Whole-brain computation of cognitive versus acoustic errors in music: A mismatch negativity study

L. Bonetti [a,b,c,d,*], F. Carlomagno [a,e], M. Kliuchko [i], B.P. Gold [j], S. Palva [h,k], N.T. Haumann [a], M. Tervaniemi [f,g], M. Huotilainen [f,g], P. Vuust [a], E. Brattico [a,e,**]

[a] Center for Music in the Brain, Department of Clinical Medicine, Aarhus University & the Royal Academy of Music Aarhus/Aalborg, Denmark
[b] Centre for Eudaimonia and Human Flourishing, Linacre College, University of Oxford, United Kingdom
[c] Department of Psychiatry, University of Oxford, United Kingdom
[d] Department of Psychology, University of Bologna, Italy
[e] Department of Education, Psychology, Communication, University of Bari Aldo Moro, Italy
[f] Cognitive Brain Research Unit, Faculty of Medicine, University of Helsinki, Helsinki, Finland
[g] Centre of Excellence on Music, Mind, Body and Brain, Faculty of Educational Sciences, University of Helsinki, Finland
[h] Helsinki Institute of Life Sciences, Neuroscience Center, University of Helsinki, Finland
[i] Department of Health Technology, Technical University of Denmark, Denmark
[j] Department of Electrical and Computer Engineering, Vanderbilt University, Nashville, TN, USA
[k] Centre for Cognitive Neuroscience, Institute of Neuroscience and Psychology, University of Glasgow, United Kingdom

A B S T R A C T

Previous studies have evidenced how the local prediction of physical stimulus features may affect the neural processing of incoming stimuli. Less known are the effects of cognitive priors on predictive processes, and how the brain computes local versus cognitive predictions and their errors. Here, we determined the differential brain mechanisms underlying prediction errors related to high-level, cognitive priors for melody (rhythm, contour) versus low-level, local acoustic priors (tuning, timbre). We measured with magnetoencephalography the mismatch negativity (MMN) prediction error signal in 104 adults having varying levels of musical expertise. We discovered that the brain regions involved in early predictive processes for local priors were primary and secondary auditory cortex and insula, whereas cognitive brain regions such as cingulate and orbitofrontal cortices were recruited for early melodic errors in cognitive priors. The involvement of higher-level brain regions for computing early cognitive errors was enhanced in musicians, especially in cingulate cortex, inferior frontal gyrus, and supplementary motor area. Overall, the findings expand knowledge on whole-brain mechanisms of predictive processing and the related MMN generators, previously mainly confined to the auditory cortex, to a frontal network that strictly depends on the type of priors that are to be computed by the brain.

1. Introduction

According to predictive coding theory, audition is an active process where models of expectations for the incoming sounds are constantly updated based on expectations (also termed priors) when errors occur (Friston and Kiebel, 2009; Garrido et al., 2009). Recent neuroimaging studies provide empirical support for this theory with electroencephalography (EEG) or magnetoencephalography (MEG) recordings of the mismatch negativity (MMN), which indexes early predictive errors of acoustic features (i.e., deviations from prior expectations within the first 120–250 ms from the onset of the deviant stimulation) (Näätänen et al., 2007; Näätänen, 1995). These studies also substantiate the existence of ascending, forward connections in the auditory cortex that convey these prediction errors to higher-order brain areas, signalling the 'new' information in external stimuli, and of backward connections from higher-order areas of the auditory cortex to predict the activity of lower-order areas (Koelsch et al., 2019; Wacongne et al., 2012).

However, most studies that have measured the MMN do so for simple acoustic feature errors and analyze only the MEEG sensor signal and parameters: only a minority of MMN studies have provided a clear reconstruction of the neural sources. These studies returned a network of active brain areas that were mainly localized in the auditory cortex and especially in Heschl's gyrus and the superior and middle temporal gyri (L. Bonetti, Bruzzone, et al., 2021; Fitzgerald and Todd, 2020;

Marco-Pallarés et al., 2005; Waberski et al., 2001). Additional, weaker generators of the MMN were localized in the inferior frontal cortex and cingulate gyrus (Fitzgerald and Todd, 2020; Marco-Pallarés et al., 2005; Waberski et al., 2001). Functional magnetic resonance imaging (fMRI) studies have further confirmed the involvement of superior temporal gyrus and right inferior and middle frontal gyri in the generation of the MMN (Molholm et al., 2005; Schall et al., 2003). Taken together, the current literature supports the hypothesis that auditory cortex is the main generator of the MMN elicited in response to errors of acoustic priors, with frontal generators also involved – perhaps in the process of involuntary "attention switching" and prior updating (Heilman & Van Den Abell, 1980; Kastner and Ungerleider, 2000; Korzyukov et al., 2003; Mesulam and Marchsel, 1981; Posner, 2016).

Yet the predictive coding theory holds that forward and backward connections should also convey and predict, respectively, the cognitive features of stimuli beyond their low-level (e.g., acoustic) features (Friston, 2018; Friston and Kiebel, 2009; Koelsch et al., 2019). Within this framework, music listening is a peculiar case, involving the prediction of both lower-level acoustic features using knowledge (priors) accumulated from life-long exposure to all kinds of sounds and, at the same time, music-specific priors based on exposure to a specific musical culture (Brattico and Pearce, 2013a, 2013b; Koelsch et al., 2019). Together, these predictions allow us to detect changes and mistakes (e. g., in tonality, harmony, transposition, or rhythm) that make music either interesting and pleasurable or, conversely, boring, and dissonant. However, the brain areas that generate these cognitive prediction errors, and whether they differ from those underlying the MMNs for low-level acoustic features (i.e., by recruiting more frontal sources) are, thus far, open research questions.

Despite the rich cognitive information contained in musical sequences, most musical MMN experiments (including the fMRI ones) have used simple auditory oddball paradigms in which the acoustic features inserted in sequences of coherent sounds (e.g., pitch, rhythm, location, timbre) are broken by sudden, infrequent deviant sounds (Risto Näätänen, 2018; Risto Näätänen et al., 2011). These studies have revealed the automatic predictive processes for sounds that rely on feedforward and backward projects from and to the auditory cortex, but they bear little resemblance to the variety of sounds and features encountered in music. Accordingly, the newer "multi-feature" paradigm (Risto Näätänen, Pakarinen, Rinne, & Takegata, 2004) (Fisher et al., 2008; Pakarinen et al., 2007), introduces a deviation in a single feature into every second sound of a musical pattern, thus allowing for the recording of several prediction errors – including cognitive ones. In one version of this paradigm, six deviants were used (pitch, slide, duration, timbre, location, and intensity), obtaining reliable MMNs for each (L. Bonetti et al., 2018; L. Bonetti, Haumann, Vuust, Kliuchko and Brattico, 2017; Mu et al., 2016; Vuust et al., 2011a; Vuust et al., 2012). Similarly, in the latest "MusMelo" version of this paradigm, six deviants are inserted in a loop of one elaborated musical melody (Mari Tervaniemi, Huotilainen, & Brattico, 2014)' (Putkinen et al., 2014), crucially including two distinct categories of deviants: cognitive or high-level deviants and acoustic or low-level ones. Cognitive deviants refer to changes in the melodic line (melodic contour) of the melody, altering the meaning of the music since they give rise to a varied version of the original melody. Conversely, acoustic deviants sound merely like "mistakes" during the musical performance without producing any actual change of the melodic line. Hence, the MusMelo paradigm offers a unique possibility of measuring the neural indexes of cognitive versus acoustic priors and their related prediction error signals, and of locating the subservient neural sources.

To summarize, much is known on the MMN prediction error signal and its neural substrate in the auditory cortex for acoustic features. However, research on the role of frontal MMN generators is still scarce and inhomogeneous. This is particularly true with regard to music perception, which draws so heavily on both acoustic and cognitive features and is dependent on the fast, automatic predictive processes

that are indexed by MMN. Related work has been conducted on the early right-anterior negativity (ERAN) which is typically elicited when participants attentively listen and detect chord violating the conventions of Western harmony (Garza Villarreal, Brattico, Leino, Østergaard and Vuust, 2011; Koelsch et al., 2001, 2019), and which relies on the right homologue of infero-frontal Broca's area and parietal regions. However, such attentive processes would not explain how we can very promptly grasp even abstract, culture-dependent violations of musical conventions without the need of focusing on listening and how understanding of musical sounds automatically unfolds over time without any conscious effort.

In this study, we therefore wished to localize the automatic predictive processes throughout the whole-brain that are responsible for the fast (120–250 ms from the deviant stimulation) generation of error signals during music listening, and to determine whether these signals differ when the error is computed against acoustic versus cognitive priors. To this goal, we investigated in a large sample of over 100 participants the neural sources of acoustic versus cognitive errors of musical melodies, as indexed by MMN. We hypothesized to observe stronger frontal MMN generators for cognitive deviants, and increased responses in the auditory cortex to MMNs elicited by acoustic deviants.

We also assessed how the activity in MMN generators were modulated by musical expertise, for both acoustic and cognitive errors. The MMN has been repeatedly connected to cognitive abilities (Antonio Criscuolo, Bonetti, Särkämö, Kliuchko, & Brattico, 2019; Franklin et al., 2008), musicianship, musical learning, and musical cognitive abilities (Brattico et al., 2009; Koelsch et al., 1999; Mari Tervaniemi, Just, Koelsch, Widmann, & Schröger, 2005; Vuust et al., 2011b, 2012). For instance, Putkinen and colleagues (Putkinen et al., 2014) showed enhanced MMNs in children exposed to musical training, especially for melody modulation, mistuning and timbre that were not existent before exposure to the musical training. Similarly, Kliuchko and colleagues (Kliuchko et al., 2019) discovered an overall stronger MMN to timbre, pitch and slide for jazz compared to classical musicians and non-musicians and amateurs. Tervaniemi and colleagues (Mari Tervaniemi, Janhunen, Kruck, Putkinen, & Huotilainen, 2016) showed that MMNs were enhanced for tuning deviants in classical musicians, for timing deviants in classical and jazz musicians, and for transposition deviant in jazz musicians. Moreover, fMRI evidence has consistently demonstrated how musical expertise refines cognitive priors, for instance, allowing musicians to notice subtle harmonic changes and their action-observation areas and higher-order frontal areas to become more activated during mere music listening (Gold et al., 2019; Salimpoor et al., 2015). For all of these reasons, we hypothesized that musical expertise would increase both acoustic and cognitive MMNs at the sites of their generators.

2. Methods

2.1. Participants

Participants were volunteers recruited with fliers in social media or posted in academies and universities, and they were compensated for the time spent in the lab with vouchers that could be used for culture and sport activities (e.g. museums, concerts, swimming pools, etc.). Prior to the beginning of the experiment, participants filled in the informed consent. The experimental procedures, included in the wide research protocol named "Tunteet" ("Emotions" in Finnish), complied with the Declaration of Helsinki – Ethical Principles for Medical Research, and were approved by the Ethics Committee of the Hospital District of Helsinki and Uusimaa (approval number: 315/13/03/00/11, dated 11th March 2012). The "Tunteet" protocol consisted of two days. In the first day, the experiment comprised 60 min of MEEG recordings and consisted of approximately 3 h spent at the Biomag laboratory (including welcoming, preparation, instructions, questionnaires and forms filling, and dismissal). In the second day, MRI recordings for obtaining the

Fig. 1. Melodic multi-feature (MusMelo) paradigm.
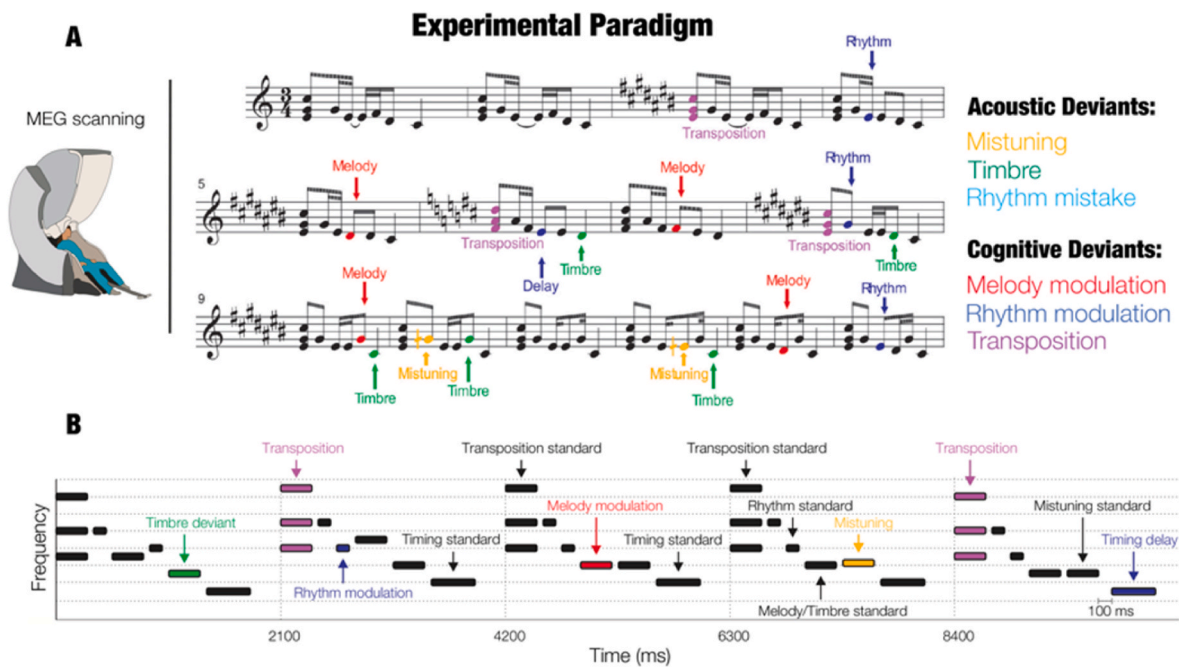
structural images of each participant were conducted, no later than two weeks from the MEEG recordings. Previous studies from our group used different and independent parts of this large dataset which involved slightly different subgroups of the participants who took part in the "Tunteet" protocol. This occurred because a few participants did not complete the full data collection, as it usually happens when conducting large experimental protocols like this one. Extensive details on the previous works published using data from "Tunteet" are described in Kliuchko and colleagues (Kliuchko et al., 2019), Haumann and colleagues (Haumann et al., 2016), Criscuolo and colleagues (Criscuolo et al., 2019), and Bonetti and colleagues (Bonetti et al., 2018, 2017). In the current study, we focused on the MEG data collected while participants were presented with the "Musmelo" paradigm, which is described in detail in the next paragraphs. Moreover, it is important to state that this data has never been reported in any previous paper. The sample of participants who were presented with the "Musmelo" paradigm and therefore analysed in the current study consisted of 104 volunteers: 44 males and 60 females (age range: 18–51 years old, mean age: 28.24 ± 7.92 years). All participants declared to be healthy and reported no current or previous drug nor alcohol abuse. In addition, they were not under any kind of medication, they did not have any neurological or psychiatric disorder, and declared to have normal hearing. Finally, their educational, economic, and social statuses were homogeneous, as studied and reported in Criscuolo and colleagues (Criscuolo et al., 2019).

Since musicianship has been connected to modulation of MMN responses [17,39,50–52], we recruited participants with different levels of musical expertise. Specifically, the average formal musical training received by our participants was 5.88 ± 7.12 years (ranging from 0 to 28 years of musical training). Indeed, our samples comprised musicians who obtained a professional musical education or graduated from Sibelius Academy and University of Helsinki, amateur musicians who had only few years of formal musical training, and non-musicians.

2.2. Experimental design and stimuli

To detect the brain predictive responses to cognitive and acoustic deviants, we used the Melodic Multifeature paradigm (MusMelo) introduced by Tervaniemi and colleagues (Tervaniemi et al., 2014) and Putkinen and colleagues (Putkinen et al., 2014) while participants' brain activity was recorded by means of magnetoencephalography (MEG).

The MusMelo paradigm consisted of brief recursive melodies composed by the author Minna Huotilainen. These melodies were played with the standard timbre correspondent to digital piano tones (McGill University Master Samples) and followed typical Western tonal musical harmonies and configurations.

The melodies started with a triad (duration of 300 ms), followed by four tones of different length, plus an ending tone (duration of 575 ms). Subsequent tones were always separated by 50 ms silence. Additionally, a 125 ms silence after the ending tone was inserted. Thus, one melody lasted for 2100 ms in total. Such melodies were presented for 15 min in a looped, recursive manner. Within these repeated melodies, six different deviants (changes) were inserted. Importantly, these deviants occurred in a random order and at random places within the melody, thus it was not possible for the participants to predict when the deviant occurred. For this reason, the inter-melody gap of 125 ms did not represent an issue for the detection of the MMN. To be noted, all deviants occurred in random places within the melody but did not alter the onset of the sounds. The only deviant which represented an exception was the Rhythm mistake which was created by introducing a silent gap of 100 ms before the deviant sound.

The six deviants were divided into low-level, acoustic deviants and high-level, cognitive deviants.

The key difference between the two categories of deviants is that low-level, acoustic deviants did not alter the melodic contour of the musical stimuli, but introduced acoustic mistakes (e.g., small variations in pitch or rhythm that were perceived as mistakes and not drastic changes of the melodies). Conversely, high-level, cognitive deviants operated a profound change in the melodies that were perceived as proper variations. One melody could contain several changes, as illustrated in Fig. 1. Finally, participants were not required to pay attention to the stimuli because the MMN is a signature of the automatic brain prediction error and is almost completely independent on attention. As commonly done in MMN studies (see, for example, Bonetti et al., 2021a, b,c; Bonetti et al., 2018; Brattico et al., 2006; Escera et al., 2000; Fisher et al., 2008; Garrido et al., 2009; Näätänen et al., 2007; Näätänen et al., 2004; Putkinen et al., 2014; Tervaniemi et al., 2014; Tervaniemi et al., 2016), at the same time of the sound stimulation, participants were presented with a muted, captioned, movie that they chose before the
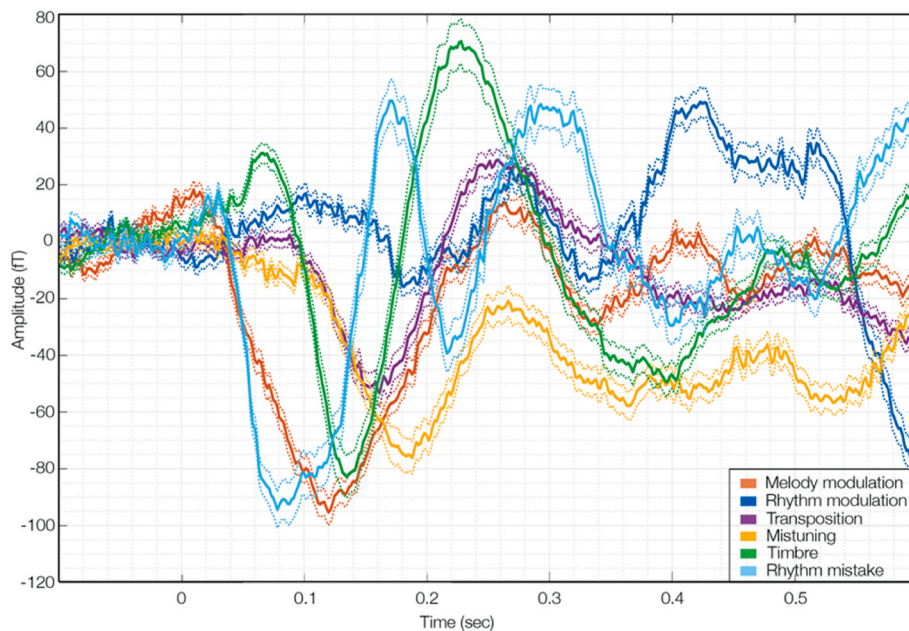
Fig. 2. MMN to all deviants (MEG channel 1341). Waveform depicting the MMN responses (deviant minus standard) to the six deviants occurring in the MusMelo paradigm (melody modulation, rhythm modulation, transposition, mistuning, timbre, and rhythm mistake). The time series were recorded at the MEG magnetometer channel 1341, which is a typical channel shown in MMN studies. Dash lines show the standard error, computed independently for each time-point t using the formula $\frac{std_t}{\sqrt{n}}$. Where std_t is the standard deviation at time t, and n is the number of participants. The grey area highlights the different peaks of the MMN to the six deviants included in the study. X-axis shows time (in seconds), while y-axis amplitude of the signal in fT. Time 0 indicates the specific onset of the deviant stimulation (e.g. for Melody modulation, it indicates the first tone where the melody changed; for Rhythm modulation the first tone where the rhythm of the melody changed; for Transposition the first chord where the tonality changed; for Mistuning the tone that was slightly edited in terms of pitch; for Timbre the tone that introduced a different timbre; for Rhythm mistake the tone that introduced the mistake (delay fo 100 ms) in terms of rhythm).

beginning of the experiment.

As follows, we provided details on the acoustic deviants:

1. *Mistuning* (half of a semitone upwards, up to 3% of the fundamental frequency of the sound). It occurred in the 14% of the melodies and could happen in the first, second or fourth tone of the melody.
2. *Timbre deviant* (flute timbre instead of the standard piano timbre). It occurred in the 8% of the melodies and could happen in the first, third or fourth tone of the melodies.
3. *Rhythm mistake* (also known as *Timing delay*) (100 ms silent gap). It occurred in the 8% of the melodies. It could happen in the first, second or third tone.

Conversely, these were the characteristics of the cognitive deviants:

1. *Melody mo*dulation occurred in the 12% of the melodies. It consisted of a pitch change of the third or fourth tone. It endured until a new melody modulation was introduced.
2. *Rhythm modulation* occurred in the 7% of the melodies and could happen in the second or third tone. There were two possible alternatives for rhythm modulation, either a short tone was replaced by a long tone (tone lengthening) or a long tone was replaced by a short one (tone shortening).
3. *Transposition* occurred in the 16% of the melodies and could occur in the first triad. In this case, after introducing the chord transposition the following melodies kept the converted key until a new chord transposition was presented.

Finally, all cognitive deviants were musically plausible, both when the change involved the melodic contour and the rhythm contour.

To be noted, when depicting the MMN waveform (e.g. Fig. 2) 'time 0' indicates the specific onset of the deviant stimulation (e.g. for Melody modulation, it indicates the first tone where the melody changed; for Rhythm modulation the first tone where the rhythm of the melody changed; for Transposition the first chord where the tonality changed; for Mistuning the tone that was slightly edited in terms of pitch; for Timbre the tone that introduced a different timbre; for Rhythm mistake the tone that introduced the mistake (100 ms delay) in terms of rhythm).

The stimuli were presented using Presentation software (Neuro-behavioural Systems, Berkeley, CA). Importantly, before starting the experiment, we have conducted an objective test to determine the

hearing threshold of each ear of each participant. Once detected those thresholds, we have adjusted the volume of the stimuli by setting it to 50 dB above the measured hearing thresholds. In this way, we have made sure that the volume perception of the sound stimulation was exactly the same for every participant in both ears. Apart from the adjustment of the volume in case the hearing thresholds of the two ears were diverse, there were no differences between the presentation of the sounds in the two ears.

Finally, in a separate session, the structural images of the participants' brain were acquired by using magnetic resonance imaging (MRI).

2.3. Data acquisition

MEG data was collected at the Biomag Laboratory of the Helsinki University Central Hospital. The measurements were conducted in a magnetically shielded room (ETS-Lindgren Euroshield, Eura, Finland) with Vectorview™ 306-channel MEG scanner (Elekta Neuromag®, Elekta Oy, Helsinki, Finland). The MEG scanner had 102 sensor elements. Specifically, it had 102 orthogonal pairs of planar gradiometer SQUID sensors and 102 axial magnetometer SQUID sensors. We placed electrodes above and below the left eye and close to the external eye corners on both sides of the face of the participants to record horizontal and vertical eye movements. Furthermore, we recorded the continuous head position of the participants by using the head position indicator (HPI) coils that were placed on the forehead and behind the ears of participants. Moreover, for each participant we recorded the fiducial points corresponding to nasion and to the preauricular anatomical landmarks by using the Isotrack 3D digitizer (Polhemus, Colchester, VT, USA). The HPI coils and fiducial points were necessary to perform co-registration between MEG and MRI data at a later stage of analysis. Finally, the MEG data was registered with a sampling rate of 600 Hz.

The recorded MRI data was the structural T1, required for the source reconstruction of the MEG signal. The MRI scanning was conducted using a 3 T MAGNETOM Skyra whole-body scanner (Siemens Healthcare, Erlangen, Germany), plus a standard 20-channel head-neck coil. The measurements were done at the Advanced Magnetic Imaging (AMI) Center (Aalto University, Espoo, Finland). Details of the T1-weighted structural images are reported as follows: 176 slices; matrix = 256 × 256; field of view = 256 × 256 mm; pulse sequence = MPRAGE; slice thickness = 1 mm; interslice skip = 0 mm. Later in the analysis pipeline, we co-registered each individual T1-weighted MRI scan to the standard

MNI brain template through an affine transformation. Then, we referenced such image to the MEG sensors space by employing the Polhemus head shape data and the three fiducial points collected prior to start the MEG recording.

2.4. Data pre-processing

We preprocessed the raw MEG sensor data by using the signal space separation solution implemented in MaxFilter (Taulu and Simola, 2006) which attenuated the interference originated outside the scalp.

Afterwards, we converted the data into the SPM format and further analysed it in Matlab (MathWorks, Natick, Massachusetts, United States of America) by employing OSL (OHBA Software Library), a freely available toolbox that relies on a combination of FSL (Woolrich et al., 2009), Fieldtrip (Oostenveld et al., 2011), SPM (Penny et al., 2007), as well as in-house-built functions.

First, a few segments of the data contaminated by large artefacts were removed after visual inspection. Second, we corrected the brain data for the interference of eyeblinks and heart-beat artefacts by using independent component analysis (ICA). We used the ICA implemented in the OSL software by the OHBA group of the University of Oxford, UK, written in Matlab. This procedure decomposed the original signal in independent components. Then, we identified and discarded the components that picked up the eyeblink and heart-beat activities. Finally, we rebuilt the signal by using the remaining components (Mantini et al., 2011). After the preprocessing steps, we epoched the signal in 4130 trials (one for each sound) lasting 700 ms each (with 100 ms of pre-stimulus time that was used for baseline correction). To be noted, in a few cases the number of trials was lower than 4130. This happened when a few segments of the data were previously discarded due to the presence of large artefacts.

2.5. MEG sensor analysis

Although our focus was on the MEG source reconstructed brain data, a first analysis on MEG sensors data was computed, in accordance with state-of-the-art guidelines about best practice in MEG analysis (Gross et al., 2013).

Thus, according to a large number of MEG and electroencephalography (EEG) task-related studies (L. Bonetti, Brattico, Carlomagno, et al., 2021; Brattico et al., 2006; Gross et al., 2013), we averaged the trials over conditions, and we combined planar gradiometers by sum-root square. Then, we assessed whether the deviant stimuli elicited a clear MMN and P300 signal by contrasting the brain responses to our six categories of deviants against the standard stimuli. Since this contrast has been done for each deviant (six), each time point (182, ranging from 0 to 300 ms from the onset of the deviant stimuli), and each MEG combined gradiometer channel (102), we have corrected for multiple comparisons by using Bonferroni correction and thus lowering the *p-value* to 4.5e-07 (0.05/(6 * 182 * 102)). In this analysis, we used combined gradiometers only because of their better signal-to-noise ratio than magnetometers when performing analysis on the MEG sensor level (Gross et al., 2013). The results showed that the MMN was strongly elicited among all deviants, are illustrated in Fig. 2. The detailed statistical results showing significant time-points and channels for each deviant are reported in **Table ST1**.

2.6. Source reconstruction

We reconstructed the neural sources of the brain activity recorded on the scalp by the MEG channels, applying the widely adopted procedure named beamforming. Here, we used the OSL implementation consisting of a local-sphere forward model and a beamformer approach as the inverse method (Hillebrand and Barnes, 2005; Huang et al., 2004; Li et al., 2004). The local-sphere forward model considers the MNI-co-registered anatomy as a simplified geometric model, and it fits a sphere separately

for each sensor (Nolte, 2003). Then, the beamforming employs a different set of weights sequentially applied to the source locations to isolate the contribution of each source to the activity recorded by the MEG channels for each different time point. Importantly, we used all MEG channels (both magnetometers and non-combined planar gradiometers) for performing the source reconstruction. Moreover, the covariance matrix necessary for computing the weights required for the beamforming was calculated using the data for all the experimental conditions (i.e. all deviants and standards). Finally, we used a three-dimensional eight-mm grid which resulted in a brain parcellation of 3559 dipoles (sources).

2.7. Neural sources of MMN peaks

We computed an independent GLM sequentially for each time point at each dipole location, where we contrasted each deviant category against the standard stimuli. This procedure, computed independently for each participant, allowed us to detect the contrast of parameter estimates (COPEs) for the brain activity specifically associated with the detection of the deviant stimuli (i.e. the MMN in source space). These results were then submitted to a second-level (group) analysis, using one-sample t-tests with spatially smoothed variance obtained with a Gaussian kernel (full-width at half-maximum: 50 mm).

Although the analysis was computed for each time-point in the epoch, we were only interested in the brain sources of the MMNs peak, since the focus of the current study was on the early automatic prediction error indexed by MMN and elicited by cognitive and acoustic deviants. Thus, first we detected the peak MMN activity independently for each deviant. To this aim, we have used the large corpus of previous studies of MMN (see, for example, Bonetti et al., 2021a,b,c; Bonetti et al., 2018; Brattico et al., 2006; Escera et al., 2000; Fisher et al., 2008; Garrido et al., 2009; Näätänen et al., 2007; Näätänen et al., 2004; Putkinen et al., 2014; Tervaniemi et al., 2014; Tervaniemi et al., 2016) which clearly showed that MMN peaks between 120 and 250 ms from the onset of the deviant stimulus. Thus, we have searched for the minimum amplitude (since the MMN is a negative component) within that time-window, independently for each deviant. In this way, we were sure to identify the MMN peaks based both on the indications from the previous literature and on objective methods specifically related to our stimuli. Second, after identifying the MMN peaks, we extracted and averaged the group-level results around those peaks (considering a small time-window of ±25 ms around each MMN peak). This procedure returned the strength of the MMNs to the six deviants for each brain dipole. To correct for multiple comparisons, we performed a cluster-based permutation test with 5000 permutations which allowed us to isolate the clusters of brain activity underlying the generation of the MMNs. Since we computed six tests (one for each deviant), we have used an α level of 0.0017 (0.05/6), corresponding to a cluster forming threshold t-value = 3.3.

2.8. MMNs to cognitive versus acoustic deviants

After detecting the sources of the brain signals underlying the MMNs peak, we performed a further analysis to assess whether such sources differed according to the category of deviants. Specifically, we were interested in assessing whether cognitive deviants (transposition, melody modulation and rhythm modulation) elicited an MMN with different brain sources than acoustic deviants (mistuning, timbre, rhythm mistake). Thus, first we averaged together the neural activity of the three deviants forming the two categories. We conducted this procedure independently for each participant. Second, we computed a t-test for each brain dipole comparing the brain activity underlying cognitive versus acoustic deviants. Finally, to correct for multiple comparisons, we performed cluster-based Monte-Carlo simulations (MCS) (Bonetti, Brattico, Bruzzone, et al., 2021; Bonetti, Brattico, Carlomagno, et al., 2021; Bonetti et al., 2020; Kroese et al., 2011). Specifically, the MCS

consisted of detecting the spatial clusters of significant dipoles (dipoles whose test had a *p-value* lower than the MCS *α* level) in the original data and assessing whether they were significant or occurred by chance. First, spatial clusters were identified in the original data. Then, we permuted the original data and detected the clusters in this new permuted set of brain values. We computed this procedure 1000 times, obtaining a reference distribution of cluster sizes detected for each permutation. Finally, the original cluster sizes were compared to the reference distribution and considered significant if they were bigger than 99.9% (MCS *p-value* of .001) of the permuted cluster sizes. In this case, we computed two MCS, one for the significant dipoles where cognitive deviants were stronger than acoustic ones, and another one for the dipoles where the acoustic deviants were stronger than the cognitive ones. Remarkably, while cognitive versus acoustic deviants returned a significant cluster only with a standard cluster-forming threshold *p-value* = .05, acoustic versus cognitive deviants returned a significant cluster even when lowering the cluster-forming threshold *p-value* to 1.0e-04, indicating a very large significant difference. Details of the outcomes of these analyses are reported in the Results section.

2.9. Cognitive, acoustic deviants and musicianship

The last step of our analysis pipeline was to assess whether there was a relationship between musical expertise and the brain areas activated during the perception of cognitive and acoustic deviants. Thus, we computed Pearson's correlations for each brain dipole between the participants' years of music playing and their brain activity underlying deviant detection. Afterwards, we corrected for multiple comparisons employing an MCS analogous to the one described above. In this case, since we computed two independent MCS analyses, one for the cognitive and one for the acoustic deviants, we used a cluster-forming threshold *p-value* = .01 and an MCS *p-value* = .001. To strengthen the reliability of our results, we computed an additional analysis by assessing whether participants' age and sex affected the relationship between the neural responses to cognitive and acoustic deviants and the participants' musical expertise. As done before, we computed one statistical test for each brain dipole. In this case, we performed one multiple linear regression for each brain voxel inserting the neural data as dependent variable and musical training, age, and sex as independent variables. Then, we corrected for multiple comparisons using the same cluster-based Monte-Carlo simulations that we described above (cluster-forming threshold *p-value* = .01 and an MCS *p-value* = .001).

2.10. Deviants with different spectral content

Finally, four out of six deviants presented a relevant change in spectral content (i.e. Mistuning, Melody modulation, Transposition and Timbre deviant), while the other two did not (Rhythm mistake and Rhythm modulation). Thus, we computed an additional analysis by averaging the neural sources of the MMN peaks for those two categories of deviants (on the one hand Mistuning, Melody modulation, Transposition and Timbre deviant and on the other hand Rhythm mistake and Rhythm modulation) and contrasting them by computing t-tests for each brain source, mirroring the procedure described above for cognitive versus acoustic deviants.

3. Results

3.1. Experimental design and MMNs detection

Our study had three main aims: reconstructing the neural sources of the deviants inserted in a melodic multifeature paradigm (i), assessing whether such neural sources differed across cognitive and acoustic deviants (ii), investigating the relationship between the neural sources of cognitive and acoustic deviants and the musical expertise of the participants (iii).

We employed the Melodic Multifeature paradigm (MusMelo), which was introduced by Tervaniemi and colleagues (Tervaniemi et al., 2014) and Putkinen and colleagues (Putkinen et al., 2014) and consists of a series of deviants breaking cognitive (transposition, melody modulation and rhythm modulation) or acoustic (mistuning, timbre, and rhythm mistake) musical features (Fig. 1). Thus, it is the ideal paradigm to assess whether the MMNs neural sources vary depending on the characteristic of the deviants (cognitive versus acoustic). To address our research questions, we presented our 104 participants with the MusMelo paradigm while we collected their brain activity using MEG.

A) Graphical depiction of the MusMelo paradigm, used during magnetoencephalography (MEG) recording. The MusMelo consists of brief recursive melodies played consecutively in a loop. In these melodies, six different deviants have been inserted. The deviants belonged to two categories: acoustic deviants (mistuning, timbre, rhythm mistake) and cognitive deviants (melody modulation, rhythm modulation, transposition). **B**) Graphical schema showing the beginning of the Musmelo paradigm. This example shows that the melodies started with a triad (duration of 300 ms), followed by four tones of different length, plus an ending tone (duration of 575 ms). Subsequent tones were always separated by 50 ms silence. Additionally, a 125 ms silence after the ending tone was inserted. Thus, one melody lasted for 2100 ms in total. Such melodies were presented for 15 min in a looped, recursive manner. Within these repeated melodies, the above-described six deviants were inserted. Importantly, as illustrated in the figure, these deviants occurred in a random order and at random places within the melody, thus it was not possible for the participants to predict when the deviants occurred.

Although our focus was on the neural sources of the MMNs elicited by the six deviants of the MusMelo, prior to computing the analysis in MEG source space, we verified that deviants and standards had elicited a significantly different neural signal, as recorded by the MEG sensors. We computed one-sample t-tests for each MEG-combined gradiometer channel (102), each time point (152, ranging from 50 to 300 ms after the onset of the stimuli), and each deviant, comparing the brain response to the deviant versus the standard stimuli. We corrected for multiple comparisons by using Bonferroni correction which resulted in an adjusted *p-value* of 5.3e-07 (0.05/ (6 * 152 * 102)). Here, we used combined gradiometers because their signal-to-noise ratio is usually better than magnetometers when computing analysis at MEG sensor level (in Bonetti et al. (2018) and Haumann et al. (2016)) quantitative measures of signal-to-noise ratio for this same dataset are illustrated). Our results showed that significant differences in the brain activity elicited by deviants versus standards were clearly identified among several MEG channels and time points ($p < 5.3e\text{-}07$). T-values showing the effect size of this difference for each time-point, as well as the percentage of significant MEG channels over time, are reported in detail in **Table ST1**. Moreover, Fig. 2 illustrates a representative MEG magnetometer channel which shows the polarity of the differences between the activity evoked by deviants minus standards. The plot clearly illustrates the MMN peaks for the six deviants that we used in the subsequent analysis.

3.2. Neural sources of MMN peaks

After verifying the reliability of our paradigm in detecting clear MMN signals, we reconstructed the sources of the neural signal by combining the MEG and MRI data of each participant. Specifically, as widely done in the field, we used a local-sphere forward model and a beamformer approach as the inverse method (see Methods for details). Our procedure returned a time series describing the strength of the neural signal over time for each category of stimuli (deviant and standard) and for each of the reconstructed 3559 brain sources (dipoles).

Then, we contrasted the neural activity underlying deviants versus standards, focusing on the brain sources generating the MMN peaks (considering a time window of ±25 ms around each of the MMN peaks). We corrected for multiple comparisons using a cluster-based
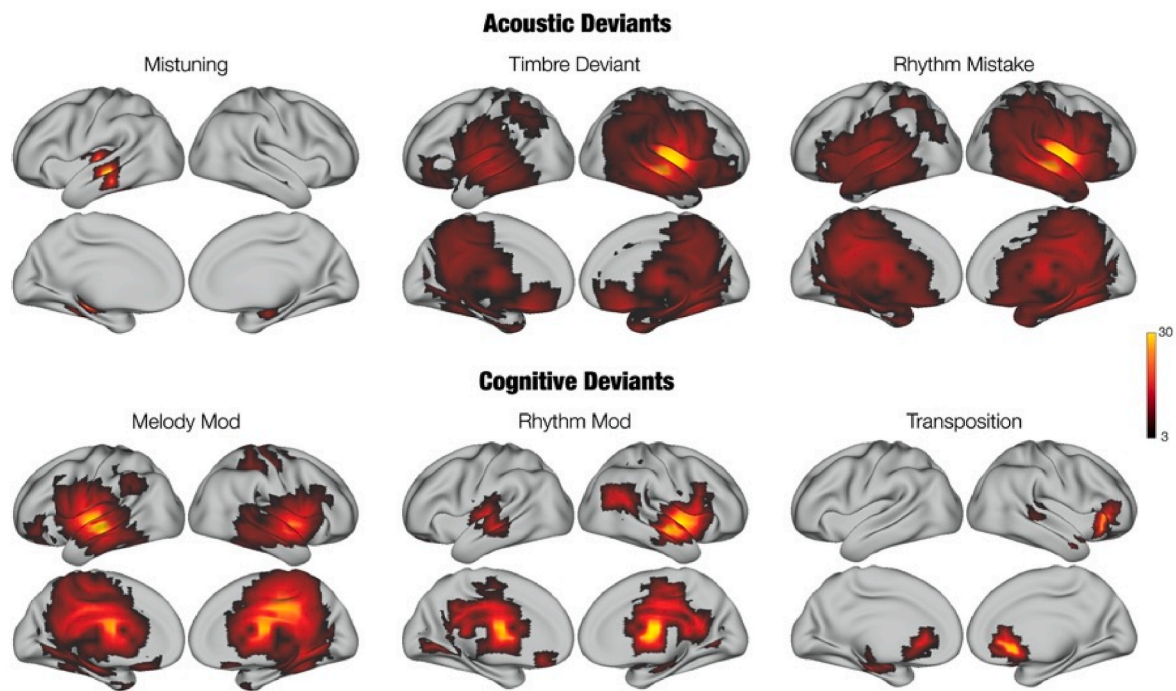
Acoustic Deviants

Mistuning　　　　Timbre Deviant　　　　Rhythm Mistake

Cognitive Deviants

Melody Mod　　　　Rhythm Mod　　　　Transposition

Fig. 3. Brain sources of the MMN to all deviants depicted in brain templates. The colorbar indicates t-values obtained by contrasting the brain response to deviant versus standard sounds considering the full sample of 104 participants. The top row illustrates acoustic deviants, while the bottom row depicts cognitive deviants. Overall, acoustic deviants show strong activity in the auditory cortex, while cognitive deviants highlight the contribution of cingulate and frontal brain areas to the generation of the MMN.

permutation test [65] with 5000 permutations which allowed us to isolate the significant clusters of brain activity underlying the generation of the MMNs. Since we computed six tests (one for each deviant), we have used an α level of 0.0017 (0.05/6), corresponding to a cluster forming threshold t-value = 3.3.

As depicted in Fig. 3 and reported in detail in **Table ST2**, these analyses ($p < .0017$) returned a main involvement of the primary and secondary auditory cortices, especially for timbre, rhythm mistake, melody modulation, and mistuning. Remarkably, medial cingulate gyrus and hippocampal regions were also strongly activated by the presentation of the deviant stimuli. This result was particularly evident for melody and rhythm modulation, rhythm mistake, and timbre. Finally, transposition, which is a rather cognitive and complex deviant, elicited activity mainly localized in the anterior part of the cingulate and in the inferior frontal gyrus.

3.3. MMNs to cognitive versus acoustic deviants

After detecting the sources of the brain signals underlying the MMNs peak, we performed a further analysis to assess whether these sources differed when comparing cognitive (transposition, melody modulation, and rhythm modulation) versus acoustic (mistuning, timbre, and rhythm mistake) deviants. Thus, first we averaged together the neural activity of the three deviants in each category. Second, we computed a t-test for each brain dipole comparing the brain activity underlying cognitive versus acoustic deviants. Finally, to correct for multiple comparisons, we performed cluster-based MCS (MCS p-value $< .001$).

When using a cluster-forming threshold p-value $< .05$ (see Methods for details), we identified a small, but significant cluster of activity where cognitive deviants had a stronger neural signal than acoustic ones. This cluster was localized in the medial cingulate gyrus.

Remarkably, when computing the MCS to identify the clusters where the brain activity was stronger for acoustic versus cognitive deviants, we observed a large cluster which was largely significant (cluster forming threshold p-value $< 1.0e$-04). This cluster mainly originated in the right primary auditory cortex, but extended to secondary auditory cortex, insula, frontal operculum, and hippocampal regions (Fig. 4A and **Table ST3**).

3.4. Cognitive, acoustic deviants, musicianship, and spectral features

Finally, we wished to assess whether there was a relationship between musical expertise and the neural sources of the MMNs elicited by cognitive and acoustic deviants. Thus, we computed Pearson's correlations between the participants' years of music playing and their brain activity underlying deviant detection. This analysis was computed for each brain source originating the peak of the MMNs. We corrected for multiple comparisons employing an MCS analogous to the one described above (cluster-forming threshold p-value $< .01$ and MCS p-value $< .001$).

This analysis showed significant clusters of positive correlations between musical expertise and neural response to deviants (Fig. 4B and **C, Table ST4**). Interestingly, such relationship was particularly evident for the cingulate, inferior frontal gyri, and supplementary motor area.

We complemented the analysis using Pearson's correlations by employing multiple linear regressions to assess whether participants' age and sex affected the relationship between the neural responses to cognitive and acoustic deviants and the participants' musical expertise. The results corrected for multiple comparisons (using the same cluster-based MCS described above) confirmed the relationship between higher-order brain areas and musical expertise, especially with regards to cognitive deviants (Fig. S1 and **Table ST4**).

Finally, since four out of six deviants presented a relevant change in spectral content (i.e. Mistuning, Melody modulation, Transposition and Timbre deviant), we computed an additional analysis by contrasting them against the ones with no spectral changes (i.e. Rhythm modulation and Rhythm mistake). As reported in **Table ST5** and illustrated in Fig. S2, this analysis showed that deviants characterized by spectral change presented a stronger activity in the right auditory cortex and in the precuneus.
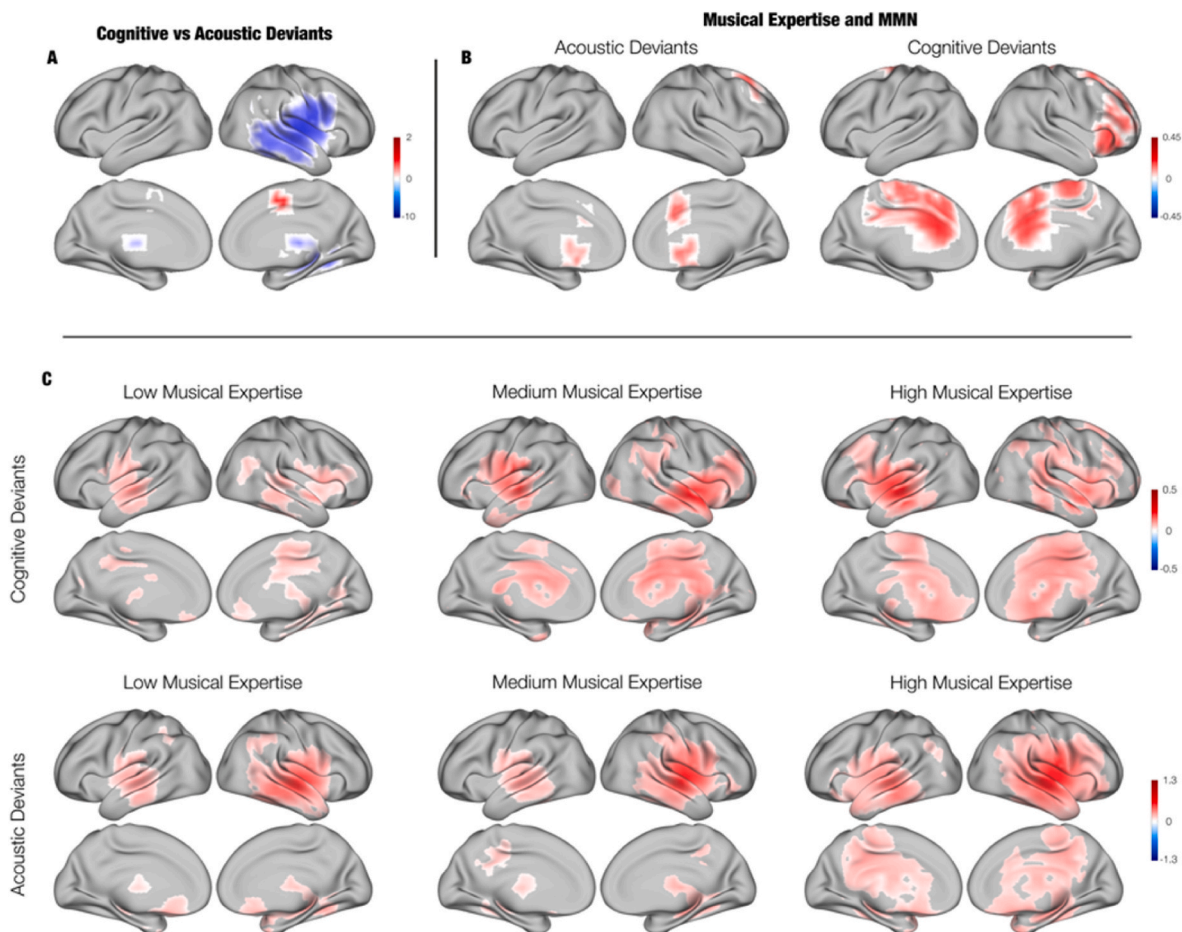
Fig. 4. MMN to acoustic versus cognitive deviants and musical expertise

(A) Depiction in brain templates of the contrast between cognitive versus acoustic deviants considering the full sample of 104 participants. The colorbar shows the t-values emerged from the contrast. Specifically, the red colour indicates the brain areas that were more active for cognitive versus acoustic deviants, while the blue colour shows the brain areas that were more active for acoustic versus cognitive deviants. This plot indicates that the medial cingulate gyrus was more active for the cognitive deviants, while a large network of brain areas around the right auditory cortex was more engaged during processing of acoustic versus cognitive deviants. **(B)** Depiction in brain templates of the correlation between musical expertise and MMN to acoustic (left) and cognitive (right) deviants. The colorbar shows the r-value obtained from the correlations. Specifically, the red colour indicates positive correlations between MMN and musical expertise. No negative correlations (that would have been indicated in blue) were observed in this case. The plot shows that musical expertise is positively associated with the neural sources of the MMN, especially in the medial cingulate gyrus and right orbito-frontal cortex, and particularly in response to cognitive deviants. **(C)** Depiction in brain template of cognitive (top row) and acoustic (bottom row) deviants, independently for three categories of participants, characterized by low (up to two years of musical training, n = 42), medium (between three and ten years of musical training, n = 27), and high (more than 10 years of musical training, n = 32) musical expertise. The colorbar shows the strength of the neural activity weighted by the intra-trial and -individual variance. This plot confirms that higher musical expertise is associated with stronger neural activity in higher-order brain areas, especially with regards to the cognitive deviants.

4. Discussion

In this study, we aimed to assess the brain automatic prediction error indexed by the MMN elicited by cognitive and acoustic deviants inserted in a musical context. Results revealed that the neural sources of the MMNs were mainly localized in the auditory cortex. However, core clusters of significant activity were also clearly localized in the cingulate gyrus, hippocampal, and frontal areas. In addition, weaker cluster of significant activity were observed in postcentral gyrus and caudate. However, these last, weaker clusters may be the results of the artificial source leakage which is normally associated with MEG source reconstruction of strong neural activity. Notably, the contrast between cognitive versus acoustic deviants showed stronger activity within the cingulate gyrus for the cognitive deviants. Conversely, the acoustic deviants elicited stronger responses in the auditory cortex. At last, we revealed that musical expertise modulated the sources of the brain prediction error indexed by MMN to both categories of deviants. Notably, such modulation was stronger for cognitive deviants and

involved especially the cingulate, inferior frontal gyri, and supplementary motor area. Finally, these results occurred independently on the attention of the participants since while they were presented with the sound stimulation they were asked to watch and focus on a silent movie. Thus, our study suggests that even when behavioral responses are not required, the brain is showing a rather strong ability to automatically discriminate the fine-grained difference between deviants that are characterized by cognitive or acoustic features.

The brain sources which generated the MMNs were coherent with the sources reported by previous literature. Specifically, several studies showed that auditory cortex, and especially Heschl's gyrus together with superior and middle temporal gyri, and the hippocampal area are primarily implicated in the generation of the MMN (Fitzgerald and Todd, 2020; Marco-Pallarés et al., 2005; Tervaniemi et al., 2006a,b; Waberski et al., 2001). In addition, we detected sources in the medial and anterior cingulate gyrus, hippocampal areas, and frontal operculum/inferior frontal gyrus. This is also supported by previous research which highlighted frontal generators of MMN (Waberski et al., 2001), proposing

that they were necessary for the process of switching attention to the deviant stimulation (Heilman & Van Den Abell, 1980; Kastner and Ungerleider, 2000; Mesulam and Marchsel, 1981; Posner, 2016). Moreover, apart from Melody modulation, all the other deviants presented a hemispheric lateralization. In fact, Timbre, Rhythm mistake, Rhythm modulation and Transposition deviants were associated with neural sources that were mainly localized in the right hemisphere. This result is very common when dealing with fast processing of auditory stimulation. Indeed, several studies (Fitzgerald and Todd, 2020; Marco-Pallarés et al., 2005; Posner, 2016; Waberski et al., 2001) showed a stronger involvement of the right hemisphere, especially with regards to auditory cortex and inferior-frontal regions. On the contrary, it is interesting to observe that the Mistuning deviant was associated with both left and right hippocampus and mainly left auditory cortex. This may be a true effect suggesting that modulation of melodies (as it happens for Mistuning) may be more related to the left hemisphere (as a matter of fact, Melody modulation also showed a slight predominance of left-over right hemisphere although the difference was subtle). However, it should also be noted that the Mistuning deviant used in this study was elicited by a very small mistuning of the standard sounds and presented one of the weakest MMNs among the six deviants. Thus, the weak MMN to Mistuning may also be directly related to the small strength of the deviation that did not allow the recording of a strong signal in both hemispheres.

Interestingly, we detected a dissociation between the sources underlying processing of cognitive versus acoustic deviants. Indeed, while auditory cortex was primarily recruited by the processing of acoustic deviants, a higher-order area such as the medial cingulate gyrus was stronger for the cognitive deviants. Moreover, insula and frontal brain areas such as anterior cingulate gyrus and inferior frontal gyrus exhibited greater activity than auditory cortex when observing the MMN sources of two of our cognitive deviants, namely rhythm modulation and transposition. In particular, the transposition deviant is thought to be the most cognitive deviant of the paradigm. Indeed, Putkinen and colleagues (Putkinen et al., 2014) showed that transposition was the only deviant not evoking larger MMN in music-trained children than in control ones. These findings broadened our understanding of MMN sources and auditory prediction error. In fact, there was no evidence in favour of a higher switching of attention for the cognitive versus acoustic deviants. Thus, the higher involvement of frontal brain regions observed in our study for the cognitive deviants should not be connected to the "attention switching" hypothesis (Heilman & Van Den Abell, 1980; Kastner and Ungerleider, 2000; Mesulam and Marchsel, 1981; Posner, 2016) mentioned above. Conversely, we argue that cognitive and acoustic deviants elicit two diverse types of auditory prediction error. As a matter of fact, acoustic deviants are perceived as "mistakes" occurring in the music. On the contrary, cognitive deviants are actual changes of the musical information carried by the melodies. In other words, in the first case the brain may simply notice an impaired quality of the musical information, while in the second scenario, the prediction error operated by the brain would be more complex, leading to the understanding that musical information has actually changed.

Notably, even though automatic and independent of a participant's attention, the prediction error associated with changes in musical information were generated by higher-order brain areas usually associated with language processing and conscious cognitive abilities, such as inferior frontal gyrus (Fadiga, Craighero, & D'Ausilio, 2009; Kotz et al., 2002; Tyler et al., 2011) and cingulate gyrus (Apps et al., 2016; Bach et al., 2008; Hampson et al., 2006). Conversely, our findings suggest that musical "mistakes" such as imprecise rhythms, small mistunings, or sudden variations in timbre would not require such complex processing. Indeed, in this case, the recruitment of the auditory cortex would be enough to detect the changes in the physical, acoustic features of the sounds.

Among our results, of particular interest is the role of cingulate gyrus which has been previously connected to several functions, including

prediction error. For example, Alexander and colleagues (Alexander and Brown, 2019) highlighted the role of anterior cingulate cortex (ACC) in processing behavioral error and signalling deviations between expected and observed events, describing it within the framework of reinforcement learning. Similarly, an activation likelihood estimation (ALE) meta-analysis investigated the neural correlates of prediction error in reinforcement learning. Authors found that ACC, medial prefrontal cortex (mPFC) and striatum were the key brain areas underlying prediction error, in studies that used both rewarding and aversive reinforcers (Garrison et al., 2013). Another fMRI study investigated the brain activity underlying a numerical Stroop task, reporting activity in the ACC when participants processed errors in the task (Critchley et al., 2005). Along this line, another contribution claimed that ACC learnt to predict error likelihood in each context, even for trials in which there was no error (Brown and Braver, 2005). A simulation study on mPFC and ACC provided modelling evidence in support of the role of these brain structures for error likelihood, signalling mistakes, and reward, concluding that they are central for learning and predicting the likely outcomes of actions whether good or bad (Alexander and Brown, 2011). Furthermore, Bonetti and colleagues (Bonetti, Brattico, Bruzzone, et al., 2021; Bonetti, Brattico, Carlomagno, et al., 2021; Bonetti et al., 2020) showed that cingulate gyrus is of primary importance for both active encoding and recognition of auditory sequences, and that its involvement positively correlate with the strength of the recognition (Fernandez Rubio et al., 2022a; Fernandez Rubio et al., 2022b). Their findings revealed that the cingulate is more central within the whole-brain network when encoding sounds than when resting (Bonetti, Brattico, Carlomagno, et al., 2021). Moreover, they found that recognition of previously learned compared to novel melodies was associated with stronger cingulate activity (Bonetti, Brattico, Bruzzone, et al., 2021; Bonetti et al., 2020). Along this line, a recent meta-analysis (Pando-Naude et al., 2021) on music perception, imagery and production highlighted the involvement of cingulate gyrus when participants were asked to do a variety of different tasks concerning music listening and production, and mental manipulation of sounds. Taken together, this evidence supports the idea that cingulate gyrus may be a key structure for extracting information from musical sequences and signalling variations from the previously learned melodies.

Conversely, acoustic "mistakes" involving basic acoustic features of musical sounds may recruit a more restricted network of auditory brain areas. This evidence is supported by previous studies employing simpler oddball and multi-feature paradigms which highlighted the primary role of auditory cortex in the MMN generations. For instance, Marco-Pallares (Marco-Pallarés et al., 2005) and colleagues reconstructed the main sources of MMN measured with EEG within supratemporal and middle temporal cortex, bilaterally. Similarly, Waberski and colleagues (Waberski et al., 2001) found the main generators of MMN in supratemporal brain regions. They also reported secondary sources, with a longer latency, localized in the cingulum and right inferior temporal gyrus. Notably, this conclusion was reached even in intracranial electroencephalography (iEEG) recording, where MMN sources were observed in Brodmann areas 21 and 42, corresponding to middle temporal gyrus and posterior transverse temporal cortex, respectively (KROPOTOV et al., 1995). Moreover, additional evidence pointed out that the auditory cortex is mainly implicated in the processing of basic acoustic features of sounds and music. For instance, in a classic work, Zatorre and colleagues (Zatorre et al., 2002) argued that auditory cortices in the two hemispheres are specialized to extract fundamental acoustic features of both music and speech such as temporal and spectral content of sounds. Specifically, they reported that temporal resolution was better in left auditory cortical regions while spectral resolution of the sounds was greater in right auditory cortical regions. In a more recent review, King and colleagues (King et al., 2018) highlighted the complexity of the auditory cortex and its important role also for high-level cognitive processes. Still, they reiterated that auditory cortex shows selectivity for sound features, which is likely at the basis of

processing of natural sounds, such as during speech and in real-life listening scenarios.

Finally, we revealed that musical expertise modulated the brain sources of the prediction error signal elicited by cognitive and acoustic deviants. Notably, this modulation was primarily evident in high-order brain areas such as the cingulate, inferior frontal gyri, and supplementary motor area. Moreover, this relationship was primarily evident for the cognitive deviants. This finding is coherent with a large corpus of studies which showed that the brains of musicians are different from non-musicians'. Indeed, the musician's brain has been suggested as a model of neuroplasticity (Münte et al., 2002), being shaped by long-lasting musical training. This hypothesis was further supported by several longitudinal studies showing structural brain changes, especially in children, after exposure to musical training (Habibi et al., 2018; Hyde et al., 2009). Likewise, a recent meta-analysis revealed that structural and functional brain differences emerged when comparing the brain of musicians versus non-musicians (Criscuolo et al., 2022). Back to MMN research, several works reported a stronger MMN activity recorded in brains of participants with higher musical expertise (Brattico et al., 2009; Koelsch et al., 1999; Tervaniemi et al., 2005; Vuust et al., 2011b, 2012). Additionally, Vuust and colleagues (Vuust et al., 2012) found different brain responses even across diverse categories of musicians. For instance, they revealed that the brain of jazz versus classical musicians was more sensitive to pitch and pitch-sliding deviants, features which are particularly involved in jazz training. In light of previous findings, our results provide additional evidence that musical expertise is associated with higher-level processing of music in the brain. Further, our study suggests that to outperform non-musicians when extracting varied information from musical melodies, musicians rely on stronger activity of higher-order brain areas such as cingulate and inferior frontal gyri, and supplementary motor area.

In conclusion, our study showed that the brain employs different strategies for processing cognitive and acoustic auditory early prediction error as indexed by MMN, and that musical expertise modulates such mechanisms. Future research is called to investigate auditory prediction error in a wider array of cognitive and acoustic deviants and assess whether similar results arise when performing attentive tasks which require a conscious elaboration of the musical information. In addition, while in this study we specifically focused on the neural sources underlying MMN responses to cognitive and acoustic deviants, future studies may investigate later brain mechanisms to assess whether the neural processing of cognitive and acoustic deviations is also discriminable 300–600 ms after the onset of the sound irregularities. Likewise, our study provided a preliminary result related to deviants which were characterized by spectral change, pointing to a stronger involvement of right auditory cortex and precuneus cortex for deviants with greater spectral change. Future research may develop novel experimental paradigms specifically designed to better investigate such topic. Finally, as previously done by Tervaniemi and colleagues (Tervaniemi et al., 2006a,b) and Pulvermüller and Shtyrov (2006), an additional indication for future studies consists of investigating acoustic and cognitive deviants in contexts different from music, such as employing linguistic experimental design and investigating speech sound MMNs.

CRediT authorship contribution statement

L. Bonetti: Conceptualization, Methodology, Software, Validation, Formal analysis, Data curation, Writing – original draft, Writing – review & editing, Funding acquisition, Visualization. **F. Carlomagno:** Methodology, Software, Formal analysis, Writing – review & editing, Visualization. **M. Kliuchko:** Data curation, Methodology. **B.P. Gold:** Data curation, Methodology, Writing – review & editing. **S. Palva:** Methodology, Writing – review & editing. **N.T. Haumann:** Methodology, Software. **M. Tervaniemi:** Data curation, Methodology, Writing – review & editing, Funding acquisition, Supervision. **M. Huotilainen:** Data curation, Methodology, Funding acquisition, Supervision. **P.**

Vuust: Writing – review & editing, Funding acquisition, Supervision. **E. Brattico:** Conceptualization, Resources, Data curation, Writing – review & editing, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request. The codes used for the analysis are reported in the following GitHub repositories: https://github.com/leonardob92/Whole_brain_computation_of_cognitive_vs_acoustic_errors_in_music_NeuroImageReports.git; https://github.com/leonardob92/LBPD-1.0.git.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ynirp.2022.100145.

References

Alexander, W.H., Brown, J.W., 2011. Medial prefrontal cortex as an action-outcome predictor. Nat. Neurosci. https://doi.org/10.1038/nn.2921.

Alexander, W.H., Brown, J.W., 2019. The role of the anterior cingulate cortex in prediction error and signaling surprise. Topics in Cognitive Science. https://doi.org/10.1111/tops.12307.

Apps, M.A.J., Rushworth, M.F.S., Chang, S.W.C., 2016. The anterior cingulate gyrus and social cognition: tracking the motivation of others. Neuron. https://doi.org/10.1016/j.neuron.2016.04.018.

Bach, D.R., Grandjean, D., Sander, D., Herdener, M., Strik, W.K., Seifritz, E., 2008. The effect of appraisal level on processing of emotional prosody in meaningless speech. Neuroimage. https://doi.org/10.1016/j.neuroimage.2008.05.034.

Bonetti, L., Brattico, E., Bruzzone, S.E.P., Donati, G., Deco, G., Pantazis, D., et al., 2021a. Temporal pattern recognition in the human brain: a dual simultaneous processing. bioRxiv.

Bonetti, L., Brattico, E., Carlomagno, F., Donati, G., Cabral, J., Haumann, N.T., et al., 2021b. Rapid encoding of musical tones discovered in whole-brain connectivity. Neuroimage. https://doi.org/10.1016/j.neuroimage.2021.118735.

Bonetti, L., Bruzzone, S.E.P., Sedghi, N.A., Haumann, N.T., Paunio, T., Kantojärvi, K., et al., 2021c. Brain predictive coding processes are associated to COMT gene Val158Met polymorphism. Neuroimage. https://doi.org/10.1016/j.neuroimage.2021.117954.

Bonetti, L., Haumann, N.T., Brattico, E., Kliuchko, M., Vuust, P., Sarkamo, T., Naatanen, R., 2018. Auditory sensory memory and working memory skills: association between frontal MMN and performance scores. Brain Res. 1700, 86–98. https://doi.org/10.1016/j.brainres.2018.06.034.

Bonetti, L., Haumann, N.T., Vuust, P., Kliuchko, M., Brattico, E., 2017. Risk of depression enhances auditory Pitch discrimination in the brain as indexed by the mismatch negativity. Clin. Neurophysiol. 128 (10), 1923–1936. https://doi.org/10.1016/j.clinph.2017.07.004.

Bonetti, Leonardo, Brattico, E., Carlomagno, F., Cabral, J., Stevner, A., Deco, G., et al., 2020. Spatiotemporal brain dynamics during recognition of the music of Johann Sebastian Bach. bioRxiv. https://doi.org/10.1101/2020.06.23.165191.

Brattico, E., Pallesen, K.J., Varyagina, O., Bailey, C., Anourova, I., Jarvenpaa, M., et al., 2009. Neural discrimination of nonprototypical chords in music experts and laymen: an MEG study. J. Cognit. Neurosci. 21 (11), 2230–2244. https://doi.org/10.1162/jocn.2008.21144.

Brattico, E., Pearce, M., 2013a. The neuroaesthetics of music. Psychology of Aesthetics, Creativity, and the Arts. https://doi.org/10.1037/a0031624.

Brattico, E., Pearce, M., 2013b. The neuroaesthetics of music overture : from neuroscience to neuroaesthetics. Psychology of Aesthetics, Creativity, and the Arts. https://doi.org/10.1037/a0031624.

Brattico, E., Tervaniemi, M., Näätänen, R., Peretz, I., 2006. Musical scale properties are automatically processed in the human auditory cortex. Brain Res. 1117 (1), 162–174. https://doi.org/10.1016/j.brainres.2006.08.023.

Brown, J.W., Braver, T.S., 2005. Learned predictions of error likelihood in the anterior cingulate cortex. Science. https://doi.org/10.1126/science.1105783.

Criscuolo, A., Bonetti, L., Sarkamo, T., Kliuchko, M., Brattico, E., 2019. On the association between musical training, intelligence and executive functions in adulthood. Front. Psychol. https://doi.org/10.3389/fpsyg.2019.01704.

Criscuolo, A., Pando-Naude, V., Bonetti, L., Vuust, P., Brattico, E., 2022. An ALE meta-analytic review of musical expertise. Scientific reports 12 (1), 1–17.

Critchley, H.D., Tang, J., Glaser, D., Butterworth, B., Dolan, R.J., 2005. Anterior cingulate activity during error and autonomic response. Neuroimage. https://doi.org/10.1016/j.neuroimage.2005.05.047.

Escera, C., Alho, K., Schröger, E., Winkler, I., 2000. Involuntary attention and distractibility as evaluated with event-related brain potentials. Audiol. Neuro. Otol. 5 (3–4), 151–166. https://doi.org/10.1159/000013877. PMID: 10859410.

Fadiga, L., Craighero, L., D'Ausilio, A., 2009. Broca's area in language, action, and music. In: Annals of the New York Academy of Sciences. https://doi.org/10.1111/j.1749-6632.2009.04582.x.

Fernandez Rubio, G., Brattico, E., Kotz, S.A., Kringelbach, M.L., Vuust, P., Bonetti, L., 2022a. The spatiotemporal dynamics of recognition memory for complex versus simple auditory sequences. bioRxiv.

Fernandez Rubio, G., Carlomagno, F., Vuust, P., Kringelbach, M.L., Bonetti, L., 2022b. Associations between abstract working memory abilities and brain activity underlying long-term recognition of auditory sequences. PNAS Nexus 1 (4), pgac216. https://doi.org/10.1093/pnasnexus/pgac216.

Fisher, D.J., Labelle, A., Knott, V.J., 2008. The right profile: mismatch negativity in schizophrenia with and without auditory hallucinations as measured by a multi-feature paradigm. Clin. Neurophysiol. https://doi.org/10.1016/j.clinph.2007.12.005.

Fitzgerald, K., Todd, J., 2020. Making sense of mismatch negativity. Front. Psychiatr. https://doi.org/10.3389/fpsyt.2020.00468.

Franklin, M.S., Moore, K.S., Yip, C.-Y., Jonides, J., Rattray, K., Moher, J., 2008. The effects of musical training on verbal memory. Psychol. Music 36 (3), 353–365. https://doi.org/10.1177/0305735607086044.

Friston, 2018. Does predictive coding have a future? Nat. Neurosci. https://doi.org/10.1038/s41593-018-0200-7.

Friston, K., Kiebel, S., 2009. Predictive coding under the free-energy principle. Phil. Trans. Biol. Sci. https://doi.org/10.1098/rstb.2008.0300.

Garrido, M.I., Kilner, J.M., Stephan, K.E., Friston, K.J., 2009. The mismatch negativity: a review of underlying mechanisms. Clin. Neurophysiol. https://doi.org/10.1016/j.clinph.2008.11.029.

Garrison, J., Erdeniz, B., Done, J., 2013. Prediction error in reinforcement learning: a meta-analysis of neuroimaging studies. Neurosci. Biobehav. Rev. https://doi.org/10.1016/j.neubiorev.2013.03.023.

Garza Villarreal, E.A., Brattico, E., Leino, S., Østergaard, L., Vuust, P., 2011. Distinct neural responses to chord violations: a multiple source analysis study. Brain Res. 1389, 103–114. https://doi.org/10.1016/j.brainres.2011.02.089.

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. Proc. Natl. Acad. Sci. U.S.A. https://doi.org/10.1073/pnas.1809855116.

Gross, J., Baillet, S., Barnes, G.R., Henson, R.N., Hillebrand, A., Jensen, O., et al., 2013. Good practice for conducting and reporting MEG research. Neuroimage. https://doi.org/10.1016/j.neuroimage.2012.10.001.

Habibi, A., Damasio, A., Ilari, B., Veiga, R., Joshi, A.A., Leahy, R.M., et al., 2018. Childhood music training induces change in micro and macroscopic brain structure: results from a longitudinal study. Cerebr. Cortex. https://doi.org/10.1093/cercor/bhx286.

Hampson, M., Driesen, N.R., Skudlarski, P., Gore, J.C., Constable, R.T., 2006. Brain connectivity related to working memory performance. J. Neurosci. https://doi.org/10.1523/JNEUROSCI.3408-06.2006.

Haumann, N.T., Parkkonen, L., Kliuchko, M., Vuust, P., Brattico, E., 2016. Comparing the performance of popular MEG/EEG artifact correction methods in an evoked-response study. Comput. Intell. Neurosci. 1–10. https://doi.org/10.1155/2016/7489108.

Heilman, K.M., Van Den Abell, T., 1980. Right hemisphere dominance for attention: the mechanism underlying hemispheric asymmetries of inattention (neglect). Neurology. https://doi.org/10.1212/wnl.30.3.327.

Hillebrand, A., Barnes, G.R., 2005. Beamformer analysis of MEG data. Int. Rev. Neurobiol. https://doi.org/10.1016/S0074-7742(05)68006-3.

Huang, M.X., Shih, J.J., Lee, R.R., Harrington, D.L., Thoma, R.J., Weisend, M.P., et al., 2004. Commonalities and differences among vectorized beamformers in electromagnetic source imaging. Brain Topogr. https://doi.org/10.1023/B:BRAT.0000019183.92439.51.

Hyde, K.L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A.C., Schlaug, G., 2009. The effects of musical training on structural brain development: a longitudinal study. In: Annals of the New York Academy of Sciences. https://doi.org/10.1111/j.1749-6632.2009.04852.x.

Kastner, S., Ungerleider, L.G., 2000. Mechanisms of visual attention in the human cortex. Annu. Rev. Neurosci. https://doi.org/10.1146/annurev.neuro.23.1.315.

King, A.J., Teki, S., Willmore, B.D.B., 2018. Recent advances in understanding the auditory cortex. F1000Research. https://doi.org/10.12688/f1000research.15580.1.

Kliuchko, M., Brattico, E., Gold, B.P., Tervaniemi, M., Bogert, B., Toiviainen, P., Vuust, P., 2019. Fractionating auditory priors: a neural dissociation between active and passive experience of musical sounds. PLoS One. https://doi.org/10.1371/JOURNAL.PONE.0216499.

Koelsch, S., Gunter, T.C., Schröger, E., Tervaniemi, M., Sammler, D., Friederici, A.D., 2001. Differentiating ERAN and MMN: an ERP study. Neuroreport. https://doi.org/10.1097/00001756-200105250-00019.

Koelsch, S., Schröger, E., Tervaniemi, M., 1999. Superior pre-attentive auditory processing in musicians. Neuroreport. https://doi.org/10.1097/00001756-199904260-00029.

Koelsch, S., Vuust, P., Friston, K., 2019. Predictive processes and the peculiar case of music. Trends Cognit. Sci. https://doi.org/10.1016/j.tics.2018.10.006.

Korzyukov, O.A., Winkler, I., Gumenyuk, V.I., Alho, K., 2003. Processing abstract auditory features in the human auditory cortex. Neuroimage. https://doi.org/10.1016/j.neuroimage.2003.08.014.

Kotz, S.A., Cappa, S.F., Von Cramon, D.Y., Friederici, A.D., 2002. Modulation of the lexical-semantic network by auditory semantic priming: an event-related functional MRI study. Neuroimage. https://doi.org/10.1006/nimg.2002.1316.

Kroese, D.P., Taimre, T., Botev, Z.I., 2011. Handbook of Monte Carlo methods. Handbook of Monte Carlo Methods. https://doi.org/10.1002/9781118014967.

Kropotov, J.D., Näätänen, R., Sevostianov, A.V., Alho, K., Reinikainen, K., Kropotova, O. V., 1995. Mismatch negativity to auditory stimulus change recorded directly from the human temporal cortex. Psychophysiology. https://doi.org/10.1111/j.1469-8986.1995.tb01226.x.

Li, W., Huang, X., Leung, H., 2004. Performance evaluation of digital beamforming strategies for satellite communications. IEEE Trans. Aero. Electron. Syst. https://doi.org/10.1109/TAES.2004.1292139.

Mantini, D., Penna, S. Della, Marzetti, L., De Pasquale, F., Pizzella, V., Corbetta, M., Romani, G.L., 2011. A signal-processing pipeline for magnetoencephalography resting-state networks. Brain Connect. https://doi.org/10.1089/brain.2011.0001.

Marco-Pallarés, J., Grau, C., Ruffini, G., 2005. Combined ICA-LORETA analysis of mismatch negativity. Neuroimage. https://doi.org/10.1016/j.neuroimage.2004.11.028.

Mesulam, M., Marchsel, 1981. A cortical network for directed attention and unilateral neglect. Ann. Neurol. https://doi.org/10.1002/ana.410100402.

Molholm, S., Martinez, A., Ritter, W., Javitt, D.C., Foxe, J.J., 2005. The neural circuitry of pre-attentive auditory change-detection: an fMRI study of pitch and duration mismatch negativity generators. Cerebr. Cortex. https://doi.org/10.1093/cercor/bhh155.

Mu, Z., Chang, Y., Xu, J., Pang, X., Zhang, H., Liu, X., et al., 2016. Pre-attentive dysfunction of musical processing in major depressive disorder: a mismatch negativity study. J. Affect. Disord. 194, 50–56. https://doi.org/10.1016/j.jad.2016.01.028.

Münte, T.F., Altenmuller, E., Jäncke, L., 2002. The musician's brain as a model of neuroplasticity. Nat. Rev. Neurosci. https://doi.org/10.1038/nrn843.

Näätänen, R., Paavilainen, P., Rinne, T., Alho, K., 2007. The mismatch negativity (MMN) in basic research of central auditory processing: a review. Clin. Neurophysiol. https://doi.org/10.1016/j.clinph.2007.04.026.

Näätänen, Risto, 1995. The mismatch negativity: a powerful tool for cognitive neuroscience. Ear Hear. https://doi.org/10.1097/00003446-199502000-00002.

Näätänen, Risto, 2018. Attention and brain function. Attention and Brain Function. https://doi.org/10.4324/9780429487354.

Näätänen, Risto, Kujala, T., Kreegipuu, K., Carlson, S., Escera, C., Baldeweg, T., Ponton, C., 2011. The mismatch negativity: an index of cognitive decline in neuropsychiatric and neurological diseases and in ageing. Brain. https://doi.org/10.1093/brain/awr064.

Näätänen, Risto, Pakarinen, S., Rinne, T., Takegata, R., 2004. The mismatch negativity (MMN): towards the optimal paradigm. Clin. Neurophysiol. https://doi.org/10.1016/j.clinph.2003.04.001.

Nolte, G., 2003. The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors. Phys. Med. Biol. https://doi.org/10.1088/0031-9155/48/22/002.

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. https://doi.org/10.1155/2011/156869.

Pakarinen, S., Takegata, R., Rinne, T., Huotilainen, M., Näätänen, R., 2007. Measurement of extensive auditory discrimination profiles using the mismatch negativity (MMN) of the auditory event-related potential (ERP). Clin. Neurophysiol. https://doi.org/10.1016/j.clinph.2006.09.001.

Pando-Naude, V., Patyczek, A., Bonetti, L., Vuust, P., 2021. An ALE meta-analytic review of top-down and bottom-up processing of music in the brain. Sci. Rep. https://doi.org/10.1038/s41598-021-00139-3.

Penny, W., Friston, K., Ashburner, J., Kiebel, S., Nichols, T., 2007. Statistical parametric mapping: the analysis of functional brain images. Statistical Parametric Mapping: The Analysis of Functional Brain Images. https://doi.org/10.1016/B978-0-12-372560-8.X5000-1.

Posner, M.I., 2016. Orienting of attention: then and now. Q. J. Exp. Psychol. https://doi.org/10.1080/17470218.2014.937446.

Pulvermüller, F., Shtyrov, Y., 2006. Language outside the focus of attention: the mismatch negativity as a tool for studying higher cognitive processes. Prog. Neurobiol. https://doi.org/10.1016/j.pneurobio.2006.04.004.

Putkinen, V., Tervaniemi, M., Saarikivi, K., de Vent, N., Huotilainen, M., 2014. Investigating the effects of musical training on functional brain development with a novel Melodic MMN paradigm. Neurobiol. Learn. Mem. https://doi.org/10.1016/j.nlm.2014.01.007.

Salimpoor, V.N., Zald, D.H., Zatorre, R.J., Dagher, A., McIntosh, A.R., 2015. Predictions and the brain: how musical sounds become rewarding. Trends Cognit. Sci. https://doi.org/10.1016/j.tics.2014.12.001.

Schall, U., Johnston, P., Todd, J., Ward, P.B., Michie, P.T., 2003. Functional neuroanatomy of auditory mismatch processing: an event-related fMRI study of duration-deviant oddballs. Neuroimage. https://doi.org/10.1016/S1053-8119(03)00398-7.

Taulu, S., Simola, J., 2006. Spatiotemporal signal space separation method for rejecting nearby interference in MEG measurements. Phys. Med. Biol. https://doi.org/10.1088/0031-9155/51/7/008.

Tervaniemi, M., Jacobsen, T., Röttger, S., Kujala, T., Widmann, A., Vainio, M., et al., 2006a. Selective tuning of cortical sound-feature processing by language experience. Eur. J. Neurosci. https://doi.org/10.1111/j.1460-9568.2006.04752.x.

Tervaniemi, M., Szameitat, A.J., Kruck, S., Schröger, E., Alter, K., De Baene, W., Friederici, A.D., 2006b. From air oscillations to music and speech: functional magnetic resonance imaging evidence for fine-tuned neural networks in audition. J. Neurosci. 26 (34), 8647–8652.

Tervaniemi, Mari, Huotilainen, M., Brattico, E., 2014. Melodic multi-feature paradigm reveals auditory profiles in music-sound encoding. Front. Hum. Neurosci. 8 https://doi.org/10.3389/fnhum.2014.00496.

Tervaniemi, Mari, Janhunen, L., Kruck, S., Putkinen, V., Huotilainen, M., 2016. Auditory profiles of classical, jazz, and rock musicians: genre-specific sensitivity to musical sound features. Front. Psychol. https://doi.org/10.3389/fpsyg.2015.01900.

Tervaniemi, Mari, Just, V., Koelsch, S., Widmann, A., Schröger, E., 2005. Pitch discrimination accuracy in musicians vs nonmusicians: an event-related potential and behavioral study. Exp. Brain Res. https://doi.org/10.1007/s00221-004-2044-5.

Tyler, L.K., Marslen-Wilson, W.D., Randall, B., Wright, P., Devereux, B.J., Zhuang, J., et al., 2011. Left inferior frontal cortex and syntax: function, structure and behaviour in patients with left hemisphere damage. Brain. https://doi.org/10.1093/brain/awq369.

Vuust, P., Brattico, E., Glerean, E., Seppänen, M., Pakarinen, S., Tervaniemi, M., Näätänen, R., 2011a. New fast mismatch negativity paradigm for determining the neural prerequisites for musical ability. Cortex 47 (9), 1091–1098. https://doi.org/10.1016/j.cortex.2011.04.026.

Vuust, P., Brattico, E., Glerean, E., Seppänen, M., Pakarinen, S., Tervaniemi, M., Näätänen, R., 2011b. New fast mismatch negativity paradigm for determining the neural prerequisites for musical ability. Cortex 47 (9), 1091–1098. https://doi.org/10.1016/j.cortex.2011.04.026.

Vuust, P., Brattico, E., Seppanen, M., Naatanen, R., Tervaniemi, M., 2012. The sound of music: differentiating musicians using a fast, musical multi-feature mismatch negativity paradigm. Neuropsychologia 50 (7), 1432–1443. https://doi.org/10.1016/j.neuropsychologia.2012.02.028.

Waberski, T.D., Kreitschmann-Andermahr, I., Kawohl, W., Darvas, F., Ryang, Y., Gobbele, R., Buchner, H., 2001. Spatio-temporal source imaging reveals subcomponents of the human auditory mismatch negativity in the cingulum and right inferior temporal gyrus. Neurosci. Lett. https://doi.org/10.1016/S0304-3940(01)01988-7.

Wacongne, C., Changeux, J.-P., Dehaene, S., 2012. A neuronal model of predictive coding accounting for the mismatch negativity. J. Neurosci. 32 (11), 3665–3678. https://doi.org/10.1523/JNEUROSCI.5003-11.2012.

Woolrich, M.W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., et al., 2009. Bayesian analysis of neuroimaging data in FSL. Neuroimage. https://doi.org/10.1016/j.neuroimage.2008.10.055.

Zatorre, R.J., Belin, P., Penhune, V.B., 2002. Structure and function of auditory cortex: music and speech. Trends Cognit. Sci. https://doi.org/10.1016/S1364-6613(00)01816-7.