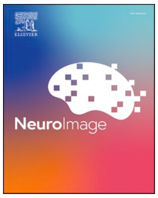
NeuroImage 318 (2025) 121365

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg



Error monitoring and correction in violin performance: An EEG study

Ángel David Blanco [a], Jordi Costa-Faidella [*,b,c,d], Alfonso Pérez Carrillo [a], David Dalmazzo [a], Rafael Ramirez-Melendez [a], Iria SanMiguel [*,b,c,d]

[a] *Music and Machine Learning Laboratory, Department of Information and Communications Technologies, Universitat Pompeu Fabra, Barcelona, Spain*
[b] *Brainlab – Cognitive Neuroscience Research Group, Departament de Psicologia Clínica i Psicobiologia, Universitat de Barcelona, Spain*
[c] *Institut de Neurociències, Universitat de Barcelona, Spain*
[d] *Institut de Recerca Sant Joan de Déu, Esplugues de Llobregat, Spain*

ARTICLE INFO

Keywords:
Violin
Error monitoring
ERN
Music
P300
Self-generation

ABSTRACT

Previous research using electroencephalography (EEG) has investigated neural error-related music production processes in expert pianists reporting a frontal event-related negativity followed by a P300 event-related potential (ERP) after mistuned notes. However, piano playing does not rely on auditory feedback and corrective movements when dealing with accurate note tuning, thus offering an incomplete picture about error monitoring processes in players of many other musical instruments, including the human voice. We here designed a setup allowing us to collect the EEG of expert bow-string players while performing short melodies on the violin, (a fretless instrument) while also manipulating in real-time the produced pitch in a controlled manner. 15 expert bow-string players were asked to 1) listen to a reference melody of four notes; 2) play that melody with a violin (Active Condition); and 3) listen passively to the replay of their performance (Replayed Condition). Randomly, the auditory feedback of one of the four melody notes was manipulated in the Active Condition by lowering or lifting the pitch by half semitone. We found that mistuned notes, independently of whether they were manipulated externally or were produced by natural playing errors, elicited an f-ERN followed by a P300. We also found that the amplitude of the f-ERN was larger for faster corrective movements than for slower ones. In addition, between the f-ERN and the P300, we also identified two negative components—one in parietal areas and another in central-right regions—whose amplitudes were greater for slower corrective movements. We suggest the existence of two error-monitoring systems: a "fast" system mediated by the medial frontal cortex that quickly corrects errors when certainty is high and a "slow" system mediated by the posterior parietal cortex that accumulates evidence under uncertainty before triggering correction.

1. Introduction

Learning to relate actions to their sensory consequences is a foundational process in human behavior, with goals that drive behavior across biological systems (Pezzulo et al., 2018). The nervous system plays a crucial role in controlling actions to generate desired sensory outcomes (Friston, 2010; Wolpert et al., 2011). The study of how individuals learn to perform actions to achieve specific sensory feedback has traditionally focused on simple motor tasks, such as reaching and grasping. However, there remain numerous unanswered questions, particularly when considering more complex sensorimotor skills that involve higher cognitive aspects of goal-directed behavior.

Musical performance offers a particularly rich context to study this relationship between action and sensory consequences. Playing music involves a complex interaction of motor, auditory, and somatosensory systems (Zatorre et al., 2007). For example, imitating melodies engages both lower and higher cognitive processes, including movement planning, execution, and action-outcome monitoring (Lahav et al., 2007; Pfordresher, 2019). These processes facilitate continuous error correction, allowing musicians to adjust their performance in real-time. A key component in musical learning is the coupling of auditory feedback with motor processes, where auditory feedback plays a critical role in establishing the associations between actions and sensory outcomes (Pfordresher and Chow, 2019; Baumann et al., 2007; Nunes-Silva et al., 2020).

Despite extensive research on musical production, much of the focus has been on keyed or fretted instruments, such as the piano and guitar (Maidhof et al., 2009; Katahira et al., 2008; Mathias et al., 2016; Loehr

---

* Corresponding authors at: P. Vall d'Hebron 171, 08035 Barcelona, Spain.
E-mail addresses: jcostafaidella@ub.edu (J. Costa-Faidella), isanmiguel@ub.edu (I. SanMiguel).

https://doi.org/10.1016/j.neuroimage.2025.121365
Received 17 March 2025; Received in revised form 13 June 2025; Accepted 9 July 2025

et al., 2013, Herrojo-Ruiz et al., 2009, Lappe et al., 2018). Such a body of research provides an incomplete view of the action-monitoring processes in other musical instruments. In particular, less is known about instruments like the violin, where pitch control is continuous and requires more precise motor movements. Research on violin performance, especially regarding pitch control and intonation monitoring, has been limited (Chen et al., 2008, 2013) possibly due to the technical challenges of using those instruments to study neural processes under experimentally controlled conditions. This gap is significant because, unlike keyed or fretted instruments, violinists must rely heavily on auditory feedback for intonation accuracy, a process similar to the pitch control in the human voice (Kleber et al., 2017).

In the context of auditory-motor interactions, the dorsal auditory pathway has been shown to play a critical role in transforming auditory information into motor commands, contributing to shared action-effect representations (Warren et al., 2005; Hickok and Poeppel, 2004; Zatorre et al., 2007). The constant monitoring of auditory feedback and the detection of prediction errors are crucial for refining sensorimotor skills in tasks such as musical performance. Research has identified various types of prediction errors, including sensory prediction errors, which relate to discrepancies between expected and observed sensory outcomes, and reward prediction errors, which signal whether goals have been achieved (Krakauer et al., 2019). Sensory prediction errors are thought to drive slow, implicit learning, while reward prediction errors are more involved in strategy learning and explicit corrections (Krakauer et al., 2019). These processes have been well studied in simple motor tasks but are less understood in the context of complex, continuous actions like those involved in playing the violin.

This study seeks to build on previous research by investigating the neural processes underlying goal-directed sensorimotor control in music performance, specifically focusing on pitch control in violin performance. By introducing errors in auditory feedback, we aim to examine the electrophysiological indices associated with action-monitoring processes, error detection, and internal model updating, including sensory and reward prediction errors. Our investigation will add to the growing body of research on audiomotor interactions and contribute to a more comprehensive understanding of how sensory feedback shapes motor performance in complex musical tasks. Here, we designed a novel setup allowing us to collect the electroencephalogram (EEG) of expert bow-string players while performing short melodies on the violin. Participants listened to a reference melody (Reference Condition), reproduced the reference melody with the violin (Active Condition), and listened to the replay of their performance (Replayed Condition). Critically, we manipulated the auditory feedback in real-time by lowering or lifting the pitch of randomly selected notes by half a semitone. Participants were required to correct online any perceived error during their performance, whether they felt it was self-generated or externally manipulated. This allowed us to identify and study the electrophysiological indices of the sensorimotor and cognitive processes involved in the ability to control behavior to achieve a specific sensory goal.

Previous research with the piano has reported prediction error signals in the ERPs following manipulated notes, specifically the elicitation of the f-ERN and the P300 components both in the active and passive condition (Maidhof et al., 2009; Katahira et al., 2008; Mathias et al., 2016; Loehr et al., 2013, Herrojo-Ruiz et al., 2009). However, because movements have to be much more precise with the violin, violinists rely much more on auditory feedback to detect errors and make continuous pitch adjustments during the performance. Thus, we hypothesize that we should find activity related to sensory prediction errors guiding pitch adjustments in the violin. Moreover, we expect to find indices of movement reprogramming together with movement-related cortical potentials (MRCP) (DiRusso et al., 2005; Fattapposta et al., 1996; Hatta et al., 2009; Kita et al., 2001) caused by both bow movement (right arm) and finger placement on the fingerboard (left arm) in the Active condition. Finally, we also expected to find the auditory P1, N1, and P2 components in both conditions and self-generation effects in the N1 amplitude as has been reported in previous studies (Ford et al., 2001; Horvath, 2015; Hughes et al., 2013; Timm et al., 2013). Suppression of the components amplitude would support the idea that sensory feedback of our actions is processed differently due to the constant scrutiny of forward models in an elaborate sensorimotor skill such as playing the violin.

2. Methods

2.1. Participants

Fifteen adult right-handed participants (7 females; mean age 28.06 (SD = 7.86)) with no self-reported history of neurological, psychiatric, or hearing impairment and with normal or corrected-to-normal visual acuity were recruited from the university campus and different music academies and conservatories of Barcelona to participate in the experiment. All participants had at least six years' experience playing a bowed string instrument: 7 violinists with 14.14 (7.81) years experience; 7 cellists with 14.57 (8.58) years experience; 1 participant who played both plus the Chinese Erhu, with 25 years' experience. Participants conceded their written consent before participation and after explaining the procedures. Procedures were approved by the Conservatoires UK Research Ethics committee, following the guidelines of the British Psychological Society and in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). The study was carried out in the Department of Clinical Psychology and Psychobiology of the University of Barcelona (Brainlab).

2.2. Procedures

Participants sat on an armless chair in an electrically shielded, sound-attenuated EEG recording chamber, holding a significantly dampened violin (Primavera Ultra Rubber mute) in vertical position between their legs as if it were a cello (see Fig. 1). The position and the specific mute were adopted to reduce the amplitude of the sound generated by the violin during playing that would directly reach the participant's ears and eliminate the possibility to hear the violin by bone conduction. Since the participants only played on the G string, cotton was introduced between the D, A, and E strings and the fingerboard to damp any resonances. Protective earmuffs 3 M PELTOR Optime III covered the participant's ears to provide an extra layer of sound attenuation, besides that furnished by the earbuds of the noise-reduction earphones. All these measures were implemented with the objective of reducing, as much as possible, the transmission of external and direct acoustic input to the participant's ears during playing, favoring instead the perception of the sounds delivered through the earphones. The development of this setup was guided by a series of pilot tests conducted with two of the study's co-authors, both professional violinists with over 20 years of experience. The configuration was considered adequate when neither of them, through an informal process, was able to reliably determine whether a note—randomly chosen at any position on the fingerboard—had been pitch-manipulated by the system. They also reported no perception of artifacts such as unintended chords, beating, or roughness that might result from the combination of the violin's acoustic sound with the manipulated tone delivered through the earphones. Special care was taken to ensure that the audio delay was short enough to be imperceptible to participants, who did not report noticing any delay during the task. The loudness was already set to the maximum self-chosen comfortable level so that pitch-shifting manipulations would not be heard mixed with the direct, non-manipulated, violin acoustic sound.

A computer screen (28 inches) was placed at 1.20 m from the participant to deliver instructions. A keyboard was placed on the floor right next to the participants' bare feet to enable them to start an experimental block by pressing the spacebar with their toes. The keyboard's location on the floor helped us minimize ample arm movements

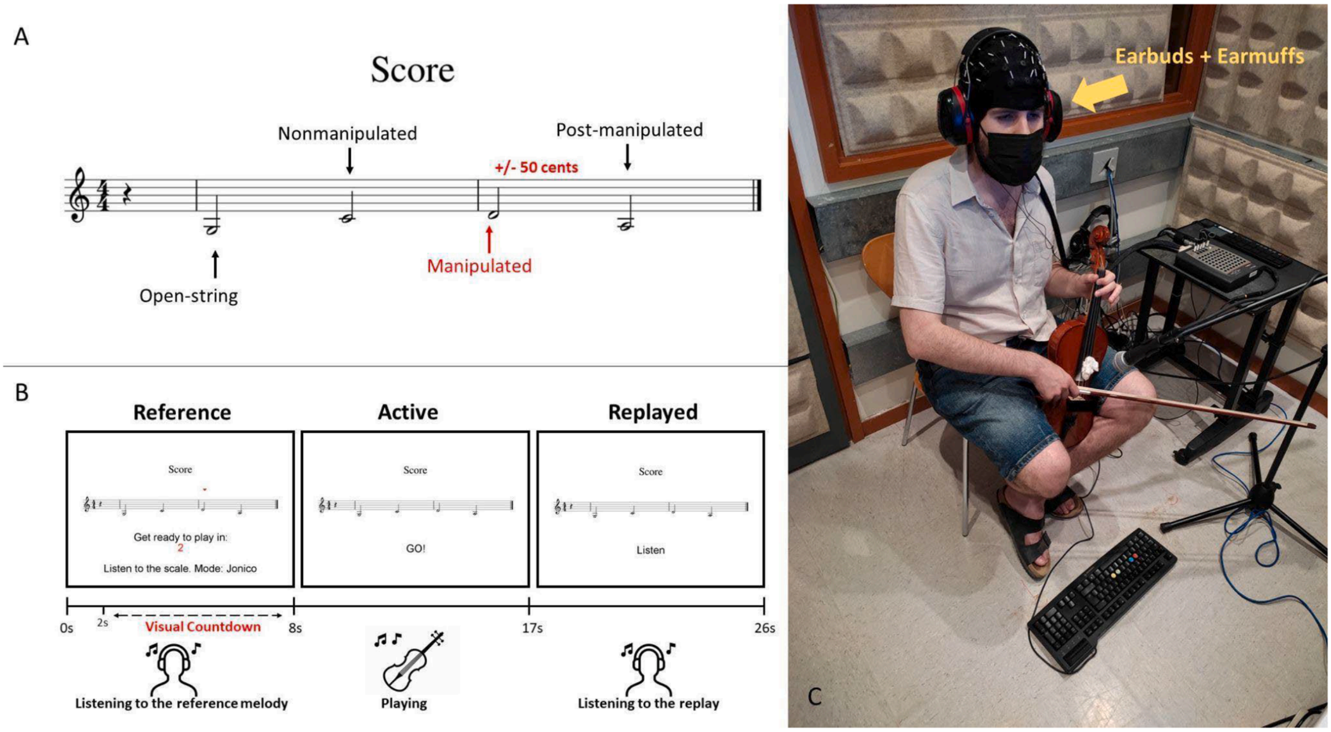


Fig. 1. A. One melody from the experiment. In this specific example, the open-string note (always G3) is followed by a non-manipulated note. The next note, and without the participant's knowledge, is shifted by 50 cents while playing (half a semitone; manipulated note). We refer to the next note as the post-manipulated note. **B.** Conditions. Each trial is composed of three different conditions: in the Reference Condition, participants hear the sounds composing the reference melody; in the Active Condition, participants play the previously heard reference melody on the violin; finally, in the Replayed Condition, participants hear the replay of their performance in the Active Condition. Participants hear all sounds (reference, self-generated, and replayed) through earphones. This section of the figure includes three rectangles, each corresponding to one of the conditions. The content of each rectangle visualizes what the participant would have seen on the screen during each respective condition. **C.** During the experiment, participants held the violin between their legs as if it were a cello. Protective earmuffs covered the participant's ears to provide an extra layer of sound attenuation, besides that furnished by the earbuds of the noise-reduction earphones.

that would be produced if the hand was displaced from the keyboard to the violin and vice versa.

The experiment was composed of 24 blocks of 6 trials each (144 trials in total; each trial lasted 26 s). A block consisted on the following sequence of events (illustrated in Fig. 1): 1) the participant pressed the spacebar on the keyboard with his/her right foot toes to start the block when desired; 2) Immediately after, the score of a short four-note melody appeared on the screen with visual instructions on the top ("Listen to the scale. Mode: Ionian" or "Listen to the scale. Mode: Aeolian") and an audio recording of a violin playing the melody was delivered through the earphones (*Reference Condition*); synchronously with each note, a visual countdown from 3 to 1 (preceded by the instruction 'Get ready in:') appeared on the screen, and a red arrow moved beneath the corresponding note in the score, highlighting each note as it was heard through the earphones.; 3) 0.5 s after the offset of the last note, the instruction "GO!" appeared on screen; from this moment, the countdown and the red arrow disappeared and the participant disposed of 9 s to play with the violin the previously heard reference melody (*Active Condition*) respecting its pitch, tempo and note duration as best as possible. In this condition, one of the notes played by the participant was randomly manipulated in real time by a pitch shift of 50 cents. Participants were required to correct online any perceived error during their performance, whether they felt it was self-generated or externally manipulated.; 4) after those 9 s, regardless of whether the participant ended his/her performance, the instruction "Listen" appeared on the screen and the replay of the recorded performance during the Active Condition was reproduced (*Replayed Condition*) This replay matched exactly what the participant had heard during their performance in the Active Condition, including the real-time pitch manipulation applied to one of the notes; 5) 9 s after the offset of the replay, a new trial started; 6) After 6 trials the screen displayed a message informing that the block had finished and that the participant could start a new block by pressing the spacebar on the keyboard when desired.

All reference melodies in a block belonged either to the Ionian or the Aeolian modes, alternating between blocks (first block modal scale counterbalanced across participants). All possible melodies were covered in the experiment (each melody appeared three times throughout the experiment: 144/(24 possible melodies x 2 modes)=3). The experiment lasted 62 min, without pauses (total experimental duration, including EEG recording preparation and pauses, about three hours). Participants could rest between blocks as desired, and a communication channel with the experimenter via CCTV remained open during the whole experiment to report any problem or discomfort.

Participants performed three training blocks, identical to the experimental blocks, supervised by the experimenters before the EEG recording preparation, to adapt to the demanded task. Participants were told that some notes would be externally manipulated and were explicitly required to correct any perceived tuning error, either self-generated or externally manipulated.

2.3. Stimuli

Individual violin notes were created using a real violin sample sound retrieved from the webpage of the Philharmonia Orchestra.[1] The violin sound was captured at a "piano" dynamic level and played "con sordino" (with a mute) in the G3 note (196.59 Hz) of the 4th string, lasting for 1.5 s. The same sound was used to generate the rest of the notes that

---

[1] www.philarmonia.co.uk

composed our stimuli set using Audacity.[2] First, the DC offset of the sound was removed, and the maximum amplitude was normalized to −1 dB. Then, the pitch of the sound was shifted to generate different notes in an equal-tempered scale. Those notes were specifically chosen to be played in the 4th string of a violin without changing the position of the hand: A3 (220.671 Hz), Bb3 (233.8 Hz), B3 (247.65 Hz), C4 (262.39 Hz), and D4 (294.441 Hz). These sounds were then arranged to create short melodies of 4 different notes (2 s onset to onset; thus, from note offset to subsequent note onset, 500 ms silent gap) using custom-made Matlab (R2020a; Mathworks) scripts. All melodies started in G3 (from now on referred to as open-string note). The following three notes (from now on referred to as target notes) were arranged according to either the Ionian mode (A3, B3, C4, D4) or the Aeolian mode (A3, Bb3, C4, D4). Target notes were chosen randomly without repetition, resulting in 24 different melodies per modal scale. From now on, we will refer to these short melodies as reference melodies and the sounds composing them as reference sounds (see Procedures below).

The remaining auditory stimuli consisted of the self-generated sounds created by each participant when playing the violin to replicate the reference melodies. These sounds were recorded at a sampling rate of 44.1 kHz by a Shure dynamic microphone (model: PPPER) located pointing at the front of the violin's body, using Max/MSP 7 software (https://cycling74.com/). We refer to these sounds as self-generated sounds when delivered to the participant in real-time while playing and as replayed sounds when delivered to the participant immediately after replicating the reference melody as a replay of the preceding performance. All sounds (reference, self-generated, and replayed) were delivered binaurally through passive noise-reduction in-ear earphones (Sony MDR-EX110LP) at the maximum comfortable level chosen by the user when listening to the reference melodies (approx. 60–75 dB range for reference and 50–65 dB range for self-generated and replayed).

During the Active Condition, one note out of the three target notes in a melody, chosen at random in each melody, was shifted by 50 cents (half semitone) above or below its pitch using a real-time pitch-shifting algorithm (see *Online Tone Manipulation System in Max MSP* in Supplementary Materials). We will refer to these target notes as *manipulated notes*; to the target note immediately following a manipulated note (when the manipulation occurred before the last target note in the melody), as the *post-manipulated note*; and to the rest of the target notes, as *non-manipulated notes*. It was expected that both manipulated and post-manipulated notes would exhibit greater mistuning than non-manipulated notes, primarily due to the adjustments participants had to make in response to the unexpected pitch shift and their predictions about the subsequent note. Note, however, that non-manipulated notes could still exhibit varying degrees of mistuning due to self-generated errors by the participants. This is primarily because, despite their years of experience, violinists were asked to play in a non-standard position, which reduced their reliance on well-established motor patterns. It is also worth noting that in some cases, a performance error made by a participant during a manipulated note could be accidentally corrected by the system in the right direction, effectively resulting in a tuned note by coincidence. Table 2 in the Supplementary Materials provides a breakdown of the number of notes of each type detected. The Replayed Condition consisted of a playback of the sounds the participant had heard during the Active Condition, preserving both the pitch manipulations and any self-generated errors.

Visual stimuli consisted of musical scores displayed at the center of the screen against a white background covering approximately 22×2 degrees of the visual field, with the word 'Score' positioned above them in Arial font, occupying about 3 × 1 degrees of the visual field. Scores were presented with a G-clef, set in the key of C-Major (which has no flats or sharps) and featured a 4/4 time signature. The notes were distributed between two bars with two half-notes in each of them. Participants received visual instructions displayed at the bottom of the screen consisting of instruction messages: "Listen to the scale. Mode: Ionian", "Listen to the scale. Mode: Aeolian", "Go!" and "Listen" (Arial; ~3 × 1° visual field). Finally, between the score and the visual instructions, a visual countdown was displayed at certain moments of the experiment (*see Procedures*) with the text "Get ready in:". The countdown went from 3 to 1 and was displayed in red with a red arrow that moved over each of the last three notes of the score. The score was present in all three conditions—Reference Condition, Active Condition, and Replayed Condition—to make the process easier for participants, without the need to memorize the melody, with the visual stimuli remaining consistent across conditions.

2.4. *Audio analysis*

The audio of each experimental block was recorded in mono to a stereo channel (audio channel 1) and stored as a stereo .wav file. We used *Tony* Software (Mauch et al., 2015) to export each note's onset, pitch, and duration from the recorded audio files. We also exported the onset, offset, pitch, and time of the sounds produced by corrective movements. Tony is based on the *pYIN* method for automatic pitch estimation and note tracking (Mauch and Dixon, 2014) and custom methods for interactive reestimation. It outputs discrete notes on a continuous pitch scale based on Viterbi-decoding of an independent Hidden Markov Model. After using Tony, we performed a visual inspection of the onsets of all events to correct the operation of the pitch detection algorithm when necessary.

In a posterior analysis in Matlab (R2020a; Mathworks), we labeled each detected onset depending on its condition (Reference, Active, and Replayed), type of note (open-string, non-manipulated, manipulated, and post-manipulated), correction (with, without), type of onset (start of the note, onset of correction, offset of correction), and absolute error in cents ([0–15], [15–30], [30–50], [50–70], [>70]).

2.5. *EEG recording and preprocessing*

EEG data were acquired from 67 Ag/AgCl electrodes and digitized at a sampling rate of 500 Hz by a Neuroscan 4.4 software and Neuroscan SynAmps RT amplifier (NeuroScan, Compumedics, Charlotte, NC, USA). Sixty-two electrodes were mounted in a nylon cap (QuickCap 64; Compumedics, Charlotte, NC, USA) following the 10–10 international system. Additionally, two electrodes were positioned over the left and right mastoids, and three electrodes were used to record the electrooculogram (EOG; 1 under the left eye; 2 at the outer canthi of the eyes). The reference electrode was located at the tip of the nose and the ground electrode at AFz. All impedances were kept below 5 kΩ during the whole recording session.

EEG recordings and audio files were synchronized by simultaneously sending, at the beginning of each experimental condition, triggers to the EEG signal and audio clicks to audio channel 2 (see *6.11 Supplementary Materials*). The sound onsets in the EEG recording were located through interpolation.

Data preprocessing was performed offline using EEGlab v2021.0 software (Delorme and Makeig, 2004) running on Matlab R2017a. EEG raw data were band-pass filtered from 1 to 50 Hz with a windowed sinc FIR filter (Hamming window) and re-referenced to the average of all channels. Then, EEG data were visually inspected, and those periods contaminated from noise and non-stereotyped muscle artifacts were rejected. Independent component analysis (ICA) decomposition was applied to remove from the signal ocular, heart rate, and muscular components identified based on their scalp topography and time course (Jung et al., 2000). Finally, EOG channels were removed.

2.5.1. *Event-related potentials (ERPs) processing*

Preprocessed EEG data were low-pass filtered at 25 Hz with a

---

[2] https://www.audacityteam.org/

windowed sinc FIR filter (Hamming window) and epoched from −600 to 800 ms time-locked to each auditory stimulus onset from all conditions (Reference, Active, and Replayed), types of notes (open-string, non-manipulated, manipulated and post-manipulated) and absolute error in cents ([0–15], [15–30], [30–50], [50–70], [>70]). No baseline correction was applied to avoid introducing motor preparation signals present in the baseline period into the poststimulus waveforms (Urbach and Kutas, 2006). Epochs from each condition containing improbable data 3 SD above or below the mean probability distribution of values across all epochs were excluded (EEGlab's function *pop_jointprob.m*).

2.5.2. *Event-related spectral perturbations (ERSPs) processing*

Preprocessed EEG data were epoched from −1500 ms to 2000 ms time-locked to each auditory stimulus onset from all conditions (Reference, Active, and Replayed), types of notes (open-string, non-manipulated, manipulated, and post-manipulated), and absolute error in cents ([0–15], [15–30], [30–50], [50–70], [>70]) without baseline correction. The ERSP was computed using Morlet wavelets (Tallon-Baudry et al., 1996) with a linearly increasing number of cycles with increasing frequency. The wavelet used in each time window began with a 3-cycle wavelet (Hanning-tapered window) and ended with a 25-cycle wavelet. The ERSP was also computed with an output of 30 frequency bins from 3 Hz to 50 Hz and 200 output time samples (EEGlab's function *newtimef.m*). Epochs from each condition containing improbable data 3 SD above or below the mean probability distribution of values across all epochs were excluded (EEGlab's function *pop_jointprob.m*). For illustration purposes, ERSPs were baseline corrected in some figures, subtracting the mean ERSP of the whole epoch for each frequency bin separately.

3. Statistical analysis

3.1. *Behavioral analysis*

In the behavioral analysis, we aimed to study the different aspects that could have influenced participants' performance. That is, if tuning error magnitude or the onset of corrective movements depended on whether the notes played were manipulated, non-manipulated or post-manipulated.

The open-string note did not require corrective movements (no finger placed on the fingerboard) and was expected to stay tuned throughout the experiment. However, because strings tend to lose tension with time, we asked participants to tune the string if they noticed that they were getting out of tune. To ascertain that the open-string note remained in tune during the experiment, we measured its pitch through the 12 experimental blocks. The error of the note was on average 6.53 cents (SD: 2.9), which is below the normal pitch discrimination threshold for musicians (Bianchi et al., 2016), and no significant changes were found across blocks (repeated measures ANOVA, $F(6.16) = 0.470$, $p = n.s.$).

To study whether target note mean tuning error differed between non-manipulated, manipulated and post-manipulated notes, we performed another repeated measures ANOVA. The error was measured in the notes heard by the participants. Thus, it included the participants' error plus the external manipulation.

We also evaluated the behavior of participants concerning their corrective movements. We expected lower errors to be more difficult to perceive than larger ones and, therefore, would have been less likely to suffer corrections. Thus, we performed a 6 × 3 repeated measures ANOVA with absolute error in cents ([0–15], [15–30], [30–50], [50–70], [>70]) and type of note (non- manipulated, manipulated and post-manipulated) as within-factors and the percentage of notes which were followed by a corrective movement as the dependent variable. We expected corrective movements to be executed in the direction that minimizes the produced error. We also expected that the magnitude of the corrective movement related to the magnitude of the error. To test this, we performed two more 6 × 3 repeated measures ANOVAs with absolute error in cents and type of note as within-factors: One with the magnitude of the correction (error after correction minus the error before correction) as the dependent variable and another with the absolute error after corrective movements.

Finally, to study the possible effects of the type of note or the magnitude of the error in the timing (onset and offset) of the corrective movements, we performed two more 6 × 3 repeated measures ANOVAs with absolute error in cents and type of note as within-subject factors and the onset and the offset (in seconds) of the corrective movements as dependent variables. Pairwise comparison tests were performed between the conditions using the Bonferroni correction.

We decided to set a minimum percentage of notes as a threshold for each type of note and level of error to consider the inclusion of the participant's results in the analysis. Eight participants did not produce >5 % of their total notes with an error higher than 70 cents, and their results were removed from that specific range. Out of the eight participants, three did not account for >5 % of their total notes within the 50 and 70 range and were subsequently excluded from the study.

3.2. *Electrophysiological analysis*

The general strategy for the analysis of the electrophysiological data was a as follows: First, we identified the main auditory, motor, and error-related ERP and ERSP components of interest, and defined the ranges used to measure each of these components in the time, frequency and spatial domains selecting a window around the peak of the component. In the spatial domain we selected the electrode in which each component was most prominent, in the temporal and spectral domains we defined a window around the peak, adjusted to the extent of the component to encompass the whole range of highest activity. Once each component of interest was clearly defined, we assessed how they were modulated as a function of performance and tuning. Specifically, we first assessed how the auditory, motor, and error-related indices were affected by production, tuning and their interactions. Then, we assessed the effects of the magnitude of the pitch error on the identified error-related components; and finally, we assessed whether both the amplitude and latency of error-related components were related to the onset latency of corrective movements. Each of these analyses are detailed in turn in the following sections.

We initially planned to perform these analyses separately for manipulated, postmanipulated, and non-manipulated notes. However, a cluster-based analysis showed no differences between manipulated, postmanipulated and non-manipulated mistuned notes neither in Active nor Replayed conditions (see Fig. 1 in Supplementary Materials) suggesting that mistuned notes were perceived equally independently on their origin. Moreover, the behavioral results showed no significant differences in the timing of the onset and offset of the corrective movements between the three different types of mistuned notes. Therefore, the influence of motor activity should also be comparable between them. Thus, in the electrophysiologycal analysis we pooled the responses to all mistuned notes independently of whether the pitch was externally altered or not and considered only the degree of mistuning as a modulating factor.

3.2.1. *Identification of auditory and motor components*

To have an overview of the electrophysiological auditory responses elicited by the individual violin notes embedded in the melody, we first inspected the responses to the notes of the reference melody played by the synthesized violin, which did not contain any tuning errors (i.e., tuning error < 15 cents; Blanco et al., 2021), averaged across participants. We then inspected the auditory responses to the passive replay of the self-performed notes (Replayed condition), selecting only those notes that were tuned. On these responses to passively heard notes, we identified by visual inspection the main typical auditory ERP (P1, N1, and P2; Bendixen et al., 2012) and ERSP (evoked auditory theta response; Schürmann and Başar, 1994 components of interest.

In the Active condition, we expected to find a superposition of auditory and motor responses caused by both bow movement (right arm) and finger placement on the fingerboard (left arm). To identify those motor components, we compared responses to tuned notes of the Active condition with those of the Replayed condition. We performed a cluster-based analysis (see *Section 3.2.6 Mass-univariate Statistics*) on the ERPs and the ERSPs of the beta (15–25 Hz) & theta (4–7 Hz) ranges for that purpose. In this contrast, we identified the main typical motor-related electrophysiological responses of interest, i.e., the centroparietal motor evoked potentials (MEPs) contralateral to the finger and arm movement (Bestmann and Krakauer, 2015) and the beta-band pre-movement ERD and post-movement beta rebound (Cardellicchio et al., 2020).

3.2.2. *Identification of error-related responses*

We considered a note was mistuned when its error was larger than the threshold of 30 cents. Notes with an error between 15 and 30 cents were removed from the electrophysiological analysis to ensure all notes included in each category would be clearly perceived as either tuned or mistuned. The exclusion range of 15–30 cents was set based on the behavioral results, which showed an average residual error of 22.66 cents (SD: 3.89; range: [13.13–32.2]) after the corrective movements, and no differences in the number of corrective movements between the 0–15 and 15–30 cents tuning error ranges (see section 6.7.1 Behavioral Results).

To assess the effects of tuning, we compared responses between tuned and mistuned notes in both the Active and Replayed conditions by means of mass-univariate statistics on both the ERPs and the ERSPs of the theta range. These contrasts were aimed at identifying well-known error-related indices possibly elicited by mistuned notes, i.e., the ERN/f-ERN (Gehring et al., 2012) and P3a and P3b (Polich, 2007) components in the ERPs; and midfrontal theta activity (Cavanagh et al., 2012) in the ERSP. Nevertheless, the conditions were contrasted in an exploratory manner throughout the whole epoch and across all electrodes to detect any other possible electrophysiological indices of pitch error detection that could be of interest for the present study.

3.2.3. *Interactions between production and tuning*

Once the main components of interest were identified, next we assessed the effects of production and their interactions with tuning on these electrophysiological indices. To this aim, we computed a 2 × 2 repeated measures ANOVA with Tuning (Tuned, Mistuned) and Condition (Active, Replayed), as within-subject factors on the amplitude of each identified component (auditory, motor or error-related).

Production effects are typically measured comparing active vs. replayed conditions after removing motor activity from the active condition by subtracting responses elicited in a motor-only control (without auditory feedback; Horváth, 2015). In our setup it was not possible to include such a control, and we expected to find a superposition of auditory and motor responses caused by both bow movement (right arm) and finger placement on the fingerboard (left arm) in the Active Condition. Nevertheless, we aimed to assess whether the well-known suppression of auditory ERPs to self-produced sounds (Schröger et al., 2015) could be replicated also in this setup. To assess production effects on the auditory ERP components while limiting the influence of possible motor confounds, we re-referenced the ERPs to the average of the mastoid electrodes and extracted from each participant's average a P1 to N1 and N1 to P2 peak-to-peak amplitude measurement from Cz (see e.g., Timm et al., 2014; SanMiguel et al., 2013; Horváth, 2015). Peak measurements corresponded to the mean value in a 20 ms window centered around the peak sample.

3.2.4. *Effects of the magnitude of the pitch error*

Subsequently, we assessed if the magnitude of the tuning error modulated the amplitudes of the motor and error-related components of interest. To this aim, we computed a 4 × 2 repeated measures ANOVA with pitch error Magnitude (15–30, 30–50, 50–70, >70) and Condition (Active, Replayed) as within-subject factors on the magnitude of the identified error-related components, and the post-movement beta rebound activity, all measured in the mistuned minus tuned difference waves. Pairwise comparisons tests were performed between the conditions using the Bonferroni correction.

3.2.5. *Effects of the latency of corrective movements*

Finally, we assessed whether the onset latency of corrective movements modulated the latency and amplitude of the error-related electrophysiological components of interest. We computed two 3 × 2 repeated measures analyses with Speed of correction (Fast (< 250 ms), Medium (250 to 350 ms), Slow (> 350 ms)), and Condition (Active, Replayed) as within-subject factors, one on the amplitude and one on the latency of each of the identified error-related components, both measured in the mistuned minus tuned difference waves.

The temporal division between the three different types of correction speeds was adjusted to have a similar proportion of epochs in each one. We also made sure that the same proportion of errors of different magnitude was found in each category. Pairwise comparisons tests were performed between the conditions using the Bonferroni correction.

The peak latencies of the error-related components were measured for each participant in the mistuned minus tuned difference waves as the latency of the maximum or minimum value in the following windows and electrodes: The latency of the f-ERN was the minimum peak between 0 and 400 ms in FCZ, the N280 the minimum value between 200 and 500 ms at C4, the N340 the minimum value between 200 and 500 ms at Pz, the P3a the maximum value between 150 and 450 at FCZ, the P3b the maximum value between 450 and 600 ms at CZ, and finally the midfrontal theta was the maximum value between 0 and 600 at FCZ.

3.2.6. *Mass-univariate statistics*

Statistical analyses for the cluster-based comparisons used to identify the components of interest were performed using a mass-univariate nonparametric randomization procedure (Maris and Oostenveld, 2007). We performed a two-dimensional (time, electrode) analysis comparing the responses elicited by the tuned notes in the Active condition vs. the Replayed condition; and between mistuned and tuned notes from each condition on the ERP amplitudes (from −600 to 800 ms) and the power estimates (from −1500 to 2500 ms) in the theta (4–7 Hz) and beta (15–25 Hz) frequency bands. We defined the neighboring electrodes using a Delaunay triangulation over a 2D projection of the electrode montage. We also established a minimum of 2 nearby electrodes per cluster. We performed a two-tailed dependent *t*-test assessed with the nonparametric Montecarlo Method for each comparison. The p-value was determined by calculating the proportion of 2D samples from 10,000 random partitions of the data. Those 2D points exceeding a significance level set to 0.05 were grouped to create the clusters. The sum of the t-statistics within every cluster was used to calculate the cluster-level statistic. The Monte Carlo method was used to assess the significance probability of the clusters. Those values of $p < .025$, corrected for two-tailed tests, were considered significant.

4. Results

4.1. *Behavioral*

4.1.1. *Mean error by type of note*

Before the onset of corrective movements, participants showed an absolute average tuning error of 33.31 cents (SD: 16.21; range: [20.59, 70.45]) when playing non-manipulated notes; of 46.88 cents (SD: 7.29; range: [35.31, 62.01]) when playing manipulated notes; and of 35.31 cents (SD: 13.32; range: [17.5, 66.8]) when playing post-manipulated notes. In Supp. Table 1, we offer the mean, range, and standard deviation of some of the most relevant measures of the analysis. These error differences were significant, $F(1.41) = 23.04$, $p < .0001$, $\eta^2 = 0.62$. As expected, manipulated notes showed a larger error level than non-manipulated notes ($p = .001$) and post-manipulated notes ($p < .0001$) due to the introduction of the external pitch-shifter. No significant differences were found between non-manipulated and post-manipulated notes.

4.1.2. *Corrective movements*

Participants were instructed to online correct any detected tuning error in their performance. Participants tended to perform more corrective movements on those notes with larger errors than on those with lower errors ($F(1.41)=55.47$; $p < .0001$; $\eta^2 = 0.90$), independently of the type of note (Fig. 2, A). Specifically, fewer corrective movements were performed for pitch errors in the 0–15 error range compared to the rest of error ranges with the exception of the 15–30 error range ([0–15] < [30–50], $p = .01$; [0–15] < [50–70], $p = .003$; [0–15] < [>70], $p < .0001$). The 30–50 range was significantly different both with the 15–30 and the 50–70 range ([30–50] > [15–30], $p = .015$; [30–50] < [50–70], $p = .018$). No significant differences were found between the 30–50 and the >70 range. The percentage of corrective movements did not differ between manipulated, post-manipulated and non-manipulated notes. In summary, notes with a perceived error below 30 cents were often ignored and corrected less frequently whereas notes with errors exceeding 30 cents were equally likely to be corrected, regardless of the type of note.

The magnitude of the corrective movements also did not differ between manipulated, post-manipulated and non-manipulated notes, but it depended on the magnitude of the pitch error (see Fig. 2, B). Larger

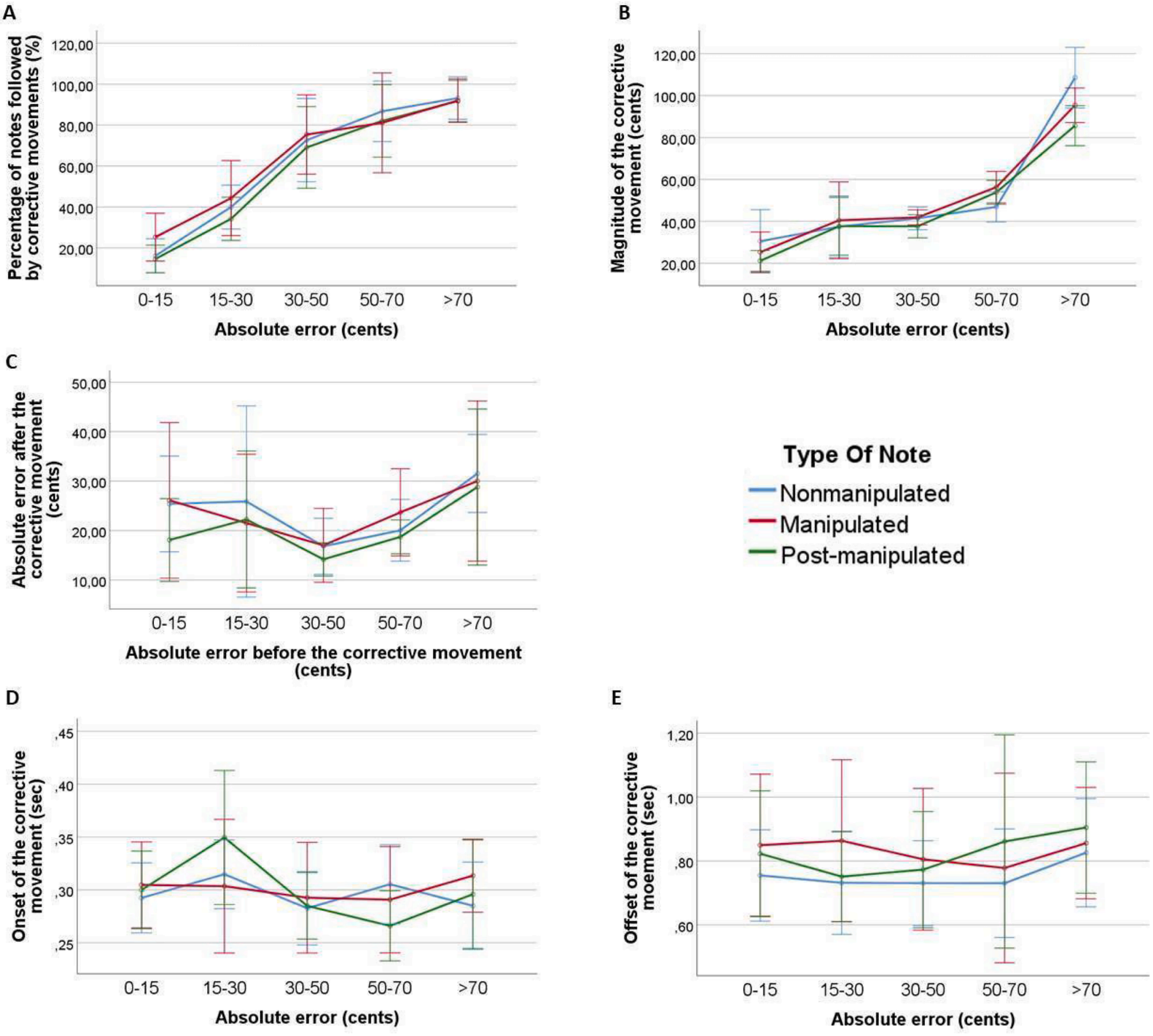


Fig. 2. A. Percentage of notes followed by corrective movements for each error range and type of note. Larger errors were followed by more corrective movements than smaller errors. No significant differences were found between the three types of notes. **B.** Magnitude of the corrective movements for each error range and type of note. Larger errors tended to be followed by larger corrective movements. No significant differences were found between the three types of notes. **C.** Absolute error after the corrective movements for each error range and type of note. No significant differences were found between the three types of notes nor in the final error after the corrective movements for each error range. **D.** Onset of corrective movements for each error range and type of note. No significant differences were found between the three types of notes or the time onset of the correction for each error range. **E.** Offset of the corrective movement for each error range and type of note. No significant differences were found between the three types of notes nor the time offset of the correction for each range of error.

pitch errors were followed by larger corrective movements ($F(1.96)= 81.972$; $p < .0001$; $\eta^2 = 0.93$). Specifically, smaller corrective movements were performed for pitch errors in the 0–15 error range, compared to errors in the 50–70 and the >70 error range ([0–15] < [50–70], $p = .003$; [0–15] < [>70], $p < .0001$). No significant differences were found between the 0–15, the 15–30, and the 30–50 error ranges. Corrective movements for notes in the >70 error range were significantly larger than in the rest of error magnitude ranges ([>70] > [50–70]; $p < .0001$; [>70] > [30–50]; $p < .0001$; [>70] > [15–30]; $p = .001$; [>70] > [0–15]; $p < .0001$].

On average, the absolute error after the correction independently of the type of note was 22.66 cents (SD: 3.89; range: [13.13–32.2]). It did not differ between non-manipulated, manipulated, and post-manipulated notes ($F(1.54) = 1.38$; $p =$ *n.s.*; $\eta^2 = 0.188$), and did not depend on the absolute error before the corrective movement ($F(1.9) = 2.38$; $p =$ *n.s.*; $\eta^2 = 0.284$) (Fig. 2, C).

On average, participants initiated their corrective movements at 299 ms (SD: 18; range: [256–342]) and ended them at 802 ms (SD: 95; range: [569–1035]) (see Fig. 2, D, and E). Again, the onset and offset of corrective movements depended neither on the type of note (onsets: $F(1.53) = 0.18$; $p =$ *n.s.*; $\eta^2 = 0.029$; offsets: $F(1.72)=2.77$; $p =$ *n.s.*; $\eta^2 = 0.316$) nor on the magnitude of the pitch error (onsets: $F(1.86) = 2.63$; $p =$ *n.s.*; $\eta^2 = 0.305$; offsets: $F(1.55) = 1.45$; $p =$ *n.s.*; $\eta^2 = 0.196$).

Summarizing, we did not find differences in the behavior of the participants after hearing manipulated, non-manipulated or post-manipulated notes. We also found that the magnitude of the error did not exert any kind of influence in the start and end of the corrective movements: both lower and higher errors were therefore detected equally fast. Finally, participants tended to initiate corrective movements when the detected error was above 30 cents, and the average error after the corrective movement was 22.66 cents.

4.2. Electrophysiological

4.2.1. Identification of auditory and motor components

In order to overview the electrophysiological auditory responses elicited by the individual violin notes embedded in the melody, we first inspected the responses to the notes of the reference melody, which did not contain any tuning errors (see Fig. 3). The reference notes elicited the prototypical frontocentrally distributed auditory ERPs: P1 peaking at ~75 ms, N1 peaking at ~120 ms, and P2 peaking at ~200 ms. The time-frequency analysis also showed a typical evoked auditory response on frontocentral electrodes with power in the Theta 4–7 Hz frequency range locked to note onset and lasting until about 500 ms. Having the electrophysiological responses to notes played by the synthesized violin as a reference, we then inspected the auditory responses to the passive replay of the self-performed notes, selecting only those performed correctly (i.e., tuning error < 15 cents). The replay of the self-performed notes that were tuned elicited the same pattern of electrophysiological responses as the synthesized notes from the reference melody (see Fig. 3), albeit with reduced amplitude, consistent with the softer volume of the Replayed notes compared to the reference melody (−15 dB approx). The reduction in amplitude might also be partially explained by habituation effects, as experienced musicians are known to show sensitivity to the repetition of melodies and tones. The implications of these potential habituation effects within our design will be further addressed in the Discussion chapter.

In order to identify the motor components, we next inspected the responses to tuned notes in the Active condition. During the performance of the melody, the electrophysiological measurements show a superposition of auditory and motor responses (see Fig. 4). The auditory P1, N1, and P2 components can be identified, showing a similar time-course and topography as during listening. The time period preceding the note's onset, however, differed markedly between playing and listening to the replay, particularly on frontal electrodes. This difference is due to the superposition of the motor evoked potentials (MEPs) related to the placement of the left finger on the board and the moving of the bow with the right arm in the playing condition, which are absent during listening. A direct contrast between the Active and Replayed conditions allowed us to isolate these MEPs. The cluster-based permutation tests on the ERPs showed significant differences between the Active and Replayed conditions over the left and right centroparietal and frontal electrodes in the time period between −130 and 28 ms (see Fig. 4, panel A). Specifically, we found two overlapping clusters: an enhanced frontal negativity for actively played notes between −130 and 28 ms ($T = -3.226$; $p = .0004$) and an enhanced positivity over left and right centroparietal electrodes between −120 and 24 ms ($T = 2.909$; $p = .0006$). The scalp maps of the Active - Replayed difference (see Fig. 4, panel B) in this time range strongly suggest a motor origin, with contributions from both the left (right arm) and right (left finger) motor cortices.

As in the Replayed condition, the ERSP response to the tuned notes in the Active condition also showed a typical evoked auditory response on frontocentral electrodes with power in the Theta 4–7 Hz frequency range locked to note onset (see Fig. 4, panel C). Again, to identify motor components in the ERSP, we compared the Active and Replayed conditions. The cluster-based analysis did not yield significant differences between conditions in the Theta frequency range. However, in the Beta range (ca. 15–25 Hz), a further motor-related induced response was observed ($T = 26.918$, $p = .00009$). The modulation of beta power was significant over the whole length of the epoch and the whole scalp, but was maximal over right centroparietal electrodes (Fig. 4, panel D), and consisted of a pre-movement event-related desynchronization (ERD) followed by a post-movement beta rebound (event-related synchronization, ERS) which peaked at 706 ms.

In summary, we identified the typical auditory and motor components in responses to tuned notes and defined the time and frequency windows and electrodes employed to measure them in subsequent analyses. In particular, we identified the auditory components in both the Active and Replayed conditions: the P1 component (henceforth computed as the average value between 25 and 60 ms in CZ), the auditory N1 component (95–124 ms, CZ), and the auditory P2 component (156–220 ms, CZ) in the ERPs; and the evoked auditory response in the theta range (100–500 ms, 4–7 Hz, FCz electrode) in the ERSP. In the Active condition, we observed the MEP 100 ms before the onset of the note in the ERPs; and we identified the pre-movement event-related desynchronization of beta power, and a post-movement beta rebound (606–806 ms, 15–25 Hz, C4 electrode) in the ERSP.

4.2.2. Identification of error-related responses

To identify the responses related to pitch error detection, we compared ERP and ERSP responses between tuned and mistuned notes using cluster-based permutation tests, separately during playing and listening (Fig. 5).

During playing, the ERP to mistuned notes differed significantly from the ERP to tuned notes between 180–600 ms (Fig. 5, panels Active Condition, A). In this time range, we found two clusters in which mistuned notes elicited more negative responses than tuned notes (negative clusters; NC1: 180–396 ms, NC2: 424–562 ms) and one cluster in which mistuned responses were more positive (positive cluster; PC1: 380–600 ms). Within these clusters, we identified several different error-related components with distinct time-courses and scalp distributions. Specifically, within NC1 ($T = -4.721$; $p = .0013$), mistuned notes elicited a frontal negativity peaking at ca. 240 ms (which we identified as f-ERN) followed by a right central and right centro-parietal negativity that started around 200 ms, maintained its amplitude until 380 ms, and was maximum at 284 ms at the C4 electrode (which we henceforth termed N-280). The N-280 was followed by a parieto-occipital negativity peaking at 340 ms, maximum at P03 (henceforth termed N-340). These negativities were followed by a broad positivity at frontocentral electrodes between ca. 380–600 ms (encompassed by PC1; $T = 4.484$; $p = .0014$), which we identified as a P3 response. However, the P3 could be

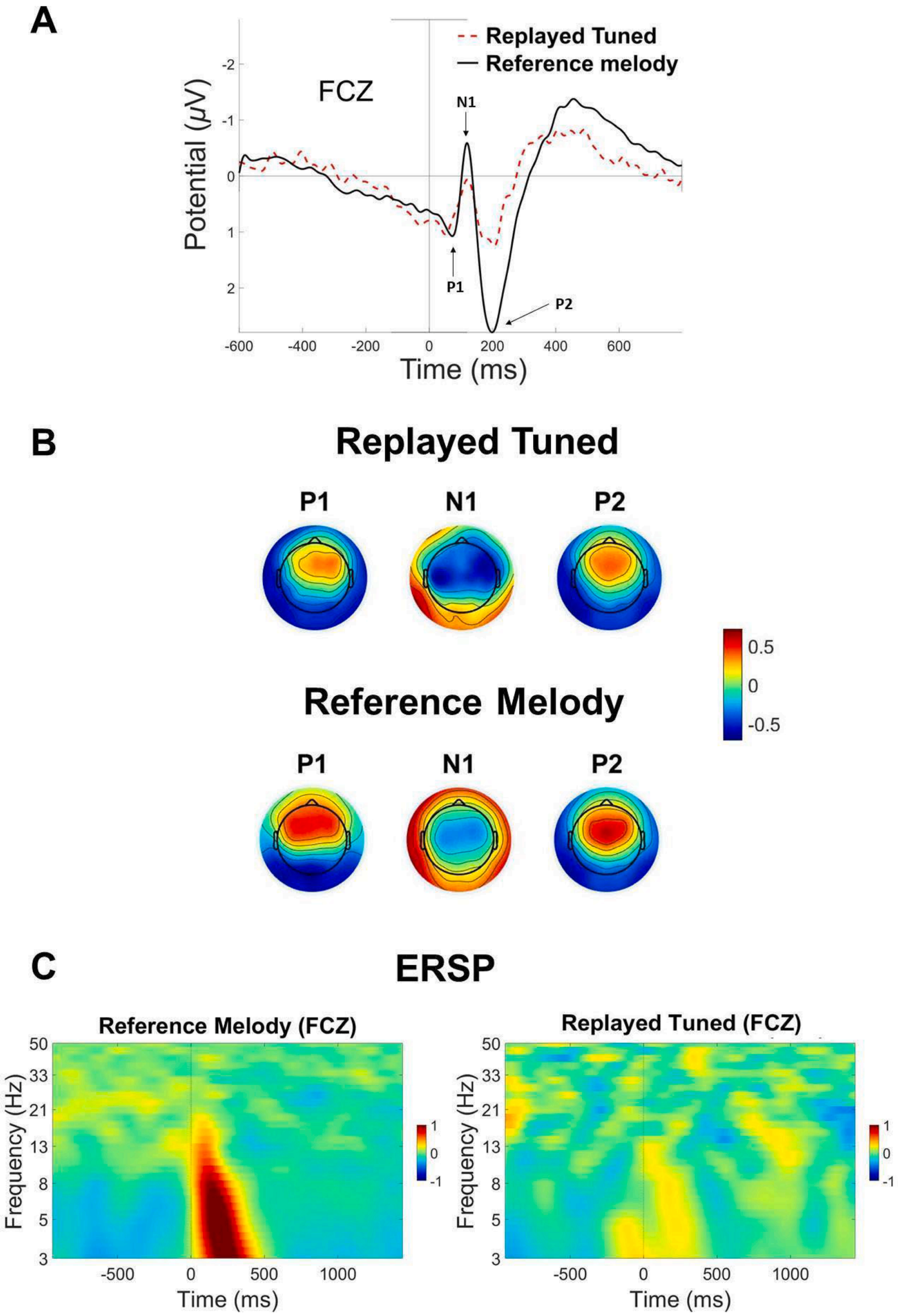


Fig. 3. Auditory responses to the notes of the reference melody and tuned notes in the passive replay. **A** ERPs at FCz elicited to reference notes (solid black line) and replayed tuned notes (dashed red line); **B** Scalp maps for the P1(25 to 60 ms), N1(95 to 124 ms) and P2(156 to 220 ms) ERP components; **C** ERSP plots at FCz showing the evoked auditory response in the theta range.

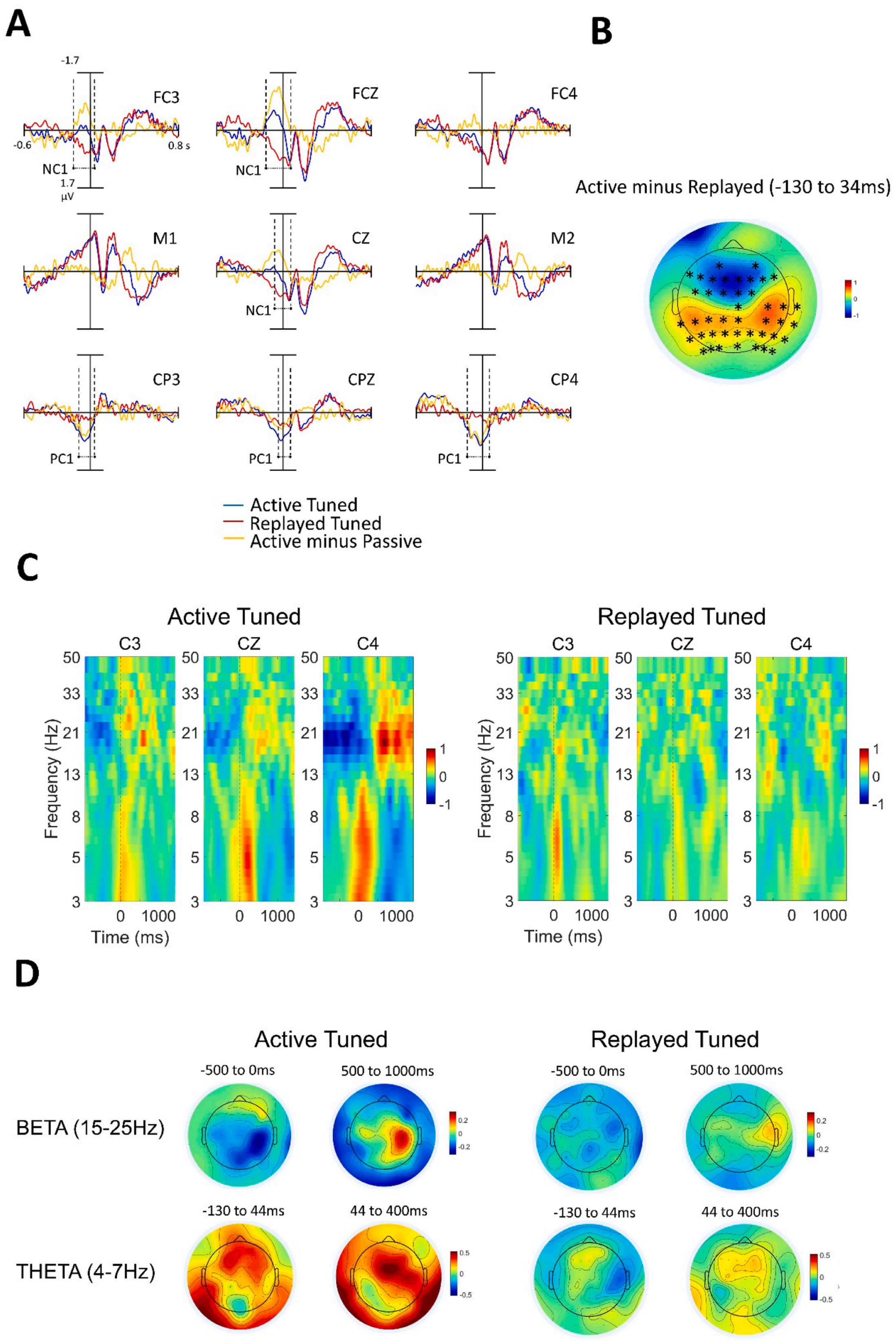
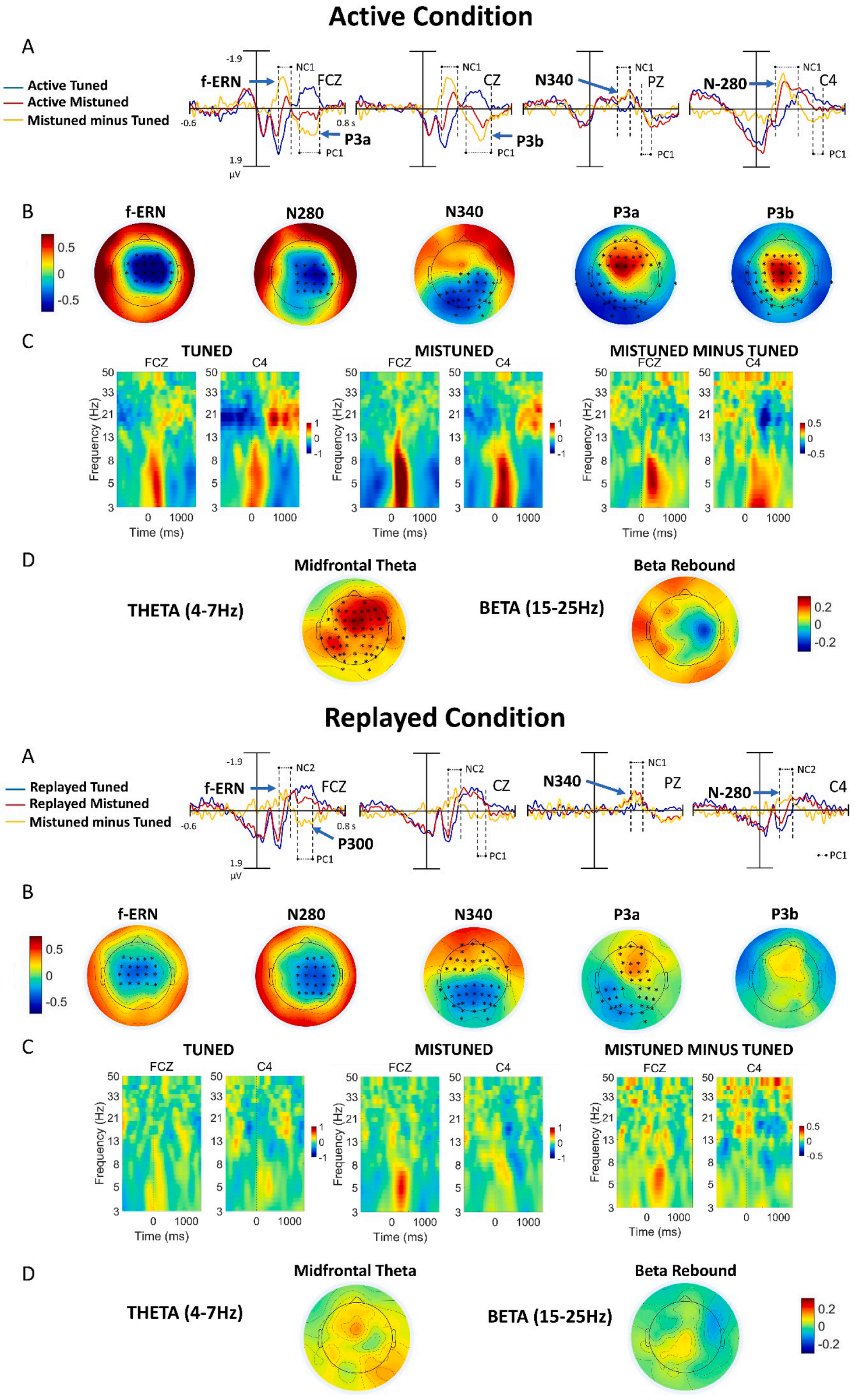


Fig. 4. Responses during playing tuned violin notes compared to listening. A ERPs elicited to active (blue) and replayed (red) notes and difference waves (yellow) for a subset of electrodes. The black lines below the ERP at each electrode show the significantly different time periods according to the cluster analysis. NC1 for Negative Cluster 1 and PC1 for Positive Cluster 1; **B** Motor-evoked potential scalp map (Active – Replayed difference from −130 to 34 ms). Asterisks correspond to the significant electrodes in this time period according to cluster analysis; **C** ERSP plots at central electrodes elicited to Replayed and Active notes; **D** Scalp plots of ERSP in theta and beta frequency ranges.



(caption on next page)

Fig. 5. Effects of producing a mistuned note (Top panel, Active Condition) and listening to the replay of a mistuned note (Bottom panel, Replayed Condition). From top to bottom: **A.** ERPs elicited to tuned (blue) and mistuned (red) notes and difference waves (yellow) for a subset of electrodes. The lines below the ERP of each electrode show the significantly different time-periods according to the cluster analysis. NC1: Negative Cluster 1. PC1: Positive Cluster 1. NC2: Negative Cluster 2. **B.** Scalp maps (mistuned – tuned difference). The marks correspond to the significant electrodes in this time period according to cluster analysis. f-ERN (180 to 260 ms), N280 (260 to 320 ms), N340 (320 to 360 ms), P3a (360 to 440 ms), P3b (460 to 560 ms). **C.** ERSP plots of tuned, mistuned, and difference on FCz and C4. **D.** Scalp plots of ERSP (mistuned – tuned difference) in theta and beta frequency ranges. The marks correspond to the significant electrodes in this time period according to cluster analysis. Midfrontal Theta (124 to 508 ms), Beta Rebound (606 to 806 ms).

decomposed into an earlier component localized in frontal electrodes between 380 and 500 ms (henceforth P3a) and a more central one between 500 and 600 ms (henceforth P3b). We also compared tuned and mistuned notes during playing in the time-frequency domain (Fig. 5, panels Active Conditions, C and D). This analysis showed an increase in theta power (4–7 Hz) for mistuned notes between 124 and 508, over the whole scalp ($T = 2.313$; $p < .001$), which was maximal in mid-frontal electrodes between 124–508 ms reaching its greatest amplitude in FCZ at 313 milliseconds. Henceforth, we refer to this error-related component as midfrontal theta. Visually, we also observed a decrease in beta power at the beta rebound location, however, no significant clusters were found in the beta range.

Similarly to during playing, listening to the replay of a mistuned note also elicited enhanced negativities followed by enhanced positivities, compared to listening to the replay of a tuned note. Specifically, we found a negative cluster in the time period between 206 and 300 ms ($T = -1.962$; $p = .0009$) in which we could again identify a f-ERN peaking at 260 ms at electrode FCz, followed by a right centro-parietal negativity that started around 200 ms, maintained its amplitude until 300 ms, and was maximum at 290 ms at the C4 electrode, which we termed the N-280. A second negative cluster followed between 332 and 446 ms ($T = 2.013$; $p < .001$) containing parietal, occipital, and left-temporal electrodes, peaking at Pz which we termed the N-340. Similarly to the Active condition, these negativities were followed by a positive cluster between 346 and 496 ms ($T = 1.877$; $p = .001$) containing frontal and central electrodes which we identified as P3a, peaking at 330 ms in electrode FCz. However, no P3b could be identified during listening, and the cluster-based analysis comparing ERSP responses between tuned and mistuned notes also yielded no significant differences.

In summary, when comparing mistuned vs tuned notes we identified the following error-related components and defined the time and frequency windows and electrodes employed to measure them in subsequent analyses: In the ERPs, we identified the f-ERN (henceforth measured at electrode FCz between 180–280 ms), the N-280 (electrode C4, 260–320 ms), the N-340 (electrode Pz, 320–360 ms), and the P3a (electrode FCz, 360–440 ms) in both the Active and Replayed conditions. The P3b (electrode Cz, 460–560 ms) was identified only in the Active condition. In the ERSP analysis, we identified the midfrontal theta in the Active condition (henceforth computed as the average value between 140 and 374 ms in the 3–7 Hz range in FCZ).

4.2.3. *Interactions between production and tuning*

Once the components of interest were clearly defined, we analyzed the effects of production and their interactions with tuning on the auditory, motor, and error-related indices. First, we analyzed the auditory components using the P1-N1 and N1-P2 peak-to-peak measurements. We did not observe any effects on the P1-N1 peak-to-peak amplitude. However, the N1-P2 peak-to-peak amplitude was reduced to mistuned notes when compared with tuned notes (main effect of Tuning, $F = 17.193$, $p = .001$, $\eta^2 = 0.551$), showing a greater reduction in the Active Condition (Condition*Tuning interaction; $F = 9.112$, $p = .009$, $\eta^2 = 0.394$).

Next, we analyzed the motor indices. In the ERSPs, the post-movement beta rebound activity was, as expected, only present in the Active and not in the Replayed condition (main effect of Condition; $F(1) = 38.61$, $p < .001$, $\eta^2 = 0.734$), and it was affected by Tuning (main effect of Tuning; $F(1) = 8.19$, $p = .01$, $\eta^2 = 0.37$). Specifically, we found a suppression of post-movement beta rebound in Mistuned compared to Tuned notes (Condition*Tuning interaction, $F(1) = 11.97$, $p = .004$, $\eta^2 = 0.46$). Thus, motor-related activity was suppressed in the Active condition following the commission of an error.

Finally, we analyzed the error-related indices. Significant main effects of Tuning confirmed that Mistuned notes elicited an f-ERN ($F(1) = 16.93$, $p = .001$, $\eta^2 = 0.54$.), an N280 ($F(1) = 45.54$, $p < .001$, $\eta^2 = 0.76$.), an N340 ($F(1) = 16.45$, $p = .001$, $\eta^2 = 0.54$.), a P3a ($F(1) = 20.50$, $p < .001$, $\eta^2 = 0.59$.), and a P3b ($F(1) = 24.77$, $p < .001$, $\eta^2 = 0.64$) in the ERPs; and an increase of midfrontal theta in the ERSPs ($F(1) = 5.14$, $p = .04$, $\eta^2 = 0.269$). In general, these error components appeared less robust during listening than during playing (Fig. 6). We found larger activity in the P3a and P3b windows in the Active compared to the Replayed condition (main effect of Condition; P3a: $F(1) = 6.34$, $p = .025$, $\eta^2 = 0.31$, P3b: $F(1) = 6.9$, $p = .02$, $\eta^2 = 0.33$). However, we found significant Condition*Tuning interactions only for the P3b ($F(1) = 21.32$, $p < .001$, $\eta^2 = 0.60$.) and the midfrontal theta ($F(1) = 5.02$, $p = .042$, $\eta^2 = 0.26$) components. Mistuned notes elicited a P3b and a midfrontal theta only in the active condition. Pairwise comparisons between tuned and mistuned notes showed significant differences in amplitude for the P3b ($p < .0001$) and the midfrontal theta component ($p = .005$). No significant differences were found between tuned and mistuned notes in the replayed condition, confirming the absence of these two components during the replay.

In summary, we found that tuning modulated auditory, motor and error-related electrophysiological indices more prominently when notes were self-produced, specifically showing a reduction of auditory (N1-P2 peak-to-peak amplitude) and motor (beta rebound) responses, and an enhancement of error-related indices (P3b and midfrontal theta) following the commission of a tuning error, compared to when listening to the replay of a mistuned note.

4.2.4. *Effects of error magnitude*

Next, we examined whether the magnitude of the motor and error-related components depended on the magnitude of the pitch error, and whether any error magnitude effects depended on whether the notes were self-produced or passively heard (see Fig. 7). Regarding the motor components, the amplitude of the post-movement beta rebound activity was not affected by error magnitude. Regarding the error-related components, the amplitude of the f-ERN, the P3b, and midfrontal theta increased with the magnitude of the error independently of Condition (main effect of Magnitude; f-ERN: $F(1.87) = 7.43$, $p = .003$, $\eta^2 = 0.21$, P3b: $F(2) = 4.91$, $p = .015$, $\eta^2 = 0.260$; midfrontal Theta: $F(2) = 6.57$, $p = .005$, $\eta^2 = 0.320$). Pairwise comparisons showed significant differences for midfrontal Theta, f-ERN and P3b between the 30–50 range and the >70 range (midfrontal Theta: $p = .006$; f-ERN: $p = .021$; P3b: $p = .003$). In contrast, the amplitude of the N340 increased with the magnitude of the error only in the Active condition (Condition*Magnitude interaction: $F(2) = 3.84$, $p = .034$, $\eta^2 = 0.215$.). Specifically, the amplitude of N340 in the Active condition was smaller for errors in the 30–50 range compared to errors in the >70 range ($p = .006$). The amplitude of the N280, and the P3a was not affected by error magnitude.

In summary, larger mistuning in either played or passively heard notes led in general to larger error-related components (f-ERN, P3b, midfrontal Theta), except for the N340, which specifically indexed error magnitude during playing only.

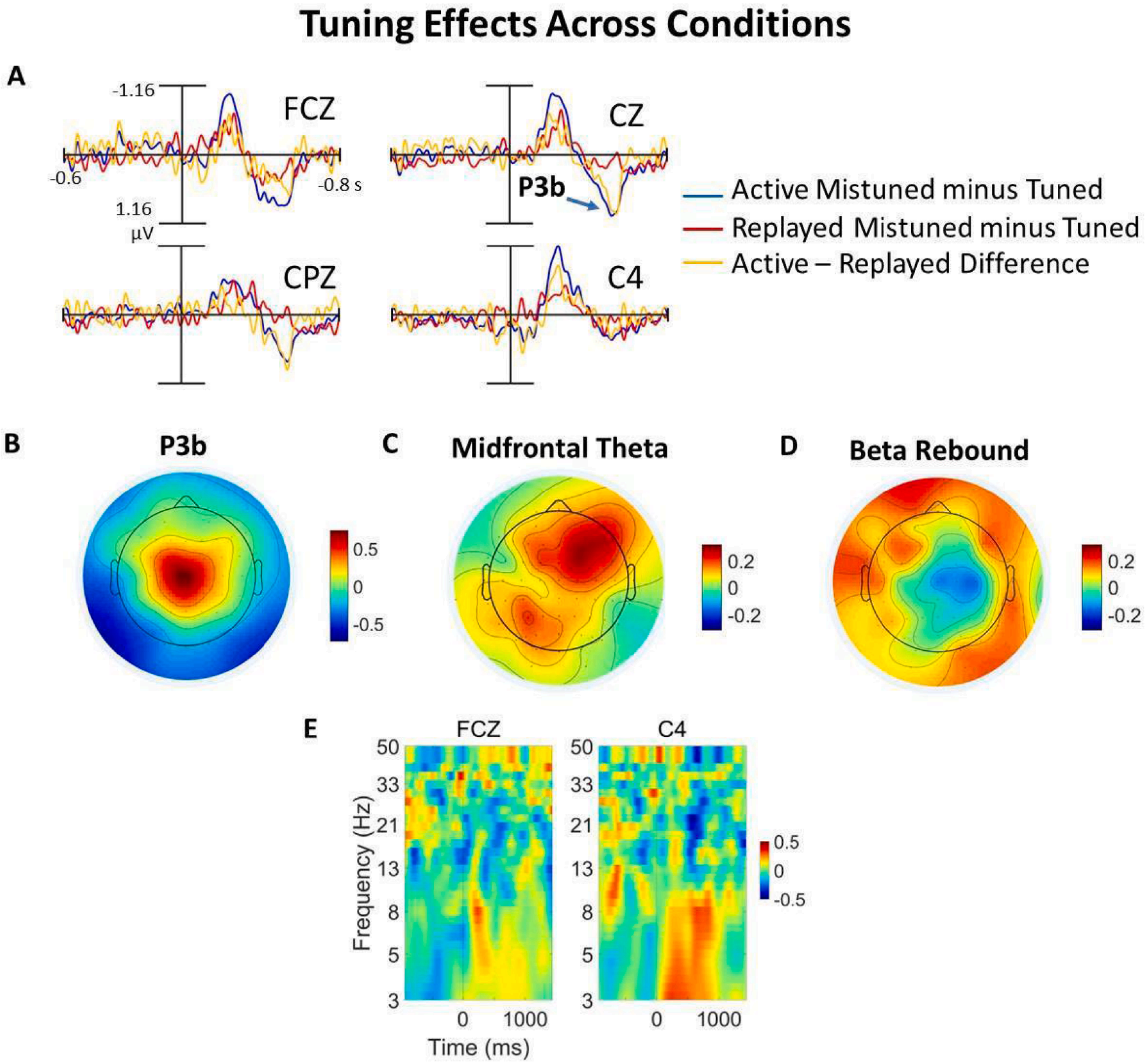


Fig. 6. Contrast of tuning effects in the Active vs. Replayed condition. A. ERPs of the tuning effects (Mistuned minus Tuned) in the Active and Replayed Conditions on electrodes FCz, Cz, CPz and C4. **B, C and D.** Scalp maps of tuning effects in Active minus tuning effects in Replayed difference in the P3b time-period (466 to 560 ms), theta power (Midfrontal Theta: 4–7 Hz, 140 to 374 ms) and beta power (Beta Rebound: 15–25 Hz, 606 to 806 ms). **E.** TF power plots of tuning effects in Active minus tuning effects in Replayed difference on FCz and C4.

4.2.5. *Effects of the latency of corrective movements*

Next, we analyzed whether the onset latency of corrective movements influenced the latency and amplitude of our error-related components of interest. We measured the latency of the peak of each error-related component elicited by mistuned notes followed by fast, medium, and slow corrections, separately for the Active and Replayed conditions. For the N280 we could not identify a clear maximum peak. It seemed to extend longer in time in slow corrections than in fast corrections and seemed to maintain its amplitude until the onset of the corrective movement (see Supplementary Materials). Regarding the rest of the error-related components, we found that the latency of corrective movements modulated the latency of the P3b component (main effect of speed: $F(1.68) = 3.75$, $p = .045$, $\eta^2 = 0.211$). Pairwise comparisons, revealed that the latency of the P3b was larger for slow than for fast corrective movements ($p = .019$). The latency of the rest of the error-related components (f-ERN, N340, and P3a) depended neither on the latency of the corrective movements, nor on whether the participants were performing the corrective movements or listening to their replay.

However, we found that the latency of corrective movements modulated the amplitude of several of the error-related components, as measured on the mistuned minus tuned difference waves (Fig. 8; the results of the repeated measures analyses for each component are summarized in Table 3, Supplementary Materials).

Both midfrontal Theta and P3b were modulated by the speed of the corrective movement independently of condition. Specifically, both the amplitude of midfrontal Theta power and the amplitude of P3b were lower for slow compared to medium corrections (midfrontal Theta: $p = .017$; P3b: $p = .004$). Conversely, f-ERN, N340 and N280 amplitudes were modulated by the speed of the correction in the Active condition only (f-ERN: $F(2) = 8.78$, $p = .001$, $\eta^2 = 0.386$; N280, $F(2) = 5.57$, $p = .009$, $\eta^2 = 262$; N340, $F(2) = 4.96$, $p = .014$, $\eta^2 = 0.262$). Pairwise comparisons between the three different speeds in the Active condition for these components are summarized in Table 1. Specifically, the amplitude of the f-ERN increased with the speed of the correction. Pairwise comparisons confirmed significantly larger f-ERN amplitudes for fast compared to slow, and medium compared to slow corrections. On the contrary, N280 and N340 amplitudes tended to be larger for slow compared to fast corrections. However, pairwise comparisons revealed only significantly increased N280 amplitudes for slow compared to fast corrections. No significant effects of speed were found in the passive condition.

In summary, only the time-course of the P3b was significantly related to the timing of the corrective movements. The amplitude of several

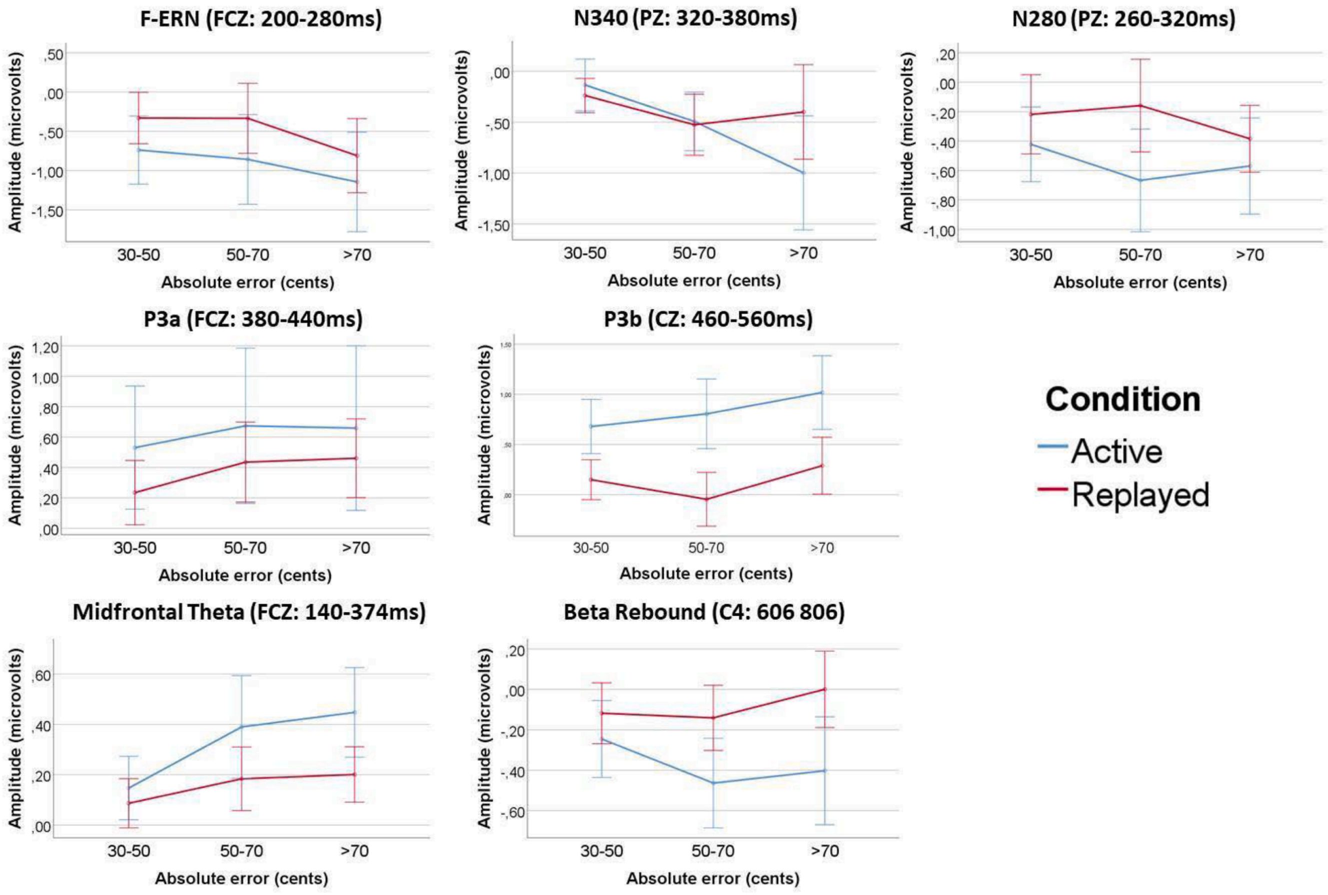


Fig. 7. ERP Component Amplitudes as a Function of Pitch Error Magnitude. Amplitude of the motor and error-related components measured on the mistuned minus tuned differences as a function of the magnitude of the pitch error.

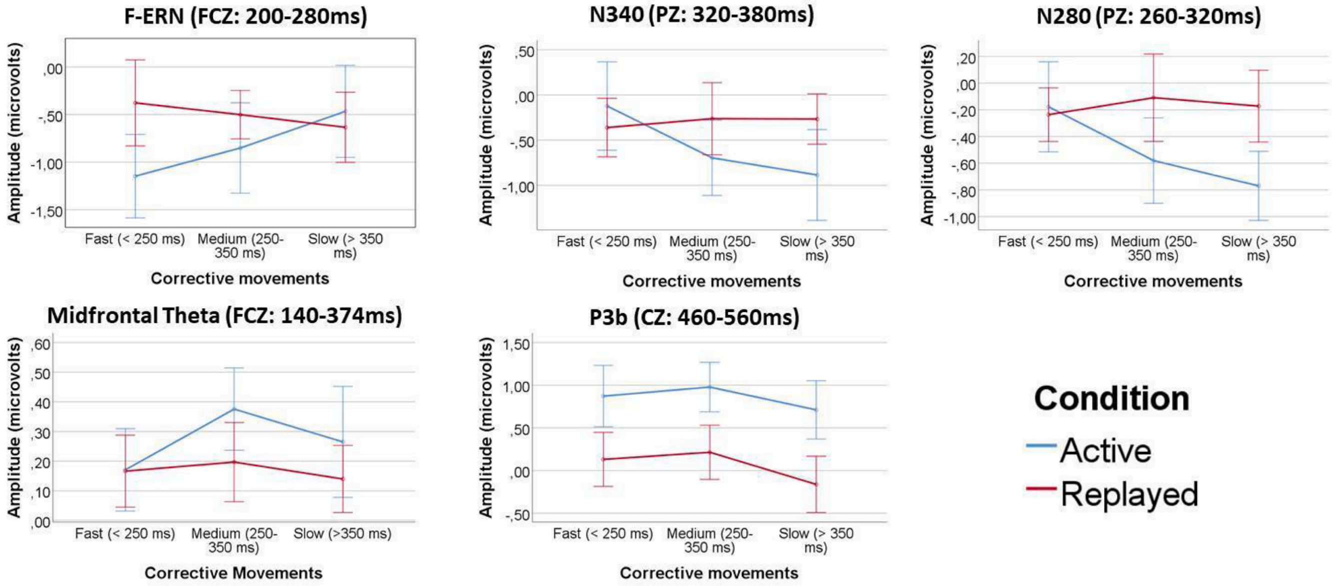


Fig. 8. Amplitude of the error-related components measured in the mistuned minus tuned difference waves as a function of the onset latency of corrective movements.

error-related indices, however, was modulated by the speed of the corrective movements. Both P3b and theta power increased with the speed of the correction regardless of condition, and f-ERN showed the same pattern in the Active condition only. Conversely, the N280 amplitude decreased with the speed of the correction in the Active condition and its amplitude extended with the timing of the corrective movements.

Table 1
Pairwise comparisons.

Component	(I)Speed	(J)Speed	Mean Diff (I-J)	Std.Error	Sig.[b]
f-ERN	Fast	Medium	−0.296	.159	n.s.
	Fast	Slow	−0.680*	.186	.008*
	Medium	Slow	−0.385*	.139	.046*
N280	Fast	Medium	.403	.212	n.s.
	Fast	Slow	.592*	.183	.018*
	Medium	Slow	.189	.142	n.s.
N340	Fast	Medium	.573	.215	n.s.
	Fast	Slow	.763	.285	n.s.
	Medium	Slow	.189	.251	n.s.

Note: Pairwise comparisons on the amplitude of the error-related components between the three different speeds in the Active Condition. Based on estimated marginal means.

* The mean difference is significant at the 0.05 level.

[b] Adjustment for multiple comparisons: Bonferroni.

5. Discussion

In this study, we investigated the neural mechanisms underlying error detection and correction in violin performance. Our findings revealed that violinists exhibit distinct neural signatures associated with pitch errors, with differences emerging between active performance and passive listening conditions. While we observed well-established components such as the f-ERN and P300, we also identified novel ERP components (N280 and N340) that, as we will discuss, appear to play a role in sensory-prediction error monitoring. These findings suggest that violin performance engages two parallel error-monitoring systems: one driven by reward-based learning and another reliant on sensory feedback mechanisms.

In the sections that follow, we first examine the ERP signatures of error detection (Section 5.1), comparing our findings with previous research on piano performance and identifying both shared and unique components of violin error monitoring. We then analyze the findings through the lens of Reinforcement Learning-ERN theory (Section 5.2).

Next, we delve into the dynamics of corrective movements and sensory-prediction errors (Section 5.3), describing how fast and slow corrections recruit distinct neural systems. We then explore indices of internal model updating (Section 5.4), discussing how the brain continuously refines its predictions based on auditory feedback and motor outcomes.

Finally, in the last section, we discuss the modulation of auditory processing during violin playing (Section 5.5), exploring how auditory perception is shaped by motor actions and real-time performance demands.

5.1. Error detection mechanisms in musical performance

5.1.1. Analysis of behavioral results

Previous research revealed that violin players exhibited an average error of 18 cents when attempting to match the pitch of a pure tone using the violin (Blanco et al., 2021). However, it is important to consider that matching tones with different timbres can pose additional challenges, potentially resulting in larger error rates (Hutchins and Peretz, 2012). To address this, we defined a note as "Tuned" when it fell within the range of 0–15 cents, and a note as "Mistuned" when it exceeded 30 cents. Interestingly, we observed that approximately half of the errors below 30 cents and above 15 cents were disregarded and left uncorrected by our participants. This finding raises questions regarding their conscious detection of such errors, prompting us to exclude them from the subsequent analysis.

In addition, we opted to combine self-generated and externally-generated performance errors into a unified category referred to as ’Mistuned’ errors. There are two reasons for this decision: Firstly, no significant differences were observed in the evoked brain activity between self-generated and externally-generated errors. Secondly, no significant differences were found in the participants’ behavior following self-generated versus externally-generated errors. This includes reaction time, percentage of corrected notes, magnitude of corrective movements, and other related measures.

5.1.2. Analysis of ERP results

Consistent with previous studies investigating error monitoring processes in piano performance, we observed a frontal negativity occurring at approximately 240 ms after mistuned notes, resembling the f-ERN (feedback-related negativity). This was followed by a P300 component in both the active and replayed conditions, which aligns with previous findings (Maidhof et al., 2010; Herrojo-Ruiz et al., 2009; Katahira et al., 2008). The P300 component can be further subdivided into two components: an earlier frontal component referred to as the P3a and a later central component known as the P3b (Ruchsow et al., 2005; Veen and Carter, 2002). While both components were observed in the active condition, only P3a activity was present in the replayed condition. The absence of the P3b component in the replayed condition may be attributed to the fact that our replayed condition was not task-relevant. This finding is consistent with the study conducted by Maidhof et al. (2010), who observed that the presence of a P3b was dependent on participants’ requirement to detect and count mistuned notes when listening to melodies.

On the other hand, similar to the findings reported by Maidhof et al., larger f-ERN amplitudes were also noted in the active condition compared to the replayed condition. In our study, although there appeared to be a higher average amplitude of the f-ERN in the active condition compared to the replayed condition, this difference did not reach statistical significance. In contrast, we did observe a more pronounced reduction of the N1-P2 component in the active condition compared to the replayed condition, indicating a greater decrease in the P2 component. This larger reduction of the P2 component may be attributed to its overlap with the f-ERN, leading to a more prominent impact on the N1-P2 complex.

Nevertheless, notable disparities emerged when comparing our findings with the existing literature on piano performance. We uncovered novel components that had not been previously reported in piano-related studies. Specifically, we identified in the ERP a negativity at right central areas around 280 ms, which we labeled as N280, as well as a negativity at parietal areas around 340 ms, which we labeled as N340. As we will elaborate further in the subsequent discussion, our interpretation of these results is predicated on the joint operation of two error monitoring systems that function in concert to effectively identify and rectify prediction errors originating from the auditory consequences of musical performance. These systems encompass the processing of reward and sensory-prediction errors (Krakauer et al., 2019).

5.2. Reinforcement learning-ern theory in violin performance

Earlier investigations involving piano performance have shown that the error-related negativity (ERN) can exhibit anticipatory responses, occurring as early as 100 ms prior to the onset of an incorrect note in the performance of skilled pianists (Maidhof et al., 2009, 2013; Ruiz et al., 2011; Herrojo-Ruiz et al., 2009). Maidhof et al., for example, found that self-generated errors elicited a pre-error negativity (pre-ERN), whereas externally introduced errors resulted in a feedback error-related negativity (f-ERN). In the present study we did not observe this distinction between self-generated and externally introduced errors. The rationale behind this disparity will now be elucidated.

While it is commonly believed that the error-related negativity (ERN) and the feedback-related negativity (f-ERN) may have a shared underlying generator in the Anterior Cingulate Cortex (ACC) (Gehring and Willoughby, 2004; Hall et al., 2007; Trujillo and Allen, 2007), the primary distinction between them, as proposed by the Reinforcement-Learning-ERN theory (Holroyd and Coles, 2002), lies in their eliciting factors. The ERN is triggered by an reference copy of the motor command (internal feedback), whereas the f-ERN is evoked by external feedback. This theoretical framework provides a plausible explanation for the divergent findings observed in prior piano literature concerning externally generated errors (manipulated notes) and self-generated errors. However, in our study, we did not observe such a distinction between manipulated and self-generated errors. Instead, we found the elicitation of an error-related negativity (f-ERN) in both types of errors. We attribute this discrepancy to two primary factors.

Firstly, the inter-onset interval (IOI) employed in previous piano studies typically ranged from 100 to 200 ms, whereas in our experiment, it was extended to 2 s. The significantly slower tempo used in our study made it less likely for participants' self-generated errors to arise due to the same reasons observed in piano experiments. On one hand, our participants had ample time to plan their intended movements between notes. On the other hand, each note in the melody corresponded to a distinct ballistic movement separated from the others. In contrast, when tempos are fast, as in the aforementioned piano experiments, a single ballistic movement encompasses a sequence of several notes, facilitating the detection of errors in advance through internal feedback.

Secondly, as previously mentioned, the violin is an instrument that heavily relies on external feedback for accurate execution. The nature of the violin necessitates continuous fine-tuning adjustments, leading us to anticipate that a significant portion of self-generated errors at lower tempos, such as those used in our experiment, would elicit an f-ERN rather than a pre-ERN.

It has been proposed that the ERN and P3a might be components of a single oscillatory potential within the theta frequency range (Herrmann et al., 2004; van Veen and Carter, 2002). However, in our study, we observed distinct patterns in the components when considering both the magnitude of the error and the latency of corrective movements. Interestingly, the f-ERN and midfrontal theta showed modulation based on the magnitude of the error in both the active and replayed conditions, while the P3a did not exhibit such sensitivity. Furthermore, we identified contrasting behaviors between the f-ERN and midfrontal theta. Specifically, slower corrective movements were associated with reduced f-ERN amplitudes but larger midfrontal theta responses, indicating that these components may have separate underlying sources. This finding aligns with research on online response correction, which also demonstrated divergent behaviors between the ERN and midfrontal theta (Kieffaber et al., 2016). Specifically, they reported that the ERN remained consistent across fast and slow error corrections, while the latency of midfrontal theta was linked to the initiation of corrective movements. In our study, however, we did not observe latency effects in either the f-ERN or midfrontal theta components.

An objection that may arise is the possibility that errors of greater magnitude were detected more quickly, leading to larger f-ERN amplitudes in fast corrective movements. However, in our study, we did not observe any association between the magnitude of the error and the speed of corrective movements. Even if such a relationship existed, we would have anticipated larger midfrontal theta responses in fast corrective movements as well.

5.3. Corrective movements and sensory-prediction errors in violin performance

Krigolson et al. (2007, 2008) introduced a distinction between two types of motor errors that rely on external feedback for their occurrence: outcome errors and target errors. This distinction aligns with the recent research in motor learning that differentiates between reward errors and sensory prediction errors (SPE) (Krakauer et al., 2019). Generally, reward errors are processed within the medial-frontal cortex, involving the anterior cingulate cortex (ACC) and the basal ganglia, and they represent the failure to achieve the intended movement goal (Krigolson and Holroyd, 2006). However, recent studies using intracranial recordings suggest that the Supplementary Motor Area (SMA) may play a more leading role in generating these signals (Bonini et al., 2014; Fu et al., 2019). On the other hand, sensory prediction errors represent a discrepancy between the actual motor command and the appropriate motor command and involve the posterior parietal cortex (PPC) and the cerebellum. The PPC is believed to play a crucial role in evaluating sensory prediction errors and facilitating subsequent adjustments to the motor command, while the cerebellum also contributes to the processing of sensory prediction errors and motor adaptation (Culham et al., 2003; Desmurget et al., 1999, 2001; Desmurget and Grafton, 2000; Diedrichsen et al., 2005). Recent neuroimaging studies have offered support for the role of the PPC for adaptive modification of behavior in music production with a string instrument (Segado et al., 2021). In that study, participants were required to ignore or compensate for pitch-shifting manipulations of their produced tone. They found that compensating manipulations recruited parietal areas related to auditory-motor integration, particularly the intraparietal sulcus (IPS) and the supramarginal gyrus (SMG).

In the context of violin performance, it is anticipated that the detection of an incorrect note would give rise to reward errors, as observed in piano playing. However, due to the heightened requirement for precise movements in violin playing, violinists heavily rely on auditory feedback to detect errors and continuously adjust the pitch during their performance. This unique characteristic led us to hypothesize that in addition to reward errors, we would also observe activity associated with sensory prediction errors.

In this study, we identified novel ERP components which we believe are linked to sensory-prediction errors and the involvement of the PPC. Specifically, we observed a parietal negativity emerging around 340 ms after mistuned notes, which we termed the N340. Notably, similar studies investigating error correction in visual-tracking tasks (Krigolson and Holroyd, 2006, 2007) have also reported a parietal negativity peaking around 360 ms, associated with sensory-prediction errors and PPC activity. In addition to the N340, we identified a right-central negativity peaking around 280 ms, which we labeled the N280.

Of note, these components were observed in both the replayed and active conditions, but their amplitude were modulated solely in the active condition. Similar to the f-ERN, these components also exhibited an increase in magnitude based on the error magnitude. However, in contrast to the f-ERN and similar to the midfrontal theta, their amplitude decreased following fast corrective movements and increased following slow corrective movements.

Interestingly, despite the latency of the N280 component remained unaffected by the onset of corrective movements we could observe how it appeared to have a longer duration in slow corrections compared to fast corrections and maintained its amplitude until the onset of the corrective movement (see Supplementary Materials). Previous studies have reported similar behavior for the ERN, which has been likened to an "alarm system" that remains active until error correction begins (Burle et al., 2008; Kieffaber et al., 2016). Given that this activity was contralateral to the hand in charge of corrective movements it could be hypothesized a similar functional behavior for this component. However, based on its timing (~280 ms) and spatial distribution over right-central areas contralateral to the correcting hand, the N280 component is most plausibly linked to a motor readiness potential preceding the corrective movement. Its presence in the Replayed condition may reflect strong auditory-motor coupling in musicians, particularly when listening to their own performance. Notably, in this condition, the potential's duration appeared to extend beyond the typical onset of the corrective movement.

Similar studies report a comparable negativity in voice pitch-shift paradigms, where participants sustain vocal tones and hear pitch perturbations. In these cases, a central or right-central negative component peaking around 100–200 ms post-perturbation has been identified as the Auditory Awareness Negativity (AAN) (Suchý et al., 2025). The AAN is thought to reflect a neural correlate of auditory awareness, emerging when pitch deviations are detected. However, the authors did not investigate whether the timing of this component was related to the onset of corrective vocal responses, leaving open the question of whether such negativities reflect awareness, motor preparation, or a combination of both. Future research should address this gap by jointly assessing auditory awareness and motor readiness, to better distinguish between these possible interpretations of the N280.

Finally, we found that corrective movements influenced both the latency and amplitude of the P3b component. Slower corrective movements were associated with longer latencies of the P3b component compared to faster corrective movements. Additionally, the amplitude of the P3b component increased with error magnitude but decreased with faster corrective movements. Overall, it can be summarized that, apart from the f-ERN, the amplitude of the other components tended to be greater for slow corrective movements and lower for fast corrective movements.

One plausible interpretation of these findings is that we have identified distinct components that may be associated with the functioning of two error monitoring systems. The "fast" system, primarily mediated by the medial frontal cortex and reward errors, is responsible for promptly detecting errors with a certain level of certainty, as reflected by the amplitude of the f-ERN. When the certainty is high, this system initiates the corrective action, rendering the activity of the "slow" system redundant, thereby resulting in lower amplitudes of midfrontal theta, N280 and N340. However, in situations where the uncertainty is high (indicated by reduced f-ERN amplitude), other processes come into play, accumulating evidence of an error until the "slow" system, mediated by the posterior parietal cortex and sensory-prediction errors, responds by producing larger amplitudes of midfrontal theta, N280 and N340, eventually initiating the corrective movements.

5.4. *Indices of updating of internal model*

A compelling interpretation suggests that the P300 component represents the updating of an internal model of the environment (Palidis et al., 2019; Donchin, 1981; Donchin and Coles, 1988; Krigolson et al., 2008; MacLean et al., 2015; Polich, 2007). Specifically, it has been established that the first component, P3a, is associated with attention, while the second component, P3b, is linked to memory processing (Polich, 2007). This perspective provides an explanation for the absence of a P3b in the replayed condition: since it is not a task-relevant condition, there is no need to update the internal model to enhance future performance in the task.

Notably, we also observed that the latency of the P3b was closely related to the onset of corrective movements. This finding supports our earlier suggestion that corrective movements begin only after accumulating sufficient evidence that an error has occurred. Under conditions of uncertainty, the need to update the internal model arises only once the error has been reliably recognized, allowing for adjustments based on new evidence.

In addition, our study revealed the presence of beta activity associated with motor-related potentials in the Active Condition. Notably, recent investigations by Tan et al. (2014) and Torrecillos et al. (2015) examined the impact of perturbed movements in a force field task on the modulation of post-movement beta-rebound and correction beta-enhancement, with varying degrees of perturbation. In our experiment, we observed beta-suppression activity, predominantly localized in the right-lateralized C4, just prior to the onset of movement. Additionally, we identified beta-rebound activity (also in C4) at approximately 600 and 800 ms, which appeared to align with the average timing of corrective movements (~800 ms). Interestingly, we found that the magnitude of the beta rebound was influenced by the magnitude of the committed error. However, neither the latency nor amplitude of the beta rebound demonstrated modulation in response to the onset of corrective movements. Torrecillos et al. (2015) similarly reported that the beta rebound appeared to be independent of online movement correction and exhibited insensitivity to both goal and sensorimotor errors. Some interpretations suggest that the beta rebound may be linked to the process of updating a forward model (Cao and Hu, 2016). According to this perspective, a higher modulation of beta rebound may reflect the active maintenance of an existing forward model, while a lower beta rebound may be associated with the process of updating the forward model.

To summarize, the observed activity related to the parietal P300 and beta rebound in our experiment may be linked to processes involved in internal model updating. The parietal P300, occurring approximately 500 ms after the onset of corrective movements (~300 ms), but before their completion (~800 ms), may reflect the updating of the internal model. Similarly, the beta rebound, initiating around 600 ms and persisting until the end of corrective movements at approximately 800 ms, may also be associated with the updating process. Both the parietal P300 and beta rebound were influenced by the magnitude of errors.

5.5. *Modulation of auditory processing during violin playing*

N1 suppression of self-generated sounds refers to a phenomenon observed in the auditory system, where the brain attenuates or suppresses the neural response to sounds that are self-generated compared to sounds that are externally generated (Ford et al., 2001). When we produce a voluntary action that generates a sound, such as speaking or playing a musical instrument, the brain predicts and expects the sensory consequences of our actions. This prediction is thought to lead to a suppression or reduction in the neural response to the self-generated sound, resulting in a smaller N1 component in the ERPs recorded from the scalp. Previous research found that this N1 reduction becomes weak when auditory feedback is altered during speaking (Heinks-Maldonado et al., 2005; Jones et al., 2013; Scheerer and Jones, 2014, 2018), supporting the existence of a forward model modulating cortical responses to self-generated sounds. Thus, in this study we were expecting a reduction of the N1 component in the active condition compared to the replayed and the altered feedback condition.

Despite this experimental design allowed us to clearly identify the three auditory components P1, N1 and P2 in the ERPs in both active and replayed conditions, we did not find self-generation effects in N1-P1 amplitudes. We found that mistuned notes elicited a greater suppression of the N1-P2 amplitudes in both active and replayed conditions, and that this suppression was larger for the active condition. However, this is probably due to the effect of the f-ERN that peaked between 180 and 260 ms during mistuned notes.

To our knowledge, no previous study investigated the effects of N1-suppression in music performance. One exception is Katahira et al. (2008) who did not find either self-generation effects in N1 between an active and a passive condition in the piano (Katahira et al., 2008).

The reason why we do not see a reduction in N1 may be due to multiple factors (Horvath, 2015; Hughes et al., 2013). In most studies where the attenuation effect has been found, self-generated sounds were not relevant for the subjects. Several studies have reported that attention in self-generated sounds generated an enhancement effect of the N1 that could overlap or interact with the suppression effect (Timm et al., 2013; Kok et al., 2012). However, the absence of clear self-generation effects may be attributed to a reduction of the N1 component in the Replayed Condition caused by repetition-related attenuation. Experienced musicians are highly sensitive to melodic and tonal repetition, and in our design, the four-tone melody was repeated three times per trial. This repetition likely led to habituation effects, particularly in the third (Replayed) condition, diminishing ERP amplitudes. Thus, more research is needed to elucidate if the N1-suppression effect can also be identified in musical performance.

5.6. *Limitations*

Several limitations should be acknowledged. First, the inter-onset interval (IOI) used in our study was relatively long (2 s), which may have influenced the absence of anticipatory pre-ERN responses observed in previous piano studies. Second, our study focused primarily on EEG responses, which provide high temporal resolution but limited spatial resolution. However, recent neuroimaging findings support the role of the posterior parietal cortex (PPC) in adaptive modifications of behavior in music production with string instruments. For instance, Segado et al. (2021) demonstrated that compensatory responses to pitch-shifting manipulations engaged parietal regions involved in auditory-motor integration, particularly the intraparietal sulcus (IPS) and the supramarginal gyrus (SMG). Given that our findings also point to PPC involvement—particularly in sensory-prediction errors—future research could further investigate its specific role in violin performance, distinguishing its contributions from those of medial-frontal and subcortical structures. Third, the repetition of the same four-note melody across the three conditions within each trial may have introduced habituation or repetition effects, particularly affecting ERP responses in the third condition—the Replayed condition. Indeed, we observed a general reduction in ERP amplitude when comparing the Replayed condition to the Reference condition, which could be partly attributed to such habituation. We acknowledge this as a limitation of our current design, as it may complicate direct comparisons of ERP amplitudes between the Active and Replayed conditions. However, if habituation were a major factor, we would expect a consistent and significant decrease in ERP amplitude in the Replayed condition relative to the Active condition. Although we observed a trend in that direction, the differences were not statistically significant. This suggests that, while repetition may have exerted some influence, it was not sufficient to systematically suppress auditory or performance-monitoring ERP components. Notable exceptions include the P3b, the beta rebound, and midfrontal theta. For the P3b and beta rebound, their reduction was expected regardless of habituation. In contrast, the significant reduction in midfrontal theta might indeed reflect, in part, a habituation effect. Overall, while we cannot fully exclude a contribution of habituation, the pattern of results suggests it was not the primary factor driving ERP amplitude differences across Tuning conditions.

Finally, a key limitation of our design is the absence of a motor-only condition, which would have allowed us to isolate motor-related activity from auditory and error-monitoring responses. Nevertheless, with the exception of the P3b and beta rebound, most components were present in both the Active and Replayed conditions. Among these, the N280 is the most likely to reflect motor-related activity, given its timing (corresponding to the ~300 ms average onset of corrective movements) and its topography (right-central areas, consistent with left-hand motor activity in violinists). While its latency did not appear strongly tied to correction onset, its extended duration—visible in Supplementary Figure 2—suggests a sustained process that may involve motor adjustments. However, we did not observe this extended duration in the Replayed condition (see Supplementary Figure 4), supporting the idea that this component may at least partially reflect active motor processes. One possibility is that, due to the strong coupling between auditory and motor systems in trained musicians, hearing their own performance may activate similar motor representations as during actual playing. This could partially account for the persistence of motor-related components in the Replayed condition, even in the absence of movement.

5.7. Summary and functional interpretation of results

To synthesize the main findings and their functional implications, we propose the following sequential interpretation of the EEG markers associated with pitch error detection and correction during violin performance: when a violinist makes a pitch error unintentionally and detects it retrospectively, the first response is the elicitation of reward error signals mediated by the medial frontal cortex—specifically the f-ERN. This response is modulated both by the magnitude of the error and the player's awareness of having made it. At that point, a motor preparation process begins, involving motor areas contralateral to the hand performing the correction. This preparation determines both the direction and magnitude of the corrective movement and is reflected in a gradually increasing negativity that peaks around the onset of the correction as well as in the midfrontal theta. Once the corrective movement is initiated, regions of the PPC involved in auditory-motor integration—such as the IPS and SMG—are recruited to process sensory prediction errors. These may be reflected in the EEG as a parietal negativity around 340 ms. Like the f-ERN, this parietal component appears to be modulated by both the magnitude of the error and the certainty of the correction. When the performer is confident about the presence of the error, as well as the direction and size of the correction needed, the f-ERN signal dominates and the amplitude of later components is reduced. In contrast, when the error is less clear and the correction is more uncertain, the f-ERN signal is weaker, but later components show increased amplitude—possibly compensating for the lack of early clarity. Finally, the internal model must be updated to prevent the same mistake from recurring. This process is reflected in the emergence of the P3b component and a beta rebound following the completion of the corrective movement.

6. Conclusion

In this study, we examined the electrophysiological responses associated with pitch errors in violin performance, identifying distinct neural markers of error detection and correction. Our results revealed the presence of both established ERP components, such as the f-ERN and P300, and novel components (N280 and N340) that appear to be associated with sensory-prediction errors. The distinction between fast and slow corrective movements provided further evidence for the coexistence of two error-monitoring systems: a medial-frontal mechanism sensitive to reward prediction errors and a parietal mechanism involved in sensory-prediction errors. In addition, we successfully implemented a setup to capture EEG data on musical performance, specifically using a fretless instrument like the violin. This innovative approach holds promise for future research endeavors. By employing this setup, we were able to explore the electrophysiological correlates of error monitoring in musical performance, an aspect that has often been overlooked due to the predominant use of keyed instruments like the piano. Keyed instruments offer less reliance on auditory feedback and lack the capacity for real-time corrections of movements to rectify errors. This research opens doors for expanding our understanding of error monitoring processes in more ecologically valid musical tasks, contributing to advancements in our comprehension of the mechanisms underlying musical learning

CRediT authorship contribution statement

Ángel David Blanco: Writing – original draft, Visualization, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jordi Costa-Faidella:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Alfonso Pérez Carrillo:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **David Dalmazzo:** Writing – review & editing, Software. **Rafael Ramirez-Melendez:** Writing – review & editing, Supervision, Resources, Funding acquisition. **Iria SanMiguel:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Ethics statement

The studies involving human participants were reviewed and approved by the Conservatoires UK Research Ethics committee on 04/04/2017, following the guidelines of the British Psychological Society. The patients/participants provided their written informed consent to participate in this study.

Funding

This work was partly sponsored by the Musical AI project (PID2019-111403GB-I00) funded by MCIN/AEI/ 10.13039/501100011033 and by "ERDF A way of making Europe"; by the Project PID2021-128790NB-I00, funded by MCIN/AEI/ 10.13039/501100011033 / FEDER UE; and by the European Union Horizon 2020 research and innovation program under grant agreement No. 688269 (TELMI project). It has been additionally supported by the Excellence Research Group 2017SGR-974 funded by the Secretaria d'Universitats i Recerca del Departament d'Empresa i Coneixement de la Generalitat de Catalunya and by the MDM-2017-0729-18-2M Maria de Maeztu Center of Excellence UBNeuro, funded by MCIN/AEI/ 10.13039/501100011033. ISM was supported by grant RYC-2013-12577, funded by MCIN/AEI/ 10.13039/501100011033 and by "ESF Investing in your future". Open access publication fees were partly covered by the University of Barcelona (Ajudes de la Universitat de Barcelona per a publicar en accés obert).

Acknowledgments

We sincerely thank all participants who voluntarily gave their time to carry out this experiment.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2025.121365.

Data availability

The datasets and code generated for this study can be found in https://doi.org/10.5281/zenodo.5090674.

References

Baumann, S., Koeneke, S., Schmidt, C.F., Meyer, M., Lutz, K., Jancke, L., 2007. A network for audio–motor coordination in skilled pianists and non-musicians. Brain Res. 1161, 65–78. Behringer (2022). Behringer. Retrieved from. https://www.behringer.com/.

Bendixen, A., SanMiguel, I., Schröger, E., 2012. Early electrophysiological indicators for predictive processing in audition: a review. Int. J. Psychophysiol. 83 (2), 120–131. https://doi.org/10.1016/j.ijpsycho.2011.08.003.

Bestmann, S., Krakauer, J.W., 2015. The uses and interpretations of the motor-evoked potential for understanding behaviour. Exp. Brain Res. 233 (3), 679–689. https://doi.org/10.1007/s00221-014-4183-7.

Bianchi, F., Santurette, S., Wendt, D., Dau, T., 2016. Pitch Discrimination in Musicians and Non-Musicians: Effects of Harmonic Resolvability and Processing Effort. J. Assoc. Res. Otolaryngol. 17 (1), 69–79. https://doi.org/10.1007/s10162-015-0548-2.

Blanco, A.D., Tassani, S., Ramirez, R., 2021. Effects of visual and auditory feedback in violin and singing voice pitch matching tasks. Front. Psychol. 12. https://doi.org/10.3389/fpsyg.2021.684693.

Bonini, F., Burle, B., Liégeois-Chauvel, C., Régis, J., Chauvel, P., Vidal, F., 2014. Action monitoring and medial frontal cortex: leading role of supplementary motor area. Science 343 (6173), 888–891. https://doi.org/10.1126/science.1247412.

Burle, B., Roger, C., Allain, S., Vidal, F., Hasbroucq, T., 2008. Error negativity does not reflect conflict: a reappraisal of conflict monitoring and anterior cingulate cortex activity. J. Cogn. Neurosci. 20 (9), 1637–1655. https://doi.org/10.1162/jocn.2008.20110.

Cao, L., Hu, Y.M., 2016. Beta rebound in visuomotor adaptation: still the status quo? J. Neurosci. 36 (24), 6365–6367. https://doi.org/10.1523/JNEUROSCI.1007-16.2016.

Cardellicchio, P., Hilt, P.M., Dolfini, E., Fadiga, L., D'Ausilio, A., 2020. Beta Rebound as an Index of Temporal Integration of Somatosensory and Motor Signals. Front. Syst. Neurosci. 14, 63. https://doi.org/10.3389/fnsys.2020.00063.

Cavanagh, J.F., Zambrano-Vazquez, L., Allen, J.J., 2012. Theta lingua franca: a common mid-frontal substrate for action monitoring processes. Psychophysiology 49 (2), 220–238. https://doi.org/10.1111/j.1469-8986.2011.01293.x.

Chen, J., Woollacott, M.H., Pologe, S., Moore, G.P., 2008. Pitch and space maps of skilled cellists: accuracy, variability, and error correction. Exp. Brain Res. 188 (4), 493–503. https://doi.org/10.1007/s00221-008-1380-2.

Chen, J., Woollacott, M.H., Pologe, S., Moore, G.P., 2013. Stochastic aspects of motor behavior and their dependence on auditory feedback in experienced cellists. Front. Hum. Neurosci. 7, 419. https://doi.org/10.3389/fnhum.2013.00419.

Culham, J.C., Danckert, S.L., De Souza, J.F., Gati, J.S., Menon, R.S., Goodale, M.A., 2003. Visually guided grasping produces fMRI activation in dorsal but not ventral stream brain areas. Exp. Brain Res. 153 (2), 180–189. https://doi.org/10.1007/s00221-003-1591-5.

Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134 (1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009.

Desmurget, M., Epstein, C.M., Turner, R.S., Prablanc, C., Alexander, G.E., Grafton, S.T., 1999. Role of the posterior parietal cortex in updating reaching movements to a visual target. Nat. Neurosci. 2 (6), 563–567. https://doi.org/10.1038/9219.

Desmurget, M., Grafton, S., 2000. Forward modeling allows feedback control for fast reaching movements. Trends Cogn. Sci. (Regul. Ed.) 4 (11), 423–431. https://doi.org/10.1016/S1364-6613(00)01537-0.

Desmurget, M., Gréa, H., Grethe, J.S., Prablanc, C., Alexander, G.E., Grafton, S.T., 2001. Functional anatomy of nonvisual feedback loops during reaching: a positron emission tomography study. J. Neurosci. 21 (8), 2919–2928. https://doi.org/10.1523/JNEUROSCI.21-08-02919.2001.

Diedrichsen, J., Verstynen, T., Lehman, S.L., Ivry, R.B., 2005. Cerebellar involvement in anticipating the consequences of self-produced actions during bimanual movements. J. Neurophysiol. 93 (2), 801–812. https://doi.org/10.1152/jn.00662.2004.

Di Russo, F., Pitzalis, S., Aprile, T., Spinelli, D., 2005. Effect of practice on brain activity: an investigation in top-level rifle shooters. Med. Sci. Sports Exerc. 37 (9), 1586. https://doi.org/10.1249/01.mss.0000177458.71676.0d.

Donchin, E., 1981. Surprise!… surprise? Psychophysiology 18 (5), 493–513. https://doi.org/10.1111/j.1469-8986.1981.tb01815.x.

Donchin, E., Coles, M.G., 1988. Is the P300 component a manifestation of context updating? Behav. Brain Sci. 11 (3), 357–374. https://doi.org/10.1017/S0140525X00058027.

Fattapposta, F., Amabile, G., Cordischi, M.V., Di Venanzio, D., Foti, A., Pierelli, F., Morrocutti, C., 1996. Long-term practice effects on a new skilled motor learning: an electrophysiological study. Electroencephalogr. Clin. Neurophysiol. 99 (6), 495–507. https://doi.org/10.1016/S0013-4694(96)96560-8.

Ford, J.M., Mathalon, D.H., Kalba, S., Marsh, L., Pfefferbaum, A., 2001. N1 and P300 abnormalities in patients with schizophrenia, epilepsy, and epilepsy with schizophrenialike features. Biol. Psychiatry 49 (10), 848–860. https://doi.org/10.1016/S0006-3223(00)01051-9.

Friston, K., 2010. The free-energy principle: a unified brain theory? Nat. Rev. Neurosci. 11 (2), 127–138. https://doi.org/10.1038/nrn2787.

Fu, Z., Wu, D.A.J., Ross, I., Chung, J.M., Mamelak, A.N., Adolphs, R., Rutishauser, U., 2019. Single-neuron correlates of error monitoring and post-error adjustments in human medial frontal cortex. Neuron 101 (1), 165–177. https://doi.org/10.1016/j.neuron.2018.11.016.

Gehring, W.J., Liu, Y., Orr, J.M., Carp, J., 2012. The error-related negativity (ERN/Ne). In: Luck, S.J., Kappenman, E.S. (Eds.), The Oxford Handbook of Event-Related Potential Components. Oxford University Press, pp. 231–291.

Gehring, W.J., Willoughby, A.R., 2004. Are all medial frontal negativities created equal? Toward a richer empirical basis for theories of action monitoring. Errors Confl. Brain. Curr. Opin. Perform. Monit. 14, 20. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.532.2756&rep=rep1&type=pdf.

Hall, J.R., Bernat, E.M., Patrick, C.J., 2007. Externalizing psychopathology and the error-related negativity. Psychol. Sci. 18 (4), 326–333. https://doi.org/10.1111/j.1467-9280.2007.01899.x.

Hatta, A., Nishihira, Y., Higashiura, T., Kim, S.R., Kaneda, T., 2009. Long-term motor practice induces practice-dependent modulation of movement-related cortical potentials (MRCP) preceding a self-paced non-dominant handgrip movement in kendo players. Neurosci. Lett. 459 (3), 105–108. https://doi.org/10.1016/j.neulet.2009.05.005.

Heinks-Maldonado, T.H., Mathalon, D.H., Gray, M., Ford, J.M., 2005. Fine-tuning of auditory cortex during speech production. Psychophysiology 42 (2), 180–190. https://doi.org/10.1111/j.1469-8986.2005.00272.x.

Herrmann, M.J., Römmler, J., Ehlis, A.C., Heidrich, A., Fallgatter, A.J., 2004. Source localization (LORETA) of the error-related-negativity (ERN/Ne) and positivity (Pe). Cogn. Brain Res. 20 (2), 294–299. https://doi.org/10.1016/j.cogbrainres.2004.02.013.

Herrojo-Ruiz, M., Jabusch, H.C., Altenmüller, E., 2009. Detecting wrong notes in advance: neuronal correlates of error monitoring in pianists. Cereb. Cortex 19 (11), 2625–2639. https://doi.org/10.1093/cercor/bhp021.

Herrojo-Ruiz, M., Strübing, F., Jabusch, H.C., Altenmüller, E., 2011. EEG oscillatory patterns are associated with error prediction during music performance and are altered in musician's dystonia. Neuroimage 55 (4), 1791–1803. https://doi.org/10.1016/j.neuroimage.2010.12.050.

Hickok, G., Poeppel, D., 2004. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition 92 (1–2), 67–99. https://doi.org/10.1016/j.cognition.2003.10.011.

Holroyd, C.B., Coles, M.G., 2002. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. Psychol. Rev. 109 (4), 679. https://doi.org/10.1037/0033-295X.109.4.679.

Horváth, J., 2015. Action-related auditory ERP attenuation: paradigms and hypotheses. Brain Res. 1626, 54–65. https://doi.org/10.1016/S0006-3223(00)01051-9.

Hughes, G., Desantis, A., Waszak, F., 2013. Attenuation of auditory N 1 results from identity-specific action-effect prediction. Eur. J. Neurosci. 37 (7), 1152–1158.

Hutchins, S.M., Peretz, I., 2012. A frog in your throat or in your ear? Searching for the causes of poor singing. J. Exp. Psychol. Gen. 141 (1), 76–97. https://doi.org/10.1037/a0025064.

Jones, A., Hughes, G., Waszak, F., 2013. The interaction between attention and motor prediction. An ERP study. Neuroimage 83, 533–541. https://doi.org/10.1016/j.neuroimage.2013.07.004.

Jung, T.P., Makeig, S., Humphries, C., et al., 2000. Removing electroencephalographic artifacts by blind source separation. Psychophysiology 37 (2), 163–178.

Katahira, K., Abla, D., Masuda, S., Okanoya, K., 2008. Feedback-based error monitoring processes during musical performance: an ERP study. Neurosci. Res. 61 (1), 120–128. https://doi.org/10.1016/j.neures.2008.02.001.

Kieffaber, P.D., Hershaw, J., Sredl, J., West, R., 2016. Electrophysiological correlates of error initiation and response correction. Neuroimage 128, 158–166. https://doi.org/10.1016/j.neuroimage.2015.12.046.

Kita, Y., Mori, A., Nara, M., 2001. Two types of movement-related cortical potentials preceding wrist extension in humans. Neuroreport 12 (10), 2221–2225. https://doi.org/10.1097/00001756-200107200-00035. PMID: 11447338.

Kleber, B., Friberg, A., Zeitouni, A., Zatorre, R., 2017. Experience-dependent modulation of right anterior insula and sensorimotor regions as a function of noise-masked auditory feedback in singers and nonsingers. Neuroimage 147, 97–110.

Kok, P., Rahnev, D., Jehee, J.F., Lau, H.C., De Lange, F.P., 2012. Attention reverses the effect of prediction in silencing sensory signals. Cereb. Cortex 22 (9), 2197–2206. https://doi.org/10.1093/cercor/bhr310.

Krakauer, J.W., Hadjiosif, A.M., Xu, J., Wong, A.L., Haith, A.M., 2019. Motor learning. Compr. Physiol. 9 (2), 613–663. https://doi.org/10.1002/cphy.c170043.

Krigolson, O.E., Holroyd, C.B., 2006. Evidence for hierarchical error processing in the human brain. Neuroscience 137 (1), 13–17. https://doi.org/10.1016/j.neuroscience.2005.10.064.

Krigolson, O.E., Holroyd, C.B., 2007. Hierarchical error processing: different errors, different systems. Brain Res. 1155, 70–80. https://doi.org/10.1016/j.brainres.2007.04.024.

Krigolson, O.E., Holroyd, C.B., Van Gyn, G., Heath, M., 2008. Electroencephalographic correlates of target and outcome errors. Exp. Brain Res. 190 (4), 401–411. https://doi.org/10.1007/s00221-008-1482-x.

Lahav, A., Saltzman, E., Schlaug, G., 2007. Action representation of sound: audiomotor recognition network while listening to newly acquired actions. J. Neurosci. 27 (2), 308–314. https://doi.org/10.1523/JNEUROSCI.4822-06.2007.

Lappe, C., Lappe, M., Keller, P.E., 2018. The influence of pitch feedback on learning of motor -timing and sequencing: a piano study with novices. PLoS One 13 (11), e0207462. https://doi.org/10.1371/journal.pone.0207462.

Loehr, J.D., Kourtis, D., Vesper, C., Sebanz, N., Knoblich, G., 2013. Monitoring individual and joint action outcomes in duet music performance. J. Cogn. Neurosci. 25 (7), 1049–1061. https://doi.org/10.1162/jocn_a_00388.

MacLean, S.J., Hassall, C.D., Ishigami, Y., Krigolson, O.E., Eskes, G.A., 2015. Using brain potentials to understand prism adaptation: the error-related negativity and the P300. Front. Hum. Neurosci. 9, 335. https://doi.org/10.3389/fnhum.2015.00335.

Maidhof, C., 2013. Error monitoring in musicians. Front. Hum. Neurosci. 7, 401. https://doi.org/10.3389/fnhum.2013.00401.

Maidhof, C., Rieger, M., Prinz, W., Koelsch, S., 2009. Nobody is perfect: ERP effects prior to performance errors in musicians indicate fast monitoring processes. PLoS One 4 (4), e5032. https://doi.org/10.1371/journal.pone.0005032.

Maidhof, C., Vavatzanidis, N., Prinz, W., Rieger, M., Koelsch, S., 2010. Processing expectancy violations during music performance and perception: an ERP study. J. Cogn. Neurosci. 22 (10), 2401–2413. https://doi.org/10.1162/jocn.2009.21332.

Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. J. Neurosci. Methods 164 (1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024.

Mathias, B., Tillmann, B., Palmer, C., 2016. Sensory, cognitive, and sensorimotor learning effects in recognition memory for music. J. Cogn. Neurosci. 28 (8), 1111–1126. https://doi.org/10.1162/jocn_a_00958.

Mauch, M., Cannam, C., Bittner, R., Fazekas, G., Salamon, J., Dai, J., et al., 2015. Computer-aided melody note transcription using the tony software: accuracy and efficiency. In: Proceedings of the First International Conference on Technologies for Music Notation and Representation (Paris).

Mauch, M., Dixon, S., 2014. PYIN: A fundamental frequency estimator using probabilistic threshold distributions. In: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (Florence), pp. 659–663. https://doi.org/10.1109/ICASSP.2014.6853678.

Nunes-Silva, M., Janzen, T.B., Rodrigues, R.G., da Luz, A.R., 2020. Sensory feedback in music performer–instrument interactions. Psychol. Music 49 (5), 1285–1302. https://doi.org/10.1177/0305735620928397.

Palidis, D.J., Cashaback, J.G.A., Gribble, P.L., 2019. Neural signatures of reward and sensory error feedback processing in motor learning. J. Neurophysiol. 121 (4), 1561–1574. https://doi.org/10.1152/jn.00792.2018.

Pezzulo, G., Rigoli, F., Friston, K.J., 2018. Hierarchical active inference: a theory of motivated control. Trends Cogn. Sci. (Regul. Ed.) 22 (4), 294–306. https://doi.org/10.1016/j.tics.2018.01.009.

Pfordresher, P.Q., Chow, K., 2019. A cost of musical training? Sensorimotor flexibility in musical sequence learning. Psychon. Bull. Rev. 26 (3), 967–973. https://doi.org/10.3758/s13423-018-1535-5.

Polich, J., 2007. Updating P300: an integrative theory of P3a and P3b. Clin. Neurophysiol. 118 (10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019.

Ruchsow, M., Spitzer, M., Grön, G., Grothe, J., Kiefer, M., 2005. Error processing and impulsiveness in normals: evidence from event-related potentials. Cogn. Brain Res. 24 (2), 317–325. https://doi.org/10.1016/j.cogbrainres.2005.02.003.

SanMiguel, I., Todd, J., Schröger, E., 2013. Sensory suppression effects to self-initiated sounds reflect the attenuation of the unspecific N1 component of the auditory ERP. Psychophysiology 50 (4), 334–343. https://doi.org/10.1111/psyp.12024.

Scheerer, N.E., Jones, J.A., 2014. The predictability of frequency-altered auditory feedback changes the weighting of feedback and feedforward input for speech motor control. Eur. J. Neurosci. 40 (12), 3793–3806. https://doi.org/10.1111/ejn.12734.

Scheerer, N.E., Jones, J.A., 2018. Detecting our own vocal errors: an event-related study of the thresholds for perceiving and compensating for vocal pitch errors. Neuropsychologia 114, 158–167. https://doi.org/10.1016/j.neuropsychologia.2017.12.007.

Schröger, E., Marzecová, A., SanMiguel, I., 2015. Attention and prediction in human audition: a lesson from cognitive psychophysiology. Eur. J. Neurosci. 41 (5), 641–664. https://doi.org/10.1111/ejn.12816.

Schürmann, M., Başar, E., 1994. Topography of alpha and theta oscillatory responses upon auditory and visual stimuli in humans. Biol. Cybern. 72 (2), 161–174. https://doi.org/10.1007/BF00205980.

Segado, M., Zatorre, R.J., Penhune, V.B., 2021. Effector-independent brain network for auditory-motor integration: fMRI evidence from singing and cello playing. Neuroimage 237, 118128. https://doi.org/10.1016/j.neuroimage.2021.118128.

Suchý, D., Behroozmand, R., Railo, H., 2025. Conscious and unconscious perception of pitch shifts in auditory feedback during vocalization: behavioral functions and event-related potential correlates. Neuroimage, 121269. https://doi.org/10.1016/j.neuroimage.2025.121269.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., Pernier, J., 1996. Stimulus specificity of phase-locked and non-phase-locked 40 hz visual responses in human. J. Neurosci. 16 (13), 4240–4249. https://doi.org/10.1523/JNEUROSCI.16-13-04240.1996.

Tan, H., Jenkinson, N., Brown, P., 2014. Dynamic neural correlates of motor error monitoring and adaptation during trial-to-trial learning. J. Neurosci. 34 (16), 5678–5688. https://doi.org/10.1523/JNEUROSCI.4739-13.2014.

Timm, J., SanMiguel, I., Keil, J., Schröger, E., Schönwiesner, M., 2014. Motor intention determines sensory attenuation of brain responses to self-initiated sounds. J. Cogn. Neurosci. 26 (7), 1481–1489. https://doi.org/10.1162/jocn_a_00552.

Timm, J., SanMiguel, I., Saupe, K., Schröger, E., 2013. The N1-suppression effect for self-initiated sounds is independent of attention. BMC Neurosci. 14 (1), 1–11. https://doi.org/10.1186/1471-2202-14-2.

Torrecillos, F., Alayrangues, J., Kilavik, B.E., Malfait, N., 2015. Distinct modulations in sensorimotor postmovement and foreperiod β-band activities related to error salience processing and sensorimotor adaptation. J. Neurosci. 35 (37), 12753–12765. https://doi.org/10.1523/JNEUROSCI.1090-15.2015.

Trujillo, L.T., Allen, J.J., 2007. Theta EEG dynamics of the error-related negativity. Clin. Neurophysiol. 118 (3), 645–668. https://doi.org/10.1016/j.clinph.2006.11.009.

Urbach, T.P., Kutas, M., 2006. Interpreting event-related brain potential (ERP) distributions: implications of baseline potentials and variability with application to amplitude normalization by vector scaling. Biol. Psychol. 72 (3), 333–343. https://doi.org/10.1016/j.biopsycho.2005.11.012.

Veen, V.V., Carter, C.S., 2002. The timing of action-monitoring processes in the anterior cingulate cortex. J. Cogn. Neurosci. 14 (4), 593–602. https://doi.org/10.1162/08989290260045837.

Warren, J.E., Wise, R.J., Warren, J.D., 2005. Sounds do-able: auditory-motor transformations and the posterior temporal plane. Trends Neurosci. 28 (12), 636–643. https://doi.org/10.1016/j.tins.2005.09.010.

Wolpert, D.M., Diedrichsen, J., Flanagan, J.R., 2011. Principles of sensorimotor learning. Nat. Rev. Neurosci. 12 (12), 739–751. https://doi.org/10.1038/nrn3112.

Zatorre, R., Chen, J., Penhune, V., 2007. When the brain plays music: auditory–motor interactions in music perception and production. Nat. Rev. Neurosci. 8, 547–558. https://doi.org/10.1038/nrn2152.