalistic Stimuli in Affective Neuroimaging: A Review. *Frontiers in Human Neuroscience*, *15*, 675068. https://doi.org/10.3389/fnhum.2021.675068

Sacheli, L. M., Tomasetig, G., Musco, M. A., Pizzi, S., Bottini, G., Pizzamiglio, L., & Paulesu, E. (2022). The unexplored link between aesthetic perception and creativity: A theory-driven meta-analysis of fMRI studies in the visual domain. *Neuroscience and Biobehavioral Reviews*, *140*, 104768. https://doi.org/10.1016/j.neubiorev.2022.104768

Sachs, M. E., Ellis, R. J., Schlaug, G., & Loui, P. (2016). Brain connectivity reflects human aesthetic responses to music. *Soc Cogn Affect Neurosci*, *11*, 884–891.

Sachs, M. E., Habibi, A., Damasio, A., & Kaplan, J. T. (2020). Dynamic intersubject neural synchronization reflects affective responses to sad music. *NeuroImage*, *218*, 116512. https://doi.org/10.1016/j.neuroimage.2019.116512

Salakka, I., Pitkäniemi, A., Pentikäinen, E., Mikkonen, K., Saari, P., Toiviainen, P., & Särkämö, T. (2021). What makes music memorable? Relationships between acoustic musical features and music-evoked emotions and memories in older adults. *PLOS ONE*, *16*(5), e0251692. https://doi.org/10.1371/journal.pone.0251692

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), Article 2.

Salimpoor, V. N., dB, 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*, 216–219.

Satpute, A. B., & Lindquist, K. A. (2019). The default mode network's role in discrete emotion. *Trends in Cognitive Sciences*, *23*(10), 851–864. https://doi.org/10.1016/j.tics.2019.07.003

Schafer, T., Zimmermann, D., & Sedlmeier, P. (2014). How we remember the emotional intensity of past musical experiences. *Front Psychol*, *5*(911).

Schlotz, W., Wallot, S., Omigie, D., Masucci, M. D., Hoelzmann, S. C., & Vessel, E. A. (2021). The Aesthetic Responsiveness Assessment (AReA): A screening tool to assess individual differences in responsiveness to art in English and German. *Psychology of Aesthetics, Creativity, and the Arts*, *15*(4), 682–696. https://doi.org/10.1037/aca0000348

Schubert, E. (2004). Modeling Perceived Emotion With Continuous Musical Features. *Music Perception*, *21*(4), 561–585.

Schubert, E. (2024). Liking music with and without sadness: Testing the direct effect hypothesis of pleasurable negative emotion. *PLOS ONE*, *19*(4), e0299115. https://doi.org/10.1371/journal.pone.0299115

Sheline, Y. I., Barch, D. M., Price, J. L., Rundle, M. M., Vaishnavi, S. N., Snyder, A. Z., Mintun, M. A., Wang, S., Coalson, R. S., & Raichle, M. E. (2009). The default mode network and self-referential processes in depression. *Proc Natl Acad Sci U S A*, *106*, 1942–1947.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420

Sihvonen, A. J., Sarkamo, T., Ripolles, P., Leo, V., Saunavaara, J., Parkkola, R., Rodriguez-Fornells, A., & Soinila, S. (2017). Functional neural changes associated with acquired amusia across different stages of recovery after stroke. *Sci Rep*, *7*(11390).

Simó, M., Rifà-Ros, X., Vaquero, L., Ripollés, P., Cayuela, N., Jové, J., Navarro, A., Cardenal, F., Bruna, J., & Rodríguez-Fornells, A. (2018). Brain functional connectivity in lung cancer population: An exploratory study. *Brain Imaging and Behavior*, *12*(2), 369–382. https://doi.org/10.1007/s11682-017-9697-8

Singer, N., Jacoby, N., Lin, T., Raz, G., Shpigelman, L., Gilam, G., Granot, R. Y., & Hendler, T. (2016). Common modulation of limbic network activation underlies musical emotions as they unfold. *Neuroimage*, *141*, 517–529.

Skov, M., & Nadal, M. (2023). *The Routledge International Handbook of Neuroaesthetics*. Routledge & CRC Press. https://www.routledge.com/The-Routledge-International-Handbook-of-Neuroaesthetics/Skov-Nadal/p/book/9781032348803

Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., Filippini, N., Watkins, K. E., Toro, R., Laird, A. R., & Beckmann, C. F. (2009). Correspondence of the brain's functional architecture during activation and rest. *Proc Natl Acad Sci U S A*, *106*, 13040–13045.

Sonkusare, S., Breakspear, M., & Guo, C. (2019). Naturalistic Stimuli in Neuroscience: Critically Acclaimed. *Trends in Cognitive Sciences*, *23*(8), 699–714. https://doi.org/10.1016/j.tics.2019.05.004

Starr, G. G. (2023a). Aesthetic experience models human learning. *Frontiers in Human Neuroscience*, *17*. https://doi.org/10.3389/fnhum.2023.1146083

Starr, G. G. (2023b). *Just in time: Temporality, aesthetic experience, and cognitive neuroscience.* MIT Press.

Taruffi, L., & Koelsch, S. (2014). The Paradox of Music-Evoked Sadness: An Online Survey. *PLOS ONE*, *9*(10), e110490. https://doi.org/10.1371/journal.pone.0110490

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. https://doi.org/10.1038/s41598-017-14849-0

Tervaniemi, M. (2023). The neuroscience of music – towards ecological validity. *Trends in Neurosciences*, *46*(5), 355–364. https://doi.org/10.1016/j.tins.2023.03.001

Tie, Y., Whalen, S., Suarez, R. O., & Golby, A. J. (2008). Group independent component analysis of language fMRI from word generation tasks. *Neuroimage*, *42*, 1214–1225.

Toiviainen, P., Alluri, V., Brattico, E., Wallentin, M., & Vuust, P. (2014). Capturing the musical brain with Lasso: Dynamic decoding of musical features from fMRI data. *Neuroimage*, *88*, 170–180.

Trost, W., Ethofer, T., Zentner, M., & Vuilleumier, P. (2012). Mapping aesthetic musical emotions in the brain. *Cereb Cortex*, *22*, 2769–2783.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., & Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, *15*, 273–289.

Upham, F., & McAdams, S. (2018). *Activity analysis and coordination in continuous responses to music*. Music Perception253-294.

Vessel, E. A., Isik, A. I., Belfi, A. M., Stahl, J. L., & Starr, G. G. (2019). The default-mode network represents aesthetic appeal that generalizes across visual domains. *Proceedings of the National Academy of Sciences*, *116*(38), 19155–19164. https://doi.org/10.1073/pnas.1902650116

Vessel, E. A., Maurer, N., Denker, A. H., & Starr, G. G. (2018). Stronger shared taste for natural aesthetic domains than for artifacts of human culture. *Cognition*, *179*, 121–131. https://doi.org/10.1016/j.cognition.2018.06.009

Vessel, E. A., & Rubin, N. (2010). Beauty and the beholder: Highly individual taste for abstract, but not real-world images. *J Vis*, *10*, 18–14.

Vessel, E. A., Starr, G. G., & Rubin, N. (2012). The brain on art: Intense aesthetic experience activates the default mode network. *Front Hum Neurosci*, *6*(66).

Vessel, E. A., Starr, G. G., & Rubin, N. (2013). Art reaches within: Aesthetic experience, the self and the default mode network. *Front Neurosci*, *7*(258).

Vuoskoski, J. K., & Eerola, T. (2017). The Pleasure Evoked by Sad Music Is Mediated by Feelings of Being Moved. *Frontiers in Psychology*, *8*, 439. https://doi.org/10.3389/fpsyg.2017.00439

Wagenmakers, E. J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., & Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*, *25*, 58–76.

Wen, O. X., & Krumhansl, C. L. (2017). *Real time responses to Stravinsky´s Symphonies of Wind Instruments: Perception of internal repetition and musical interest*. Music Perception60-76.

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), 6130. https://doi.org/10.1038/srep06130

Wu, X., Lu, J., Chen, K., Long, Z., Wang, X., Shu, H., Li, K., Liu, Y., & Yao, L. (2009). Multiple neural networks supporting a semantic task: An fMRI study using independent component analysis. *Neuroimage*, *45*, 1347–1358.

Xu, J., Potenza, M. N., & Calhoun, V. D. (2013). Spatial ICA reveals functional activity hidden from traditional fMRI GLM-based analyses. *Front Neurosci*, *7*(154).

Xu, J., Zhang, S., Calhoun, V. D., Monterosso, J., Li, C. S., Worhunsky, P. D., Stevens, M., Pearlson, G. D., & Potenza, M. N. (2013). Task-related concurrent but opposite modulations of overlapping functional networks as revealed by spatial ICA. *Neuroimage*, *79*, 62–71.

Yang, S., Reed, C. N., Chew, E., & Barthet, M. (2023). Examining Emotion Perception Agreement in Live Music Performance. *IEEE Transactions on Affective Computing*, *14*(2), 1442–1460. IEEE Transactions on Affective Computing. https://doi.org/10.1109/TAFFC.2021.3093787

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nat Methods*, *8*, 665–670.

Zatorre, R. J., & Salimpoor, V. N. (2013). From perception to pleasure: Music and its neural substrates. *Proc Natl Acad Sci U S A*, *2*, 10430–10437.