



# Neural correlates of accelerated auditory processing in children engaged in music training



Assal Habibi<sup>a,\*</sup>, B. Rael Cahn<sup>a,b</sup>, Antonio Damasio<sup>a</sup>, Hanna Damasio<sup>a</sup>

<sup>a</sup> Brain and Creativity Institute, Dornsife College of Letters, Arts and Sciences, University of Southern California, Los Angeles, CA, United States

<sup>b</sup> Department of Psychiatry, Keck School of Medicine, University of Southern California, Los Angeles, CA, United States

## ARTICLE INFO

### Article history:

Received 23 November 2015  
Received in revised form 7 April 2016  
Accepted 8 April 2016  
Available online 16 April 2016

### Keywords:

Auditory evoked potentials  
Electroencephalography  
Pitch perception  
Musical training  
Child development  
Brain plasticity  
El Sistema  
Communication

## ABSTRACT

Several studies comparing adult musicians and non-musicians have shown that music training is associated with brain differences. It is unknown, however, whether these differences result from lengthy musical training, from pre-existing biological traits, or from social factors favoring musicality. As part of an ongoing 5-year longitudinal study, we investigated the effects of a music training program on the auditory development of children, over the course of two years, beginning at age 6–7. The training was group-based and inspired by El-Sistema. We compared the children in the music group with two comparison groups of children of the same socio-economic background, one involved in sports training, another not involved in any systematic training. Prior to participating, children who began training in music did not differ from those in the comparison groups in any of the assessed measures. After two years, we now observe that children in the music group, but not in the two comparison groups, show an enhanced ability to detect changes in tonal environment and an accelerated maturity of auditory processing as measured by cortical auditory evoked potentials to musical notes. Our results suggest that music training may result in stimulus specific brain changes in school aged children.

© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

### 1.1. Background

Over the past two decades, numerous studies have reported differences in the brain and behavior of musicians when compared to non-musicians (for comprehensive reviews see: Gaser and Schlaug, 2003; Herholz and Zatorre, 2012; Levitin, 2012; Pantev and Herholz, 2011; Strait and Kraus, 2014). Music training has been found to be positively associated with superior performance on a variety of auditory tasks, including frequency discrimination (Schellenberg and Moreno, 2010), perception of pitch in spoken language (Schön et al., 2004; Wong et al., 2007), detection of minor changes of pitch in familiar (Schellenberg and Moreno, 2010) and unfamiliar melodies (Habibi et al., 2013), identification of a familiar melody when it is played at a fast or slow tempo (Andrews et al., 1998; Dowling et al., 2008) and recognition of whether a sequence of chords ends correctly based on Western classical music

rules (Koelsch et al., 2007). Compared to non-musicians, musicians also tend to show enhanced language skills including phonological awareness (Degé and Schwarzer, 2011; Moreno et al., 2009, 2011), vocabulary (Forgeard et al., 2008; Piro and Ortiz, 2009) and verbal memory (Franklin et al., 2008; Jakobson et al., 2008a,b).

Differences between musicians and non-musicians in neural structure and function have also been demonstrated, in particular in auditory (Bangert and Schlaug, 2006; Gaser and Schlaug, 2003; Herholz and Zatorre, 2012; Jäncke, 2009; Schneider et al., 2002; Tillmann et al., 2003; Zatorre, 2005) and sensorimotor areas (Gaser and Schlaug, 2003; Jäncke, 2009; Schlaug, 2001), as measured by magnetic resonance imaging (MRI), and in enhancement of auditory evoked potentials measured by electroencephalography (EEG) (Brattico et al., 2006; Fujioka et al., 2005; Musacchia et al., 2007; Shahin et al., 2004).

In spite of a growing interest in the benefits of music training and in the brain differences of musicians compared to non-musicians within the central auditory system, the interpretation of such findings remains unclear. The differences reported in cross-sectional studies, which mostly employ quasi-experimental designs, might be due to long-term regular and intensive training or might result, either partly or primarily, from pre-existing biological and genetic factors that predispose individuals to develop musical aptitude if exposed to music during a sensitive period of development. An

\* Corresponding author at: Brain and Creativity Institute, University of Southern California, 3620A McClintock Avenue, Suite 150, Los Angeles, CA 90089-2921, United States.

E-mail address: [ahabibi@usc.edu](mailto:ahabibi@usc.edu) (A. Habibi).

appropriate way toward disentangling the effects of predisposing factors from the effects of musical training, involves the longitudinal study of groups of children, of the same age with and without musical training, beginning prior to the onset of their music training. Ideally the comparison group should involve non-musical but comparably socially-interactive training, such as athletics programs.

### 1.2. Electrophysiological response to musical stimuli

Event related potentials (ERPs) have been extensively used to assess the development of the central auditory system with its developmental related natural anatomical and physiological changes during childhood (Wunderlich and Cone-Wesson, 2006). ERPs are averages of the EEG signal, time-locked to repeated stimuli that allow for the identification of sensory and cognitive processing steps in response to auditory stimuli. Given their excellent temporal resolution, ERPs provide a robust means to measure the maturation of the auditory pathway through changes in latency, amplitude and topography (Luck, 2014).

The cortical P1 component dominates the ERP response to auditory stimuli in early childhood and has a latency of approximately 100 ms; it matures during development and reaches an adult latency of 40–60 ms with a bilateral frontal-positive scalp topography by around age 18–20; it originates from the lateral portion of Heschl's gyrus (Ponton et al., 2000; Sharma et al., 1997; Wunderlich et al., 2006). The cortical P1 is followed by the vertex negative N1, which has a latency of approximately 90–110 ms, in adults, and is generated within primary and secondary auditory cortices (Näätänen and Picton, 1987). Development of the central auditory pathway is accompanied by a decrease in the amplitude and latency of the P1 component and corresponding increase in the amplitude of the N1 component, a process that is completed by young adulthood (Ponton et al., 2000; Shahin et al., 2010; Sharma et al., 1997; Tierney et al., 2015; Wunderlich and Cone-Wesson, 2006).

In relation to the impact of musical training, adult musicians have been shown to have enhanced auditory N1 amplitude (Shahin et al., 2003). The magnetic counterpart N1m has also been reported to be larger in musicians compared to non-musicians, as evoked by musical stimuli (Pantev et al., 1998). In addition, relative to non-musicians, musicians display larger mismatch negativity (MMN) in response to changes of chords, melody and rhythm (Brattico et al., 2006; Koelsch et al., 1999; Vuust et al., 2005). Recent studies have shown accelerated maturation of the cortical auditory response (P1 and N1) in high school students who underwent three years of school-based music training (Tierney et al., 2015), and enhancement of MMN in school-age children involved in music training (Chobert et al., 2014; Putkinen et al., 2014; Virtala et al., 2012).

The P2 peak has an adult latency of approximately 200 ms (it varies between about 150 and 275 ms) after the onset of an auditory stimulus. It is generated in associative auditory temporal regions with additional contributions from frontal areas (Bishop et al., 2011; Tremblay et al., 2001). Traditionally, the P2 was considered to be an automatic response, modulated only by the stimulus; but it has been shown that its latency and amplitude are sensitive to learning and attentional processes (Lappe et al., 2011). Specifically, P2 amplitude accompanying the processing of music has been reported to be larger in adult musicians compared to non-musicians (Pantev et al., 2001; Shahin et al., 2003).

The P3 is a positive potential that appears in response to target stimuli or rare, unexpected stimuli presented among standard stimuli. It reflects context updating and the orienting of attention. It has a peak latency between 250 and 700 ms and is maximally distributed at fronto-central or parietal areas of the scalp depending on the type of eliciting stimulus (Donchin and Coles,

1988; Picton, 1992). The P3 is comprised of two contributing subcomponents—the P3a and the P3b. The auditory P3a typically has a peak latency within ~300 ms, is generated primarily by the anterior cingulate cortex and displays a fronto-central distribution on the scalp. The P3a is related to the automatic orienting of attention as occurs in paradigms in which distracting stimuli engage attention without any required behavioral response. The P3b usually peaks later than 300 ms, often at 300–500 ms or later, and is generated primarily by medial temporal areas and the temporo-parietal junction, thus displaying a parietal distribution on the scalp. The P3b is elicited by stimuli that require a behavioral response or clearly match with a target stimulus template held in working memory (Polich, 2007). Larger amplitude N2b-P3 potentials in response to deviations in melody and rhythm have been reported in individuals with music training (Habibi et al., 2013; Nikjeh et al., 2008; Seppänen et al., 2012; Tervaniemi et al., 2005; Trainor, 2012).

### 1.3. Design, aims and hypotheses

Given the above knowledge on the effects of music training in the brain of adults, we used auditory evoked potentials and behavioral tasks to investigate whether the development of the auditory system was sensitive to musical training in 6–7 year old children and, if so, determine which components were affected. To address the issues related to pre-existing differences between musicians and non-musicians, we employed a longitudinal design comparing children involved with music training with two age-matched comparison groups without music involvement but with the same socio-economic and cultural background (Habibi et al., 2014).

We assessed all participants at baseline and then again two years later. We compared auditory evoked potentials elicited by violin, piano and pure tones at the two time points. Also, at year 2, using a same-different judgement design, we assessed the participants' ability to detect changes in tonal or rhythmic content of unfamiliar melodies and the associated brain processing. We hypothesized that:

1. Music trained children would show accelerated development of the P1-N1 complex. The rationale for this hypothesis is as follows: the cortical N1 component emerges between 8 and 10 years of age, while the P1 amplitude decreases; furthermore, enhanced N1 amplitude has been associated with music training in adults and adolescents, therefore, we predicted the development of the N1 would be associated with music training in children as well.
2. Music trained children would show a heightened ability to detect changes in pitch and rhythm, and would show enhanced amplitude of correlated P3 auditory cortical potentials.

## 2. Materials and methods

### 2.1. Subjects

Fifty children were recruited from public elementary schools and community music and sports programs in the greater Los Angeles area. Between induction and baseline assessment, 5 enrolled participants discontinued their participation, in their respective program or relocated. Between baseline assessment and evaluation two years later, 8 participants (2 music, 4 sports-comparison & 2 non-sports comparisons) discontinued their participation, in their respective program or the research study, or relocated and thus were not included in the final analysis (Fig. 1). Thirty-seven remaining participants formed the following three groups: Thirteen children (6 girls and 8 boys, mean age at baseline assessment = 6.68 yrs., SD = 0.6) who were beginning their participation

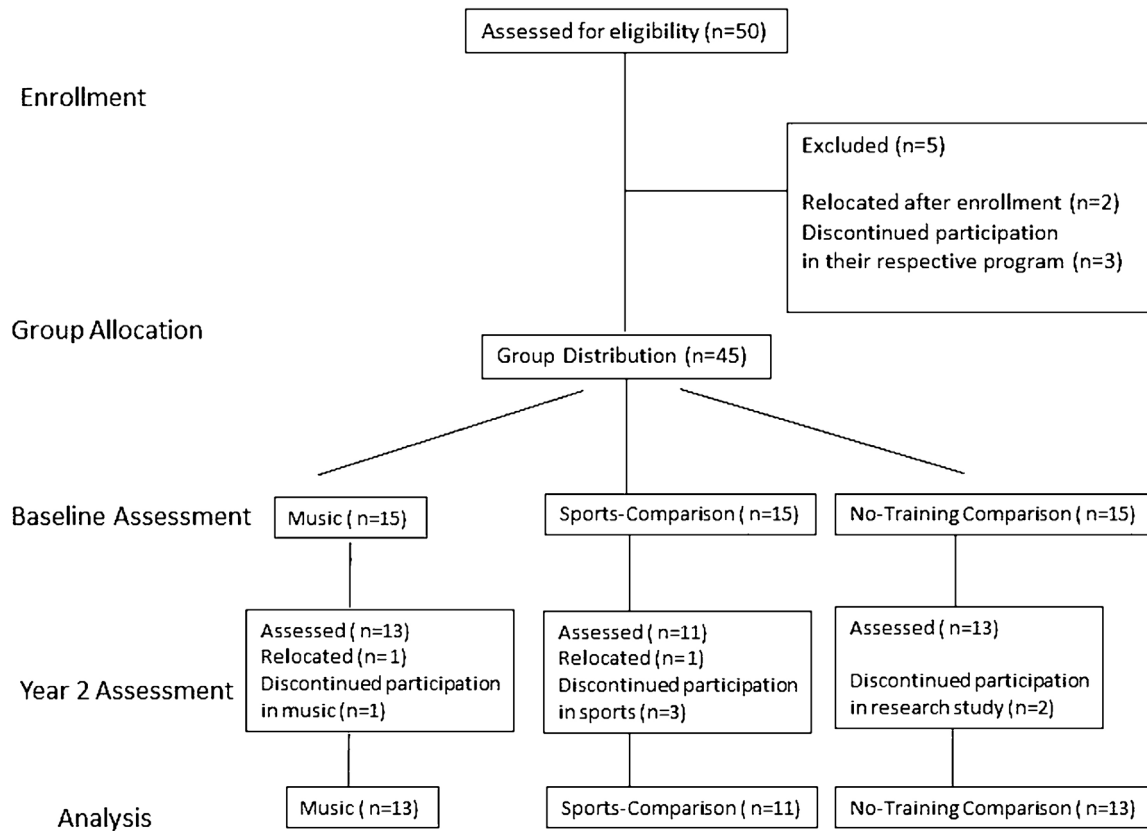


Fig. 1. Diagram depicting numbers of participants recruited, enrolled and lost to follow up.

in the Youth Orchestra of Los Angeles at Heart of Los Angeles program (hereafter called “music group”). The program is based on the Venezuelan system of musical training known as El Sistema and offers free music instruction 6–7 h weekly to children from underprivileged areas of Los Angeles. The program emphasizes ensemble practice and group performances. Children enrolled in this program are selected, by lottery, up to a maximum of 20 per year, from a list of interested families. Eleven children (4 girls and 7 boys, mean age at baseline assessment = 7.15 yrs.,  $SD = 0.62$ ) formed the first comparison group (hereafter called “sports group”) who were beginning training in a community-based soccer program and were not engaged in any musical training. This program accepts all children whose parents want to enroll them in the program. The children enrolled in the study (to match the music target group) were those that first showed interest in participating in our longitudinal study. The soccer program offers free soccer training (3 times a week for two hours each, with an additional one hour game each weekend) to children ages 6 and older. In addition, thirteen children (2 girls and 11 boys, mean age at baseline = 7.16 yrs.,  $SD = 0.52$ ) formed the second comparison group (hereafter called “no-training group”). Children in this second comparison group were recruited from public schools in the same area of Los Angeles provided they were not involved in any systematic and intense after-school program. All three cohorts came from equally underprivileged minority communities and included primarily Latino and one Korean family, in downtown Los Angeles. All children were raised in bilingual households, but all attended English speaking schools and spoke fluent English as revealed by normal performance on the verbal components of the Wechsler Abbreviated Scale of Intelligence (WASI II). Exclusion criteria included any history of psychiatric or neurologic disease in the children. At both assessment times, participants were screened by interview with their parents to ensure that they did not have any diagnosis of

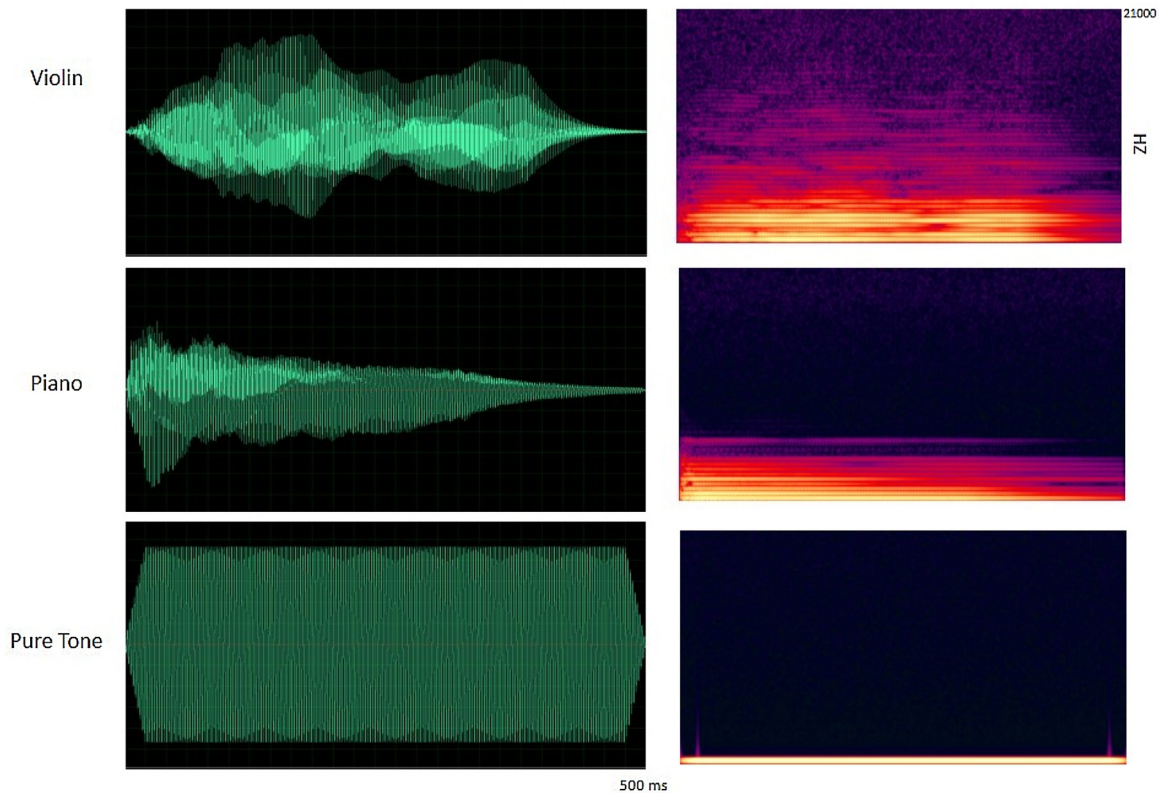
developmental or neurological disorder and were tested with Wechsler Abbreviated Scale of Intelligence (WASI-II) to assure equal and normal cognitive development across the three groups (Habibi et al., 2014).

## 2.2. Socio-economic status (SES)

Parents indicated their highest level of education and annual household income on a questionnaire. Responses to education level were scored on a 5-point scale: (1) Elementary/Middle school; (2) High school; (3) College education; (4) Master’s degree (MA, MS, MBA); (5) Professional degree (PhD, MD, JD). Responses to annual household income were scored on a 5-point scale: (0) <\$ 10,000 (1) \$10,000–\$19,999 (2) \$20,000–29,999 (3) \$30,000–39,999 (4) \$40,000–49,999 (5) >\$50,000. A final socio-economic status (SES) score was calculated as the mean of each parent’s education score and annual income.

## 2.3. Tonal perception task (passive task)

Participants were presented with violin tones, piano tones (A4 and C4, American notation) and pure tones matched in fundamental frequency to the musical tones. Tones were 500 ms in duration and were presented with an interstimulus interval of 2500 ms (offset to onset). A passive listening protocol was followed in which the children watched a silent movie while the tones were presented. All the tones were computer-generated, created in MIDI format, using Finale Version 3.5.1 (Coda Music), and were then converted to WAV files with a “Grand Piano” and “Violin” sound font, using MidiSyn Version 1.9 (Future Algorithms). Pure tones were created with a cosine envelope in Matlab and matched to the fundamental frequency of the musical tones (Fig. 2). Stimuli and paradigm design were adapted from (Shahin et al., 2004). Tones



**Fig. 2.** The temporal profile and spectrograms of violin, piano and pure tones matched in fundamental frequency (A4).

were grouped to create 9 min runs. Each run consisted of six tone classes; each was presented 60 times for a total of 360 tones in a random order. The experimental session consisted of 4 runs (lasted approximately 40 min) at baseline and 3 runs (lasted approximately 30 min) at year 2. The number of runs at year two was adjusted since children were older and had less movement artifacts and because of time limitations due to adding the active pitch discrimination task.

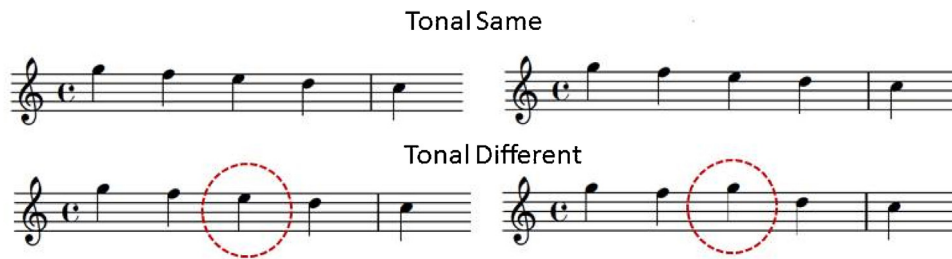
All auditory stimuli were delivered binaurally via ER-3 insert earphones (Etymotic Research) at 70 dB sound pressure level. Breaks were given in between runs when necessary. Ongoing EEG was continuously monitored for evidence of sleep or drowsiness and if either occurred, the recording and stimulus train were paused and the subject awoken and/or given the opportunity for a break before continuing.

#### 2.4. Tonal/rhythm discrimination task (active task)

This task required a same/different judgement of two short musical phrases indicated by a button press response. Five distinct pitches, corresponding to the first 5 notes of the C major scale (fundamental frequencies 261, 293, 329, 349 and 392 Hz) were used to create 24 pairs of melodies divided equally into four conditions. In the tonal category, melody pairs were presented at a steady beat (120 bpm) with tone durations of 500 ms each with varying pitches to create melodic contour. Each melody of the phrase lasted 2500 ms with a 2000 ms pause in between successive melodies. The second melody of the phrase was either identical to the first melody resulting in “*tonal same*” condition or contained a single note that was different in pitch compared to the first melody creating “*tonal different*” condition (Fig. 3). The change was restricted to the first 5 notes of C major scale. In the rhythm condition, phrases were presented with a single pitch (restricted to the first 5 notes of C major scale but was varied between trials); the duration of each

note ranged from 125 ms to 1500 ms to create rhythmic patterns. Similar to the tonal condition, the second melody of the phrase was either a duplicate of the first melody, “*rhythm same*” condition or contained a deviation resulting in a “*rhythm different*” condition. The rhythm different condition was created by changing the duration values of two adjacent notes to alter the rhythmic grouping by temporal proximity, while retaining the same meter and total number of notes. Specifically, this was accomplished by changing two eighth notes (each 500 ms long) to a dotted eighth note (750 ms long) and a sixteenth note (250 ms long). All the notes were computer-generated, created in MIDI format, using Finale Version 3.5.1 (Coda Music), and were then converted to WAV files with a marimba-like timbre to avoid any potential familiarity with the instrumental sounds. A 500 ms fixation point on a screen preceded the first note of each phrase pair to cue participants to the beginning of the stimulus and it remained on screen during stimulus presentation; following the end of the second melody, a response screen immediately replaced the fixation point and participants were given up to 3 s to make a same/different judgement. In order to ensure precise time-locking for the analysis of the data relative to the presentation of each individual note, a marker was sent by the stimulus presentation software (Matlab, Mathworks, 2012) to the EEG amplifier over the trigger channel at the onset of each melody.

Trials were grouped to create short, child appropriate runs of 8 min. Each run consisted of 48 trials – 24 melodies (6 tonal same, 6 tonal different, 6 rhythmic same and 6 rhythmic different) each repeated twice – and presented in a randomized order. The full experiment consisted of 5 runs of 8 min each. Prior to the onset of the paradigm, a practice session was provided to each subject with feedback (“Correct” or “Incorrect”) after each same/different categorization response; in the subsequent experimental session no such feedback was given. Behavioral responses were collected in addition to EEG.



**Fig. 3.** Example of melodic stimuli for pitch discrimination task, in same and different conditions.

## 2.5. Experimental procedures

Study protocols were approved by the University of Southern California Institutional Review Board. Informed consent was obtained in writing, from the parents/guardians in the preferred language, on behalf of the child participants and verbal assent was obtained from all children individually. Either the guardians or the children could end their participation at any time. Participants (parents/guardians) received monetary compensation (\$15 per hour) for their child's participation and children were awarded small prizes (e.g. toys or stickers).

All children were tested individually at our laboratory at the Brain and Creativity Institute at the University of Southern California. They were tested two times with the tonal perception task; once at the start of their participation in the longitudinal study which, for the music and sport-comparison groups coincided with the beginning of their participation in the music or sports program (baseline). They were tested again at two years ("year 2") after the initial baseline assessment. Each participant was tested only once with the tonal/rhythm discrimination task at year 2. The reason for not including the discrimination task at baseline was that the attentional demands posed by this test were deemed too difficult for the participants at age 6–7. Due to the attentional demanding nature of the pitch-discrimination task, this task was presented prior to the tonal perception task for all participants at the year 2 assessment. These tasks are part of a comprehensive battery of evaluations in an ongoing longitudinal study on the effects of music training on brain, cognitive and social development (Habibi et al., 2014).

## 2.6. EEG recording

A 64-channel Neuroscan Synamps2 recording system was used to collect electrophysiological data. Subjects were seated alone in a comfortable chair in a dark, quiet (acoustically and electrically shielded) testing room. Electrode placements included the standard 10–20 locations and intermediate sites but only 32 channels (FP1, F3, FC3, C3, CP3, P3, F7, FT7, T7, TP7, P7, FPz, Fz, FCz, Cz, CPz, Pz, FP2, F4, FP4, C4, CP4, P4, F8, FT8, T8, TP8, O1, Oz, O2, M1 & M2) were used for data collection, related to time demands and for the convenience of the subjects. Impedances were kept below 10 k $\Omega$ . Lateral and vertical eye movements were monitored using two bipolar electrodes on the left and right outer canthi and two bipolar electrodes above and below the right eye to define the horizontal and vertical electro-oculogram (EOG). Signals were digitized at 1000 Hz, amplified by a factor of 2010, and band-pass filtered on-line with cutoffs at 0.05 and 200 Hz. Eye movement effects on scalp potentials were removed offline in the continuous recording from each subject using a singular value decomposition-based spatial filter utilizing principal component analysis of averaged eye blinks for each subject (Ille et al., 2002).

Results reported below include auditory event related potential amplitude, latency and scalp topography to perception of piano tones in the tonal perception task at baseline and at year 2;

**Table 1**

Time windows for ERP quantification separately for each stimulus condition and each group for a. passive pitch perception and b. active pitch discrimination task.

a			
Stimulus Category	P1 (ms)	N1(ms)	P2(ms)
Baseline Assessment	100–130 (M & S) 95–115 (NT)	–	180–210
Year two	70–90 (M) 80–100 (S & NT)	105–125	160–190
b			
Stimulus Category	P2(ms)	N2(ms)	P3(ms)
Pitch Same	130–180 (M) 130–180 (S) 130–180 (NT)	220–300 (M) 220–290 (S) 200–270 (NT)	300–370 (M) 310–380 (S) 300–380 (NT)
Pitch Different	120–160 (M) 130–190 (S) 140–180 (NT)	220–260 (M) 220–290 (S) 200–270 (NT)	300–370 (M) 310–380 (S) 300–380 (NT)

behavioral accuracy on the discrimination task and detection of tonal changes in the discrimination task at year 2. The results regarding the effects found in response to the rhythm different trials and perception of violin and pure tones showed no significant differences in this population and will be reported separately after collecting data from further subjects.

## 2.7. Data analysis

Analyses were carried out with Brainvision Analyzer 2.0 and Matlab R2013a. For both passive and active tasks, continuous EEG data were divided into epochs starting 200 ms before and ending 500 ms after the onset of each stimulus according to the stimulus type. Channel P7 was removed from all participants' data due to excessive noise in the majority of subjects. Additional noisy channels were removed manually with no more than 4 channels removed from any individual subject's data. Epochs were average referenced (excluding EOG and other removed channels), baseline corrected (–200–0 ms prior to each note) and offline digitally filtered (band-pass 1–20 Hz). Epochs with a signal change exceeding  $\pm 150$  microvolt at any EEG electrode were rejected and not included in the averages. A one way ANOVA comparing number of removed channels between groups was non-significant for baseline and year 2 assessments: baseline (M  $\pm$  SD): no-training =  $1.76 \pm 1.16$ , music =  $1.61 \pm 0.86$ , sports =  $1.9 \pm 0.94$ ,  $p = 0.77$ ; passive tonal perception year 2: no-training =  $1.92 \pm 1.18$ , music =  $2.23 \pm 1.09$ , sports =  $1.9 \pm 1$ ,  $p = 0.71$ ; A one way ANOVA comparing number of trials between groups was non-significant for any of the conditions; passive tonal perception at baseline (M  $\pm$  SD): no-training =  $184 \pm 37$ , music =  $190 \pm 23$ , sports =  $197 \pm 23$ ,  $p = 0.53$ ; passive tonal perception year 2: no-training =  $137 \pm 29$ , music =  $147 \pm 38$ , sports =  $139 \pm 32$ ,  $p = 0.69$ ; active tonal discrimination task at year 2, tonal same condition: no-training =  $37 \pm 9$ , music =  $42 \pm 8$ , sports =  $38 \pm 7$ ,  $p = 0.25$ ;

and tonal different condition: no-training =  $36 \pm 8$ , music =  $41 \pm 10$ , sports =  $38 \pm 7$ ,  $p = 0.28$ .

For the passive tonal perception task, accepted trials were averaged according to stimulus type (piano) collapsing over C4 and A4 tones to increase the signal to noise ratio. We quantified the mean voltage of the ERPs for each stimulus category from the following electrodes, based on the scalp distribution for the peaks observed in the grand-averaged data for the population: (F3, Fz, F4) for P1; global field power (GFP) for P1/N1 ratio (relative percent amplitude difference between P1 and N1) and (F3, Fz, F4, FC3, FCz, FC4, C3, Cz, C4, CP3, CPz, CP4) for P2 in time-windows centered on the peak of the respective component in the grand average waveform. The parameters of the time-windows for peak measurements in individuals for each component are summarized in Table 1. These time-windows were chosen based on previous findings (Shahin et al., 2004) as well as the observed peak amplitude and latency of the grand average waveforms in this dataset. Peak latency for each component was measured at the peak of the GFP waveform for the same time ranges. The peak latency was marked automatically and inspected visually for accuracy. For the quantitation of the N1 incidence at year 2 in the passive task paradigm, each subject's data was visually inspected for the occurrence of a frontocentral negativity in the N1 time range, and all subjects with such an observable negative-going peak were counted as having an N1.

For the active tonal discrimination task at year 2, behavioral data from each subject were recorded and analyzed in terms of correct detection of same and different conditions; however, due to the fact that there were not enough trials for stable ERP quantitation of correct trials only in many participants, ERP averaging was performed without regard to whether the subject made correct or incorrect responses. ERPs were quantified for each subject in response to the tonal same versus tonal different category collapsed over the 6 melodies. The tonal different note refers to the pitch-changed note in the comparison melody and tonal same refers to the same note in the comparison melody for the trials wherein the target and comparison melodies were identical. We quantified the mean voltage of the ERPs for each stimulus category from (F3, Fz, F4) for P1 and (FC3, FCz, FC4) for N1, P2, N2 and P3 in time-windows centered on the peak of the respective component in the grand average waveform. We also conducted a secondary analysis for P3 amplitude on the frontal-central site FCz alone as the amplitude of the P3 was greatest at this site in the grand average waveform. The parameters of the time-windows for peak measurements in individuals are summarized in Table 1 and were chosen for analysis based on previous findings as well as the observed peak amplitude and latency of the grand average waveforms and scalp maps in this dataset.

## 2.8. Statistical analysis

For the tonal perception task, the mean amplitudes of the ERP components of interest (P1, P2) were compared using analysis of variance (ANOVA) with Group (no-training, sports, music), as between-group factor and electrode laterality (left, midline, right) and, for P2, frontality (frontal, frontocentral, central, centroparietal) as within-group factors for both the baseline data and the year 2 data separately. So as to assess the change in P1 and P2 amplitude across time, in addition, a repeated measures ANOVA analysis was conducted for both P1 and P2 with the inclusion of year of assessment (baseline, year two) as an additional within-groups factor. Lastly, P1/N1 ratio amplitude at year 2 was compared using analysis of variance (ANOVA) with Group (no-training, sports, music), as between-group factor at baseline N1 was not apparent so analysis of this component was not done.

For the pitch discrimination task, mean amplitudes of the ERP components of interest (P1, N1, P2, N2, P3) were compared using  $3 \times 3$  analysis of variance (ANOVA) with Group (no-training,

sports, music), as between-group factor, and electrode laterality (left, midline, right) as within-group factors for two pitch categories of same and different. In all statistical analyses, type I errors were reduced by decreasing the degrees of freedom with the Greenhouse-Geisser epsilon (the original degrees of freedom for all analyses are reposted throughout the paper). *Post-hoc* tests were conducted using Tukey post-hoc statistical comparisons.

## 3. Results

### 3.1. Demographics

Analysis revealed no significant differences in sex  $\chi^2$  (2,  $N = 37$ ) = 1.98,  $p = 0.37$ , age  $F$  (2, 34) = 2.84,  $p = 0.08$  and socio-economic status  $F$  (2, 3) = 1.36,  $p = 0.68$  [(M  $\pm$  SD): no-training =  $1.74 \pm 0.95$ , music =  $1.87 \pm 0.58$ , sports =  $1.69 \pm 0.52$ ] and IQ  $F$  (2, 34) = 2.57,  $p = 0.1$  [(M  $\pm$  SD): no-training =  $93.9 \pm 10.4$ , music =  $103.6 \pm 13.3$ , sports =  $100.6 \pm 8.6$ ] between the three groups and therefore these factors were not included in subsequent analysis.

### 3.2. Behavioral response to pitch same versus pitch difference notes

Participants in the music group were more accurate in detecting pitch changes in melodies compared to the two comparison groups (for pitch different trials music group accuracy  $M \pm SD = 62 \pm 15.9\%$ ; sports group accuracy =  $30.4 \pm 14.2\%$  and no-training group accuracy =  $45.1 \pm 26.2\%$ ; [ $F$  (2, 34) = 7.99,  $p = 0.001$ , partial eta squared = 0.31]). *Post-hoc* analysis showed significant differences between music and sports group ( $p = 0.001$ ) and there was a strong trend between music and no-training group ( $p = 0.07$ ). In trials wherein the target and comparison melodies were identical, there was no significant difference in accuracy between the three groups (music group =  $84.4 \pm 10.2\%$ ; sports group =  $78.1 \pm 16.7\%$  and no-training group =  $78 \pm 19.65\%$ ; [ $F$  (2, 34) = 0.66,  $p = 0.51$ ]), indicating that the improved performance in the music group on this task was related to improved abilities to detect change and that errors of commission were relatively rare across all groups.

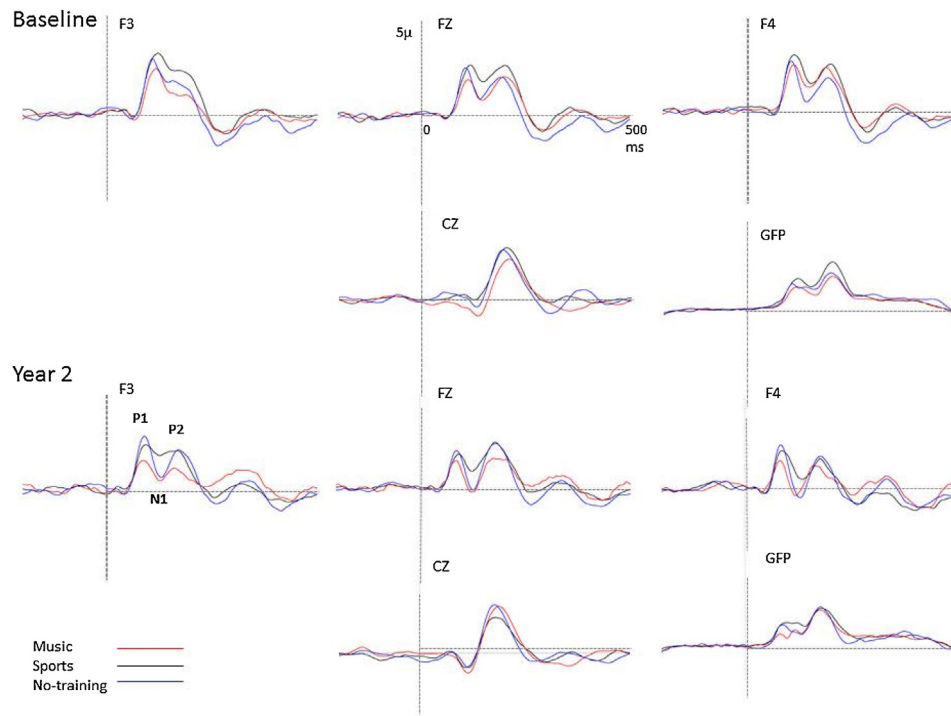
### 3.3. Event related potentials (ERPs) in tonal perception task

Fig. 4 shows auditory event related potentials elicited by piano tones for the three groups at baseline and at year two. At baseline, only P1 and P2 responses were reliably elicited for all three groups as an N1 potential was not yet present at this age, as has been previously reported (Sharma et al., 1997).

At baseline there was no significant difference in the amplitude of P1 [ $F$  (2, 34) = 0.90,  $p = 0.41$ ] or of P2 [ $F$  (2, 34) = 1.15,  $p = 0.32$ ] between the music group and the two comparison groups. There was no significant difference in laterality or frontality of P1 or P2 between the groups at baseline. Furthermore, the latencies of the P1 and P2 components elicited by the piano notes were not significantly different between the music and the two comparison groups P1, [ $F$  (2, 34) = 2.02,  $p = 0.14$ ]; P2, [ $F$  (2, 34) = 1.23,  $p = 0.3$ ].

Repeated measures ANOVA analysis including baseline and year 2 data, indicated that P1 amplitude decreased significantly for all three groups, main effect of year [ $F$  (1, 34) = 10.03,  $p = 0.003$ , partial eta squared = 0.22]. The music group showed the largest decrease from baseline to year 2, which was reflected in significant group  $\times$  year interaction [ $F$  (2, 34) = 3.01,  $p = 0.05$ , partial eta squared = 0.15]. *Post-hoc* contrasts were significant for the change in the music group P1 amplitude from baseline to year 2 ( $p = 0.02$ ) but not for the sports ( $p = 0.22$ ) or no-training group ( $p = 0.99$ ).

ANOVA analysis of year 2 data further indicated that the amplitude of P1 was smallest in the music group compared



**Fig. 4.** Event related potentials elicited by piano tones for the three groups at baseline and at year two.

to the two comparison groups, main effect of group [ $F(2, 34) = 6.46, p = 0.004$ , partial eta squared = 0.27]. Post-hoc analysis showed a significant difference between music and no-training group ( $p = 0.03$ ) and strong trend level difference between music and sports group ( $p = 0.07$ ). Although the amplitude was smallest in the midline electrode (Fz) as evidence by a significant main effect of laterality  $F(2, 68) = 5.64, p = 0.006$ , partial eta squared = 0.13, the group by laterality interaction was not significant ( $p = 0.25$ ).

The N1 component could only be reliably measured at year 2 and, even so, not for all participants. We examined individual ERP traces across each subject for reliable evidence of the N1 component, defined both as clear central negativity on scalp topography in the  $\sim 100$  ms time range and a clear and separate component at Cz and global field power waveforms between the P1 and P2 components. This analysis indicated that 76% of participants in the music group showed an N1 peak at year 2, whereas an N1 component was observed in only 46% of the participants in the sports group and 53% of participants in the no-training group. Although the incidence of N1 was higher in the music group, the difference was not significant between the music and the two comparison groups ( $\chi^2(2, N = 37) = 1.85, p = 0.39$ ). To further elucidate the relationship between the P1/N1 complex we measured the amplitude of the P1 and N1 components in the global field power waveform. We then calculated the ratio of relative percent amplitude difference between P1 and N1 amplitudes in all three groups using the formula  $\{[(P1 - N1)/P1] \times 100\}$ . The relationship between P1 and N1 differed significantly between the three groups,  $M \pm SD$  for no-training group  $+17.3 \pm 31.6\%$ , sports group  $-8 \pm 35.8\%$ , music group  $-23 \pm 43.6\%$ ; [ $F(2, 34) = 3.88, p = 0.03$ , partial eta squared = 0.18]. Post-hoc contrasts were significant between music versus no-training group ( $p = 0.02$ ) but not music versus sports group ( $p = 0.62$ ) nor sports versus no-training group ( $p = 0.2$ ).

Repeated measures ANOVA analysis of the P2 data across baseline and year 2 revealed that the amplitude of the P2 did not change significantly over time, as evidenced by a non-significant

main effect of year [ $F(1, 34) = 0.04, p = 0.83$ ] and no significant group  $\times$  year interaction [ $F(2, 34) = 1.19, p = 0.31$ ]. The interaction of Group  $\times$  Year  $\times$  Laterality reached significance [ $F(4, 68) = 4.2747, p = 0.003$ ], however subsequent post-hoc contrasts between groups at these sites were not significant reflecting that a subtle change in scalp topography between groups was driving the significance of this interaction.

ANOVA analysis of year 2 data revealed that P2 amplitude was not different between groups at year 2, main effects of Group [ $F(2, 34) = 0.56, p = 0.57$ ]. A significant Group  $\times$  Laterality interaction [ $F(4, 68) = 2.7, p = 0.02$ ] implied possible differences between groups. Subsequent post-hoc contrasts between groups were not significant for P2 amplitude at any electrode site between groups; but instead a relatively greater right sided involvement in the topography of the P2 in the music group—the significant interaction was driven by larger amplitude P2 in the mid line versus left side electrodes for the music group ( $p = 0.04$ ) in contrast to larger amplitude P2 in the mid line electrodes versus right side for no-training group ( $p = 0.009$ ). Fig. 5 shows topography maps of each potential, P1, N1 and P2, separately for each group at year two.

#### 3.4. Event related potential latency in tonal perception task

The peak latency of P1 component, as measured at the peak of the P1 component in the GFP waveform, was equivalent at baseline across groups [ $F(2, 34) = 2.2, p = 0.14$ ]. However at year 2, there was a difference between groups as evidenced by significance of  $1 \times 3$  ANOVA of the P1 GFP peak latency [ $F(2, 34) = 4.25, p = 0.02$ , partial eta squared = 0.2]. Post-hoc contrasts were significant between the music and sports group ( $p = 0.01$ ), whereas the contrast between the music and no-training groups ( $p = 0.36$ ) and between sports and no-training groups ( $p = 0.25$ ) were not significant. There was no difference between the three groups with respect to the latency of P2 at baseline or year 2 and no difference between groups for the latency of N1 at year 2. Table 2 includes peak latencies for all components at baseline and year two measurements for the three groups.

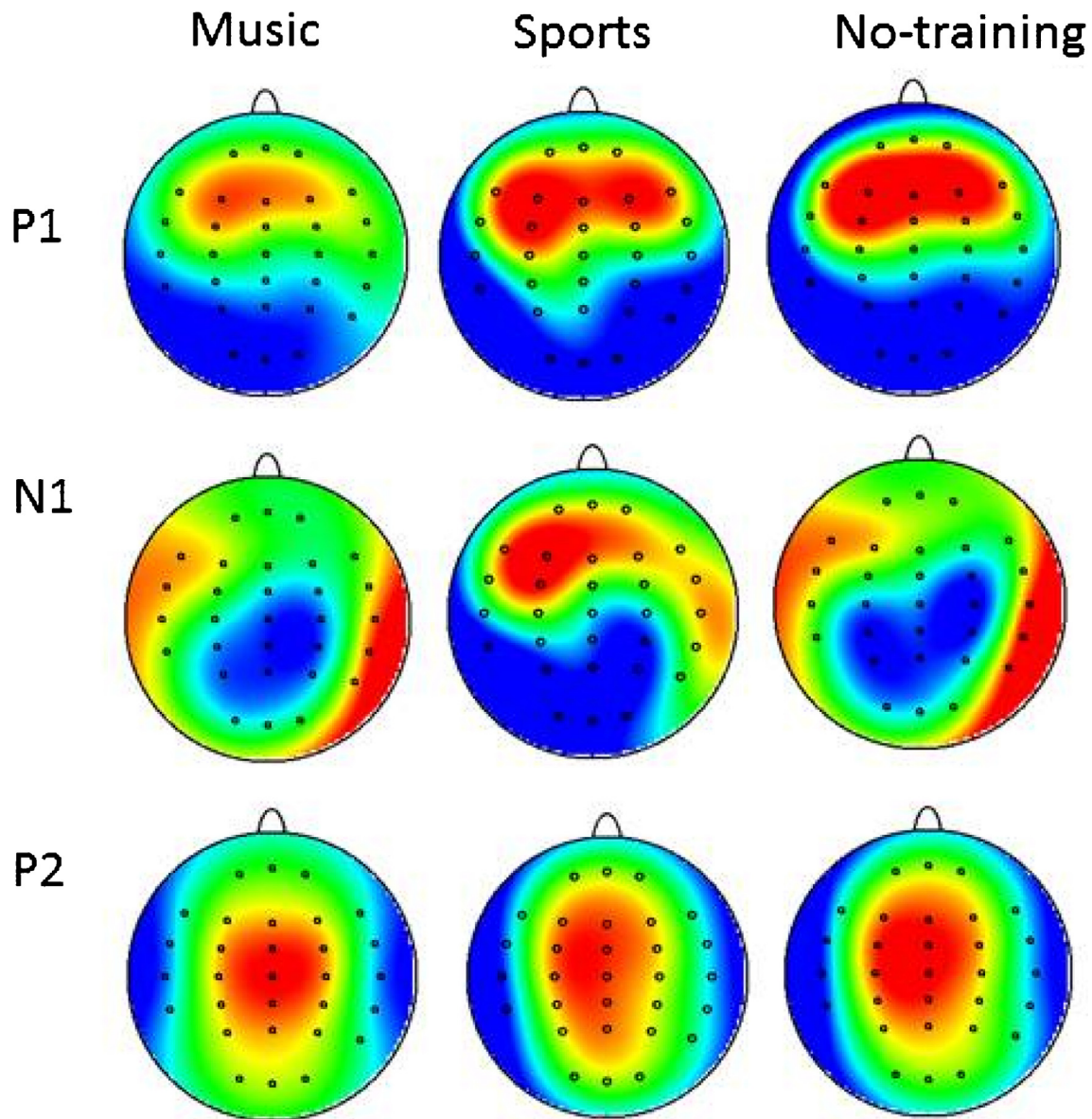


Fig. 5. Topography maps of P1, N1, P2 potentials, for music, sports and no-training groups at year two assessment.

Table 2

Average peak latency for P1, N1 and P2 auditory evoked potentials measured at baseline and year two for music, sports and no-training groups.

P1	Music	Sports	No-training
Baseline Peak Latency (ms)	119	116.6	112
Year 2 Peak Latency (ms)	84	92	87.6
N1	Music	Sports	No-training
Baseline Peak Latency (ms)	NA	NA	NA
Year 2 Peak Latency (ms)	120.7	122.6	128.8
P2	Music	Sports	No-training
Baseline Peak Latency (ms)	202	201.2	197.4
Year 2 Peak Latency (ms)	180.2	179.8	181

### 3.5. Event related potentials (ERPs) in pitch discrimination task

In the trials where the target and the comparison melodies were identical, there was no difference between the music and the two

comparison groups in the amplitude of the P1, N1, P2, N2, or P3 components (all  $p > 0.5$ ).

In the condition where the comparison melody included a changed-pitch note, *tonal different*, the amplitude of the P1 component was larger in the sports group compared to the music and no-training groups, as evidence by main effect of group [ $F(2, 34) = 4.91, p = 0.01$ , partial eta squared = 0.22]. The post-hoc contrasts indicated that the amplitude of P1 was significantly larger for the sports groups compared to the music and no-training groups. In contrast, the N1 component was smallest in the sports group relative to the music and no-training groups, as evidence by main effect of group [ $F(2, 34) = 7.28, p = 0.002$ , partial eta squared = 0.29]. Post-hoc contrasts indicated that the amplitude of N1 was significantly smaller for the sports groups compared to both music and no-training groups. Subsequently, P2 component was larger in the sports group compared to the music and no-training groups as well, as evidenced by main effect of group [ $F(2, 34) = 2.88, p = 0.06$ ] and a group x laterality interaction (largest at FCz) [ $F(4, 68) = 2.15, p = 0.08$ ].



Given that in the grand average waveform for the tonal different condition, the amplitude difference in the first 200 ms range between the sports group compared to the music and no-training groups, seemed to be mediated by the preceding enhanced amplitude of the P1 component with no subsequent return to baseline, we assessed the amplitude of N1 and P2 components via a “peak-to-peak” analysis, a common method to reduce the such biasing in ERP quantitation (Handy, 2005). In this analysis we analyzed the amplitude of N1 and P2 in relation to the amplitude of the preceding peak, e.g., by subtracting amplitude of N1 from P1 and P2 from N1 separately (Handy, 2005). There were no significant differences between the sports and the music or no-training groups in the peak-to-peak amplitudes of N1 relative to P1 [ $F(2, 34) = 0.14$ ,  $p = 0.86$ ] nor in the amplitude of P2 relative to N1 [ $F(2, 34) = 0.84$ ,  $p = 0.43$ ].

The amplitude of N2 component elicited by pitch different note was not significantly different between the three groups ( $p = 0.5$ ). The P3 component demonstrated greatest amplitude in fronto-central locations in the tonal different condition (see Fig. 6b) and a difference between groups was obtained as evidenced by a trend level main effect of Group [ $F(2, 34) = 2.73$ ,  $p = 0.07$ ] and a significant Group  $\times$  Laterality [ $F(4, 68) = 3.43$ ,  $p = 0.01$ , partial eta squared = 0.16] interaction. The subsequent *post-hoc* contrasts indicated that P3 amplitude at FCz was trend level different between music and no-training group ( $p = 0.07$ ) and not significant between music vs. sports ( $p = 0.84$ ) or sports vs. no-training groups ( $p = 0.88$ ). Fig. 6 shows auditory evoked potentials along with behavioral performance in response to the *tonal same* and *tonal different* conditions for the three groups.

A secondary analysis for P3 amplitude was conducted on the frontal-central site FCz only as the amplitude of the P3 was greatest at this site; such analysis indicated a significant difference of P3 amplitude between the three groups [ $F(2, 34) = 3.57$ ,  $p = 0.03$ , partial eta squared = 0.17]. *Post-hoc* contrasts showed that the P3 was significantly larger at FCz between the music and no-training group ( $p = 0.01$ ) and at a trend level difference between music and sports groups ( $p = 0.10$ ). Fig. 7 shows topography maps of P3 response to different note in *tonal different* condition for music, sports and no-training groups.

Finally, for the participants in the music group, the amplitude of the P3 component at FCz, correlated at a trend level with the accuracy of responses in detecting changes in pitch ( $r = 0.43$ ,  $p = 0.1$ ). This relationship was not observed in either of the two comparison groups.

Although the three groups of children did not differ on any of the demographic variables, because age showed a marginal difference between groups, we re-ran the initial ANOVAs using ANCOVA with age as covariate for the amplitude of P1, P1/N1 ratio in the passive task and P3 (in pitch different condition) in the active task. All previously reported significant main effects and interactions remained significant in the ANCOVA analysis.

#### 4. Discussion

We investigated the impact of music training on the maturation of central auditory processes by assessing auditory event related potentials and behavioral responses in children ages 6–7, prior to their participation in music training and two years after the start of training. We report two main findings: (1) In the passive pitch perception task, a decrease in P1 amplitude and an increase in the incidence of an identifiable N1 component as well as an increase in the N1/P1 ratio from baseline to year two were observed in the music group; none of these effects had emerged to the same extent in the age-matched comparison groups. Moreover, a decrease in latency of the P1 peak was present when comparing music and

sports groups. (2) In the active pitch discrimination task, children involved with music training detected deviations in pitch more accurately and showed larger P3 component amplitude in response to such changes. Next, we discuss each of these results separately and in relation to previous findings.

The P1 component is the most reliably measured component of auditory evoked potentials in children and has a fronto-central distribution (Ponton et al., 2000). Its latency decreases systematically with increasing age (Cunningham et al., 2000; Sharma et al., 1997; Wunderlich and Cone-Wesson, 2006) from peak latency of 85–95 ms in 5–6 year olds to 40–60 ms in 18–20 years old adults. This decline in P1 latency is known to be related to the ongoing increases in neural transmission speed as a result of developmental changes in myelination of underlying neural generators. Furthermore, increases in synaptic synchronization contributes to the speed of neural transmission during development (Huttenlocher, 1979; Wunderlich and Cone-Wesson, 2006). The P1 amplitude decreases with age as well, a phenomenon likely related to the emergence of N1 component and the development of underlying P1 generators (Čeponiene et al., 2002).

The most robust differential findings between groups were observed on the P1 component elicited by the piano tones in the passive task. We observed a decrease in P1 amplitude and latency that was largest in the music group compared to age-matched comparison groups after two years of training. In addition, focusing just on the year 2 data, the music group showed the smallest amplitude of P1 compared to both no-training and sports group, in combination with the accelerated development of the N1 component. These findings indicate that the pattern of auditory cortical processing was associated with faster maturation in the music group, likely related to more active sustained engagement of auditory neural circuitry (Trainor et al., 2012). This may be an indication that underlying processes of developmental myelination proceeded faster as a result of musical experience, a possibility we hope to address in our ongoing structural MRI studies of these same groups.

Of note, Shahin et al., 2004, using a similar paradigm, showed larger P1 amplitude in children trained for one year with music using the Suzuki method relative to non-music trained children. Although our findings, at the first glance appear contrary to theirs, few points are important to consider when comparing these results: First, the children participants in the Shahin et al. study were on average 5–6 years old and were assessed cross-sectionally at this one time point, not over time so as to interact with the development of the P1. Secondly, the morphological changes of the P1 peak (decreasing amplitude and latency in combination with the developing N1 peak) has been noted to take shape more prominently after age 7 (McArthur and Bishop, 2002; Ponton et al., 2000). Lastly, several recent studies (Kraus and Anderson, 2014; Slater et al., 2014) have shown that effects of music training, specifically at the neural level, are not noticeable until at least after two years of training and children in the Shahin et al. study were assessed at only one year after music training. Therefore our results of experience related plasticity of P1 amplitude and latency in the expected direction of development are in concordance with previous findings. It may well be the case that early music training increases the amplitude of the P1 component at age 4–5 as this is the predominant neural response to auditory input at that stage of brain development, but that as training proceeds, between the ages of 6–9, it tends to favor the transition of the P1 to the more fully developed P1-N1 complex as observed in the present data.

Furthermore, unlike Shahin et al., 2004, who reported an instrument specific enhancement of P2 component, we did not find any difference of P2 component between music and the two comparison groups. This may be related to the difference in musical training methods. While the music children in Shahin et al., 2004; were trained on a specific instrument (piano or violin) based on

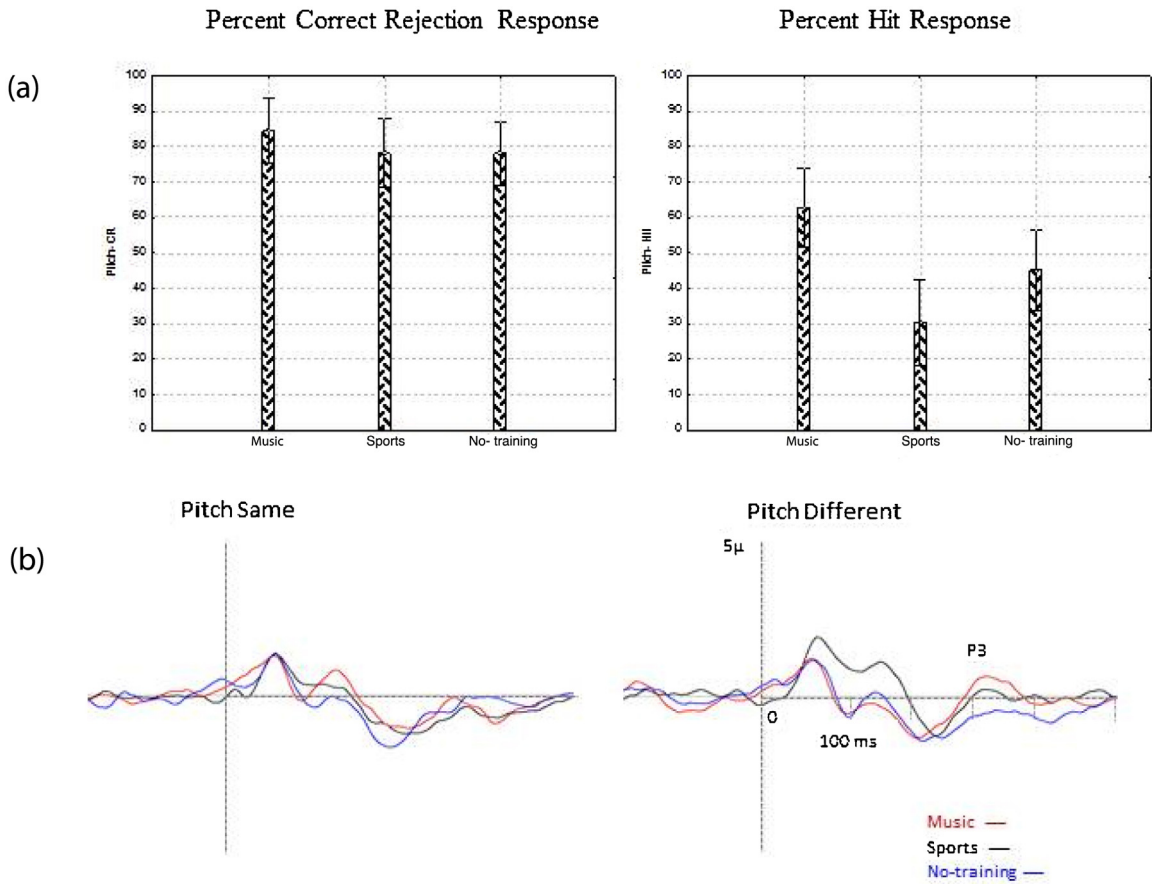


Fig. 6. (a) Behavioral results in response to tonal same and tonal different conditions for all three groups and (b) Auditory evoked potentials averaged of FC3, FCz & FC4.

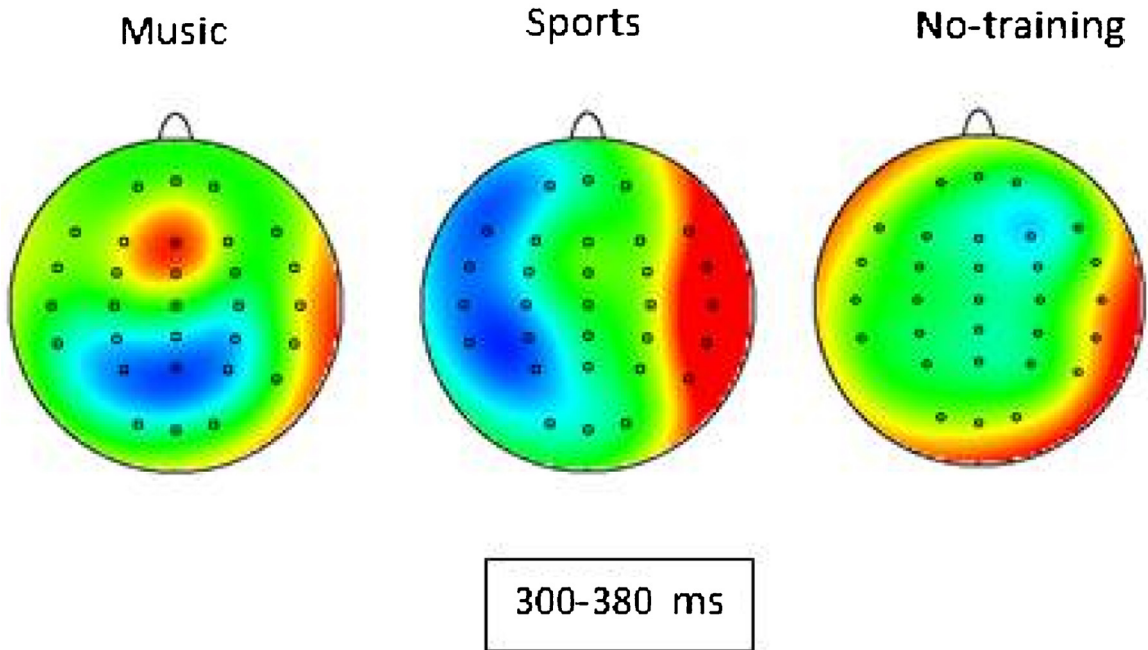


Fig. 7. Topography maps of P3 response to different note in *tonal different* condition for music, sports and no-training groups.

the Suzuki method, the musical curriculum of our participants was more diverse. It included instrument training (violin), choir, and musicianship and theory skills. The curriculum is specifically designed for group instructions and ensemble performance, not

individualized intensive instrumental training. The difference in the time spent on a musical instrument, on an individual basis, versus group musical activities, may have resulted in the lack of P2 effect in our findings.

The N1 component, which is the most dominant negative auditory ERP component in adult is absent in young children. During development, as age increases, the P1 peak narrows and decreases in amplitude and latency and the N1 peak emerges between ages 8–10, broadens and becomes increasingly negative compared to baseline. This process continues until early adulthood (between ages 18–20), by which time the P1 becomes much smaller in amplitude relative to the robust N1 and, in paradigms where it occurs, typically requiring very short duration auditory stimuli it is commonly named the P50 due to its occurring at approximately 50 ms post-stimulus (Cunningham et al., 2000; McArthur and Bishop, 2002; Ponton et al., 2000). N1 peak latency also declines with age. It is important to note that, specifically in children, the morphology of N1 depends critically on inter-stimulus-interval (ISI). Studies using a long ISI, generally longer than 2 s, have shown an N1 potential that is sensitive to age whereas the presence of N1 with ISI shorter than 2 s is not always reliable. This is the consequence of the high refractory nature of N1 generator in children (Čeponiene et al., 1998; Paetau et al., 1995) which has been shown to decrease with age (Rojas et al., 1998). To account for this phenomenon, we used a long inert-stimulus-interval (2500 ms) in our design to assure optimal conditions to detect the presence of N1.

The N1 amplitude has been specifically shown to be sensitive to experience-related plasticity—an increase in amplitude has been reported in adults without music training, after short-term training (Pantev and Herholz, 2011). In musicians, N1 amplitude has been shown to be larger, compared to non-musicians, specifically in response to musical stimuli (Habibi et al., 2013; Shahin et al., 2003, 2004). Here we find an increase in N1 amplitude relative to P1 amplitude in children, but only in the music group, over the course of two years of music training. This implies that music training likely is associated with accelerated development of central auditory pathways in children as young as age 8. Similar findings have been recently reported in adolescents with three years of training who show an increase in N1 amplitude relative to P1 amplitude when compared to an age-matched active (participants enrolled in Junior Reserve Officer Training Corp) comparison group (Tierney et al., 2015). Of note, the N1 amplitude has also been shown to be influenced by attention (Picton and Hillyard, 1974). Given that children in the music group are trained to tune their attention to the musical sounds of their instrument and other instruments while practicing and rehearsing, the faster development of N1 in this group may also be related to stimulus specific attentional mechanisms. However, the combination of the changes seen in N1 with those seen in P1, may suggest changes in the maturation of the auditory system.

The significant differences obtained in brain measures on the passive pitch perception task between groups was specific to the piano tones, and observed to a smaller but not statistically significant extent to the violin tones and not at all to the pure tones. Given that the participants in the music training group had specific training in violin this was somewhat unexpected. It is important to note, however, that in addition to their specific violin training, music group children participated in music theory, choir and ear training sessions where piano was used as the teaching instrument. Therefore, they were trained to discriminate musical features in particular with piano tones. This stimulus specific training in pitch discrimination may partly explain the lack of difference in the brain processing of pure tones as well as violin tones between the groups. In addition, the spectral complexity of the piano and violin tones, compared to the pure tones likely contributed to the increased number and/or synchronization of neurons coding for the temporal and spectral features of these stimuli and subsequently greater engagement of the auditory pathway, as has been previously demonstrated (Meyer et al., 2006). While the maturation of the P1 component towards an N1 is known to be related to

underlying brain morphological development, the effects on the maturation of the P1 to N1 observed to piano tones did not generalize across other stimulus types. This suggests that the observed effects may be related to specific musical stimuli or it is possible that longer periods of training is necessary for such an effect to translate to other auditory stimuli.

Our findings provide evidence, for the first time, that experience with music training in younger children interacts with development and is associated with accelerated cortical maturation necessary for general auditory processes such as language, speech and social interaction. This finding is especially important since we showed an equal baseline among the children in the music group and the two comparison groups on these measures prior to training. Furthermore, our design provided a reliable objective measure of auditory function, at the cortical level, in children, because the components of interest were recorded in a passive condition and were not influenced by performance related demands. Of note while the P1 amplitude diminution was greater in the music group compared to no-training and sports group, N1/P1 ratio showed no statistically significant difference between the sports and music groups. The probability of incidence of the N1 was greater in the music group compared to the two comparison groups, with the sports group showing intermediate incidence of the N1 component, possibly indicating a tendency for intensive sensorimotor engagement to contribute to auditory pathway development as well. It is worth mentioning that data collected from 15 adults in a separate experiment using the same paradigm showed evidence of N1 component in 100% of participants.

In addition to enhancement in the development of the auditory processing in the passive listening task, we showed that children involved with music training could detect deviations in pitch more accurately and showed larger amplitude P3 components in response to changes in tonal environment. The P3 component is thought to reflect neural events associated with immediate memory processes (Ladish and Polich, 1989; Polich, 2007) and is produced when subjects attend and discriminate stimulus events which are different from one another on some dimension. P3 amplitude and latency are shown to be affected by subject age: amplitude tends to increase and peak latency decreases substantially from the ages of five to puberty when it stabilizes before amplitude decrements and latency increases in older age (Brown et al., 1983; Ladish and Polich, 1989; Polich and Burns, 1990).

Previous studies have examined the effects of music training on P3 in adults (Nikjeh et al., 2008; Tervaniemi et al., 2005; Trainor Desjardins and Rockel, 1999) reporting both amplitude increases and latency decreases associated with long-term musical training. Our results replicated these findings in young children after only two years of training. We also found that there was a positive correlation, albeit not quite significant, between the magnitude of the P3 amplitude in the musical group and their enhanced accuracy in detecting deviant notes. Considering the number of participants in the music group, and the relatively strong correlation ( $r=0.43$ ), it is reasonable to expect that this correlation may reach significance once we add further subjects. The presence of this association between behavior and P3 amplitude in the music group, but not in the comparison groups, may be related to the fact that the P3 in the auditory domain matures into its adulthood more quickly as a result of childhood music training. Several previous studies have found associations between P3 amplitude and target performance in adult populations (Polich, 2007).

Given that our task involved matching two melodies – template and target – in order to detect tonal alterations, the differences of P3 amplitude between the music and comparison groups seems to reflect superior auditory template matching abilities in the music group. The fact that the neural difference between groups was specific to the P3 and not the earlier more sensory P2/N2 potentials

points towards the improved performance being mediated through higher-level attentional and working memory processes. This may be a consequence of enhanced auditory working memory as a result of music training and possibly related to prior evidence of positive impact of music training on verbal memory (Chan et al., 1998; Franklin et al., 2008; Ho et al., 2003; Jakobson et al., 2008a,b).

The relatively focal topography of the P3 component in the music group (Fig. 7) suggests the possibility that a generator close to the skull is involved, such as the premotor and/or motor cortices. A recent study indicated that early musical training is specifically linked to increased grey matter in the premotor cortices (Bailey et al., 2014). Also, it has been previously shown that in the adult musician's brain the two networks are strongly linked and even when a task involves only auditory or only motor processing, co-activation phenomena within the respective brain areas can be expected (Haueisen and Knösche, 2001).

Unexpectedly, we observed a significant difference in the amplitude of P1, N1 and P2 components between the sports and the other two groups in response to changes in pitch. Despite these differences, the sports group pitch detection accuracy was only 30%, worst among the three groups. Therefore the amplitude difference in the first 200 ms range was unrelated to the conscious detection of musical deviations. Moreover, given the appearance of the ERP we find it most likely that the differences seen in the N1 and P2 components were mediated by the preceding P1 amplitude being increased. Indeed peak to peak analysis of N1 and P2 components showed that there were no group differences in the amplitude of N1 or P2.

In relation to the significant enhancement of the P1 amplitude, we hypothesize that slower functional maturity of the auditory system in the music group compared to the other groups, may have contributed to the enhanced P1 relative to N1 amplitude. However, given that this difference is more pronounced in the tonal different versus tonal same condition, it seems to have some functional significance related to attentional engagement. Thus, alternatively, the increased P1 amplitude in the music group may be related to the intensive sensorimotor training the sports group had engaged in over the two years of the study leading them to be more readily engaged at the sensory level with encoding pitch differences. Nevertheless, without the refined auditory attentional training that the music group had, this increased sensorimotor engagement did not translate into improved performance nor enhanced P3 processing as was observed in the music group.

An inevitable limitation of this study is the absence of a randomized controlled trial design. However, such randomization is very difficult in long term longitudinal studies such as this. These are normal developing children who enrolled in their respective extracurricular programs (music or sports) by their (and their family's) own motivation. On the other hand, if children are not motivated and not emotionally engaged in an activity, it is unlikely they will continue participation for the long term period necessary for a longitudinal investigation. In addition, assigning children to specifically not engage in beneficial activity for long periods during critical times of development would be unethical. Randomized trials are of course ideal for short term and/or clinical interventions and it would be of benefit if such study design could be ethically and practically designed for the study of systematic musical training at some point in future. In the meantime, "real world" naturalistic studies such as ours, will help develop a deeper understanding of how music education can benefit auditory skills in children.

Another limitation of this study can be seen in the relative small number of participants; this was due to significant logistical challenges in recruiting and retaining the participants, especially these from low socio-economic backgrounds. Under the circumstances, and given that the three groups showed no differences on any measure of cognitive, social, emotional and neural processing at the

time of induction in our study (all  $p > 0.1$ ; see Habibi et al., 2014), we believe the results reported here provide an important contribution to the existing literature on music training and childhood development. Furthermore, given the longitudinal nature of our study, we believe that our findings are consistent with the notion of significant role of music training in child brain development; however, since allocation to training was not randomized, a causal relationship is only one possibility and other factors, such as motivation or predisposition towards music, could influence both group maintenance and changes in brain function.

In summary, we have shown accelerated development of auditory cortical potentials and greater ability to detect pitch changes in children who underwent two years of musical training compared to age-matched comparison groups. Taken together these findings provide evidence that childhood music training has a measurable impact in the development of auditory processes. Although the findings described here are restricted to auditory skills and to their neural correlates, such enhanced maturation may favor faster and more efficient development of language skills as well, given that some of the neural substrates to these different processes are shared. Our findings demonstrate that music education has an important role to play in childhood development and add to the converging evidence that music training is capable of shaping skills that are ingredients of success in social and academic development. It is of particular importance that we show these effects in children from disadvantaged backgrounds.

## Acknowledgments

The authors thank all participating children and their families. We also thank other members of our team who, in various capacities contributed throughout the execution of this project. They are Beatriz Ilari, Bronte Ficek, Jennifer Wong, Kevin Crimi, Alissa Dar Sarkissian, Cara Fesjian, Ryan Veiga, Emily He, Marth Gomez and Leslie Chinchilla.

We are also grateful to the Los Angeles Philharmonic, Youth Orchestra of Los Angeles, Heart of Los Angeles, Brotherhood Crusade and their Soccer for Success Program as well as the following elementary schools and community programs in the Los Angeles area: Vermont Avenue Elementary School, Saint Vincent School and Macarthur Park Recreation Center. This research was funded in part by an anonymous donor and by the USC's Brain & Creativity Institute Research Fund.

## References

- Čeponiene, R., Cheour, M., Näätänen, R., 1998. Interstimulus interval and auditory event-related potentials in children: evidence for multiple generators. *Electroencephalogr. Clin. Neurophysiol.* 108 (4), 345–354.
- Čeponiene, R., Rinne, T., Näätänen, R., 2002. Maturation of cortical sound processing as indexed by event-related potentials. *Clin. Neurophysiol.* 113 (6), 870–882.
- Andrews, M.W., Dowling, W.J., Bartlett, J.C., Halpern, A.R., 1998. Identification of speeded and slowed familiar melodies by younger, middle-aged, and older musicians and nonmusicians. *Psychol. Aging* 13 (3), 462–471. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9793121>.
- Bailey, J.A., Zatorre, R.J., Penhune, V.B., 2014. Early musical training is linked to gray matter structure in the ventral premotor cortex and auditory-motor rhythm synchronization performance. *J. Cogn. Neurosci.* 26 (4), 755–767. <http://dx.doi.org/10.1162/jocn>.
- Bangert, M., Schlaug, G., 2006. Specialization of the specialized in features of external human brain morphology. *Eur. J. Neurosci.* 24 (6), 1832–1834. <http://dx.doi.org/10.1111/j.1460-9568.2006.05031.x>.
- Bishop, D.V.M., Anderson, M., Reid, C., Fox, A.M., 2011. Auditory development between 7 and 11 years: an event-related potential (ERP) study. *PLoS One* 6 (5), e18993. <http://dx.doi.org/10.1371/journal.pone.0018993>.
- Brattico, E., Tervaniemi, M., Näätänen, R., Peretz, I., 2006. Musical scale properties are automatically processed in the human auditory cortex. *Brain Res.* 1117 (1), 162–174. <http://dx.doi.org/10.1016/j.brainres.2006.08.023>.
- Brown, W.S., Marsh, J.T., LaRue, A., 1983. Exponential electrophysiological aging: p3 latency. *Electroencephalogr. Clin. Neurophysiol.* 55 (3), 277–285.

- Chan, A.S., Ho, Y.C., Cheung, M.C., 1998. Music training improves verbal memory. *Nature* 396 (6707), 128.
- Chobert, J., François, C., Velay, J.-L., Besson, M., 2014. Twelve months of active musical training in 8- to 10-year-old children enhances the preattentive processing of syllabic duration and voice onset time. *Cereb. Cortex* 24 (4), 956–967. <http://dx.doi.org/10.1093/cercor/bhs377>.
- Cunningham, J., Nicol, T., Zecker, S., Kraus, N., 2000. Speech-evoked neurophysiological responses in children with learning problems: development and behavioral correlates of perception. *Ear Hear.* 21 (6), 554–568.
- Degé, F., Schwarzer, G., 2011. The effect of a music program on phonological awareness in preschoolers. *Front. Psychol.* 2 (June), 124. <http://dx.doi.org/10.3389/fpsyg.2011.00124>.
- Donchin, E., Coles, M.G., 1988. Is the P300 component a manifestation of context updating? *Behav. Brain Sci.* 11 (03), 357–374.
- Dowling, W.J., Bartlett, J.C., Halpern, A.R., Andrews, M.W., 2008. Melody recognition at fast and slow tempos: effects of age, experience, and familiarity. *Percept. Psychophys.* 70 (3), 496–502.
- Forgeard, M., Winner, E., Norton, A., Schlaug, G., 2008. Practicing a musical instrument in childhood is associated with enhanced verbal ability and nonverbal reasoning. *PLoS One* 3 (10), e3566.
- Franklin, M.S., Sledge Moore, K., Yip, C.-Y., Jonides, J., Rattray, K., Moher, J., 2008. The effects of musical training on verbal memory. *Psychol. Music* 36 (3), 353–365. <http://dx.doi.org/10.1177/0305735607086044>.
- Fujioka, T., Trainor, L.J., Ross, B., Kakigi, R., Pantev, C., 2005. Automatic encoding of polyphonic melodies in musicians and nonmusicians. *J. Cogn. Neurosci.* 17 (10), 1578–1592. <http://dx.doi.org/10.1162/089892905774597263>.
- Gaser, C., Schlaug, G., 2003. Brain structures differ between musicians and non-musicians. *J. Neurosci.* 23 (27), 9240–9245. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/14534258>.
- Habibi, A., Wirantana, V., Starr, A., 2013. Cortical activity during perception of musical pitch comparing musicians and nonmusicians. *Music Percept.* 30 (5), 463–479.
- Habibi, A., Ilari, B., Crimi, K., Metke, M., Kaplan, J.T., Joshi, A., Damasio, H., 2014. An equal start: absence of group differences in cognitive, social, and neural measures prior to music or sports training in children. *Front. Hum. Neurosci.* 8 (September), 690. <http://dx.doi.org/10.3389/fnhum.2014.00690>.
- Handy, T.C., 2005. *Event-Related Potentials: A Methods Handbook*. MIT press.
- Herholz, S.C., Zatorre, R.J., 2012. Musical training as a framework for brain plasticity: behavior, function, and structure. *Neuron* 76 (3), 486–502. <http://dx.doi.org/10.1016/j.neuron.2012.10.011>.
- Ho, Y.-C., Cheung, M.-C., Chan, A.S., 2003. Music training improves verbal but not visual memory: cross-sectional and longitudinal explorations in children. *Neuropsychology* 17 (3), 439–450. <http://dx.doi.org/10.1037/0894-4105.17.3.439>.
- Huttenlocher, P.R., 1979. Synaptic density in human frontal cortex—developmental changes and effects of aging. *Brain Res.* 163 (2), 195–205.
- Ille, N., Berg, P., Scherg, M., 2002. Artifact correction of the ongoing EEG using spatial filters based on artifact and brain signal topographies. *J. Clin. Neurophysiol.* 19 (2), 113–124.
- Jäncke, L., 2009. Music drives brain plasticity. *Biol. Rep.* 1 (October), 78. <http://dx.doi.org/10.3410/b1-78>.
- Jakobson, L.S., Cuddy, L.L., Kilgour, A.R., 2008a. Time tagging: a key to musicians' superior memory. *Music Percept.* 20 (3), 307–313.
- Jakobson, L.S., Lewycky, S.T., Kilgour, A.R., Stoesz, B.M., 2008b. Memory for verbal and visual material in highly trained musicians. *Music Percept.* 20 (3), 307–313.
- Koelsch, S., Schröger, E., Tervaniemi, M., 1999. Superior pre-attentive auditory processing in musicians. *Neuroreport* 10 (6), 1309–1313. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10363945>.
- Koelsch, S., Jentschke, S., Sammler, D., Mietchen, D., 2007. Untangling syntactic and sensory processing: an ERP study of music perception. *Psychophysiology* 44 (3), 476–490. <http://dx.doi.org/10.1111/j.1469-8986.2007.00517.x>.
- Kraus, N., Anderson, S., 2014. Community-based training shows objective evidence of efficacy. *Hear. J.* 67 (11), 46–48.
- Ladish, C., Polich, J., 1989. P300 and probability in children. *J. Exp. Child Psychol.* 48 (12), 212–223.
- Lappe, C., Trainor, L.J., Herholz, S.C., Pantev, C., 2011. Cortical plasticity induced by short-term multimodal musical rhythm training. *PLoS One* 6 (6), e21493. <http://dx.doi.org/10.1371/journal.pone.0021493>.
- Levitin, D.J., 2012. What does it mean to be musical? *Neuron* 73 (4), 633–637. <http://dx.doi.org/10.1016/j.neuron.2012.01.017>.
- Luck, S.J., 2014. *An Introduction to the Event-Related Potential Technique*. MIT press.
- McArthur, G., Bishop, D., 2002. Event-related potentials reflect individual differences in age-invariant auditory skills. *Neuroreport* 13 (8), 1079–1082.
- Meyer, M., Baumann, S., Jancke, L., 2006. Electrical brain imaging reveals spatio-temporal dynamics of timbre perception in humans. *Neuroimage* 32 (4), 1510–1523. <http://dx.doi.org/10.1016/j.neuroimage.2006.04.193>.
- Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S.L., Besson, M., 2009. Musical training influences linguistic abilities in 8-year-old children: more evidence for brain plasticity. *Cereb. Cortex* 19 (3), 712–723. <http://dx.doi.org/10.1093/cercor/bhn120>.
- Moreno, S., Bialystok, E., Barac, R., Schellenberg, E.G., Cepeda, N.J., Chau, T., 2011. Short-term music training enhances verbal intelligence and executive function. *Psychol. Sci.* 22 (11), 1425–1433. <http://dx.doi.org/10.1177/0956797611416999>.
- Musacchia, G., Sams, M., Skoe, E., Kraus, N., 2007. Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proc. Natl. Acad. Sci. U. S. A.* 104 (40), 15894–15898. <http://dx.doi.org/10.1073/pnas.0701498104>.
- Näätänen, R., Picton, T., 1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology* 24 (4), 375–425.
- Nikjeh, D.A., Lister, J.J., Frisch, S.A., 2008. Hearing of note: an electrophysiologic and psychoacoustic comparison of pitch discrimination between vocal and instrumental musicians. *Psychophysiology* 45 (6), 994–1007.
- Paetau, R., Ahonen, A., Salonen, O., Sams, M., 1995. Auditory evoked magnetic fields to tones and pseudowords in healthy children and adults. *J. Clin. Neurophysiol.* 12 (2), 177–185.
- Pantev, C., Herholz, S.C., 2011. Plasticity of the human auditory cortex related to musical training. *Neurosci. Biobehav. Rev.* 35 (10), 2140–2154. <http://dx.doi.org/10.1016/j.neubiorev.2011.06.010>.
- Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L.E., Hoke, M., 1998. Increased auditory cortical representation in musicians. *Nature* 392 (6678), 811–814.
- Pantev, C., Engelien, a, Candia, V., Elbert, T., 2001. Representational cortex in musicians. Plastic alterations in response to musical practice. *Ann. N. Y. Acad. Sci.* 930, 300–314. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11458837>.
- Picton, T.W., Hillyard, S.A., 1974. Human auditory evoked potentials. II. Effects of attention. *Electroencephalogr. Clin. Neurophysiol.* 36 (2), 191–199. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10705765>.
- Picton, T.W., 1992. The P300 wave of the human event-related potential. *J. Clin. Neurophysiol.* 9 (4), 456–479.
- Piro, J.M., Ortiz, C., 2009. The effect of piano lessons on the vocabulary and verbal sequencing skills of primary grade students. *Psychol. Music* 37 (3), 325–347. <http://dx.doi.org/10.1177/0305735608097248>.
- Polich, J., Burns, T., 1990. Normal variation of P300 in children: and head size age, memory span. 9, 237–248.
- Polich, J., 2007. Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118 (10), 2128–2148. <http://dx.doi.org/10.1016/j.clinph.2007.04.019>.
- Ponton, C.W., Eggermont, J.J., Kwong, B., Don, M., 2000. Maturation of human central auditory system activity: evidence from multi-channel evoked potentials. *Clin. Neurophysiol.* 111 (2), 220–236.
- Putkinen, V., Tervaniemi, M., Saarikivi, K., de Vent, N., Huotilainen, M., 2014. Investigating the effects of musical training on functional brain development with a novel Melodic MMN paradigm. *Neurobiol. Learn. Mem.* 110C (January), 8–15. <http://dx.doi.org/10.1016/j.nlm.2014.01.007>.
- Rojas, D.C., Walker, J.R., Sheeder, J.L., Teale, P.D., Reite, M.L., 1998. Developmental changes in refractoriness of the neuromagnetic M100 in children. *Neuroreport* 9 (7), 1543–1547.
- Schön, D., Magne, C., Besson, M., 2004. The music of speech: music training facilitates pitch processing in both music and language. *Psychophysiology* 41 (3), 341–349. <http://dx.doi.org/10.1111/1469-8986.00172.x>.
- Schellenberg, E.G., Moreno, S., 2010. Music lessons, pitch processing, and g. *Psychol. Music* 38 (2), 209–221. <http://dx.doi.org/10.1177/0305735609339473>.
- Schlaug, G., 2001. The brain of musicians. A model for functional and structural adaptation. *Ann. N. Y. Acad. Sci.* 930, 281–299. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11458836>.
- Schneider, P., Scherg, M., Dosch, H.G., Specht, H.J., Gutschalk, A., Rupp, A., 2002. Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nat. Neurosci.* 5 (7), 688–694. <http://dx.doi.org/10.1038/nn871>.
- Seppänen, M., Pesonen, A.-K., Tervaniemi, M., 2012. Music training enhances the rapid plasticity of P3a/P3b event-related brain potentials for unattended and attended target sounds. *Atten. Percept. Psychophys.* 74 (3), 600–612. <http://dx.doi.org/10.3758/s13414-011-0257-9>.
- Shahin, A., Bosnyak, D.J., Trainor, L.J., Roberts, L.E., 2003. Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *J. Neurosci.* 23 (13), 5545–5552.
- Shahin, A., Roberts, L.E., Trainor, L.J., 2004. Enhancement of auditory cortical development by musical experience in children. *Neuroreport* 15 (12), 1917–1921. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/15305137>.
- Shahin, A., Trainor, L., Roberts, L., Backer, K., Miller, L., 2010. Development of Auditory Phase-Locked Activity for Music Sounds, 218–229. <http://dx.doi.org/10.1152/jn.00402.2009>.
- Sharma, A., Kraus, N., Mcgee, T.J., Nicol, T.G., 1997. Developmental changes in P1 and N1 central auditory responses elicited by consonant-vowel syllables, 104, 540–545.
- Slater, J., Strait, D.L., Skoe, E., O'Connell, S., Thompson, E., Kraus, N., 2014. Longitudinal effects of group music instruction on literacy skills in low-income children. *PLoS One* 9 (11), e113383.
- Strait, D.L., Kraus, N., 2014. Biological impact of auditory expertise across the life span: musicians as a model of auditory learning. *Hear. Res.* 308, 109–121. <http://dx.doi.org/10.1016/j.heares.2013.08.004>.
- Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., Schröger, E., 2005. Pitch discrimination accuracy in musicians vs nonmusicians: an event-related potential and behavioral study. *Exp. Brain Res.* 161 (1), 1–10. <http://dx.doi.org/10.1007/s00221-004-2044-5>.

- Tierney, A.T., Krizman, J., Kraus, N., 2015. Music training alters the course of adolescent auditory development. *Proc. Natl. Acad. Sci.*, <http://dx.doi.org/10.1073/pnas.1505114112>, 201505114.
- Tillmann, B., Janata, P., Bharucha, J.J., 2003. Activation of the inferior frontal cortex in musical priming. *Cogn. Brain Res.* 16 (2), 145–161, [http://dx.doi.org/10.1016/S0926-6410\(02\)00245-8](http://dx.doi.org/10.1016/S0926-6410(02)00245-8).
- Trainor Desjardins, R.N., Rockel, C., 1999. A comparison of contour and interval processing in musicians and nonmusicians using event-related potentials. *Aust. J. Psychol.* 51 (3), 147–153.
- Trainor, L.J., Marie, C., Gerry, D., Whiskin, E., Unrau, A., 2012. Becoming musically enculturated: effects of music classes for infants on brain and behavior. *Ann. N. Y. Acad. Sci.* 1252, 129–138, <http://dx.doi.org/10.1111/j.1749-6632.2012.06462.x>.
- Trainor, L.J., 2012. Musical experience, plasticity, and maturation: issues in measuring developmental change using EEG and MEG. *Ann. N. Y. Acad. Sci.* 1252, 25–36, <http://dx.doi.org/10.1111/j.1749-6632.2012.06444.x>.
- Tremblay, K., Kraus, N., McGee, T., Ponton, C., Otis, B., 2001. Central auditory plasticity: changes in the N1-P2 complex after speech-sound training. *Ear Hear.* 22 (2), 79–90, Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11324846>.
- Virtala, P., Huotilainen, M., Putkinen, V., Makkonen, T., Tervaniemi, M., 2012. Musical training facilitates the neural discrimination of major versus minor chords in 13-year-old children. *Psychophysiology* 49 (8), 1125–1132.
- Vuust, P., Pallesen, K.J., Bailey, C., van Zuijen, T.L., Gjedde, A., Roepstorff, A., Østergaard, L., 2005. To musicians, the message is in the meter pre-attentive neuronal responses to incongruent rhythm are left-lateralized in musicians. *Neuroimage* 24 (2), 560–564, <http://dx.doi.org/10.1016/j.neuroimage.2004.08.039>.
- Wong, P.C.M., Skoe, E., Russo, N.M., Dees, T., Kraus, N., 2007. Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat. Neurosci.* 10 (4), 420–422, <http://dx.doi.org/10.1038/nn1872>.
- Wunderlich, J.L., Cone-Wesson, B.K., 2006. Maturation of CAEP in infants and children: a review. *Hear. Res.* 212 (1–2), 212–223, <http://dx.doi.org/10.1016/j.heares.2005.11.008>.
- Wunderlich, J.L., Cone-Wesson, B.K., Shepherd, R., 2006. Maturation of the cortical auditory evoked potential in infants and young children. *Hear. Res.* 212 (1–2), 185–202, <http://dx.doi.org/10.1016/j.heares.2005.11.010>.
- Zatorre, R., 2005. Music, the food of neuroscience? *Nature* 434 (7031), 312–315.