

Article

Attention Control and Audiomotor Processes Underlying Anticipation of Musical Themes while Listening to Familiar Sonata-Form Pieces

Chia-Wei Li ¹ and Chen-Gia Tsai ^{2,3,*} 

¹ Department of Radiology, Wan Fang Hospital, Taipei Medical University, Taipei 116, Taiwan; ffiln0628@gmail.com

² Graduate Institute of Musicology, National Taiwan University, Taipei 106, Taiwan

³ Neurobiology and Cognitive Science Center, National Taiwan University, Taipei 106, Taiwan

* Correspondence: tsaichengia@ntu.edu.tw

Abstract: When listening to music, people are excited by the musical cues immediately before rewarding passages. More generally, listeners attend to the antecedent cues of a salient musical event irrespective of its emotional valence. The present study used functional magnetic resonance imaging to investigate the behavioral and cognitive mechanisms underlying the cued anticipation of the main theme's recurrence in sonata form. Half of the main themes in the musical stimuli were of a joyful character, half a tragic character. Activity in the premotor cortex suggests that around the main theme's recurrence, the participants tended to covertly hum along with music. The anterior thalamus, pre-supplementary motor area (preSMA), posterior cerebellum, inferior frontal junction (IFJ), and auditory cortex showed increased activity for the antecedent cues of the themes, relative to the middle-last part of the themes. Increased activity in the anterior thalamus may reflect its role in guiding attention towards stimuli that reliably predict important outcomes. The preSMA and posterior cerebellum may support sequence processing, fine-grained auditory imagery, and fine adjustments to humming according to auditory inputs. The IFJ might orchestrate the attention allocation to motor simulation and goal-driven attention. These findings highlight the attention control and audiomotor components of musical anticipation.

Keywords: prediction; attention; auditory imagery; motivation; motor simulation; musical theme



Citation: Li, C.-W.; Tsai, C.-G. Attention Control and Audiomotor Processes Underlying Anticipation of Musical Themes while Listening to Familiar Sonata-Form Pieces. *Brain Sci.* **2022**, *12*, 261. <https://doi.org/10.3390/brainsci12020261>

Academic Editor: Laila Craighero

Received: 22 January 2022

Accepted: 11 February 2022

Published: 13 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Arousal and attention to reward cues play a substantial role in human reward pursuit. Spectators on the golf course focus on the long putt and almost forget to breathe. When listening to music, people tend to be excited by the musical cues prior to rewarding passages. For example, the peak pleasurable experiences of electronic dance music (EDM) are linked to two musical sections: the build-up and the ensuing drop. The rewarding drop section was found to elicit maximal skin conductance responses in listeners, while some listeners experienced chills and an urge to move their bodies during the final bars of the build-up section that announce the coming of the drop [1]. The antecedent cues of musical rewards are associated with psychological 'wanting' or the anticipatory phase of reward processing, leading to increasing activity in reward-related brain regions in the listeners [2,3].

The anticipation of a salient musical passage may entail audiomotor processes irrespective of its emotional valence. Tracking the sequence of antecedent cues and the salient passage may be performed via motor simulation such as covert humming along with the music, especially when the listener is familiar with the piece. This is analogous to the anticipation of the motion of a table tennis ball, in which anticipation performance is based on the recruitment of sensorimotor representations in the premotor cortex [4]. Although a humming voice never amounts to real music, the associated audiomotor representation

provides a prediction of the ensuing passage [5]. A salient musical passage may be the main tune of a musical piece. Given that the encoding, storage, and retrieval of tunes are all essential for a number of human social activities [6], it is likely that successful prediction based on cue-triggered auditory imagery of salient tunes is rewarded with pleasure in the brain.

One kind of salient event in music is the recurrence of a main theme, which is often easy to memorize and imagine. In Western classical music, the main theme's recurrences manifest significant structural and emotional effects. Notably, the musical cues immediately before the main theme's recurrences are designed to foster the listener's motivation and cognitive organization towards anticipating the main theme. Using functional magnetic resonance imaging (fMRI), the present study aimed to identify the neural correlates of the cued anticipation of the main themes during listening to familiar music. All musical stimuli in the present study were selected from compositions written in sonata form. Developed in the late eighteenth century, the sonata form became one of the most important forms in Western classical music. The most dramatic moment of this form is typically around the coincidence of the main theme's recurrence and the harmonic resolution at the start of the recapitulation section [7]. The passage immediately before the recapitulation, the retransition, provides a window into the nature of musical anticipation. Carl Czerny, a pianist, composer, and music educator in the nineteenth century, gave the following advice to composers: '[s]omething [the retransition] of peculiar beauty, calculated to fix the attention, may be introduced in the concluding bars leading back to the principal theme' [8]. The tension-resolution patterns around the theme's recurrence in sonata form were found to elicit distinct emotional physiological responses in listeners [9]. In the current study, half of the main themes were of a joyful character and positive valence, and half of a tragic character and negative valence. The addition of tragic stimuli served to reveal behavioral and cognitive mechanisms underlying anticipatory processing of music irrespective of its valence.

The thematic and harmonic resolution at the recapitulation's start can be regarded as a goal of the listening experience of sonata form. Smith [10] drew parallels between the first-movement sonata form and the Hero's Journey [11]; the exposition section corresponds to the hero's departure from home, the development section corresponds to the hero's expeditions, and the recapitulation section corresponds to the journey's return stage. The development section 'is the scene of conflicting moods, of restlessness, of drama, of the unexpected' [12]. Metaphorically, the main theme finds its way home in the retransition after wandering through remote keys in the development. The German name of the retransition, *Rückleitung*, means 'return path'. This path typically emphasizes the dominant of the home key (tonic key). During exposure to a retransition, listeners who are empathetic toward the main theme may recall a goal intention and expect the dominant-tonic resolution, which is the most common type of tension-resolution pattern in music. Therefore, the brain regions implicated in goal-driven attention may be engaged during the concluding bars of retransitions. Based on the literature mentioned above, we hypothesized that the antecedent cues of the main themes in the concluding bars of retransitions were associated with enhanced activity in the premotor and attention control areas.

2. Materials and Methods

2.1. Participants

The participants were recruited via a public announcement on the internet, which stated that the research goal was to examine the neural responses to sonata form and that a requirement of participants was a high familiarity with Western classical music. To evaluate this requirement, we asked volunteers to express and explain their feelings in response to the musical stimuli in the present study. An online questionnaire was used to collect these data. This questionnaire began with a brief introduction to the musical structure of the first-movement sonata form. The introduction mentioned that (1) this musical form consists of three main sections (i.e., an exposition, a development, and a

recapitulation); (2) the themes presented in the exposition recur in the recapitulation; and (3) the passage prior to the recapitulation often intensifies the moving power of the main theme's recurrence. Eight musical stimuli were presented in a fixed order, with the minor-mode excerpts alternating with the major-mode excerpts. After listening to each stimulus, the volunteers were asked to write down their feelings in response to the retransition and their feelings in response to the main theme's recurrence. They were also asked to explain their feelings in terms of musical features, such as tempo, tonality/mode, melody, rhythm, timbre, articulation, and dynamics. The volunteers were allowed to listen to the stimulus repeatedly before answering these questions. This online questionnaire was expected to familiarize the volunteers with the experimental stimuli, and the data collected were used for participant screening. The correspondence author of this article selected participants according to the following inclusion criteria: (1) more than five passages immediately before the recapitulation evoked a keen sense of anticipation; (2) the recurrence of more than five themes evoked a feeling of resolution; and (3) their feelings were appropriately explained in terms of musical features for more than five excerpts.

Thirty-one adult volunteers completed this questionnaire. Twenty-seven volunteers met our screening criteria. Twenty participants completed the fMRI experiment. They were free from neurological, psychiatric, or auditory problems and there were no abnormalities in their structural MRI. Fifteen of them had studied musical instruments for six years or more. Written informed consent was obtained from each participant prior to participation in the study. The participants were compensated with approximately USD 16 after the completion of the fMRI scan. Data from two participants were discarded because of excessive head movements. Therefore, the final sample analyzed in this study consisted of 18 participants (12 females aged 27.0 ± 6.7 years, range 20–40 years). All the research procedures were performed in accordance with a protocol approved by the Institutional Review Board of National Taiwan University (201611HM008). This study was conducted in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

2.2. Stimuli

The musical stimuli consisted of four minor-mode and four major-mode excerpts from masterpieces in Western classical music. It was presumed that the minor-mode and major-mode excerpts possess tragic and joyful characters, respectively. The correspondence author of this article (a musicologist) selected the music excerpts from the first movements of famous symphonies and concertos composed in the classical or early romantic period. The following selection criteria for the musical stimuli were used: (1) the first movement was written in sonata form; (2) the tempo was faster than 90 beats per minute; (3) the final passage of the development produced high musical tension; (4) the main theme (first theme) was played at a moderate to high intensity level; and (5) the main theme was of a joyful character (in a major-mode key) or a tragic character (in a minor-mode key). The selected musical stimuli included eight excerpts (detailed in Table S1) around the theme's recurrence with a duration of 30 s.

It should be noted that the tragic musical themes presented in this study differ from sad music in acoustic, musical, and affective characteristics. Sad music and vocal expressions of sadness such as lamenting, sighing, and weeping are typically soft and slow. On the other hand, tragic themes such as Beethoven's 'fate knocking at the door' theme in his fifth symphony are louder, heavier, and more threatening than sad music. Compared to sad themes, tragic themes are much more frequently used in the first-movement sonata form, which is typically in a moderate to fast tempo. In the present study, the arousing character of the joyful and tragic themes indicates that the arrival of such a theme is a salient musical event.

All the musical excerpts were extracted from commercially available recordings, digitized (44,100 Hz sampling rate, 16-bit stereo) and edited by GoldWave (version 5.58, Goldwave Inc., Newfoundland, Canada). They were taken from the first movements of symphonies and concertos composed between 1772 and 1858 and written in sonata form.

Each excerpt in the present study consisted of an 18 s passage of the concluding portion of the development leading back to the first theme (main theme) and a 12 s passage of the ensuing recapitulation (Figure 1). A brief introduction to these pieces can be found at <https://youtu.be/Mz-X7aPcfG8> (accessed on 10 February 2022).

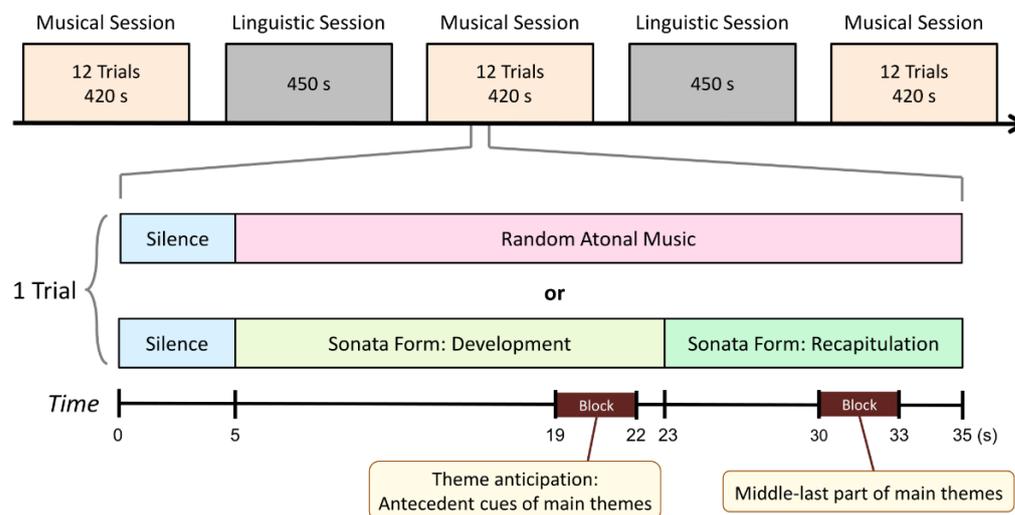


Figure 1. Schematic description of the functional scanning sessions.

We followed a prior study [13] to use random atonal sequences as stimuli for the control condition. Four atonal random sequences with duration of 30 s were generated using Reason (version 8.0, Propellerhead Software Inc., Stockholm, Sweden). The virtual musical instruments used in these atonal random sequences were the violin, cello, flute, and piano. The duration of each note in these atonal random sequences ranged from 0.1 to 6.0 s, and the fundamental frequency of each note ranged from 65.41 Hz (C2) to 1046.50 Hz (C6). All of the stimuli had a 1 s fade-in and a 1 s fade-out.

2.3. Procedure

Prior to the MRI scan, the participants listened to all of the musical stimuli in the same order as that in the online questionnaire. They were asked to rate their agreement with four statements on a seven-point Likert scale (1 = entirely disagree, 2 = mostly disagree, 3 = somewhat disagree, 4 = neither agree nor disagree, 5 = somewhat agree, 6 = mostly agree, 7 = entirely agree) after the presentation of each stimulus: (1) I am very familiar with this music; (2) I felt strong anticipation before the theme came out; (3) the theme is very tragic sounding; and (4) the theme is very joyful sounding.

There were five fMRI scanning sessions separated by 1 min rest intervals, as schematically illustrated in Figure 1. Three fMRI scanning sessions for the present musical study alternated with two fMRI scanning sessions for a linguistic study. This design was expected to minimize affective habituation that could occur with repeated exposure to the same emotional music. The methods and results of this linguistic study [14] are not mentioned further in this paper. Four major-mode music excerpts, four minor-mode music excerpts, and four atonal random stimuli (control) were presented in a pseudo-randomized order within each fMRI scanning session for the present study on musical emotion. The duration of each scanning session was approximately 7 min. Prior to the first scanning session, the first 5 s of Mozart's Clarinet Concerto were presented to allow the experimenter to adjust the auditory volume according to the participant's feedback. During this pre-scan volume-adjusting test, the participant also adapted to listening to the music in the presence of scanner noise. The stimuli were presented through scanner-compatible headphones. Participants wore earplugs to reduce the scanner noise by 15–25 dB.

2.4. MRI Data Acquisition

The experiments were performed using a Siemens 3T MRI Prisma at National Taiwan University. A gradient-echo echo-planar imaging (EPI) sequence was used in the functional data scanning. Approximately 2.5 mm slices of axial images were acquired in the functional scans using gradient-echo planar imaging with the following parameters: time to repetition = 2500 ms, echo time = 30 ms, flip angle = 87°, in-plane field of view = 192 × 192 mm, and acquisition matrix = 78 × 78 × 45, to cover all the cerebral areas. Magnetization-prepared rapid gradient echo T1-weighted imaging with an isotropic spatial resolution of 0.9 mm was acquired for each participant for spatial individual-to-template normalization.

2.5. Data Analyses

We used MIRtoolbox [15] to estimate the sound intensity and brightness of the antecedent cues of the main themes and the middle-last part of these themes in the sonata-form stimuli. Paired *t*-tests were used to examine whether the sound intensity and brightness varied significantly across the antecedent cues of the main themes and the middle-last part of the main themes.

Wilcoxon rank sum tests were used to examine whether the participants' scores regarding familiarity with musical excerpts, anticipation of all of the themes, the tragic character of the minor-mode themes, the joyful character of the minor-mode themes, the tragic character of the major-mode themes, and the joyful character of the major-mode themes differed significantly from a neutral score of 4. A Bonferroni-corrected significance threshold of $p < 0.05/6 = 0.0083$ was used to account for these comparisons.

The preprocessing and analyses of the fMRI data were performed using SPM12 (Wellcome Trust Centre for Neuroimaging, London, UK) and the Artifact Detection Tools (ART) toolkit. The first four volumes of each run were discarded to allow for magnetic saturation effects. The remaining functional images were corrected for head movement artefacts and timing differences in the slice acquisitions. The preprocessed functional images were coregistered to the individual's anatomical image, normalized to the standard MNI brain template, and resampled to a 2 mm isotropic voxel size. The normalized images were spatially smoothed using a Gaussian kernel of 5 mm full-width at half maximum to accommodate any anatomical variability across participants. The ART toolkit was used to check the spikes and motion of each dataset. Eighteen functional datasets showed head motion below 2.5 mm of the maximal translation (in any direction), below 1° of maximal rotation throughout the course of scanning, and no outliers occurred. These images were used in the following analyses.

Statistical inference was based on a random effect approach at two levels. The data of each participant were analyzed using the general linear model via fitting the time series data with the canonical hemodynamic response function (HRF) modelled at the relevant blocks in the stimulation paradigm [16]. The block of the antecedent cues of the main themes (theme anticipation) and that of the middle-last part of the main themes are shown in Figure 1. Linear contrasts were computed to characterize responses of interest, averaging across fMRI runs. The group-level analysis consisted of paired *t*-tests for (1) the contrast of theme anticipation minus control, (2) the contrast of the middle-last part of the main themes minus control, and (3) the contrast of theme anticipation minus the middle-last part of the main themes. Results were reported as significant at a voxel-wise threshold level of $p < 0.001$ uncorrected and cluster-level threshold of $p < 0.05$ false discovery rate corrected.

3. Results

The sound intensity and brightness did not significantly differ across the antecedent cues of the main themes and the middle-last part of the main themes ($p > 0.5$, detailed in Table S1). The results of the questionnaire measures are presented in Figure 2. The Wilcoxon rank-sum tests demonstrated that the scores of familiarity with the musical stimuli, anticipation of the musical themes, the joyful character of the major-mode themes,

and the tragic character of the minor-mode themes were significantly higher than a neutral score of 4. By contrast, the scores regarding the tragic character of the major-mode themes and the joyful character of the minor-mode themes were significantly lower than 4. These results showed that the participants were familiar with the musical stimuli and felt strong anticipation before the theme came out. We also confirmed that the minor- and major-mode themes were of tragic and joyful characters, respectively.

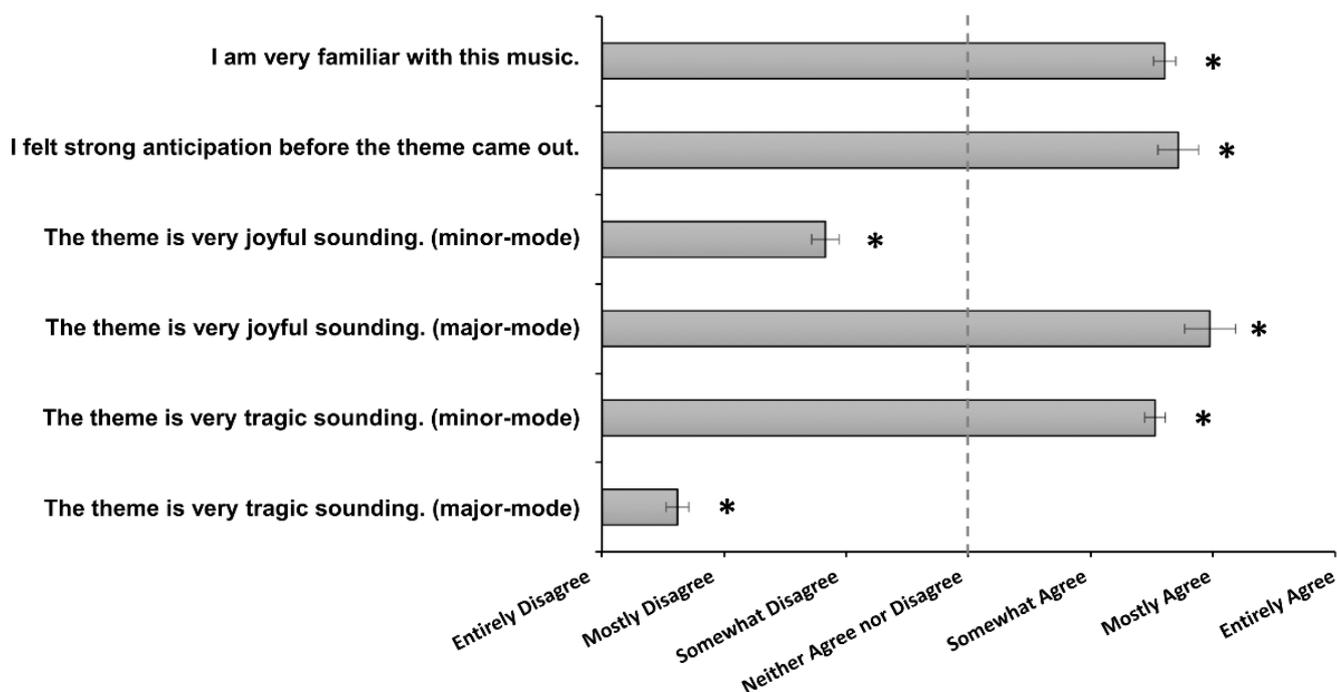


Figure 2. Participants' ratings (7-point Likert scale). Error bars indicate standard error of the mean. * $p < 0.05/6 = 0.0083$.

The analysis of the fMRI data revealed that the antecedent cues of the main themes (theme anticipation) were associated with significantly greater activity in the bilateral superior temporal gyri, right anterior insula, left putamen, bilateral pre-supplementary motor area (preSMA), bilateral middle frontal gyrus (dorsal premotor cortex), and left posterior cerebellum as compared to the control. The middle-last part of the main themes was associated with significantly greater activity in the anterior portion of the bilateral superior temporal gyri as compared to the control. The right anterior thalamus, right putamen, right caudate, bilateral preSMA, bilateral posterior cerebellum, bilateral inferior frontal junction (IFJ), and left superior temporal gyrus (auditory cortex) showed significantly increased activity for the theme anticipation condition, relative to the middle-last part of the main themes (Table 1, Figure 3). The time courses of focal brain activity in a subset of activated locations, including the anterior insula, premotor cortex, posterior preSMA, posterior cerebellum, anterior thalamus, and IFJ, are demonstrated in Figure 4.

Table 1. Activation clusters for the contrasts of theme anticipation minus control, middle-last part of themes minus control, and theme anticipation minus middle-last part of themes.

	<i>t</i> -Value	MNI Coordinates (mm)			Cluster Size (Voxel)	Brain Regions
		X	Y	Z		
Theme Anticipation > Control	8.55	−48	0	−2	190	Superior temporal gyrus
	5.90	−48	−34	12	73	
	6.87	50	6	−8	260	Superior temporal gyrus
	6.81	48	0	0		Insula
	4.33	56	10	−14		Superior temporal pole
	8.00	−24	0	6	79	Putamen
	7.45	50	−2	56	251	Middle frontal gyrus (premotor cortex)
	7.28	8	10	62	241	Pre-supplementary motor area
	5.62	−32	−56	−30	143	Cerebellar lobule VI
	5.32	−52	0	42	125	Middle frontal gyrus (premotor cortex)
Middle-last Part of Themes > Control	6.8	−50	0	−8	145	Superior temporal gyrus
	5.93	52	−2	−10	96	
Theme Anticipation > Middle-last Part of Themes	6.81	−6	2	64	364	Pre-supplementary motor area
	6.62	−46	−16	0	175	Superior temporal gyrus
	5.74	14	−72	−24	151	Cerebellar lobule VI
	5.32	12	−80	−26		Cerebellar lobule Crus I-II
	5.63	−10	−80	−28		240
	5.32	−14	−72	−26	119	Cerebellar lobule VI
	5.66	46	20	26		Inferior frontal gyrus, pars opercularis (inferior frontal junction)
	5.63	46	18	22		Inferior frontal gyrus, pars triangularis (inferior frontal junction)
	5.65	−36	14	32	289	Inferior frontal gyrus, pars opercularis (inferior frontal junction)
	5.32	−42	14	28		Inferior frontal gyrus, pars triangularis (inferior frontal junction)
	4.80	−52	8	44		Middle frontal gyrus (premotor cortex)
	5.62	8	−2	0	342	Anterior thalamus
	5.42	22	10	2		Putamen
5.39	22	10	8	Caudate		
5.05	20	−62	−24	93	Cerebellar lobule VI	

Abbreviations—MNI: Montreal Neurological Institute coordinate system.

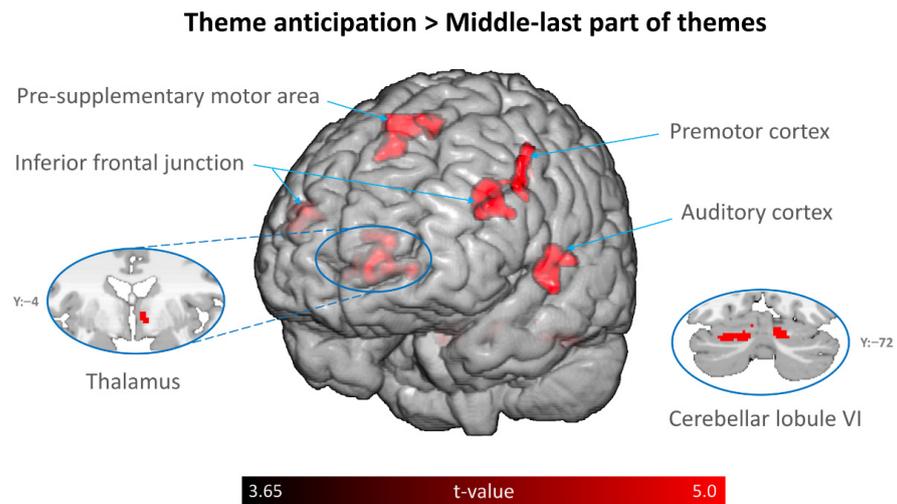


Figure 3. Group-level activation map for the contrast of theme anticipation minus middle-last part of themes.

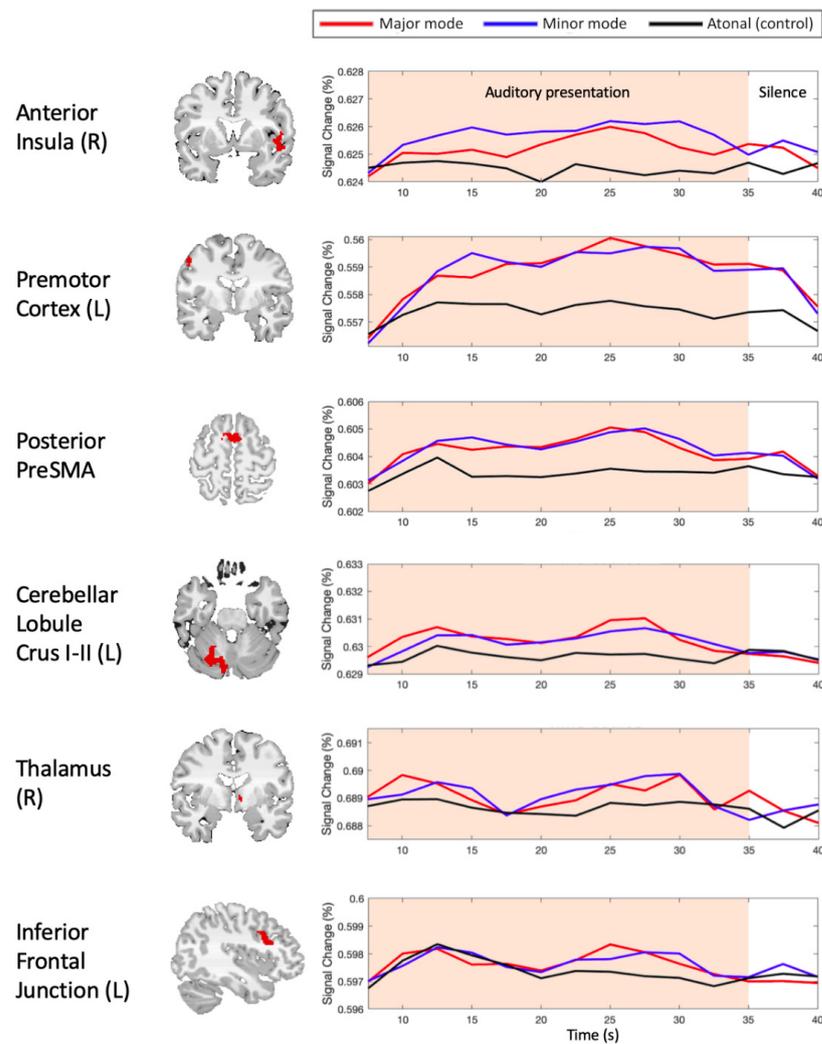


Figure 4. Time courses of focal brain activity in a subset of activated locations. **Left** panels: activation clusters. **Right** panels: averaged time courses of blood-oxygen-level dependent (BOLD) signal change.

4. Discussion

During exposure to familiar music, the listener often attentively tracks the main theme's recurrences that have structural and affective significance in musical pieces. Motor simulation and attention control processes may be engaged in the cued anticipation of the main theme's recurrences. To better understand the behavioral and cognitive mechanisms underlying this anticipation, we used fMRI to examine the neural responses to the concluding bars of retransitions in sonata form. A retransition encompasses musical cues that foster the listener's motivation and cognitive organization of the main theme's recurrence. Compared to the control condition of atonal random music, the cued anticipation of the main themes during the retransitions was associated with enhanced activity in the anterior insula, premotor regions, and posterior cerebellum. This finding bore a remarkable similarity to the network that commonly subserves predictive coding in action perception, language, and music [17]. Moreover, we observed significantly enhanced activity in the anterior thalamus, putamen, caudate, preSMA, premotor cortex, posterior cerebellum, IFJ, and auditory cortex during theme anticipation compared to the middle-last part of the main themes.

A novelty of the present study was the use of both positively and negatively valenced musical excerpts as stimuli. The greater activity in the right anterior thalamus during the anticipation of joyful and tragic main themes than during the middle-last part of these themes may reflect its role in the anticipation of salient events. The right anterior thalamus has been found to be involved in itch processing [18], anticipation of monetary reward [19–21], anticipation of aversive pictures [22], and anticipation of pain [23]. Moreover, animal experiments demonstrated that memory tasks for temporally linked items are sensitive to anterior thalamus damage. For example, lesions of the anterior thalamus in rats impaired their ability to remember the sequential order of a series of odors in order to retrieve food rewards [24]. There is emerging evidence indicating that the anterior thalamus may guide attention towards stimuli that consistently predict important outcomes [25]. Our finding of the involvement of the right anterior thalamus in the cued anticipation of salient musical themes lends support to the role of the anterior thalamus in anticipatory processing of motivationally salient events. It should be noted that emotional responses to the anticipation of joyful and tragic themes could be elusive and divergent, and therefore, we did not observe significant activation in the amygdala or nucleus accumbens for theme anticipation.

Since the participants were familiar with the sonata-form stimuli, they were likely to track the sequence of antecedent cues and theme recurrence via covert humming along with music. As expected, we observed greater activity in the premotor cortex, preSMA, and posterior cerebellum during the theme anticipation condition relative to the control stimuli. These areas have been implicated in covert or overt rehearsal of familiar music [26–28] and melody imagery [29].

The preSMA and posterior cerebellum showed greater activity for the theme anticipation condition than for the middle-last part of the main themes. It is well recognized that the preSMA supports domain-general sequence processing [30]. The posterior part of the preSMA underpins linear sequence processing in language [31] and music [32]. In the present study, each cue–theme sequence was fixed and linear. Therefore, the increased activity in the posterior ($y \sim 2$) preSMA for theme anticipation may reflect its role in linear sequence processing. On the other hand, the anterior ($y \sim 13$) preSMA has been implicated in working memory processes involved in sentence comprehension [33] and motor response selection/control [34]. In the present study, the participants may have recruited the anterior preSMA for maintenance of the ensuing theme during the final bars of the retransition, while the working memory load may have significantly decreased after the entry of this theme. Moreover, the greater activity in the anterior preSMA during theme anticipation may have reflected more vivid imagery of the music in service of attentively monitoring the theme's recurrence, compared to the middle-last part of the main themes. It is proposed that the preSMA may support the generation of rich and fine-grained sensory expectations when the task requires attentive access to auditory representations [35]. Increased activity

in the anterior preSMA has been linked to the attentive monitoring of goal progress over time [36]. Taken together, these observations allow us to suggest that the preSMA underpins the linear sequence processing and vivid fine-grained musical imagery necessary for precisely monitoring the theme's recurrence.

The cerebellum is implicated in the prediction of sensory outcomes as a consequence of motor command, and this predictive computation is known as an internal forward model [37]. During motor learning or motor control, the cerebellum is engaged in the fine modulation of the movement parameters controlled by the cortex [38]. Moreover, the posterior cerebellum provides prospective signals about the timing of perceptual events [39]. Increased activity in the posterior cerebellum during theme anticipation suggests that this region may utilize such prospective signals to make fine adjustments to covert humming along with the music. This view is consistent with a study on joint actions showing that the posterior cerebellum processed errors in synchrony between one's own actions and those of others [40]. Like joint actions, humming along with familiar music may recruit the posterior cerebellum to fine tune the way in which the listener needs to coordinate the imagined action with the auditory inputs.

An interesting finding of the present study was the involvement of the bilateral IFJ in theme anticipation. Accumulating evidence suggests that the IFJ plays an essential role in tasks requiring cognitive control such as working memory and task switching [41–43]. Two studies suggest specific aspects of the hub role of the left IFJ. De Baene and colleagues [44] indicated that the left IFJ may contribute to the integration of goal information and action-related information. Asplund and colleagues [45] pointed out that the left IFJ may orchestrate goal-directed attention control and stimulus-driven attention. In the present study, the goal-driven attention linked to the harmonic/thematic resolution at the recapitulation's start and the attention allocated to covert humming along with music might meet in the IFJ. While this hypothesis must presently be viewed as speculative, it provides a novel view on the role of the inferior frontal regions in the anticipatory processing of music. In a putative network of predictive coding, the bilateral inferior frontal gyri were reported to underpin the processing of harmonic syntax, the recognition of structural regularities, the analysis of structural relationships among events occurring within sequences, and the hierarchical organization of actions during multi-component goal-directed behavior [17]. We suggest that the coincidence of the release of harmonic tension and the main theme's recurrence at the recapitulation's start may be a goal of the listening experience of sonata form. The pursuit of this goal may engage a suite of cognitive and audiomotor processes, including goal-driven attention allocation, maintenance of the goal in working memory, and monitoring the progress toward the goal via re-enactment of the fixed sequence of the cues and the main theme. The bilateral IFJ might be a hub for driving and coordinating these processes.

When comparing the theme anticipation condition to the middle-last part of the main themes, exclusively left-sided activation was observed in the auditory cortex and dorsal premotor cortex. It has been found that responses of the human auditory cortex to rapid temporal information were weighted towards the left [46,47]. Moreover, relative to the connections of the right dorsal premotor cortex, those of the left dorsal premotor cortex were stronger with frontal and parietal areas representing action dominance based on a lead in motor intention and fine precision skills [48]. In light of these findings, the sensitivity of the left auditory cortex and left dorsal premotor cortex to the goal-directed anticipation of music might be related to their role in providing goal-relevant, fine-grained audiomotor representations on a narrower time scale. It is noteworthy that our finding of increased activity in the left auditory cortex and attention control regions for theme anticipation compared to the middle-last part of the main themes supports the idea that attention modulates early sensory processing, somewhat speaking against the idea that expectation has little influence on early sensory responses [49]. While there have been a number of studies on visual selective attention and visual expectation, the present study calls for

future studies to illustrate how goal-directed anticipation, attention, and motor resonance modulate activity in the auditory cortex in response to familiar musical passages.

In light of the incentive-sensitization theory of addiction [50], our results might provide important information for the mechanisms underlying the enjoyment and learning of music. A core of enjoyment of familiar music might be incentive sensitization to cues, namely, amplification of ‘wanting’ for the main theme triggered by its antecedent cues. Attentively monitoring the sequence of the antecedent cues and the main theme might be rewarded with pleasure in the brain when this prediction is confirmed by the auditory inputs. Importantly, the confirmation of this prediction might be more highly rewarded when this prediction is based on more precise, fine-grained audiomotor representations. During the process of familiarization with a specific musical piece, the listener might undergo a cycle where sensitization to the antecedent cues of the main theme, attentively and precisely humming along with the music, and ‘liking’ for the confirmation of the prediction of this theme enhance each other. The musical tension of the antecedent cues and the harmonic resolution at the main theme’s entry could contribute to the listener’s ‘wanting’ and ‘liking’ in this cycle, respectively (Figure 5). Given that the anteromedial thalamus may play a role in instrumental learning processes and the control of goal-directed behavior [25], our finding regarding the involvement of this region in theme anticipation provides preliminary support for this hypothesis. Future studies could consider using a paradigm of learning cue–theme sequences to test this hypothesis.

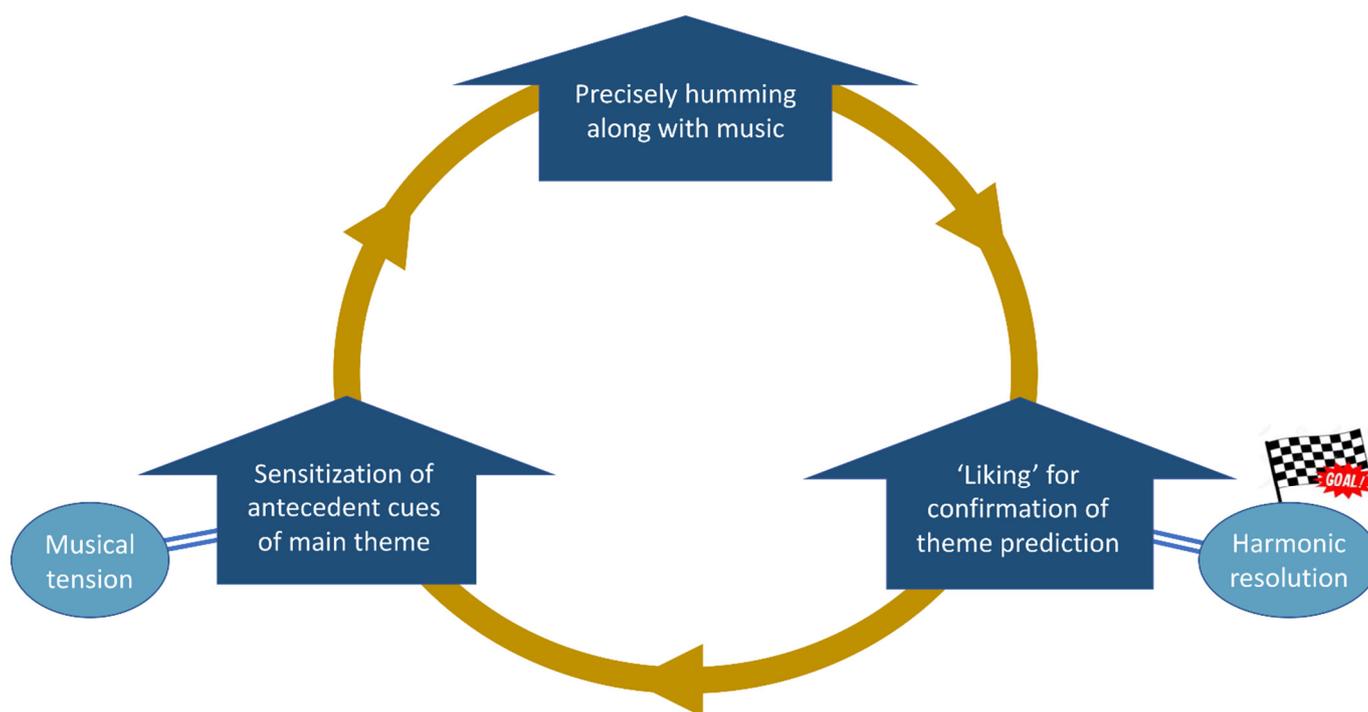


Figure 5. Putative model for the enjoyment and learning of music (inspired by the incentive-sensitization theory of addiction). During exposure to a musical piece, the antecedent cues of the main theme may increase attention to covert humming along with the music. More precise humming may lead to more intense ‘liking’ for the confirmation of theme prediction. Repeated intermittent experiences of ‘liking’ may in turn cause sensitization to the antecedent cues of the main theme. The listener’s ‘wanting’ and ‘liking’ may be facilitated by the musical tension of the antecedent cues and the harmonic resolution at the main theme’s entry, respectively.

There were several limitations of this study that need to be pointed out. First, the final sample of the fMRI data consisted of 18 participants, and the sonata-form stimuli were extracted from only eight compositions. Future research should increase the sample size of participants and musical stimuli. Second, the present study lacked behavioral data

about the participants' affective/motivational states and the vividness of the auditory imagery during fMRI scanning. We did not ask the participants to respond after stimulus presentation because we were concerned about the distracting effect of self-monitoring for rating tasks. Third, we did not analyze the IFJ's functional connectivity to the thalamus, premotor regions, posterior cerebellum, or auditory cortex. Therefore, the IFJ's specific contributions to driving and coordinating cognitive and audiomotor processes remain unclear. We suggest that future research could be devoted to the parcellation of the bilateral IFJ using musical stimuli imbued with incentive salience.

In conclusion, the present study focused on listeners' cued anticipation of the main theme's recurrence. While partially consistent with the putative prediction network underlying action perception, language, and music [17], our results demonstrate parallels between attentively monitoring the main theme's recurrence and goal pursuit. Given that the harmonic/thematic resolution at the recapitulation's start serves as a goal of the listening experience of sonata form, the anterior thalamus may play a key role in guiding attention towards the antecedent cues of this goal. The pursuit of the main theme's recurrence was associated with increased activity in the posterior cerebellum and premotor areas, which may contribute to monitoring the progress toward a goal via fine-grained motor simulation and maintaining the goal in working memory. The bilateral IFJ might orchestrate the attention allocation to motor simulation and goal-driven attention. These findings highlight how cognitive and audiomotor processes work together during the cued anticipation of motivationally salient musical events.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/brainsci12020261/s1>, Table S1: Sonata-form stimuli.

Author Contributions: Conceptualization, C.-G.T.; methodology, C.-G.T. and C.-W.L.; software, C.-W.L.; validation, C.-G.T. and C.-W.L.; formal analysis, C.-W.L.; investigation, C.-G.T. and C.-W.L.; resources, C.-G.T. and C.-W.L.; data curation, C.-G.T. and C.-W.L.; writing—original draft preparation, C.-G.T. and C.-W.L.; writing—review and editing, C.-G.T. and C.-W.L.; visualization, C.-W.L.; supervision, C.-G.T.; project administration, C.-G.T.; funding acquisition, C.-G.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by grant projects (MOST 109-2410-H-002-196 and MOST 110-2410-H-002-210) from the Ministry of Science and Technology, Taiwan.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of National Taiwan University (201611HM008, December 2016).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data are not publicly available due to privacy/ethical restrictions. The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors would like to thank Chao-Ju Chen for data collection.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Solberg, R.T.; Dibben, N. Peak Experiences with Electronic Dance Music: Subjective Experiences, Physiological Responses, and Musical Characteristics of the Break Routine. *Music Percept* **2019**, *36*, 371–389. [[CrossRef](#)]
2. Salimpoor, V.N.; Benovoy, M.; Larcher, K.; Dagher, A.; Zatorre, R.J. Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nat. Neurosci.* **2011**, *14*, 257–262. [[CrossRef](#)]
3. Li, C.W.; Chen, J.H.; Tsai, C.G. Listening to music in a risk-reward context: The roles of the temporoparietal junction and the orbitofrontal/insular cortices in reward-anticipation, reward-gain, and reward-loss. *Brain Res.* **2015**, *1629*, 160–170. [[CrossRef](#)]
4. Bischoff, M.; Zentgraf, K.; Pilgramm, S.; Stark, R.; Kruger, B.; Munzert, J. Anticipating action effects recruits audiovisual movement representations in the ventral premotor cortex. *Brain Cogn.* **2014**, *92C*, 39–47. [[CrossRef](#)]
5. Schubotz, R.I. Prediction of external events with our motor system: Towards a new framework. *Trends Cogn. Sci.* **2007**, *11*, 211–218. [[CrossRef](#)]
6. Schulkin, J.; Raglan, G.B. The evolution of music and human social capability. *Front. Neurosci.* **2014**, *8*, 292. [[CrossRef](#)]

7. Rosen, C. *The Classical Style: Haydn, Mozart, Beethoven*; W. W. Norton: New York, NY, USA, 1972.
8. Reicha, A.; Czerny, C. Vollst andiges Lehrbuch der Musikalischen Composition, oder, Ausf uhrliche und Ersch opfende Abhandlung  uber die Harmonie (den Generalbass) die Melodie, die Form und Ausarbeitung der Verschiedenen Arten von Tonst ucken, den Gebrauch der Gesangstimmen, die Gesammte Instrumentirung, den Hohern Tonsatz im Doppelten Contrapunct, die Fuge und den Canon, und  uber den Strengen Satz im Kirchenstyl. Online Resource. 1834. Available online: <https://www.worldcat.org/title/vollstaendiges-lehrbuch-der-musikalischen-composition-oder-ausfuhrliche-und-erschopfende-abhandlung-uber-die-harmonie-den-generalbass-die-melodie-die-form-und-ausarbeitung-der-verschiedenen-arten-von-tonstuecken-den-gebrauch-der-gesangstimmen-die-gesammte-instrumentirung-den-hoehern-tonsatz-im-doppelten-contrapunct-die-fuge-und-den-canon-und-uber-den-strengen-satz-im-kirchenstyl/oclc/1102225744> (accessed on 10 February 2022).
9. Tsai, C.G.; Chen, C.P. Musical Tension over Time: Listeners' Physiological Responses to the 'Retransition' in Classical Sonata Form. *J. New. Music Res.* **2015**, *44*, 271–286. [[CrossRef](#)]
10. Smith, E.L. *The Hero Journey in Literature: Parables of Poesis*; University Press of America: Lanham, MD, USA, 1996; p. 309.
11. Campbell, J. *The Hero with a Thousand Faces*; Pantheon Books: New York, NY, USA, 1949.
12. Berry, W. *Form in Music; An Examination of Traditional Techniques of Musical Structure and Their Application in Historical and Contemporary Styles*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1966; p. 191.
13. Trost, W.; Ethofer, T.; Zentner, M.; Vuilleumier, P. Mapping Aesthetic Musical Emotions in the Brain. *Cereb. Cortex* **2012**, *22*, 2769–2783. [[CrossRef](#)]
14. Tsai, C.G.; Li, C.W. Is It Speech or Song? Effect of Melody Priming on Pitch Perception of Modified Mandarin Speech. *Brain Sci.* **2019**, *9*, 286. [[CrossRef](#)]
15. Lartillot, O.; Toiviainen, P.; Eerola, T. A Matlab Toolbox for Music Information Retrieval. In *Data Analysis, Machine Learning and Applications. Studies in Classification, Data Analysis, and Knowledge Organization*; Preisach, C., Burkhardt, H., Schmidt-Thieme, L., Decker, R., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 261–268. [[CrossRef](#)]
16. Josephs, O.; Henson, R.N. Event-related functional magnetic resonance imaging: Modelling, inference and optimization. *Philos Trans. R. Soc. Lond. B Biol. Sci.* **1999**, *354*, 1215–1228. [[CrossRef](#)]
17. Siman-Tov, T.; Granot, R.Y.; Shany, O.; Singer, N.; Hendler, T.; Gordon, C.R. Is there a prediction network? Meta-analytic evidence for a cortical-subcortical network likely subserving prediction. *Neurosci. Biobehav. R.* **2019**, *105*, 262–275. [[CrossRef](#)]
18. Leknes, S.G.; Bantick, S.; Willis, C.M.; Wilkinson, J.D.; Wise, R.G.; Tracey, I. Itch and motivation to scratch: An investigation of the central and peripheral correlates of allergen- and histamine-induced itch in humans. *J. Neurophysiol.* **2007**, *97*, 415–422. [[CrossRef](#)]
19. Apaydin, N.; Ustun, S.; Kale, E.H.; Celikag, I.; Ozguven, H.D.; Baskak, B.; Cicek, M. Neural Mechanisms Underlying Time Perception and Reward Anticipation. *Front. Hum. Neurosci.* **2018**, *12*, 115. [[CrossRef](#)]
20. Kirk, U.; Brown, K.W.; Downar, J. Adaptive neural reward processing during anticipation and receipt of monetary rewards in mindfulness meditators. *Soc. Cogn. Affect. Neurosci.* **2015**, *10*, 752–759. [[CrossRef](#)]
21. Wu, C.C.; Samanez-Larkin, G.R.; Katovich, K.; Knutson, B. Affective traits link to reliable neural markers of incentive anticipation. *Neuroimage* **2014**, *84*, 279–289. [[CrossRef](#)]
22. Grupe, D.W.; Oathes, D.J.; Nitschke, J.B. Dissecting the anticipation of aversion reveals dissociable neural networks. *Cereb. Cortex* **2013**, *23*, 1874–1883. [[CrossRef](#)]
23. Reicher, P.; Wiemer, J.; Gerdes, A.B.M.; Schulz, S.M.; Pauli, P.; Wieser, M.J. Anxious anticipation and pain: The influence of instructed vs conditioned threat on pain. *Soc. Cogn. Affect. Neurosci.* **2017**, *12*, 544–554. [[CrossRef](#)]
24. Wolff, M.; Gibb, S.J.; Dalrymple-Alford, J.C. Beyond spatial memory: The anterior thalamus and memory for the temporal order of a sequence of odor cues. *J. Neurosci.* **2006**, *26*, 2907–2913. [[CrossRef](#)]
25. Nelson, A.J.D. The anterior thalamic nuclei and cognition: A role beyond space? *Neurosci. Biobehav. Rev.* **2021**, *126*, 1–11. [[CrossRef](#)]
26. Tsai, C.G.; Chen, C.C.; Chou, T.L.; Chen, J.H. Neural mechanisms involved in the oral representation of percussion music: An fMRI study. *Brain Cogn.* **2010**, *74*, 123–131. [[CrossRef](#)]
27. Schulze, K.; Zysset, S.; Mueller, K.; Friederici, A.D.; Koelsch, S. Neuroarchitecture of verbal and tonal working memory in nonmusicians and musicians. *Hum. Brain Mapp.* **2011**, *32*, 771–783. [[CrossRef](#)]
28. Schon, D.; Gordon, R.; Campagne, A.; Magne, C.; Astesano, C.; Anton, J.L.; Besson, M. Similar cerebral networks in language, music and song perception. *Neuroimage* **2010**, *51*, 450–461. [[CrossRef](#)]
29. Herholz, S.C.; Halpern, A.R.; Zatorre, R.J. Neuronal correlates of perception, imagery, and memory for familiar tunes. *J. Cogn. Neurosci.* **2012**, *24*, 1382–1397. [[CrossRef](#)]
30. Cona, G.; Semenza, C. Supplementary motor area as key structure for domain-general sequence processing: A unified account. *Neurosci. Biobehav. R.* **2017**, *72*, 28–42. [[CrossRef](#)]
31. Alario, F.X.; Chainay, H.; Lehericy, S.; Cohen, L. The role of the supplementary motor area (SMA) in word production. *Brain Res.* **2006**, *1076*, 129–143. [[CrossRef](#)]
32. Leaver, A.M.; Van Lare, J.; Zielinski, B.; Halpern, A.R.; Rauschecker, J.P. Brain Activation during Anticipation of Sound Sequences. *J. Neurosci.* **2009**, *29*, 2477. [[CrossRef](#)]
33. Bonhage, C.E.; Fiebach, C.J.; Bahlmann, J.; Mueller, J.L. Brain Signature of Working Memory for Sentence Structure: Enriched Encoding and Facilitated Maintenance. *J. Cogn. Neurosci.* **2014**, *26*, 1654–1671. [[CrossRef](#)]

34. Barber, A.D.; Caffo, B.S.; Pekar, J.J.; Mostofsky, S.H. Effects of Working Memory Demand on Neural Mechanisms of Motor Response Selection and Control. *J. Cogn. Neurosci.* **2013**, *25*, 1235–1248. [[CrossRef](#)]
35. Lima, C.F.; Krishnan, S.; Scott, S.K. Roles of Supplementary Motor Areas in Auditory Processing and Auditory Imagery. *Trends Neurosci.* **2016**, *39*, 527–542. [[CrossRef](#)]
36. Benn, Y.; Webb, T.L.; Chang, B.P.; Sun, Y.H.; Wilkinson, I.D.; Farrow, T.F. The neural basis of monitoring goal progress. *Front. Hum. Neurosci.* **2014**, *8*, 688. [[CrossRef](#)]
37. Tanaka, H.; Ishikawa, T.; Lee, J.; Kakei, S. The Cerebro-Cerebellum as a Locus of Forward Model: A Review. *Front. Syst. Neurosci.* **2020**, *14*, 19. [[CrossRef](#)] [[PubMed](#)]
38. Proville, R.D.; Spolidoro, M.; Guyon, N.; Dugue, G.P.; Selimi, F.; Isope, P.; Popa, D.; Lena, C. Cerebellum involvement in cortical sensorimotor circuits for the control of voluntary movements. *Nat. Neurosci.* **2014**, *17*, 1233–1239. [[CrossRef](#)] [[PubMed](#)]
39. O'Reilly, J.X.; Mesulam, M.M.; Nobre, A.C. The cerebellum predicts the timing of perceptual events. *J. Neurosci.* **2008**, *28*, 2252–2260. [[CrossRef](#)] [[PubMed](#)]
40. Kokal, I.; Keyser, C. Granger causality mapping during joint actions reveals evidence for forward models that could overcome sensory-motor delays. *PLoS ONE* **2010**, *5*, e13507. [[CrossRef](#)]
41. Crone, E.A.; Wendelken, C.; Donohue, S.E.; Bunge, S.A. Neural evidence for dissociable components of task-switching. *Cereb. Cortex* **2006**, *16*, 475–486. [[CrossRef](#)]
42. Linden, D.E.J.; Bittner, R.A.; Muckli, L.; Waltz, J.A.; Kriegeskorte, N.; Goebel, R.; Singer, W.; Munk, M.H.J. Cortical capacity constraints for visual working memory: Dissociation of fMRI load effects in a fronto-parietal network. *Neuroimage* **2003**, *20*, 1518–1530. [[CrossRef](#)]
43. Ravizza, S.M.; Carter, C.S. Shifting set about task switching: Behavioral and neural evidence for distinct forms of cognitive flexibility. *Neuropsychologia* **2008**, *46*, 2924–2935. [[CrossRef](#)]
44. De Baene, W.; Albers, A.M.; Brass, M. The what and how components of cognitive control. *Neuroimage* **2012**, *63*, 203–211. [[CrossRef](#)]
45. Asplund, C.L.; Todd, J.J.; Snyder, A.P.; Marois, R. A central role for the lateral prefrontal cortex in goal-directed and stimulus-driven attention. *Nat. Neurosci.* **2010**, *13*, 507–512. [[CrossRef](#)]
46. Zatorre, R.J.; Belin, P. Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* **2001**, *11*, 946–953. [[CrossRef](#)]
47. Zaehle, T.; Wustenberg, T.; Meyer, M.; Jancke, L. Evidence for rapid auditory perception as the foundation of speech processing: A sparse temporal sampling fMRI study. *Eur. J. Neurosci.* **2004**, *20*, 2447–2456. [[CrossRef](#)] [[PubMed](#)]
48. van der Hoorn, A.; Potgieser, A.R.; de Jong, B.M. Transcallosal connection patterns of opposite dorsal premotor regions support a lateralized specialization for action and perception. *Eur. J. Neurosci.* **2014**, *40*, 2980–2986. [[CrossRef](#)] [[PubMed](#)]
49. Rungratsameetaweemana, N.; Serences, J.T. Dissociating the impact of attention and expectation on early sensory processing. *Curr. Opin. Psychol.* **2019**, *29*, 181–186. [[CrossRef](#)] [[PubMed](#)]
50. Robinson, T.E.; Berridge, K.C. The neural basis of drug craving: An incentive-sensitization theory of addiction. *Brain Res. Brain Res. Rev.* **1993**, *18*, 247–291. [[CrossRef](#)]