

The background of the cover features a stylized brain composed of various colored segments (yellow, orange, red, purple, blue, green) arranged in a circular pattern. A network of white lines connects nodes, resembling a neural network or a complex graph, overlaid on the brain segments. The top half of the cover has a blue background, while the bottom half is white.

MUSIC TRAINING, NEURAL PLASTICITY, AND EXECUTIVE FUNCTION

EDITED BY: Claude Alain, Assal Habibi and Paul J. Colombo

PUBLISHED IN: Frontiers in Neuroscience, Frontiers in Integrative Neuroscience
and Frontiers in Psychology



frontiers Research Topics



frontiers

Frontiers eBook Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence.

The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714

ISBN 978-2-88966-047-6

DOI 10.3389/978-2-88966-047-6

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

MUSIC TRAINING, NEURAL PLASTICITY, AND EXECUTIVE FUNCTION

Topic Editors:

Claude Alain, Rotman Research Institute (RRI), Canada

Assal Habibi, University of Southern California, Los Angeles, United States

Paul J. Colombo, Tulane University, United States

Citation: Alain, C., Habibi, A., Colombo, P. J., eds. (2020). Music Training, Neural Plasticity, and Executive Function. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88966-047-6

Table of Contents

05	<i>Editorial: Music Training, Neural Plasticity, and Executive Function</i> Paul J. Colombo, Assal Habibi and Claude Alain
08	<i>Electrical Neuroimaging of Music Processing in Pianists With and Without True Absolute Pitch</i> Sélim Yahia Coll, Noémi Vuichoud, Didier Grandjean and Clara Eline James
24	<i>On the Association Between Musical Training, Intelligence and Executive Functions in Adulthood</i> Antonio Criscuolo, Leonardo Bonetti, Teppo Särkämö, Marina Kliuchko and Elvira Brattico
36	<i>Auditory T-Complex Reveals Reduced Neural Activities in the Right Auditory Cortex in Musicians With Absolute Pitch</i> Masato Matsuda, Hironaka Igarashi and Kosuke Itoh
45	<i>Comparing the Effects of Rhythm-Based Music Training and Pitch-Based Music Training on Executive Functions in Preschoolers</i> Ulrike Frischen, Gudrun Schwarzer and Franziska Degé
56	<i>Musicians Show Better Auditory and Tactile Identification of Emotions in Music</i> Andréanne Sharp, Marie-Soleil Houde, Benoit-Antoine Bacon and François Champoux
63	<i>Dynamic Orchestration of Brains and Instruments During Free Guitar Improvisation</i> Viktor Müller and Ulman Lindenberger
75	<i>Absolute Pitch and Musical Expertise Modulate Neuro-Electric and Behavioral Responses in an Auditory Stroop Paradigm</i> Vivek V. Sharma, Michael Thaut, Frank Russo and Claude Alain
89	<i>Uncertain Emotion Discrimination Differences Between Musicians and Non-musicians is Determined by Fine Structure Association: Hilbert Transform Psychophysics</i> Francis A. M. Manno, Raul R. Cruces, Condon Lau and Fernando A. Barrios
102	<i>Auditory Categorization of Man-Made Sounds Versus Natural Sounds by Means of MEG Functional Brain Connectivity</i> Vasiliki Salvari, Evangelos Paraskevopoulos, Nikolas Chalas, Kilian Müller, Andreas Wollbrink, Christian Dobel, Daniela Korth and Christo Pantev
114	<i>Effects of Music Training on Inhibitory Control and Associated Neural Networks in School-Aged Children: A Longitudinal Study</i> Sarah L. Hennessy, Matthew E. Sachs, Beatriz Ilari and Assal Habibi
130	<i>Selectively Enhanced Development of Working Memory in Musically Trained Children and Adolescents</i> Katri Annukka Saarikivi, Minna Huotilainen, Mari Tervaniemi and Vesa Putkinen
142	<i>New Perspectives on Music in Rehabilitation of Executive and Attention Functions</i> Yuko Koshimori and Michael H. Thaut

- 148** *Short-Term Choir Singing Supports Speech-in-Noise Perception and Neural Pitch Strength in Older Adults With Age-Related Hearing Loss*
Ella Dubinsky, Emily A. Wood, Gabriel Nespoli and Frank A. Russo
- 166** *The Effects of Bimanual Coordination in Music Interventions on Executive Functions in Aging Adults*
Jennifer A. Bugos
- 179** *Music Training, Working Memory, and Neural Oscillations: A Review*
Kate A. Yurgil, Miguel A. Velasquez, Jenna L. Winston, Noah B. Reichman and Paul J. Colombo
- 196** *Mechanisms of Timing, Timbre, Repertoire, and Entrainment in Neuroplasticity: Mutual Interplay in Neonatal Development*
Joanne Loewy and Artur C. Jaschke



Editorial: Music Training, Neural Plasticity, and Executive Function

Paul J. Colombo^{1*}, Assal Habibi² and Claude Alain³

¹ Department of Psychology, School of Science and Engineering, Tulane University, New Orleans, LA, United States,

² Brain and Creativity Institute, Department of Psychology, University of Southern California, Los Angeles, CA, United States,

³ Rotman Research Institute, Department of Psychology, University of Toronto, Toronto, ON, Canada

Keywords: executive function, cognitive control, neural plasticity, music training, prefrontal cortex, neural oscillation, absolute pitch, longitudinal study

Editorial on the Research Topic

Music Training, Neural Plasticity, and Executive Function

Music training is rapidly emerging as an important model system for investigating experience-dependent brain plasticity, and also as the basis for robust sensory, motor, and cognitive therapeutic interventions. This special topic contains articles describing theoretical and methodological advancements that further our understanding of relationships between musicianship, music training and executive function, and many of the articles delve into the neural mechanisms that drive these relationships. It is well-established that music training is associated with sensory, motor, and cognitive benefits. As described in the articles in this collection, investigators are characterizing how elements of musical abilities (e.g., absolute pitch and relative pitch processors) and training (e.g., amateur vs. professional), are related to components of executive function including inhibitory control, working memory, and cognitive flexibility. In several reports, new analytical methods are being implemented to probe neural mechanisms of plasticity associated with music training and performance.

MUSIC TRAINING INTERVENTIONS IMPROVE EXECUTIVE FUNCTIONS IN RELATION TO BRAIN PLASTICITY: CHILDHOOD DEVELOPMENT AND OLDER ADULTS

OPEN ACCESS

Edited and reviewed by:

Elizabeth B. Torres,
Rutgers, The State University of
New Jersey, United States

*Correspondence:

Paul J. Colombo
pcolomb@tulane.edu

Received: 02 June 2020

Accepted: 30 June 2020

Published: 07 August 2020

Citation:

Colombo PJ, Habibi A and Alain C
(2020) Editorial: Music Training, Neural
Plasticity, and Executive Function.
Front. Integr. Neurosci. 14:41.
doi: 10.3389/fnint.2020.00041

Systematic manipulations of musical experience are the key to making causal statements about their effects on sensory, motor, and cognitive outcomes, and several investigators used this approach. For example, Dubinsky et al. report that short-term choir singing improves speech-in-noise perception and pitch discrimination among older adults with hearing loss, and improvement is related to the strength of the frequency following response, a neural representation of auditory stimuli. In another study, older adults received different types of musical experience, including music listening, piano training, or percussion training. Both of the active music groups outperformed the listening group in bimanual synchronization and visual scanning/working memory, and piano training significantly improved motor synchronization skills in comparisons with the percussion training or listening groups (Bugos). Frischen et al. also tested effects of different types of music training by assessing several measures of executive functions, then randomly assigning preschoolers to rhythm-based or pitch-based music training. Inhibition improved from pre- to post-test among children who received rhythm-based training, but not pitch-based training or a sports control, and similar numeric differences were found for measures of set shifting and visuospatial working memory. Taken together, these studies demonstrate that music training causes improvement of sensory, motor, and cognitive control processes among children

and older adults, and are not merely reflections of pre-existing neurocognitive differences. Moreover, they begin to dissociate components of musical training that may lead to improvement of specific cognitive processes. With regard to very early childhood development, Loewy and Jaschke identify how parameters of music such as timing, timbre, and repetition may influence cognitive development and neural plasticity in therapeutic interventions with neonates. The advancements noted above are bolstered by results of longitudinal studies of musically trained children in which working memory (Saarikivi et al.) and inhibitory control (Hennessy et al.) were assessed in relation to years of training. Musically trained children outperformed controls on the trails A and B tests, and forward digit span, but not backward digit span, suggesting that music training may selectively influence working memory capacity and maintenance more than with updating (Saarikivi et al.). Hennessy et al. report that children with 3 years of music training chose a larger, delayed reward in place of a smaller, immediate reward when compared to children without music training, indicating enhancement of delayed-gratification measures of inhibition. In the flanker task, children in the music group improved their performance accuracy parallel to increasing years of training, while such improvements were not observed in the groups without music training. The groups were matched at the onset of the study to have no differences among them in cognitive capacities, providing evidence that systematic music-based training accelerates development of inhibitory control in children.

COMPARISONS BETWEEN ADULT MUSICIANS AND NON-MUSICIANS REVEAL MECHANISMS OF ENHANCED EXECUTIVE FUNCTIONS

The articles cited above provide new evidence for behavioral and neural mechanisms of cognitive benefits caused by music training during early childhood development, and among older adults. A parallel approach, taken by several contributors, compared behavioral performance on control processing tasks and neural measures of activity among adult musicians and non-musicians (Manno et al.; Sharp et al.; Sharma et al.; Coll et al.; Criscuolo et al.). This approach yielded insights into the mechanisms by which musical experience may enhance cognitive processes. For example, musician advantages in emotion processing are related to greater use of temporal fine structure information (Manno et al.), and to better identification of complex emotional content in both the auditory and tactile sensory modalities (Sharp et al.). Sharma et al. presented non-musicians, and musicians with either relative or absolute pitch, with three different versions of an auditory Stroop task to measure conflict resolution. The pitch-label association ranged from simple semantic associations (i.e., “Low” or “High”) to intermediate verbal encodings with no obvious semantic properties (i.e., “Doh” or “Soh”) to more abstract semiotic associations (i.e., “C” and “G”). The neural activity

indexing conflict detection for abstract pitch label (i.e., musical notation) was present only in musicians with absolute pitch, consistent with a strong automaticity in retrieving the pitch-label association. Coll et al. also showed greater brain electrical source activity in left temporal junctions in musicians with absolute pitch, which could play a part in the automatic retrieval of pitch-label associations. In addition, Criscuolo et al. show that associations between musical experience and enhanced cognitive abilities, which are frequently reported among children, are also evident among adults after controlling for potential confounding variables including age, education, socio-economic status, and personality variables. Musicians show higher general intelligence, verbal intelligence, working memory, and attention than non-musicians, while amateur musicians score in between. It is notable that the reported correlations between years of musical playing and cognitive abilities support the hypothesis that musical practice is associated with intelligence and executive functions.

NEURAL MECHANISMS OF SENSORY AND COGNITIVE PROCESSING AMONG MUSICIANS

In addition to innovative behavioral and cognitive assessments, several investigators reported neurophysiological effects of engagement in auditory and musical processes among musicians. In one striking example, Müller and Lindenberger developed a method to study intra- and inter-brain synchronization, or so-called extended hyper-brain networks, by measuring phase synchronization between transformed acoustic recordings of guitar signals and raw EEG signals of guitarists freely improvising in duet. Of importance, this form of time-frequency analysis incorporates the dynamic interaction of both musicians’ production and responses to music. As reviewed by, Yurgil et al. the study of neural oscillations is ideal for investigating neural processes occurring over durations of time spanning seconds to minutes, and particularly suitable for investigating components of executive function such as temporal stages of working memory. In addition to studies of neural oscillations in relation to musical experience and executive functions, event-related potentials yield more temporally discrete information about cortical functions. As an example, Matsuda et al. report a dissociation of auditory cortical potentials in relation to music training and absolute pitch, indicating distinct stages of pitch processing in addition to hemispheric specialization of auditory cortical functions. Salvati et al. examined the functional connectivity among regions of the auditory pathway in processing natural, musical, and artificial non-speech sounds, by means of MEG. The different categories of sounds produced differences in activation and interconnection in prefrontal areas, anterior-superior temporal gyrus, posterior cingulate cortex, and supramarginal gyrus, demonstrating an extended brain network of human-made musical and artificial sounds when compared to natural sounds. A summary of behavioral, neuroimaging, and neurophysiological studies in musicians, non-musicians,

and clinical populations is taken up in the Perspective article by Koshimori and Thaut in which they discuss potential for executive function and attentional processing stimulation and neurorehabilitation.

SUMMARY

Taken together, the articles in this Research Topic represent significant advances in our understanding of relationships between musical training and control processes that comprise executive function. They illustrate the range of research questions that are being addressed by application of neuroscientific methods and technology to the universal cultural phenomenon of musical experience. They represent development of a model system for investigating long-term experience-dependent neuronal plasticity, and they bring into focus components of musical experience that may be targeted for specific therapeutic interventions aiming to benefit individuals across their lifespan.

AUTHOR CONTRIBUTIONS

All authors drafted, revised, and approved the final version of the manuscript for submission.

FUNDING

This work was supported by the Phyllis M. Taylor Center for Social Innovation and Design Thinking at Tulane University.

Conflict of 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.

Copyright © 2020 Colombo, Habibi and Alain. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Electrical Neuroimaging of Music Processing in Pianists With and Without True Absolute Pitch

Sélim Yahia Coll¹, Noémi Vuichoud¹, Didier Grandjean¹ and Clara Eline James^{1,2,3*}

¹ Neuroscience of Emotion and Affective Dynamics Laboratory, Faculty of Psychology and Educational Sciences and Swiss Centre for Affective Sciences, University of Geneva, Geneva, Switzerland, ² School of Health Sciences Geneva, HES-SO University of Applied Sciences and Arts Western Switzerland, Geneva, Switzerland, ³ Geneva Neuroscience Center, University of Geneva, Geneva, Switzerland

OPEN ACCESS

Edited by:

Claude Alain,
Rotman Research Institute (RRI),
Canada

Reviewed by:

Peter Schneider,
Universität Heidelberg, Germany
Lutz Jäncke,
University of Zurich, Switzerland

*Correspondence:

Clara Eline James
clara.james@hesge.ch

Specialty section:

This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Neuroscience

Received: 06 December 2018

Accepted: 07 February 2019

Published: 13 March 2019

Citation:

Coll SY, Vuichoud N, Grandjean D and
James CE (2019) Electrical
Neuroimaging of Music Processing in
Pianists With and Without True
Absolute Pitch.
Front. Neurosci. 13:142.
doi: 10.3389/fnins.2019.00142

True absolute pitch (AP), labeling of pitches with semitone precision without a reference, is classically studied using isolated tones. However, AP is acquired and has its function within complex dynamic musical contexts. Here we examined event-related brain responses and underlying cerebral sources to endings of short expressive string quartets, investigating a homogeneous population of young highly trained pianists with half of them possessing true-AP. The pieces ended regularly or contained harmonic transgressions at closure that participants appraised. Given the millisecond precision of ERP analyses, this experimental plan allowed examining whether AP alters music processing at an early perceptual, or later cognitive level, or both, and which cerebral sources underlie differences with non-AP musicians. We also investigated the impact of AP on general auditory cognition. Remarkably, harmonic transgression sensitivity did not differ between AP and non-AP participants, and differences for auditory cognition were only marginal. The key finding of this study is the involvement of a microstate peaking around 60 ms after musical closure, characterizing AP participants. Concurring sources were estimated in secondary auditory areas, comprising the planum temporale, all transgression conditions collapsed. These results suggest that AP is not a panacea to become a proficient musician, but a rare perceptual feature.

Keywords: complex music, electrical source imaging, ERP microstate analysis, trained pianists, true absolute pitch

INTRODUCTION

The present study applied electrical neuroimaging to investigate the influence of true absolute pitch on processing of complex classical music in pianists trained in the western classical repertoire. The research is a follow-up of 3 previous studies that showed progressive processing changes with musical training intensity and proficiency using the same musical stimuli (Oechslin et al., 2013; James et al., 2017; Jenni et al., 2017) and allows to answer the question whether true-absolute pitch offers additional advantages for music processing.

Absolute pitch (AP) perception is the rare faculty to identify or produce instantaneously any musical tone or other sound according to the 12-tone equal temperament, in the strictest sense

with semitone precision, without external reference (Miyazaki, 1988; Takeuchi and Hulse, 1993; Levitin and Rogers, 2005; Elmer et al., 2015). In AP possessors, this tone identification is performed automatically and cannot be suppressed (Akiva-Kabiri and Henik, 2012).

Levitin and Rogers (2005) propose an alternative more clement model of “latent” or quasi-absolute pitch (QAP). These authors could show that almost half of the non-musician Western population is able to sing pitches of familiar pop songs from memory with at least whole tone precision (Levitin, 1994). Although this finding is highly interesting in the context of musicality as a human universal, the current research focused exclusively on semitone precision AP with very high accuracy. This extreme or “true” form of AP also called AP-1 (Baharloo et al., 1998, 2000) or “genuine” AP (Bachem, 1937, 1955) constitutes a distinct and truly exceptional capacity, with an “all-or-nothing” quality (Profita and Bidder, 1988) that should be distinguished from QAP. Moreover, true-AP could be anatomically dissociated from both Relative Pitch (RP) and QAP, with larger left-right planum temporale asymmetry in true-AP possessors (Wilson et al., 2009).

Except for tone-deaf amusics (Ayotte et al., 2002; Tillmann et al., 2015), the rest of the population rather relies on RP and processes musical idiom contextually, comparing tones with surrounding ones. However, according to recent observations (Wengenroth et al., 2014; Ziv and Radin, 2014), both AP and RP capacities can be considered as continuous distributions, with a distinctive independent position in both distributions for each individual. So, in the end, any normal individual possesses both AP and RP to some extent (Zatorre, 2003; Wengenroth et al., 2014; Ziv and Radin, 2014).

Whether possessing AP is an advantage or even the panacea to become an exceptional musician is a matter of debate. AP can represent an advantage for developing musical abilities by strengthening internal representations, facilitating musical memory and for score reading, musical dictation or composing (Dooley and Deutsch, 2010). However, the faculty can also hamper proper functioning, as the automatic AP response may cause interference in the context of special tunings, ensemble playing or transpositions (Miyazaki and Rakowski, 2002; Zatorre, 2003; Wilson et al., 2009). According to Ziv and Radin (2014), AP may increase reaction times in the context of global processing of music as opposed to local processing.

AP only develops in those who receive musical tonal training from an early and before a critical age (Gregersen et al., 2001; Vitouch, 2003). As a consequence, classically reported percentages of AP in the general population (0.01%; Bachem, 1955; Ward, 1999) are probably underestimated, as no consensus exists on how to appraise latent AP in people lacking musical education. Finally, training is not sufficient to develop AP, which is also a function of genetic predisposition, often occurring in families (Baharloo et al., 2000; Zatorre, 2003). The estimated proportion is larger in highly trained Western musicians (Baharloo et al., 1998; Gregersen et al., 1999; up-to 15% and over), and much larger in Chinese musicians speaking Mandarin, a tone language (up-to 60% with semitone precision; Deutsch et al., 2006). These observations probably reflect the influence

of intensive tonal experience from birth on AP acquisition. However, some influence of genetic factors cannot be excluded in Asian populations (Gregersen et al., 1999).

Some authors propose a 2-level model of AP functioning (Levitin, 1994; Zatorre, 2003; Levitin and Rogers, 2005). The first level involves the representation of pitches in long-term memory and is essentially perceptual in nature. The second level, principally cognitive, involves the categorization or labeling of pitch according to the 12-tone temperament. These 2 levels may rely on distinct cerebral networks (Elmer et al., 2015). The left planum temporale might foster long-term pitch representation, allowing matching incoming spectro-temporal patterns with a template (Ohnishi et al., 2001; Griffiths and Warren, 2002). As a matter of fact, degree of left planum temporale activity correlated significantly with absolute pitch ability (Ohnishi et al., 2001). The left posterior dorsolateral prefrontal cortex would then perform the pitch-verbal associations (Zatorre et al., 1998). The bridge between these networks may be assured by functional connectivity in the left hemisphere that could be shown by theta band synchronization, likely via the arcuate fasciculus (Elmer et al., 2015). Wengenroth et al. (2014), using a unique AP test excluding the intervention of RP during the AP task, rather found an association between AP and the volume of the right Heschl's gyrus and proposed that a right-hemispheric network mediates AP perception. However, this study also accepted whole tone precision -albeit with weighted scoring-, and did thus not focus exclusively on true-AP. A recent publication explained observed differences concerning AP by the diversity of its definitions, and the lack of consensus concerning the tasks used to measure this faculty (Hou et al., 2016). In the current experiment we focused exclusively on true-AP possessors as compared to all other forms of pitch processing.

Most research on auditory processing in AP and non-AP (NAP) populations focused on the processing of isolated tones, chords or on oddball paradigms (McLachlan et al., 2013; Wengenroth et al., 2014; Ziv and Radin, 2014; Rogenmoser et al., 2015; Greber et al., 2018). However, it is in the context of complex dynamic musical contexts that AP is acquired and has its function. In the current study, we examined event-related brain responses to expressive polyphone musical stimuli, investigating a homogeneous population of young highly trained pianists with half of them possessing true-AP.

The musical stimuli were short expressive pieces for string quartet with 3 levels of harmonic transgression at closure that participants appraised, presented while high density EEG was recorded. We previously used these stimuli successfully to show the impact of different levels of musical training intensity on brain and behavioral responses using fMRI (Oechslin et al., 2013) and EEG (James et al., 2017), also exclusively in pianists. Both the fMRI and the EEG study results disclosed progressive changes in music processing as a function of musical training intensity and proficiency. The present study used EEG neuro-imaging to investigate whether differences in processing of these pieces would also occur as a function of absolute pitch possession within a homogeneous population of highly trained pianists. EEG neuro-imaging allowed to compare AP and NAP for Event Related Potentials (ERPs), functional ERP microstates

and underlying brain sources (Michel and Murray, 2012), providing millisecond precision information on the time course of brain processing.

Microstates are stable voltage topographies over time, lasting tens to hundreds of milliseconds, reflecting time windows of coherent synchronized activation of large-scale neuronal networks that represent basic physiological units of cognition (Lehmann et al., 1987; Murray et al., 2008; Brunet et al., 2011; James et al., 2017). Microstate analysis or spatio-temporal ERP analysis, compared to the isolated amplitude analysis of an array of single electrode sites, has the advantage of conserving the major part of the variance in the data by extracting the microstates in a multivariate way including all electrodes, groups and conditions in one unified analysis (Brunet et al., 2011; Michel and Murray, 2012; James et al., 2017). Classical ERP analysis, studying an array of isolated ERPs, although most effective in differentiating experimental conditions and groups at specific electrode sites (Kutas and Federmeier, 2011; Koelsch et al., 2013), only explores a small part of the variance in the data and does not explore the spatial dimensions of high-density EEG (Michel and Murray, 2012). Especially in the context of intrinsically different populations like in the current study, studying all electrodes and their underlying sources provides a more comprehensive view on the functioning of the brain. Moreover, changes in the spatial configuration of the microstates, indicate a change in the underlying cerebral sources (Vaughan, 1982; Michel and Murray, 2012), making them reliable precursors for EEG source imaging, thus allowing comparisons with other neuroimaging results. Although source localizations via EEG only allow centimeter precision, the combined temporal and spatial analyzes may shed new light on the time course of cerebral processing. Rigorous statistical thresholds may ensure reliability of these computations.

Some indications exist for possible cognitive advantages of AP outside the music domain, particularly for auditory span or short-term memory (STM; Deutsch and Dooley, 2013). In order to verify the possible influence of true-AP on general cognition, we also compared measures of auditory STM and working memory (WM) as well as the results of an auditory figure-ground perception (“hearing in noise”) test between the groups.

This experimental plan allowed us to examine the following hypothesis.

We anticipated that AP musicians would outperform NAP musicians for harmonic transgression sensitivity, following their automatic and infallible pitch analysis (Hypothesis 1), and that specific ERP patterns and microstates would occur in response to the harmonic transgressions. As this has never been investigated before in this population in a dynamic ongoing musical context, we cannot foresee the type of ERPs that may occur.

We expected both early perceptual differences in the ERPs and later more cognitive differences (Hypothesis 2), reflecting the 2-level model of AP functioning (Elmer et al., 2015).

We predicted influence of AP on general auditory cognition, with advantages for short term and working memory confirming the results of Deutsch and Dooley (2013) and possibly disadvantages for hearing in noise, as AP musicians automatically process all pitches (Hypothesis 3).

Finally, we hypothesized to find distinct brain sources, possibly early in time comprising the left planum temporale and later on in the prefrontal cortex in AP musicians (Hypothesis 4; Ohnishi et al., 2001).

MATERIALS AND METHODS

Participants

Twenty-four right-handed pianists accepted to participate in this study on a voluntary basis. They granted written informed consent, and received financial compensation after participation. The pianists were recruited from the Conservatoires of Geneva, Lausanne, and Neuchâtel. Among them 12 possessed true-AP (3 men; $M = 24.18$ years, $SD = 4.91$) and 12 did not (6 men; $M = 24$ years, $SD = 4.81$; see the Results section “Absolute Pitch Test” below for mean scores and confidence intervals of both groups). One AP participant was eliminated due to EEG artifacts. Final analyses were made on 11 true-AP and 12 NAP participants. Most of them also practiced other instruments or sang, 9 in the AP group and 7 in the NAP group. Secondary practices concerned mainly wind instruments and singing, but not string instruments. As tested by Student independent sample *t*-tests and chi-squared tests for independence (Table 1), true-AP and NAP participants did not statistically differ for age, gender, socio-economic status, age onset of musical practice, total number of years of piano training, instrumental level on the piano, practice of secondary instruments and years of practice on the secondary instruments.

Concerning the instrumental level (see Table 1), 13 pianists (5 AP, 8 NAP) obtained a certificate from the amateur sections of the above-mentioned Conservatoires, thus reaching the highest amateur level. The certificate implies up to 14 years of piano instruction, public final exams, 9 years of music theory closed by examinations, and 3 years of supplementary music instruction like composition or chamber music. The 8 pre-professionals (4 AP and 4 NAP) also obtained the certificate and were admitted in the preparatory classes of the Conservatoires and expected to enter the professional section in maximum 3 years. Finally, 2 AP participants already entered the professional section of one of the Conservatoires. No significant difference existed between the instrumental levels of AP and NAP participants [see Table 1; $\chi^2(2) = 4.23, P = 0.16$].

In the current experiment, the musical proficiency of the participants was confirmed by their elevated *d*-prime scores for detecting both apparent and subtle transgressions in the musical test (see section “Musical Test”).

All participants reported normal hearing and presented no history of neurological illnesses. The experiment was conducted at the Brain and Behavior Laboratory in the “Centre Médical Universitaire” in Geneva and approved by the local ethics committee.

Materials and Trial Sequence

Participants realized 6 tasks described in order of execution below.

TABLE 1 | Biographical data.

	AP	NAP	$t/\chi^2(df)$	<i>P</i>
Age	24.18 (5.15)	24 (5.03)	$t_{(21)} = -0.09$	0.93
Age onset of piano practice	6.86 (3.67)	7.29 (3.32)	$t_{(21)} = 0.29$	0.77
Total number of years of piano training	17.14 (7.87)	15.88 (3.86)	$t_{(21)} = -0.49$	0.63
Gender	Men: 3 Women: 8	Men: 6 Women: 6	$\chi^2(1) = 1.24$	0.26
Instrumental level	Certificated amateur: 5 Preprofessional: 4 Professional: 2	Certificated amateur: 8 Preprofessional: 4 Professional: 0	$\chi^2(2) = 4.23$	0.16
Socio-economic status	Compulsory school: 0 Matura/Baccalaureate: 2 University/Higher education: 9	Compulsory school: 1 Matura/Baccalaureate: 2 University/Higher education: 9	$\chi^2(2) = 2.65$	0.27
Practice of secondary instruments	Yes: 9 No: 2	Yes: 7 No: 5	$\chi^2(1) = 1.50$	0.22
Total number of years of secondary instruments' practice	7 (5.74)	11.64 (7.70)	$t_{(14)} = 1.39$	0.19

Listed are the means of the variables of both groups and the *p*-values (*t*-tests for the quantitative variables and chi-square test for independence for the qualitative variables). SDs are reported in brackets.

Edinburgh Handedness Inventory

Developed by Oldfield (1971) this questionnaire precisely assessed the laterality of each participant, providing a score between -100 (100% left-handed) and 100 (100% right-handed). This inventory comprises questions concerning handedness preference for 10 tasks, like writing or drawing.

Absolute Pitch Test

Created by Oechslin et al. (2010), this test is composed of 36 piano tones ($A = 440$ Hz), all repeated 3 times. Sinus tones are reported to be harder to detect than piano tones (Baharloo et al., 1998). However, Baharloo et al. found very similar scores in true-AP people for pure tones vs. piano tones. This can be explained by the “all-or-nothing” quality of pitch labeling of true-AP possessors (Profita and Bidder, 1988), for whom the timbre of the stimuli will not impact their pitch perception. Moreover, we consider piano tones to possess higher ecological validity (Elmer et al., 2013) and therefore more appropriate in this study on music processing. The resulting 108 tones were presented in pseudo-random order, with 18 Ω Klipsch Image S4 earphones. The tones were situated in the third, fourth and fifth piano octave reaching from C3 to H5. All 12 semitones of each octave were presented. Tones lasted 1 s and were separated by a 4 s inter-stimulus interval of brown noise. Tones and brown noise were generated using the “Cool Edit Pro” software (<http://cool-edit-pro.soft32.com/>). After each tone, participants immediately wrote down its exact label (with semitone precision), during the following 4 s window of brown noise. Each correct answer counted for 1 point, allowing a maximum score of 108 points.

At first, we recruited participants based on auto-reporting of semitone precision AP or the absence of it. For the final composition of the groups, after testing for AP, we retained the NAP group participants only if their score respected an upper limit of maximum 65% of correct answers for the 95% CI (Confidence Interval). This threshold was chosen as this percentage corresponds to the level at which adult individuals

without AP may temporarily bring their semitone precision detection of tones or learn to identify one single tone following intensive training (Brady, 1970; Ward, 1999; Van Hedger et al., 2015). For the true-AP group participants, the lower limit of the 95% CI was set to 90% (Baharloo et al., 1998). Based on previous studies, AP participants who scored more than 90% with semitone precision on an AP test appeared to be a special group (Miyazaki, 1988; Baharloo et al., 1998; Itoh et al., 2005; Schulze et al., 2009). The other AP possessors that were called “mid AP” (Itoh et al., 2005) had scores around 60% on an AP test but they never attained more than 90% of correct answers. Therefore, with a lower limit of the 95% CI fixed at 90% of correct answers for our AP group we can be confident that it represents true-AP possessors (for the sake of simplicity, from now on AP in the text).

Musical Test

While recording high density EEG (electroencephalogram), 90 different pieces for string quartet (2 violins, 1 alto and 1 cello), with a 10 s mean duration, were presented via EEG compatible earphones (Etymotic Research, Elk Grove Village, IL, USA) to all participants. We used string quartets in order to cancel out the “own instrument effect”, consisting in enhanced cortical responses to the timbre of a well-mastered instrument (Pantev et al., 2001; Margulis et al., 2009). A professional composer (see acknowledgments) developed the musical pieces specifically for our experiments, equally distributed over all 24 minor and major tonalities. The “Sibelius” (<http://www.sibelius.com/>) and “Logic Pro” (<http://www.apple.com/logic-pro/>) software served to create the compositions; natural instruments' timbres were implemented from the “Garritan Personal Orchestra” plug-in (<http://www.garritan.com/>). Musical styles of the pieces ranged from the Baroque to the Romantic period. The 90 original pieces were presented each with 3 distinct versions of the ending of the piece (see **Figure 1**), resulting in 270 stimuli. These 3 versions represent the 3

FIGURE 1 | Example of a stimulus used in the musical test, a string quartet in B-flat major. The 3 different closures are shown: R corresponds to a regular ending (degree I), T^{sub} to a subtle transgression (degree I^6) and T^{app} to an apparent transgression (degree IV^6). Matching sound-files are provided in the **Supplementary Information**. For more examples, we refer to Oechslin et al. (2013), James et al. (2017), and Jenni et al. (2017).

levels of our experimental condition “Transgression”: (1) tonic chord (I; regular ending, no transgression = R), (2) first inversion of the tonic chord (I^6 ; subtle transgression = T^{sub}) and (3) first inversion of the sub-dominant chord (IV^6 ; apparent transgression = T^{app}). Participants appraised whether closure of the pieces was appropriate, using a left mouse click if this was the case and a right mouse click otherwise. The stimuli were presented along with a fixation cross. All terminal chords of the pieces were cut off at 1,400 ms after onset and faded linearly over the last 150 ms (using Adobe Audition 3.0, Adobe Systems Corp., California, USA). The response screen appeared 1,900 ms after the onset of the terminal chord (stimulus onset) and 500 ms after the disappearance of sound, to avoid EEG contamination by the motor response. The order of presentation was fixed, but there were 2 versions of the task: Direct order with pseudo-randomized presentation, and the reversed order of the latter (in order to cope with possible order effects). Half of the participants of each pitch group completed the direct order and half the reverse order. The task consisted of 5 blocks of 54 pieces, separated by breaks in order to prevent fatigue.

The harmonic transgressions used here, that involve in-key chords that would make perfectly sense if the musical stimuli continued, clearly involve cognitive processing. They represent refined grammatical errors in the musical idiom and are by no means dissonances or out of-key elements that would principally involve perceptual processing. The 2 chords used for transgression are just as consonant (or dissonant, consonance and dissonance are 2 extremes of a continuum) as a tonic chord, both being composed of thirds, sixths, fifths and fourths).

For more comprehensive information on the stimuli and more musical examples we refer to 3 previous publications (Oechslin et al., 2013; James et al., 2017; Jenni et al., 2017).

Auditory Selective Attention Test

We created this test based on the figure-ground principle. Participants’ task consisted in comparing 2 series of sounds (“melodies”) lasting 500 ms, separated by one-second of brown noise, and reporting whether they were identical by means of a left (“identical”) or right (“different”) mouse click. The series of sounds, the “figures,” consisted of 7 tones, drawn at random (sampling without replacement) from the chosen range: The 12 semitones of the fourth octave of the piano. Together with each tone of the figure, a cluster composed of all 12 semitones played simultaneously (ground), with slightly less amplitude than the figure tones. Therefore, the participants had to compare the series in the presence of background noise, demanding auditory selective attention, even more so as the series were not arranged within a tonality, but randomly drawn. This random arrangement limits the grouping of several notes, hampering memorization of the series. In total, the test comprised 50 different series of 2 times 7 tones, of which each appeared in an “identical” and “different” version resulting in 100 stimuli. In a “different” version, 1 tone among the 7 differed from the first vs. the second series, a semitone higher (23 trials) or lower (27 trials). The modification was applied to one of the middle 5 tones, excluding the first and last one of the series (to avoid facilitating serial position effects). The series of piano-like tones were generated with the “Cool Edit Pro” software (<http://cool-edit-pro.soft32.com>). Participants received 1 point for each correct answer for a total of 100 points. This test can be succeeded at 100% by means of relative pitch alone, as the participant has to compare interval series.

WAIS-IV Digit Span

In order to test auditory short-term memory and working memory, participants accomplished the “digit span” subtest of the Wechsler adult intelligence scale 4th edition (Wechsler, 2011)

in direct and reverse order. During these tests, participants listened to spoken series of digits with increasing length (from 2 to 8 digits), and their task consisted in repeating them orally first in direct (testing primarily auditory short-term memory) then in reverse order (testing primarily auditory working memory). Two series of digits (1 for each task), progressively increasing by level of difficulty (length), were presented. Spoken digits were recorded beforehand at a rate of 1 digit per second to guarantee identical conditions for all participants. Participants completed first the direct order, then the reverse order task. They proceeded through the trials until they made 2 successive mistakes at the same level of difficulty. Each correct answer was given 1 point for a total of 16 points per task.

Loudness/Intensity of the Auditory Stimuli

We normalized all auditory stimuli to an identical amplitude envelope (−5 dB) with the “Cool Edit Pro” software (<http://cool-edit-pro.soft32.com>). Participants determined the optimal loudness individually during stimulus presentation, therefore we do not report intensity/dB values for the presentation of the stimuli. As the size and shape of the ear cup and canal varies strongly between individuals, inserting the earphones more or less deeply in the ear canal impacts the distance to the eardrum and thus influences perceived loudness. As the participants were all trained musicians, they are used to optimize hearing using earphones.

EEG Acquisition and Raw Data Processing

Participants were comfortably seated in the EEG room in front of a computer screen with approximate 1 m distance from the eyes. Continuous EEG was acquired using the “Biosemi” system (Amsterdam, The Netherlands) comprising an ActiveTwo amplifier system AD-Box with 64 active AG/AgCL electrodes, sampled at 1,024 Hz in a bandwidth filter of 0–268 Hz. The average of all electrodes was used as an online and offline reference. Meanwhile we recorded electro-oculography (EOG) and applied the voltage difference of 2 horizontal (HEOG) electrodes fixed at the outer canthi side of both eyes, to detect horizontal eye movements. Two additional electrodes were placed above and under the right eye to measure the voltage difference (VEOG) caused by blinks.

Raw data pre-processing was performed using the “BrainVision Analyzer 2.0” software (Brain Products GmbH, Munich, Germany). Offline, data were band-pass filtered between 0.25 and 30 Hz within a Butterworth zero phase filter. An additional 50 Hz notch filter removed noise from the power-line. Finally, given that the low-pass filter was set to 30 Hz, data were down-sampled to 256 Hz to reduce data size, thus processing time. Artifact rejection was applied to exclude voltages above 100 μ V and under −100 μ V. Independent component analysis was used to remove ocular artifacts and reduce deformation at the edge of the epochs. Finally, a baseline correction of −200 ms to stimulus onset was performed. On average 8.82% ($SD = 11.38\%$) of the trials per experimental condition were removed. Data of one AP participant were excluded due to excessive artifacts.

Grand-Averages for the AP group were based on 80.00 ± 13.97 SD epochs for R, 81.00 ± 12.55 SD epochs for T^{sub}, and 77.73 ± 15.82 SD epochs for T^{app}. For the NAP group, 84.42 ± 5.23 SD epochs were kept for the Grand-Averages for R, 84.25 ± 3.49 SD epochs for T^{sub}, and 84.33 ± 4.42 SD epochs for T^{app}.

Noisy channels were interpolated, by means of a 3-D spherical spline, for on average 0.34% ($SD = 1.05\%$) of the electrodes.

ERP Analyses

The evoked potentials were analyzed in 3 stages described in detail below.

Classical Waveform Analysis

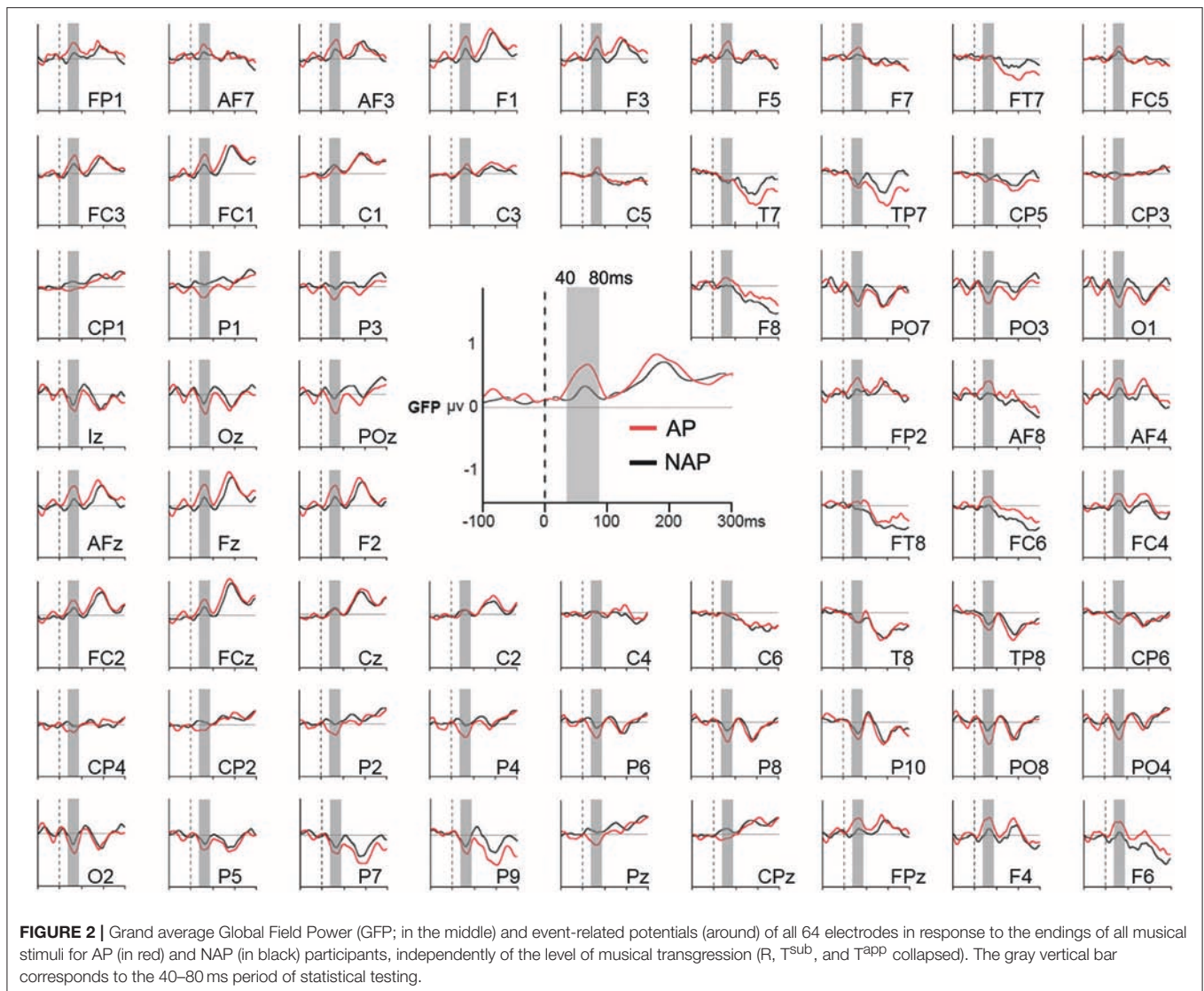
Our main interest lies in the microstate and source analyses that provide more comprehensive information, as they comprise all electrodes. Therefore, as multiple testing over our period of analysis (0–1,000 ms) only showed marginal differences for individual electrodes between AP and NAP for all 3 transgression conditions, we limited our waveform statistics to an analysis of Global Field Power (GFP), all 3 transgression conditions collapsed. GFP provides one single, always positive, reference free value representing the neural response strength throughout the brain, incorporating absolute voltages of all electrodes (Murray et al., 2008). This choice was made because during a very early time period, a generalized effect of pitch group manifested at almost all electrodes in all conditions (Figure 2).

We did not compute difference waves because they do not allow to compute microstates and source estimations. Moreover, as only marginal differences occurred for all 3 transgression conditions between AP and NAP, we collapsed all conditions to study the above mentioned early generalized brain response.

Microstate Analysis

Evoked potentials and spontaneous EEG display over time, as scalp voltage topographies remaining relatively stable for tens to hundreds of seconds, and represent microstates of information processing “atoms of thought” (Pascual-Marqui et al., 1995; Murray et al., 2008; Brunet et al., 2011).

Microstate analysis or spatio-temporal ERP analysis consists of 2 stages. During the first stage, a *k*-means cluster analysis applied to the grand-means of both conditions and groups defined the most dominant scalp topographies or microstates (Murray et al., 2008). In order to study robust effects, we chose to group microstates sharing over 92% of variance and lasting more than 20 ms. To define the optimal number of microstates we relied on the cross-validation index (Pascual-Marqui et al., 1995), which minimizes the residual variance, and on a modified “Krzanowski–Lai criterion,” a quality measure for clustering (Michel et al., 2004; Tibshirani and Walther, 2005; Murray et al., 2008; Michel and Brandeis, 2009; Brunet et al., 2011). The resulting microstate series represent an a priori hypothesis, to be statistically verified in the second stage. This second stage consists in “fitting” the microstates back over time across all individual subjects, by means of a spatial correlation analysis, resulting in a parameter of presence, Duration (in ms) and of Global Explained Variance (GEV; in %) of each microstate for a period of interest.



Statistical testing of these parameters then verifies differential microstate presence over time comparing groups or conditions.

We started by comparing microstates between the 2 groups (AP vs. NAP) for all 3 conditions of musical transgression (R , T^{sub} , and T^{app}). One single precocious distinct microstate occurred in all 3 conditions principally in the AP group. No other statistically significant differences occurred as a function of one of the transgression conditions over the further time period of analysis (up to 1,000 ms after stimulus onset). Therefore, we focused on this early microstate dissociating AP and NAP groups, averaging the data over all 3 transgression conditions, and recomputing a microstate analysis only on the Grand-Average ERPs of the 2 groups (R , T^{sub} , and T^{app} collapsed). As the aim of this experiment was to study ERP differences as a function of pitch group, we will only focus on the latter analysis in the present paper.

The microstate analysis was performed with the Cartool software, developed by Denis Brunet (<http://brainmapping.unige.ch/cartool>).

Source Analysis

Unlike the direct solution that consists in calculating voltage configurations on the scalp from known underlying sources, the inverse problem of computing the underlying hypothetical sources from the measured voltages on the scalp is not unique or rather “ill-posed.” To deal with this “ill-posed” inverse problem with surface EEG, distributed source analysis should be based on well-defined a priori constraints (Michel et al., 2004; Grech et al., 2008). Our a priori constraints derive from the microstate analysis that allowed determining a period during which distinct microstates appeared in both groups. Differences in scalp topographies over time indicate a change in the underlying generators (Lehmann et al., 1987; Murray et al., 2008), which relates to changes of the brain’s functional state (Michel et al., 1999).

As our main interest lies in exploring differences between the AP and NAP group for auditory processing of music, we focused on ERP sources regarding the primary outcome of the microstate analysis, i.e., the early time period during which 2 spatially

different microstates appeared in each group. These microstates appeared during the first 100 ms of processing of musical closure, in all experimental conditions. Microstates reflect stable scalp topographies and underlying brain networks, and maps' stability is maximum during the global field power ("GFP") peak (Koenig et al., 2014). In addition, the signal-to-noise ratio increases with the strength of the GFP (Lehmann et al., 2005). Therefore, we limited the time window of analysis around the GFP peak of the microstate characterizing the AP group (40–80 ms), the NAP group showing no marked GFP peak. This approach of centering source analyses on the GFP peak has been successfully used in different settings (Plomp et al., 2013; James et al., 2015).

We estimated the intracranial distribution (microampere per cubic millimeter, $\mu\text{A}/\text{mm}^3$) of the averaged ERP for each subject across all experimental conditions over the 40–80 ms period of interest, using an inverse linear solution of depth-weighted minimum norm (Hamalainen and Ilmoniemi, 1994; Michel et al., 2004; "WMN"). The WMN algorithm compensates the tendency of minimum standard algorithms to favor weak and superficial sources by using a weighted matrix (Grech et al., 2008). In combination with a statistical parametric mapping (SPM) approach, consisting of a statistical comparison of the underlying sources between groups, a considerable amount of noise is annihilated. In addition, the SPM method avoids the domination of maxima of the source that are identical in both groups (Michel et al., 2004). The intracranial distribution was calculated within a discrete 3-D grid of 3005 solution points, evenly distributed through the volume of gray matter of the Montreal Neurological Institute ("MNI") standard brain. By means of a homogeneous transformation that adapted the volume to the most appropriate sphere ("SMAC model"; Spinelli et al., 2000) a 3-shell spherical model was applied to calculate the lead field for the 64 electrodes and the inverse solution based on the constraints of the WMN. Then, to eliminate noise and to compare the groups, we applied the SPM model using 2-sided *t*-tests at each node, by means of the "Statistical Toolbox for Electrical Neuroimaging" (STEN; <https://doi.org/10.5281/zenodo.1164037>; courtesy of Jean-François Knebel). Results of the SPM consist of stronger current density in certain brain voxels for either group (positive vs. negative *t*-values).

In order to cope with multiple comparisons, we used conservative statistical thresholds and only considered as significant globular clusters of at least 20 contiguous nodes each with a *p*-value of ≤ 0.01 (Knebel et al., 2011; De Meo et al., 2015). The spatial criterion of 20 nodes was determined through the "AlphaSim" software (<http://afni.nimh.nih.gov>) with a 10 mm FWHM smoothing and 10 mm radius of connected poles, together with the node *p*-value of < 0.005 that we applied. Following a bootstrap procedure (10,000 Monte Carlo iterations) a 21-node cluster appeared, with a 0.001 cluster-level probability and a 0.00001 *p*-value corresponding at the node level. Spatial accuracy of these source estimations was thoroughly investigated in the past and showed that these approaches provide similar results compared to fMRI studies, only with a lesser precision range of 1–2 cm (Martuzzi et al., 2009; Plomp et al., 2010; Birot et al., 2014). The WMN source analysis was performed with the Cartool software (brainmapping.unige.ch/cartool).

A common misunderstanding interpreting these analysis techniques is that the source computations are applied to the microstates. The microstate analysis is a separate analysis stage, and also serves to determine the appropriate time periods for the source analysis, that is then applied to the ERPs. The latter are genuine neurophysiological data, whereas the microstate data are the fruit from a statistical cluster analysis.

RESULTS

For the sake of brevity and transparency statistical results can be found mainly in the tables, and as little as possible in the text. Note that in order to compute effect sizes for our generalized linear mixed model (McCulloch, 2003; Bolker et al., 2009; GLMM) analyses, we used Nakagawa and Schielzeth's method (Nakagawa and Schielzeth, 2013) implemented in the "MuMin" R package. Their method is based on 2 indicators: A marginal (R^2_m) and a conditional (R^2_c) R^2 . R^2_m is the variance explained by the fixed factors, whereas R^2_c is the variance explained by the entire model (both fixed and random effects).

Behavioral Results

Edinburgh Handedness Inventory

The Edinburgh Handedness Inventory confirmed that all participants were right-handed. Moreover, as verified by a Student independent sample *t*-test comparing the 2 groups (AP vs. NAP) there was no significant difference between the AP ($M = 79.43$; $SD = 23.45$) and NAP ($M = 86.75$, $SD = 10.80$) group concerning handedness level [$t_{(21)} = 0.98$, $P = 0.34$, $d = 0.40$].

Absolute Pitch Test

Concerning the absolute pitch test, we applied a GLMM specifying a binomial model with Group (AP vs. NAP) as fixed factor, and participants' performance as random factor. These analyses demonstrated that the AP group was significantly more accurate than the NAP group (see **Figure 3**, **Table 2**). The test

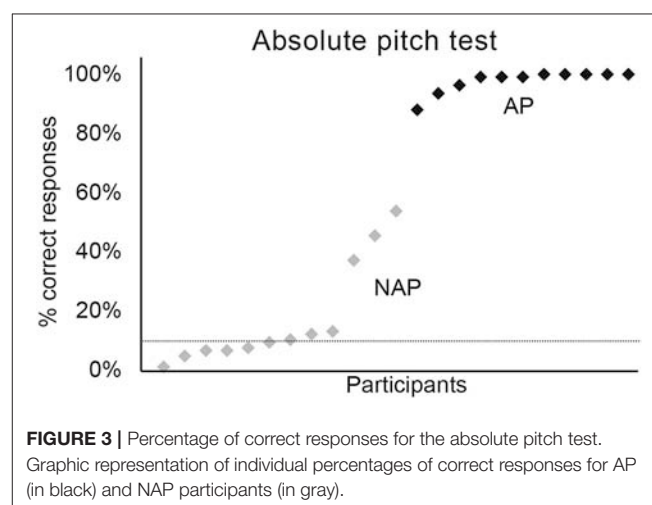


TABLE 2 | Percentage of correct responses for the absolute pitch test.

Effect	AP	NAP	χ^2	<i>P</i>	R^2_m	R^2_c
	Mean accuracy (95% CI)	Mean accuracy (95% CI)				
Group	99% ([97, 100])	12% ([6, 23])	42%	<0.001	0.69	0.80

yielded a large effect size [$R^2_m = 0.69$; (Cohen, 1988; Calin-Jageman, 2018)]. The NAP groups showed a less homogeneous performance than the AP group. Indeed, 9 NAP scored <20% on average and 3 between 30 and 60%. As we intended to oppose true-AP possessors against all other types of pitch perception, the inclusion of 3 quasi-absolute pitch (QAP) possessors seems legitimate (Wilson et al., 2009).

Musical Test

In order to investigate differences in performance at the musical test between AP and NAP participants, we calculated the rater sensitivity index “d-prime” for T^{sub} and T^{app} . The d-prime index is a statistic used in signal detection theory (Macmillan and Creelman, 1997), preventing impact of response biases, by comprising both “hits” (correctly detecting that T^{sub} and T^{app} endings are not appropriate) and “false alarms” (incorrectly detecting that R endings are not appropriate) in the calculation of task performance. The d-prime index thus incorporates the responses to the R (regular) endings. A higher d-prime value reflects better discrimination between transgressed and regular endings. Comparison of the d-prime parameter for AP ($M = 2.28$; $SD = 1.19$) and NAP ($M = 2.22$; $SD = 1.31$) participants for T^{sub} , using a Student independent sample *t*-test, revealed no significant difference between the groups [$t_{(21)} = -0.11$, $P = 0.91$, $d = 0.05$]. The same analysis showed no significant difference between AP ($M = 3.77$; $SD = 1.26$) and NAP ($M = 3.82$; $SD = 1.07$) participants for T^{app} [$t_{(21)} = 0.11$, $P = 0.91$, $d = 0.04$].

The high d-prime scores in both groups confirm the musical and specifically harmonic skills of our participants. Compared to our previous study comparing expert pianists, amateur pianists and non-musicians (James et al., 2017) mean d-prime scores of the population in the current study for the subtle (difficult to detect) transgressions for the AP ($M = 2.28$; $SD = 1.19$) and NAP group ($M = 2.22$; $SD = 1.31$) are situated in between those of amateur ($M = 1.54$; $SD = 1.00$) and professional pianists ($M = 4.19$; $SD = 0.78$) in James et al. (2017).

Auditory Selective Attention Test

GLMMs with the group (AP vs. NAP) as fixed factor and participants’ performance as random factor were run. The accuracy at the auditory selective attention test was not significantly different between the AP and NAP group (see Table 3). Two Student one sample *t*-tests demonstrated that the score for the AP group [$t_{(10)} = 6.25$, $P < 0.001$, $d = 6.26$] and NAP group [$t_{(11)} = 5.36$, $P < 0.001$, $d = 1.55$] was significantly different (i.e., above) from chance level (50%). Before analyzing reaction time data (ms), a log base 10 transformation of them was applied to correct for a floor effect and equalize the variances.

Results showed that NAP participants were marginally faster than the AP participants (see Table 3).

WAIS-IV Digit Span

A GLMM with the group (AP vs. NAP) as fixed factor and participants’ performance as random factor was performed on the scores of the direct order digit span subtest. The analysis revealed that AP participants’ scores were not significantly different from those of the NAP participants (see Table 3). There was again no significant difference between the AP and NAP group concerning the indirect order digit span subtest (see Table 3).

EEG Results

Classical Waveform Analysis

A 2-tailed unpaired *t*-test comparing mean GFP over the time window 40–80 ms between AP and NAP groups, all 3 transgression conditions collapsed (i.e., R, T^{sub} , and T^{app}), yielded a significant result, with greater strength for AP ($M = 0.71 \mu\text{V}$, $SD = 0.18$) than for NAP ($M = 0.49 \mu\text{V}$, $SD = 0.23$) participants [$t_{(21)} = 2.47$, $P = 0.02$, $d = 1.07$]. With $d = 1.07$, the effect size is large (Cohen, 1988; Calin-Jageman, 2018).

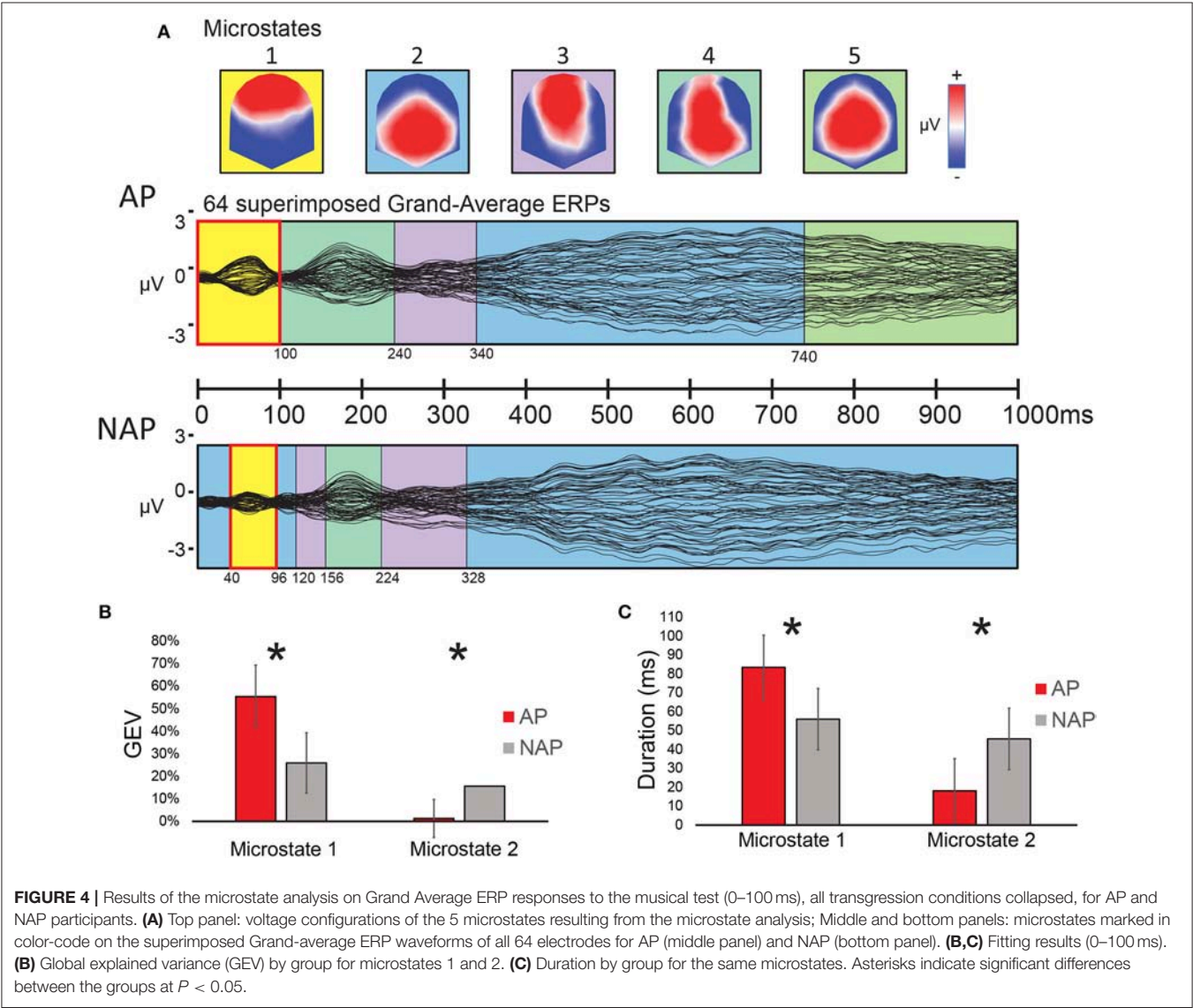
As we collapsed all 3 conditions (R, T^{sub} and T^{app}), the GFP Grand-Averages of AP and NAP were based on 239.18 ± 40.19 SD epochs for AP participants, and 253.75 ± 13.20 SD epochs for NAP participants.

Microstate Analysis

A spatio-temporal analysis between 0 and 100 ms on the Grand-Average ERPs of the 2 groups (stage 1, see “Classical Waveform Analysis”) yielded a solution with 5 stable microstate map configurations over time, independently of experimental condition (i.e., R, T^{sub} , and T^{app} collapsed). The series of microstates appearing over time are depicted in Figure 4A. These 5 microstates explained 96% of the ERP data variance. Visual inspection of this microstate series at the grand mean level showed a similar sequence of microstates for both groups except from 0 to 156 ms and from 740 to 1,000 ms. Only for the early period significant statistical differences could be found. Between 0 and 100 ms after stimulus onset, a frontal positivity (microstate 1) characterized the AP participants, whereas for NAP participants, both microstate 1 and microstate 2, the latter showing posterior positivity, occurred. Interestingly microstates 1 and 2 manifested almost opposite voltage configurations. Statistical analyses (stage 2, “fitting” see “Microstate Analysis”) on the microstate parameters Duration and GEV between 0 to 100 ms, a period of difference also observed in our ERP data (see Figure 2), reached significance. Because the microstate

TABLE 3 | Summary of the Group results (AP vs. NAP) at the auditory selective attention test and WAIS IV direct and indirect order subtests.

	AP	NAP	χ^2	P	R^2_m	R^2_c
	Mean accuracy (95% CI)	Mean accuracy (95% CI)				
Auditory selective attention test	66% ([61; 71])	66% ([60; 71])	0.01	0.91	<0.001	0.30
WAIS IV direct order	80% ([72; 86])	78% ([71; 84])	0.12	0.73	<0.001	<0.001
WAIS IV indirect order	78% ([70; 84])	77% ([70; 83])	0.01	0.91	<0.001	<0.001
	Mean RT (ms)-log10 (95% CI)	Mean RT (ms)-log10 (95% CI)	χ^2	P	R^2_m	R^2_c
Auditory selective attention test	6.23 ([6.04; 6.42])	5.97 ([5.78; 6.16])	3.34	0.07	0.03	0.19



Durations were not normally distributed, we opted for non-parametric testing for both parameters GEV and Duration. A Kruskal-Wallis statistical comparison based on the fitting of these maps into the individual data showed a significant difference of the parameter GEV between the groups (AP vs. NAP) for microstate 1 [$H(1, N = 23) = 4.91, P < 0.05, d = 0.95$] and microstate 2 [$H(1, N = 23) = 4.46, P < 0.05, d = 0.89$; see **Figure 4B**]. A significant difference between groups

was also found for Duration in ms for microstate 1 [$H(1, N = 23) = 4.34, P < 0.05, d = 0.87$] and microstate 2 [$H(1, N = 23) = 4.34, P < 0.05, d = 0.87$; see **Figure 4C**]. These results disclosed that, in the 0–100 ms window of interest, microstate 1 principally characterized AP participants, whereas microstate 2 characterized NAP participants. Effect sizes, varying between $d = 0.87$ and $d = 0.95$ are large (Cohen, 1988; Calin-Jageman, 2018).

Source Analysis

The a priori constraint for determining the time period for comparing AP and NAP participants for underlying brain sources of the surface EEG was the simultaneous occurrence of 2 distinct microstates (1 and 2; see **Figure 4A**) in both groups in all experimental conditions in the first 100 ms after stimulus onset. These microstates derive from the microstate analysis comparing the 2 groups, over all experimental conditions (R, T^{sub}, and T^{app} collapsed). The scalp configurations of these 2 microstates representing each essentially 1 pitch group, displayed quasi-opposite voltage configurations at the scalp level, with a clear GFP peak for AP possessors for microstate 1. Therefore, this time period provides a strong precursor for computing inverse solutions. To cancel out as much noise as possible, we centered our analysis around the GFP peak of microstate 1 in the AP group (40–80 ms), and applied SPM, comparing the groups with 2-tailed *t*-tests for all 3,005 nodes of the 3-D grid, using the STEN software.

One single 21-node globular cluster passed our statistical thresholds (see section “Source Analysis”) characterizing the AP group (positive *t*-values), in left secondary auditory areas located around the temporo-parietal junction, encompassing part of the planum temporale (**Figure 5**). The peak *t*-value was found in the gray matter of the left posterior superior temporal gyrus [$t_{(21)} = 3.65, P < 0.002, d = 0.76$; **Figure 5**, middle upper panel], the other areas composing the cluster comprised parts of the supramarginal gyrus and the rolandic operculum. With $d = 0.76$ we report a close to large effect size here (Cohen, 1988; Calin-Jageman, 2018).

DISCUSSION

The key and novel finding of this study is that a precocious ERP microstate, peaking as early as 60 ms after closure onset in expressive music, dissociated AP pianists from NAP pianists, in response to all musical stimuli, independently of the harmonic congruence of the musical closure (partially confirming Hypothesis 2). Later more cognitive cerebral processing of harmonic transgressions in music turned out alike in AP and NAP pianists. Identical rater sensitivity (*d*-prime) for musical transgressions and equal STM and WM scores underpin this observation.

Despite the absence of a behavioral difference for the harmonic transgression test, the robust behavioral difference for the AP test evidences that we are dealing with 2 intrinsically distinct populations, that process the musical stimuli differently, resulting, however, in a similar level of transgression detection proficiency.

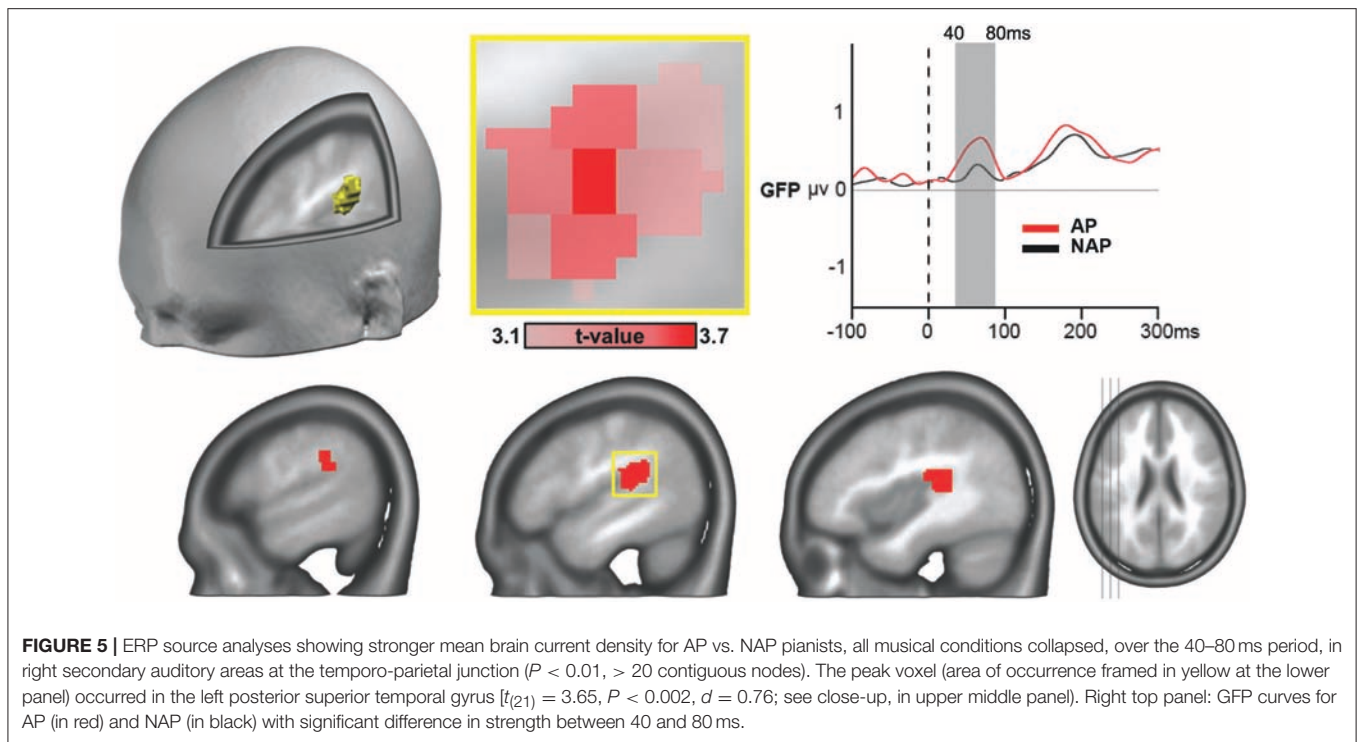
According to the 2-level model of AP functioning (Levitin, 1994; Zatorre, 2003; Levitin and Rogers, 2005), this precocious microstate involves the perceptual representation of pitches in long-term memory and not the semantic labeling. Given its very early appearance, this stage of information processing apparently consists in a highly automatized pitch analysis capacity, which can be dissociated from later more cognitive processing of music. Our source analyses revealed that this microstate concurred in AP musicians with involvement of left secondary auditory areas at the temporo-parietal junction, comprising the planum temporale, corresponding to those described in the literature distinguishing AP possessors (Schlaug et al., 1995; Ohnishi et al., 2001; Jancke et al., 2012; partially confirming Hypothesis 4). Moreover, Elmer et al. (2015) hypothesized that specifically the left planum temporale (PT), that is part of the cluster the current study localized, fosters the first perceptual component of AP according to the 2-level model, matching incoming spectrotemporal patterns with a template (Ohnishi et al., 2001). These left auditory related areas may function as a “hub of information flow” toward inferofrontal areas (Elmer et al., 2015, p. 369).

These results also constitute another confirmation of the validity of source analysis from scalp EEG (Plomp et al., 2010; Birot et al., 2014), while adding a temporal precision that other imaging techniques cannot provide (Michel et al., 2004; Meyer-Lindenberg, 2010; Michel and Murray, 2012).

To the best of our knowledge, our study is the first to highlight such an early difference between AP and NAP participants. The earliest differences reported in previous studies concerned the N150 ERP component (Itoh et al., 2005), and the MEG M100 (Hirata et al., 1999). Given the component’s time of occurrence, and the fact that it appears much earlier than the MMN (mismatch negativity) which reflects pitch memorizing, and the P3a, the latter probably reflecting the labeling process (Rogenmoser et al., 2015), we suppose that it highlights an automatic and unconscious sensorial process.

The fact that we were able to identify specific sources underlying AP function in left secondary auditory areas, including the PT at this early point in time, is noteworthy. The assumed high degree of automation of this special pitch processing and identification capacity is strongly supported by this result. This processing stage thus precedes cognitive processing of complex music.

The literature showed that rather pruning in the right PT than development of the left PT explains the observed left-right PT asymmetry in AP possessors (Keenan et al., 2001). This underdevelopment of the right PT then favors functional dominance of the left PT during early development in AP individuals. The here observed pitch processing likely performed by the left PT and surrounding areas undoubtedly represents only part of the AP faculty, i.e., the acoustical part. Left frontal areas, more precisely the left posterior dorsolateral prefrontal areas, may carry the “linguistic labeling” of the acoustically identified pitches (Zatorre et al., 1998; Ohnishi et al., 2001). Our study did not highlight any frontal sources, which can be explained by the SPM analysis technique that annihilates all common activations in both groups and only shows differences. The NAP musicians



in the current study, that are also highly trained musicians, likely also rely on left dorsolateral prefrontal areas when making relative pitch judgments of chords, relevant for the task in the current experiment (Zatorre et al., 1998; James et al., 2017). Likewise, differences in right auditory cortex activation did not reach statistical significance either as AP and NAP participants probably both relied on these brain areas for chord processing.

Two studies investigated brain connectivity in AP possessors. Enhanced structural brain connectivity in AP people was described by Loui et al. (2011) and involved tracts connecting more strongly the left superior temporal gyrus with the left middle temporal gyrus. Volumes of these tracts predicted AP performance. Furthermore these authors observed hyperconnectivity in bilateral superior temporal lobe structures in AP possessors. A recent study by Brauchli et al. (2019) observed enhanced fine-grained functional connectivity in the left auditory cortex in AP musicians using multi-voxel pattern analysis. These studies do not refute our findings.

Our study deviates however from 2 important publications in the domain of comparison between AP and non-AP musicians. First, we did not find right Heschl's involvement in AP processing like Wengenroth et al. (2014). This may be explained in the first place by the different time window studied with fMRI, not allowing to detect a precocious effect like we could in the current study. Secondly, the AP definition slightly differed, as the authors applied an innovative AP test, excluding any RP involvement, moreover, also accepting whole tone precision -albeit with weighted scoring. In the test used here, relative pitch perception is involved. Additionally, in Wengenroth et al. (2014), tones and not expressive

music served as stimuli, and the participants played different instruments. As stated before, observed differences concerning AP may be explained by the diversity of its definitions, and by the absence of consensus concerning its measurement (Hou et al., 2016).

Within a passive listening Oddball paradigm, Rogenmoser et al. (2015) found identical MMN responses in AP and NAP musicians but different P3a responses, concluding that early auditory processing did not differ between the groups, but later more cognitive processing did, with AP showing smaller P3a amplitudes. We report opposite results, with a very early difference arising from secondary auditory areas at 60 ms and no later differences in dorsolateral prefrontal areas. This difference in observations may be explained by the different stimuli used, and also by the criteria for AP. Rogenmoser et al. (2015) tested their participants on out-of-tune tones in a passive listening Oddball paradigm. In that case the automatic searching for the correct label seems triggered in AP possessors, because out of tune notes may induce doubt on the exact pitch of the note, despite the passive listening, as AP people process pitch automatically and involuntary. In our active harmonic transgression task, searching for the correct label is not prominent, searching for the correct syntactical function of the chord is, implicating also relative pitch processing. Moreover, in Rogenmoser et al. (2015), AP possessors provided at least 40% of correct responses in the AP test, vs. 90% in our study. So several of our NAP participants would have been categorized as AP in their study and their results did not restrain to true-absolute pitch processors. Finally, in a replication study of Rogenmoser et al. (2015), performed by Greber et al. (2018) on a larger sample of

musicians, the findings on a weaker P3a in AP possessors were not confirmed.

An earlier EEG study by our research group compared participants with 3 levels of training intensity for the 3 transgression conditions (James et al., 2017). In that study, non-musicians, amateur pianists and expert pianists showed pronounced differences in ERP responses and source estimations between the 3 transgression conditions. Interestingly, in the current experiment with 2 groups not differing in training intensity, no such differences were found, although we expected this initially. As, moreover, no differences manifested for transgression detection at the behavioral level, these results suggest that no important differences exist between equally trained AP and NAP pianists for cognitive music processing as involved in our task (infirming Hypothesis 1).

We initially hypothesized that AP participants would perform better than NAP participants at auditory STM and WM tasks. Concerning STM, we expected to replicate Deutsch and Dooley's results (2013), who used similar stimuli and task: Series of digits that participants had to repeat in direct order, for which they observed an advantage for AP participants. In contrast, we observed similar scores in both groups for STM and WM (partially infirming Hypothesis 3). These contradictory results may be explained by differential recruiting of participants and the tests used. Deutsch and Dooley tested 7 APs against 20 NAPs. The instruments played were not specified, thus probably varied. Handedness was not reported. In the current experiment groups were of similar size and composed of right-handed pianists only. Finally, we used the WAIS-IV normed subtests to assess STM and WM, whereas Deutsch and Dooley used a test that they created for their experiment. Interestingly, Benassi-Werke et al. (2012) using the direct and indirect order digit span subtests of the WAIS-III like in the current experiment, did not report any significant differences either between two groups of professional singers with and without AP and a group of amateur NAP singers. However, when testing on tones instead of digits, the AP group performed significantly better. STM and WM advantages possibly only manifest clearly when the test material consists of tones.

As already explained in the introduction, possessing AP may also yield disadvantages in musical contexts. Although not present at the level of correct responses, AP participants responded marginally later than NAP participants in our in-house auditory selective attention test (partially confirming Hypothesis 3). This may be explained by the fact that because AP possessors automatically process all pitches they hear (Akiva-Kabiri and Henik 2012), they are unable to ignore the "ground," therefore slowing down the processing of relevant information for the task (comparing the tone series). Defining and confirming more precisely domains of disadvantage of the AP skill represents an interesting domain for further research.

LIMITATIONS

As the study population consisted exclusively of pianists, all conclusions also restrain to this particular population. Pianists perform highly complex polyphone scores as compared to

melody instrumentalists like string or wind players. Therefore, NAP pianists are also very proficient in the analyses of chords, which may partially explain why no differences occurred for music syntactic processing as a function of AP.

Concerning source computations, since we did not use digitized electrode positions nor individual MRIs, our source localizations are approximately 10% less sensitive (Brodbeck et al., 2011; James et al., 2017) relative to studies that used such individualized techniques (Biro et al., 2014; Megevand et al., 2014; Plomp et al., 2015).

Finally, the sample size of this study is relatively small. From our point of view there is no reason to suspect that the current study is underpowered due to a lack of sensitivity. As we report large effect sizes for all results (absolute pitch test, classical waveform analysis on GFP, microstate analysis and source analysis), we fulfill the recommendations of Whitehead et al. (2016), with at least 10 participants per group for studies with large effect sizes, in order to obtain 90% power and two-sided 5% significance.

CONCLUSIONS

To summarize, the key and novel finding of this study is the revelation of a precocious microstate (peaking around 60 ms after stimulus onset) characterizing AP participants in response to musical closure in complex polyphone music. Given its very early occurrence and underlying brain sources at this time point in secondary auditory areas, this stage of information processing clearly involves the acoustical features of AP functioning. Specific cerebral sources for AP possessors contributing to this microstate were localized in secondary auditory regions at the temporo-parietal junction, including parts of the PT. These brain areas correspond to those described in the literature to distinguish AP possessors at the anatomical and functional level (Schlaug et al., 1995; Keenan et al., 2001; Itoh et al., 2005). In contrast, AP did not facilitate later musical cognition or general cognition. These findings therefore demystify to some extent the ingrained beliefs on the major impact of AP on musical skills.

We consider our findings particularly interesting, because it is the first study to compare such homogeneous populations for investigating AP. Indeed, as AP and NAP participants in our study did not differ for age, age of onset of musical practice, total number of years of music training, gender, instrumental level and socio-economic status, they exclusively differed for AP skill, making our results stand out from previous research in the domain.

AUTHOR CONTRIBUTIONS

SYC, NV, DG, and CEJ conceived and designed the study. NV recruited the participants. SYC and NV welcomed the participants and administered the EEG recordings. SYC, NV, and CEJ analyzed the data, designed the figures and wrote the main manuscript text. All authors were involved in the discussions about the interpretation of the results and the proofreading of the manuscript.

FUNDING

This work was supported by the Swiss National Science Foundation (grant number 100014–125050 to CEJ) and is part of a multimodal brain imaging project entitled Behavioral, neuro-functional, and neuro-anatomical correlates of experience dependent music perception.

ACKNOWLEDGMENTS

The Cartool software (brainmapping.unige.ch/cartool) has been programmed by Denis Brunet, from the Functional Brain Mapping Laboratory, Geneva, Switzerland, and is supported by the Center for Biomedical Imaging (CIBM) of Geneva and Lausanne. The STEN toolbox (<https://doi.org/10.5281/zenodo.1164037>) has been programmed by Jean-François Knebel, from the Laboratory for Investigative Neurophysiology (the LINE),

Lausanne, Switzerland, and is supported by the Center for Biomedical Imaging (CIBM) of Geneva and Lausanne and by the National Center of Competence in Research, project SYNAPSY—The Synaptic Bases of Mental Disease; project no. 51AU40_125759. We would like to thank Bruno Pietri for the beautiful compositions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2019.00142/full#supplementary-material>

Audio 1 | Regular closure (degree I).

Audio 2 | Subtle transgression of closure (T^{sub} , degree I⁶).

Audio 3 | Apparent transgression of closure (T^{app} , degree IV⁶).

REFERENCES

- Akiva-Kabiri, L., and Henik, A. (2012). A unique asymmetrical stroop effect in absolute pitch possessors. *Exp. Psychol.* 59, 272–278. doi: 10.1027/1618-3169/a000153
- Ayotte, J., Peretz, I., and Hyde, K. (2002). Congenital amusia: a group study of adults afflicted with a music-specific disorder. *Brain* 125, 238–251. doi: 10.1093/brain/awf028
- Bachem, A. (1937). Various types of absolute pitch. *J. Acoust. Soc. Am.* 9, 146–151. doi: 10.1121/1.1915919
- Bachem, A. (1955). Absolute pitch. *J. Acoust. Soc. Am.* 27, 1180–1185. doi: 10.1121/1.1908155
- Baharloo, S., Johnston, P. A., Service, S. K., Gitschier, J., and Freimer, N. B. (1998). Absolute pitch: an approach for identification of genetic and nongenetic components. *Am. J. Hum. Genet.* 62, 224–231. doi: 10.1086/301704
- Baharloo, S., Service, S. K., Risch, N., Gitschier, J., and Freimer, N. B. (2000). Familial aggregation of absolute pitch. *Am. J. Hum. Genet.* 67, 755–758. doi: 10.1086/303057
- Benassi-Werke, M. E., Queiroz, M., Araújo, R. S., Bueno, O. F. A., and Oliveira, M. G. M. (2012). Musicians' working memory for tones, words, and pseudowords. *Q. J. Exp. Psychol.* 65, 1161–1171. doi: 10.1080/17470218.2011.644799
- Biró, G., Spinelli, L., Vulliemoz, S., Megevang, P., Brunet, D., Seeck, M., et al. (2014). Head model and electrical source imaging: a study of 38 epileptic patients. *Neuroimage Clin.* 5, 77–83. doi: 10.1016/j.nicl.2014.06.005
- Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., et al. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol. Evol.* 24, 127–135. doi: 10.1016/j.tree.2008.10.008
- Brady, P. T. (1970). Fixed-scale mechanism of absolute pitch. *J. Acoust. Soc. Am.* 48, 883–887. doi: 10.1121/1.1912227
- Brauchli, C., Leipold, S., and Jancke, L. (2019). Univariate and multivariate analyses of functional networks in absolute pitch. *Neuroimage* 189, 241–247. doi: 10.1016/j.neuroimage.2019.01.021
- Brodbeck, V., Spinelli, L., Lascano, A. M., Wissmeier, M., Vargas, M. I., Vulliemoz, S., et al. (2011). Electroencephalographic source imaging: a prospective study of 152 operated epileptic patients. *Brain* 134, 2887–2897. doi: 10.1093/brain/awr243
- Brunet, D., Murray, M. M., and Michel, C. M. (2011). Spatiotemporal analysis of multichannel EEG: CARTOOL. *Comput. Intell. Neurosci.* 2011:813870. doi: 10.1155/2011/813870
- Calin-Jageman, R. J. (2018). The new statistics for neuroscience majors: thinking in effect sizes. *J. Undergrad. Neurosci. Educ.* 16, E21–E25. doi: 10.31234/osf.io/zvm9a
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd edition. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
- De Meo, R., Bourquin, N. M., Knebel, J. F., Murray, M. M., and Clarke, S. (2015). From bird to sparrow: learning-induced modulations in fine-grained semantic discrimination. *Neuroimage* 118, 163–173. doi: 10.1016/j.neuroimage.2015.05.091
- Deutsch, D., and Dooley, K. (2013). Absolute pitch is associated with a large auditory digit span: a clue to its genesis. *J. Acoust. Soc. Am.* 133, 1859–1861. doi: 10.1121/1.4792217
- Deutsch, D., Henthorn, T., Marvin, E., and Xu, H. (2006). Absolute pitch among American and Chinese conservatory students: prevalence differences, and evidence for a speech-related critical period. *J. Acoust. Soc. Am.* 119, 719–722. doi: 10.1121/1.2151799
- Dooley, K., and Deutsch, D. (2010). Absolute pitch correlates with high performance on musical dictation. *J. Acoust. Soc. Am.* 128, 890–893. doi: 10.1121/1.3458848
- Elmer, S., Rogenmoser, L., Kuhn, J., and Jancke, L. (2015). Bridging the gap between perceptual and cognitive perspectives on absolute pitch. *J. Neurosci.* 35, 366–371. doi: 10.1523/JNEUROSCI.3009-14.2015
- Elmer, S., Sollberger, S., Meyer, M., and Jancke, L. (2013). An empirical reevaluation of absolute pitch: behavioral and electrophysiological measurements. *J. Cogn. Neurosci.* 25, 1736–1753. doi: 10.1162/jocn_a_00410
- Greber, M., Rogenmoser, L., Elmer, S., and Jancke, L. (2018). Electrophysiological correlates of absolute pitch in a passive auditory oddball paradigm: a direct replication attempt. *eneuro* 5:ENEURO.0333-0318.2018. doi: 10.1523/ENEURO.0333-18.2018
- Grech, R., Cassar, T., Muscat, J., Camilleri, K. P., Fabri, S. G., Zervakis, M., et al. (2008). Review on solving the inverse problem in EEG source analysis. *J. Neuroeng. Rehabil.* 5:25. doi: 10.1186/1743-0003-5-25
- Gregersen, P. K., Kowalsky, E., Kohn, N., and Marvin, E. W. (1999). Absolute pitch: prevalence, ethnic variation, and estimation of the genetic component. *Am. J. Hum. Genet.* 65, 911–913. doi: 10.1086/302541
- Gregersen, P. K., Kowalsky, E., Kohn, N., and Marvin, E. W. (2001). Early childhood music education and predisposition to absolute pitch: teasing apart genes and environment. *Am. J. Med. Genet.* 98, 280–282. doi: 10.1002/1096-8628(20010122)98:3<280::AID-AJMG1083>3.0.CO;2-6
- Griffiths, T. D., and Warren, J. D. (2002). The planum temporale as a computational hub. *Trends Neurosci.* 25, 348–353. doi: 10.1016/S0166-2236(02)02191-4
- Hamalainen, M. S., and Ilmoniemi, R. J. (1994). Interpreting magnetic fields of the brain: minimum norm estimates. *Med. Biol. Eng. Comput.* 32, 35–42. doi: 10.1007/BF02512476
- Hirata, Y., Kuriki, S., and Pantev, C. (1999). Musicians with absolute pitch show distinct neural activities in the auditory cortex. *Neuroreport* 10, 999–1002. doi: 10.1097/00001756-199904060-00019

- Hou, J., Chen, A. C., Song, B., Sun, C., and Beauchaine, T. P. (2016). Neural correlates of absolute pitch: a review. *Musicae Scientiae* 21, 287–302. doi: 10.1177/1029864916662903
- Itoh, K., Suwazono, S., Arai, H., Miyazaki, K., and Nakada, T. (2005). Electrophysiological correlates of absolute pitch and relative pitch. *Cereb. Cortex* 15, 760–769. doi: 10.1093/cercor/bhh177
- James, C. E., Cereghetti, D. M., Roulet Tribes, E., and Oechslin, M. S. (2015). Electrophysiological evidence for a specific neural correlate of musical violation expectation in primary-school children. *Neuroimage* 104, 386–397. doi: 10.1016/j.neuroimage.2014.09.047
- James, C. E., Oechslin, M. S., Michel, C. M., and De Pretto, M. (2017). Electrical neuroimaging of music processing reveals mid-latency changes with level of musical expertise. *Front. Neurosci.* 11:613. doi: 10.3389/fnins.2017.00613
- Jancke, L., Langer, N., and Hanggi, J. (2012). Diminished whole-brain but enhanced peri-sylvian connectivity in absolute pitch musicians. *J. Cogn. Neurosci.* 24, 1447–1461. doi: 10.1162/jocn_a_00227
- Jenni, R., Oechslin, M. S., and James, C. E. (2017). Impact of major and minor mode on EEG frequency range activities of music processing as a function of expertise. *Neurosci. Lett.* 647, 159–164. doi: 10.1016/j.neulet.2017.03.022
- Keenan, J. P., Thangaraj, V., Halpern, A. R., and Schlaug, G. (2001). Absolute pitch and planum temporale. *Neuroimage* 14, 1402–1408. doi: 10.1006/nimg.2001.0925
- Knebel, J. F., Javitt, D. C., and Murray, M. M. (2011). Impaired early visual response modulations to spatial information in chronic schizophrenia. *Psychiatry Res.* 193, 168–176. doi: 10.1016/j.psychres.2011.02.006
- Koelsch, S., Rohrmeier, M., Torrecuso, R., and Jentschke, S. (2013). Processing of hierarchical syntactic structure in music. *Proc. Natl. Acad. Sci. U.S.A.* 110, 15443–15448. doi: 10.1073/pnas.1300272110
- Koenig, T., Stein, M., Grieder, M., and Kottlow, M. (2014). A tutorial on data-driven methods for statistically assessing ERP topographies. *Brain Topogr.* 27, 72–83. doi: 10.1007/s10548-013-0310-1
- Kutas, M., and Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annu. Rev. Psychol.* 62, 621–647. doi: 10.1146/annurev.psych.093008.131123
- Lehmann, D., Faber, P. L., Galderisi, S., Herrmann, W. M., Kinoshita, T., Koukoku, M., et al. (2005). EEG microstate duration and syntax in acute, medication-naïve, first-episode schizophrenia: a multi-center study. *Psychiatry Res. Neuroimaging* 138, 141–156. doi: 10.1016/j.psychres.2004.05.007
- Lehmann, D., Ozaki, H., and Pal, I. (1987). EEG alpha map series: brain microstates by space-oriented adaptive segmentation. *Electroencephalogr. Clin. Neurophysiol.* 67, 271–288. doi: 10.1016/0013-4694(87)90025-3
- Levitin, D. J. (1994). Absolute memory for musical pitch: evidence from the production of learned melodies. *Percept. Psychophys.* 56, 414–423. doi: 10.3758/BF03206733
- Levitin, D. J., and Rogers, S. E. (2005). Absolute pitch: perception, coding, and controversies. *Trends Cogn. Sci.* 9, 26–33. doi: 10.1016/j.tics.2004.11.007
- Loui, P., Li, H. C., Hohmann, A., and Schlaug, G. (2011). Enhanced cortical connectivity in absolute pitch musicians: a model for local hyperconnectivity. *J. Cogn. Neurosci.* 23, 1015–1026. doi: 10.1162/jocn.2010.21500
- Macmillan, N. A., and Creelman, C. D. (1997). d'plus: a program to calculate accuracy and bias measures from detection and discrimination data. *Spat. Vis.* 11, 141–143.
- Margulis, E. H., Mlsna, L. M., Uppunda, A. K., Parrish, T. B., and Wong, P. C. (2009). Selective neurophysiologic responses to music in instrumentalists with different listening biographies. *Hum. Brain Mapp.* 30, 267–275. doi: 10.1002/hbm.20503
- Martuzzi, R., Murray, M. M., Meuli, R. A., Thiran, J. P., Maeder, P. P., Michel, C. M., et al. (2009). Methods for determining frequency- and region-dependent relationships between estimated LFPs and BOLD responses in humans. *J. Neurophysiol.* 101, 491–502. doi: 10.1152/jn.90335.2008
- McCulloch, C. E. (2003). “Generalized linear mixed models,” in *NSF-CBMS Regional Conference Series in Probability and Statistics* vol. 7, (San Francisco, CA: Institute of Mathematical Statistics), 1–84.
- McLachlan, N. M., Marco, D. J., and Wilson, S. J. (2013). Pitch and plasticity: insights from the pitch matching of chords by musicians with absolute and relative pitch. *Brain Sci.* 3, 1615–1634. doi: 10.3390/brainsci3041615
- Megevand, P., Spinelli, L., Genetti, M., Brodbeck, V., Momjian, S., Schaller, K., et al. (2014). Electric source imaging of interictal activity accurately localises the seizure onset zone. *J. Neurol. Neurosurg. Psychiatr.* 85, 38–43. doi: 10.1136/jnnp-2013-305515
- Meyer-Lindenberg, A. (2010). From maps to mechanisms through neuroimaging of schizophrenia. *Nature* 468, 194–202. doi: 10.1038/nature09569
- Michel, C. M., and Brandeis, D. (2009). “Data acquisition and pre-processing standards for electrical neuroimaging,” in *Electrical Neuroimaging*, eds C. M. Michel, T. Koenig, D. Brandeis, L. R. R. Gianotti and J. Wackermann (Cambridge: Cambridge University Press), 79–92.
- Michel, C. M., and Murray, M. M. (2012). Towards the utilization of EEG as a brain imaging tool. *Neuroimage* 61, 371–385. doi: 10.1016/j.neuroimage.2011.12.039
- Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., and Grave de Peralta, R. (2004). EEG source imaging. *Clin. Neurophysiol.* 115, 2195–2222. doi: 10.1016/j.clinph.2004.06.001
- Michel, C. M., Seck, M., and Landis, T. (1999). Spatiotemporal dynamics of human cognition. *News Physiol. Sci.* 14, 206–214. doi: 10.1152/physiologyonline.1999.14.5.206
- Miyazaki, K. (1988). Musical pitch identification by absolute pitch possessors. *Percept. Psychophys.* 44, 501–512. doi: 10.3758/BF03207484
- Miyazaki, K., and Rakowski, A. (2002). Recognition of notated melodies by possessors and nonpossessors of absolute pitch. *Percept. Psychophys.* 64, 1337–1345. doi: 10.3758/BF03194776
- Murray, M. M., Brunet, D., and Michel, C. M. (2008). Topographic ERP analyses: a step-by-step tutorial review. *Brain Topogr.* 20, 249–264. doi: 10.1007/s10548-008-0054-5
- Nakagawa, S., and Schielzeth, H. (2013). A general and simple method for obtaining R² from generalized linear mixed-effects models. *Methods Ecol. Evol.* 4, 133–142. doi: 10.1111/j.2041-210x.2012.00261.x
- Oechslin, M. S., Meyer, M., and Jancke, L. (2010). Absolute pitch–functional evidence of speech-relevant auditory acuity. *Cereb. Cortex* 20, 447–455. doi: 10.1093/cercor/bhp113
- Oechslin, M. S., Van De Ville, D., Lazeyras, F., Hauert, C. A., and James, C. E. (2013). Degree of musical expertise modulates higher order brain functioning. *Cereb. Cortex* 23, 2213–2224. doi: 10.1093/cercor/bhs206
- Ohnishi, T., Matsuda, H., Asada, T., Aruga, M., Hirakata, M., Nishikawa, M., et al. (2001). Functional anatomy of musical perception in musicians. *Cereb. Cortex* 11, 754–760. doi: 10.1093/cercor/11.8.754
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4
- Pantev, C., Roberts, L. E., Schulz, M., Engelien, A., and Ross, B. (2001). Timbre-specific enhancement of auditory cortical representations in musicians. *Neuroreport* 12, 169–174. doi: 10.1097/00001756-200101220-00041
- Pascual-Marqui, R. D., Michel, C. M., and Lehmann, D. (1995). Segmentation of brain electrical activity into microstates: model estimation and validation. *IEEE Trans. Biomed. Eng.* 42, 658–665. doi: 10.1109/10.391164
- Plomp, G., Hervais-Adelman, A., Astolfi, L., and Michel, C. M. (2015). Early recurrence and ongoing parietal driving during elementary visual processing. *Sci. Rep.* 5:18733. doi: 10.1038/srep18733
- Plomp, G., Michel, C. M., and Herzog, M. H. (2010). Electrical source dynamics in three functional localizer paradigms. *Neuroimage* 53, 257–267. doi: 10.1016/j.neuroimage.2010.06.037
- Plomp, G., Roinishvili, M., Chkonia, E., Kapanadze, G., Kereselidze, M., Brand, A., et al. (2013). Electrophysiological evidence for ventral stream deficits in schizophrenia patients. *Schizophr. Bull.* 39, 547–554. doi: 10.1093/schbul/sbr175
- Profita, J., and Bidder, T. G. (1988). Perfect pitch. *Am. J. Med. Genet.* 29, 763–771. doi: 10.1002/ajmg.1320290405
- Rogenmoser, L., Elmer, S., and Jancke, L. (2015). Absolute pitch: evidence for early cognitive facilitation during passive listening as revealed by reduced P3a amplitudes. *J. Cogn. Neurosci.* 27, 623–637. doi: 10.1162/jocn_a_00708
- Schlaug, G., Jancke, L., Huang, Y., and Steinmetz, H. (1995). *In vivo* evidence of structural brain asymmetry in musicians. *Science* 267, 699–701. doi: 10.1126/science.7839149
- Schulze, K., Gaab, N., and Schlaug, G. (2009). Perceiving pitch absolutely: comparing absolute and relative pitch possessors in a pitch memory task. *BMC Neurosci.* 10:106. doi: 10.1186/1471-2202-10-106
- Spinelli, L., Andino, S. G., Lantz, G., Seck, M., and Michel, C. M. (2000). Electromagnetic inverse solutions in anatomically constrained

- spherical head models. *Brain Topogr.* 13, 115–125. doi: 10.1023/A:1026607118642
- Takeuchi, A. H., and Hulse, S. H. (1993). Absolute pitch. *Psychol. Bull.* 113, 345–361. doi: 10.1037/0033-2909.113.2.345
- Tibshirani, R., and Walther, G. (2005). Cluster validation by prediction strength. *J. Comput. Graph. Stat.* 14, 511–528. doi: 10.1198/106186005X59243
- Tillmann, B., Albouy, P., and Caclin, A. (2015). Congenital amusias. *Handb. Clin. Neurol.* 129, 589–605. doi: 10.1016/B978-0-444-62630-1.00033-0
- Van Hedger, S. C., Heald, S. L., Koch, R., and Nusbaum, H. C. (2015). Auditory working memory predicts individual differences in absolute pitch learning. *Cognition* 140, 95–110. doi: 10.1016/j.cognition.2015.03.012
- Vaughan, H. G. Jr. (1982). The neural origins of human event-related potentials. *Ann. N. Y. Acad. Sci.* 388, 125–138. doi: 10.1111/j.1749-6632.1982.tb50788.x
- Vitouch, O. (2003). Absolutist models of absolute pitch are absolutely misleading. *Music Perception* 21, 111–117. doi: 10.1525/mp.2003.21.1.111
- Ward, W. D. (1999). “Absolute pitch,” in *The Psychology of Music*, ed D. Deutsch (New York, NY: Academic Press), 265–298. doi: 10.1016/B978-012213564-4/50009-3
- Wechsler, D. (2011). *Wechsler Adult Intelligence Scale: Fourth Edition (WAIS-IV)*. [Assessment instrument], San Antonio, TX: Pearson.
- Wengenroth, M., Blatow, M., Heinecke, A., Reinhardt, J., Stippich, C., Hofmann, E., et al. (2014). Increased volume and function of right auditory cortex as a marker for absolute pitch. *Cereb. Cortex* 24, 1127–1137. doi: 10.1093/cercor/bhs391
- Whitehead, A. L., Julious, S. A., Cooper, C. L., and Campbell, M. J. (2016). Estimating the sample size for a pilot randomised trial to minimise the overall trial sample size for the external pilot and main trial for a continuous outcome variable. *Stat. Methods Med. Res.* 25, 1057–1073. doi: 10.1177/0962280215588241
- Wilson, S. J., Lusher, D., Wan, C. Y., Dudgeon, P., and Reutens, D. C. (2009). The neurocognitive components of pitch processing: insights from absolute pitch. *Cereb. Cortex* 19, 724–732. doi: 10.1093/cercor/bhn121
- Zatorre, R. J. (2003). Absolute pitch: a model for understanding the influence of genes and development on neural and cognitive function. *Nat. Neurosci.* 6, 692–695. doi: 10.1038/nn1085
- Zatorre, R. J., Perry, D. W., Beckett, C. A., Westbury, C. F., and Evans, A. C. (1998). Functional anatomy of musical processing in listeners with absolute pitch and relative pitch. *Proc. Natl. Acad. Sci. U.S.A.* 95, 3172–3177. doi: 10.1073/pnas.95.6.3172
- Ziv, N., and Radin, S. (2014). Absolute and relative pitch: global versus local processing of chords. *Adv. Cogn. Psychol.* 10, 15–25. doi: 10.5709/acp-0152-7

Conflict of Interest Statement: 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.

Copyright © 2019 Coll, Vuichoud, Grandjean and James. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



On the Association Between Musical Training, Intelligence and Executive Functions in Adulthood

Antonio Criscuolo¹, Leonardo Bonetti², Teppo Särkämö³, Marina Kliuchko² and Elvira Brattico^{2*}

¹ Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, Netherlands, ² Center for Music in the Brain, Department of Clinical Medicine, Aarhus University – The Royal Academy of Music, Aarhus/Aalborg, Aarhus, Denmark, ³ Cognitive Brain Research Unit, Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Helsinki, Finland

OPEN ACCESS

Edited by:

Claude Alain,
Rotman Research Institute (RRI),
Canada

Reviewed by:

Alexandra Parbery-Clark,
Swedish Neuroscience Institute (SNI),
Swedish Medical Center,
United States
Cyrille Magne,
Middle Tennessee State University,
United States

*Correspondence:

Elvira Brattico
elvira.brattico@clin.au.dk

Specialty section:

This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Psychology

Received: 13 December 2018

Accepted: 08 July 2019

Published: 30 July 2019

Citation:

Criscuolo A, Bonetti L, Särkämö T,
Kliuchko M and Brattico E (2019) On
the Association Between Musical
Training, Intelligence and Executive
Functions in Adulthood.
Front. Psychol. 10:1704.
doi: 10.3389/fpsyg.2019.01704

Converging evidence has demonstrated that musical training is associated with improved perceptual and cognitive skills, including executive functions and general intelligence, particularly in childhood. In contrast, in adults the relationship between cognitive performance and musicianship is less clear and seems to be modulated by a number of background factors, such as personality and socio-economic status. Aiming to shed new light on this topic, we administered the Wechsler Adult Intelligence Scale III (WAIS-III), the Wechsler Memory Scale III (WMS-III), and the Stroop Test to 101 Finnish healthy adults grouped according to their musical expertise (non-musicians, amateurs, and musicians). After being matched for socio-economic status, personality traits and other demographic variables, adult musicians exhibited higher cognitive performance than non-musicians in all the mentioned measures. Moreover, linear regression models showed significant positive relationships between executive functions (working memory and attention) and the duration of musical practice, even after controlling for intelligence and background variables, such as personality traits. Hence, our study offers further support for the association between cognitive abilities and musical training, even in adulthood.

HIGHLIGHTS

- Musicians show higher general intelligence (FSIQ), verbal intelligence (VIQ), working memory (WMI) and attention skills than non-musicians. Amateurs score in between.
- Significant positive correlations between years of musical playing and cognitive abilities support the hypothesis that long-term musical practice is associated with intelligence and executive functions.

Keywords: musical training, cognition, intelligence quotient, working memory, attention, executive functions

INTRODUCTION

Musical Training Relies on Executive Functions

Musical training is a multisensory experience engaging multiple cognitive functions and underlying neural networks. Indeed, reading, listening, understanding and performing polyphonic music require the simultaneous processing of sounds and rhythms, higher order perceptual processing and fine sensory-motor coordination (Münste et al., 2002). Long-term musical training engages and trains all those functions on a daily basis and, as a result, musicians seem to improve not only music-related abilities, but also domain-general skills. Hence, musicians show increased auditory perception and production abilities, such as enhanced capacity to detect deviations in complex regularities and tone patterns (Tervaniemi, 2001, 2009; Fujioka et al., 2004; Zuijzen et al., 2004; Van Zuijzen et al., 2005; Bangert and Schlaug, 2006; Herholz et al., 2009) as well as fine motor control (Krings et al., 2000; Koeneke et al., 2004; Vuust et al., 2005; Kleber et al., 2013; Burunat et al., 2015).

Besides improving listening and sensorimotor abilities closely linked to the musical practice (Schellenberg, 2011), there is also evidence in favor of the *far transfer effect* to non-musical functions. In the literature, *far transfer effects* relates to the influence of musical training on general (not confined to the auditory domain) cognitive functions, such as spatial (Gromko and Poorman, 1998; Rauscher, 2002; Brochard et al., 2004; Sluming et al., 2007), mathematical (Cheek and Smith, 1999), and non-verbal (Forgeard et al., 2008) abilities. Among these, working memory (WM) refers to the ability to retrieve, monitor, analyze, integrate, chunk and recall within a short time span both auditory and non-auditory information (Herholz et al., 2009; Hansen et al., 2013; for reviews, see, e.g., Kraus and Chandrasekaran, 2010; Reybrouck and Brattico, 2015; Schlaug, 2015). In music processing, WM integrates sound events, recollects information from memory systems, links sounds to meaning and to memories, and supports the generation of emotional reactions (Burunat et al., 2014).

Along with cognitive flexibility, response inhibition and interference control, WM is considered one of the fundamental executive functions in humans (Diamond, 2013). Executive functions designate a set of abilities related to updating and manipulating relevant information (WM), inhibiting automatic responses, shifting attention to mental tasks (selective attention), planning, reasoning and decision making (Guare and Dawson, 2004; Perrotin et al., 2008; Garner, 2009; Collins and Koechlin, 2012). Improvements in executive functions and cognitive flexibility by musical training have been observed in Finnish school-age children and these improvements positively correlated to enhanced neural sound discrimination (Saarikivi et al., 2016).

Effects of Musical Training on General Intelligence

Previous evidence suggests that long-term engagement in musical activities modulates not only executive functions but also general intelligence or *g* (Hansen et al., 2013; see Schellenberg and

Weiss, 2013 for a review). In psychological science, *g* has been defined in several ways and assessed using a variety of behavioral tests. Gottfredson (1997, PAG-13) described it as a “very general mental capability that involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and from experience.” *G* seems to rely on similar neural substrates to WM, comprising a network of prefrontal and subcortical regions, along with the anterior cingulate, parietal and premotor regions (Duncan et al., 2000). For similarities at both functional and anatomical levels, Salthouse and Pink (2008) suggested that WM is closely related to *g*. Moreover, the benefits of *far transfer-effects* of musical training extend to the domain of *g* and its quantitative measure, namely intelligence quotient (IQ). For instance, musically trained children show higher IQ as compared to non-trained children (Schellenberg, 2004, 2011).

Nature or Nurture?

Findings in favor of the link between musical training and cognitive functions led to the notorious debate of nature vs. nurture: i.e., are the observed differences in cognitive abilities only associated with pre-existing neurocognitive differences which predispose people to engage in musical activities (nature), or is the cognitive training promoted by musical activities able to influence cognitive abilities (nurture)? On one side, numerous experimental studies have shown a link between musical training and cognitive development. For instance, in Schellenberg (2004) IQ of a sample of 144 6-year-old children were assessed before and after 1 year of music or drama classes by Wechsler Intelligence Scale for Children–Third Edition (WISC-III). The authors demonstrated that, despite nodifferences in the pre-test for WISC scores, there were greater improvements in WISC scores for children of the music group as compared to the control groups, along with improvements in some of its subscales, such as Verbal Comprehension and Perceptual Organization Indices. Similarly, in Schellenberg (2011) 106 children aged 9 to 12 (half musically trained and half untrained) were tested with the Wechsler Abbreviate Scale of Intelligence (WASI): trained children showed higher IQ scores than their untrained counterparts. Lastly, beyond IQ assessed with psychological tests, music training in childhood is associated with positive academic achievement (Schellenberg, 2006) and improvements in language-related abilities (Moreno, 2009; François et al., 2013).

Together, these findings highlight an association between musical training and general cognitive abilities in childhood. However, when taking in consideration background variables other than musical training, some authors have shown that pre-existing differences in cognitive abilities, together with differences in children's and their parents' personality traits, may contribute in the choice of engaging in musical training and in the duration of such training. In turn, this may ultimately account for differences in cognitive performances in adulthood (Schellenberg, 2006; Corrigan et al., 2013; Corrigan and Schellenberg, 2015). Furthermore, Bonetti and Costa (2017) showed associations between fluid intelligence and music tasks in children aged 4–6 years old with no previous musical training, suggesting a possible innate connection between some musical skills and intelligence that could potentially lead to a higher

probability of engaging in musical studies for children with higher IQ. Lastly, by showing that genetic and environmental factors interact in determining music behaviors, such as musical practice and music enjoyment, Butkovic et al. (2015) have further highlighted the need for new investigations to clarify the complex association between music and cognitive development.

The Current Study

While a consistent corpus of research has focused on child populations, findings in adults are sparse (Brandler and Rammsayer, 2003; Helmbold et al., 2005; Schellenberg and Moreno, 2010). Existing evidence suggests that *far transfer effects* of musical training to general cognitive skills might be related to confounding variables that are usually neglected (Sala and Gobet, 2017), such as personality traits (Corrigall et al., 2013; Corrigall and Schellenberg, 2015).

Aiming to test the hypothesis that long-term musical practice is associated with improved cognitive abilities in adulthood, we assessed intelligence and executive functions of adults with different levels of musical expertise while controlling for background variables such as socio-economic status (SES), age, years of education and personality traits. Differently from previous studies (Brandler and Rammsayer, 2003; Helmbold et al., 2005; Schellenberg and Moreno, 2010; Swaminathan et al., 2017), we adopted the Wechsler Adult Intelligence Scale III (WAIS-III) as intelligence test. WAIS belongs to the family of Wechsler tests, the most used to assess intelligence in the psychological literature (Weiss et al., 2016). To investigate executive functions, we used the Wechsler Memory Scale III (WMS-III) for WM and the Stroop test for selective attention, cognitive flexibility and processing speed. Lastly, personality was assessed by administering participants with the Big Five Inventory questionnaire (BFI) as it was previously done in Corrigall et al. (2013) and in Corrigall and Schellenberg (2015). Our sample includes 101 highly educated Finnish adults (representative of the high education level in Finland; oecd.org) with a mean IQ higher than the average Finnish population [comparing the individual scores with the WAIS norms; (Wechsler, 1997a)].

In line with previous studies (Schellenberg, 2006; Swaminathan et al., 2017, 2018), we expected (i) expert musicians to report higher intelligence and executive functions than non-musicians and (ii) to verify that the positive relationship between the duration of musical practice and cognitive performances would hold even after controlling for potential confounding variables.

MATERIALS AND METHODS

Participants

The participants were part of the broad “Tunteet” research protocol involving a multi-dimensional dataset of brain measures, psychological tests and behavioral data on audition, emotion and musical behavior. The dataset was obtained from 140 participants recruited among university students or qualified professionals. Further details on this protocol

can be found in Burunat et al. (2015, 2016, 2018), Kliuchko et al. (2015, 2016, 2018), Alluri et al. (2017), Bonetti et al. (2017, 2018), and Saari et al. (2018), where some of the participants involved in this study were included. All experimental procedures for this protocol were approved by the Coordinating Ethics Committee of the Hospital District of Helsinki and Uusimaa (approval number: 315/13/03/00/11, obtained on March the 11th, 2012). Furthermore, all procedures were conducted in agreement with the ethical principles of the Declaration of Helsinki.

For the purpose of the current study, we selected only participants who completed psychological and cognitive testing with a trained psychologist ($N = 114$). The other participants did not take part in the testing because their native language was not Finnish, or they did not have enough time to dedicate to the study, and thus other measurements were prioritized. We obtained information on their musical expertise crossing details derived from both a paper and pencil questionnaire (used in previous studies: e.g., Brattico et al., 2009) and an online survey called Helsinki Inventory for Music and Affect Behavior or HIMAB (Gold et al., 2013). Based on those details, subjects were divided into three groups according to levels of musical expertise (or “musicianship” from now on): non-musicians, amateurs, and musicians. Participants were considered musicians when they reported more than 5 years of music practice and considered themselves as musicians. In addition to this, for entering the musicians’ group two criteria had to be matched: a final degree at a music academy or monetarily compensation for their music performance or teaching activities. Participants not matching these parameters, though involved in music activities, were classified as amateurs, and all participants with less than 3 years of musical training entered the group of non-musicians. For the scope of this study, we decided to combine the duration of musical training and the years of musical practice in a comprehensive variable named “years of music playing.” Out of 114 participants, 13 were further excluded because they deviated from normal distribution in one or more background variables (age, years of music playing, years of education, SES). Therefore, only 101 participants were included in this research: 45 were males (44.5%) and 56 were females (55.6%) within the age range of 18–55 years (mean age 28.44 ± 8.26 SD). The SES of participants was assessed by means of the Hollingshead Four-Factor Index (Hollingshead, 1975) included in HIMAB. Details on the participants’ SES, age, gender, years of education, and years of music playing are reported in **Table 1**, together with personality indices of neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness, as measured by the BFI. Musicians’ and amateurs’ musical background information are provided in **Table 2**, which includes the starting age of musical training and musical practice, together with the average of weekly hours spent in practicing and in listening to music and years of music playing.

All participants took part in the experiment on a voluntary basis and they were compensated with vouchers to use for culture and sport purposes (e.g., museums, concerts, or swimming pools). All of them were healthy and declared to have no history

TABLE 1 | Participants' background information according to musicianship.

	Non-musicians	mateurs	Musicians	P-value	Beta	95% CI LB UB	Partial correlation
<i>N of subjects</i>	51	27	23				
<i>Gender</i>	22 M + 26 F	11 M + 16 F	12 M + 14 F				
<i>Handedness</i>	5 L	0 L	2 L				
<i>Age</i>	29.08 ± 8.91	28.19 ± 8.60	30.17 ± 8.65	0.38	−0.047	−0.022 0.012	−0.096
<i>Years of music playing</i>	0.55 ± 1.62	6.35 ± 3.28	21.35 ± 7.36	<0.001	0.66	0.374 0.485	0.928
<i>Years of education</i>	17.77 ± 3.51	16.65 ± 3.21	18.95 ± 4.17	0.59	−0.014	−0.029 0.024	−0.03
<i>SES</i>	35.12 ± 3.19	23.08 ± 4.60	42.89 ± 3.66	0.96	−0.001	−0.007 0.007	−0.003
<i>Neuroticism</i>	−10.94 ± 11.03	−10.83 ± 10.58	−9.95 ± 10.21	0.45	0.033	−0.01 0.015	0.072
<i>Extraversion</i>	8.00 ± 9.21	10.48 ± 12.05	9.25 ± 10.41	0.24	0.076	−0.008 0.022	0.106
<i>Openness</i>	19.19 ± 8.86	19.43 ± 10.24	24.55 ± 6.86	0.59	0.015	−0.014 0.017	0.031
<i>Agreeableness</i>	15.44 ± 8.54	18.43 ± 7.93	17.10 ± 8.73	0.49	−0.025	−0.018 0.012	−0.059
<i>Conscientiousness</i>	13.31 ± 9.92	14.43 ± 8.30	10.60 ± 10.05	0.43	0.038	−0.010 0.016	00.08

In order, listed are participants' gender, handedness, mean and SD for age, duration of musical playing, years of general education, SES (socio-economic status) and personality indices of neuroticism, extraversion, openness to experience, agreeableness and conscientiousness, as measured by the Big Five Inventory questionnaire (BFI). These variables were used as predictors of musicianship in a regression model; associated p-value, standardized Beta, 95% confidence interval (CI; lower bound, LB, and upper bound, UB) and partial correlation for each variable are provided. Years of musical playing was the only significant predictor of musicianship and its statistical values are reported in bold.

TABLE 2 | Musical background information for amateurs and musicians.

	Training onset	Practice onset	Weekly practice hours	Listening hours	Years of practice
<i>Amateurs</i>	14.04 ± 7.91	11.88 ± 7.21	1.72 ± 3.59	21.85 ± 20.86	6.35 ± 3.28
<i>Musicians</i>	7.42 ± 4.61	5.36 ± 2.11	15.02 ± 11.12	23.74 ± 15.82	21.35 ± 7.36
<i>P-value</i>	0.964	0.48	0.042	0.311	<0.001
<i>Beta</i>	0.007	−0.122	0.285	0.127	0.54
<i>95% CI LB UB</i>	−0.031 0.033	−0.056 0.027	0.001 0.026	−0.008 0.025	0.129 0.346
<i>Partial correlation</i>	0.008	−0.123	0.346	0.176	0.612

In order, means and standard errors of the starting age (in years) of musical training and practice, weekly hours spent in practicing and in listening to music, years of musical practice. p-Value, standardized Beta, 95% confidence interval (CI; lower bound, LB, and upper bound, UB) and partial correlation for each variable as reported from a regression model are also provided. Significant statistical values are reported in bold.

of neurological or psychiatric disorders. All participants signed an informed consent before the beginning of the experiment and a researcher was present and available for assistance at any time.

Psychological Tests

Wechsler Adult Intelligence Scale III

The WAIS-III is a widely used test for the assessment of adults' and old adolescents' intelligence (Wechsler, 1997a). In this study, we used the following eight WAIS-III subtests: Vocabulary, Similarities, Information, Picture Completion, Block Design, Matrix Reasoning, Digit–Symbol Coding, and Symbol Search. The Vocabulary, Similarities, and Information subtests were used to calculate the Verbal Comprehension Index (VCI). The Picture Completion, Block Design, and Matrix Reasoning subtests were used to calculate the Perceptual Organization Index (POI). The Digit–Symbol Coding and Symbol Search subtests were used to calculate the Processing Speed Index (PSI). In addition, these subtests and the Letter–Number Sequencing subtest from WMS-III (which is the same as in WAIS-III) were adopted to estimate the Verbal

Intelligence Quotient (VIQ), Performance IQ (PIQ), and full-scale IQ (FSIQ). More details on the tests can be found in Table 3.

Stroop Test

The Stroop effect is measured with the Stroop test and refers to the interference in the reaction time of a task providing conflicting cues. The Stroop effect is used to assess cognitive abilities, such as selective attention, cognitive flexibility and processing speed and, in general, executive functions (Strauss et al., 2006, pp. 477–499). Stroop test scores are calculated based on performances in word reading and color naming tasks. The word reading and color naming are control tasks where the subject is asked to just (i) read the color words printed in black ink and (ii) name the colors of given bars printed in different inks. In the third task, the subject is shown the color words printed in different ink (conflicting the semantic meaning of the word) and is asked to name the colors in which the words are printed. Since word reading is an automatic process, performance in this task requires the subject to inhibit the reading while focusing on the color naming.

TABLE 3 | Description of the task and the required abilities for the psychological tests administered in the study.

Test	Task	Required abilities
WECHSLER ADULT INTELLIGENCE SCALE III (WAIS-III)		
INFORMATION SUBTEST	Answer general questions	Ability to acquire and retrieve general factual information
SIMILARITIES SUBTEST	Define how two words are similar	Verbal concept formation and reasoning
VOCABULARY SUBTEST	Define the meaning of a word	Word knowledge and verbal concept formation
COMPREHENSION INDEX	Combined score of Information, Similarities and Vocabulary	Verbal reasoning
PICTURE COMPLETION SUBTEST	Find what detail is missing in a picture	Perception and recognition of essential visual information
BLOCK DESIGN SUBTEST	Arrange colored blocks to form visual patterns	Ability to analyze and synthesize abstract visual stimuli
MATRIX REASONING SUBTEST	Chooses the missing part that completes the design	Spatial and classification ability, fluid intelligence
PERCEPTUAL ORGANIZATION INDEX	Combined score of Picture Completion, Block Design and Matrix Reasoning	Non-verbal/spatial reasoning
DIGIT SYMBOL-CODING SUBTEST	Draw symbols that match numbers paired with the symbols	Visuomotor coordination, psychomotor speed, short-term memory
SYMBOL SEARCH SUBTEST	Match symbols with targets	Visuomotor coordination, psychomotor speed, short-term memory
PROCESSING SPEED INDEX	Combined score of Digit Symbol-Coding and Symbol Search	Speed of mental processing
WECHSLER MEMORY SCALE (WMS-III) LETTER-NUMBER SEQUENCING SUBTEST	Recall and mentally rearrange a sequence of digits and letters	Verbal working memory
SPATIAL SPAN SUBTEST	Recall and tap a sequence of spatial positions	Spatial working memory
WORKING MEMORY INDEX	Combined score of Letter-Number Sequencing and Spatial span	Working memory
WORD LISTS I SUBTEST	Recall a list of 12 words presented 4 times	Verbal/episodic memory and learning
WORD LISTS II SUBTEST	Delayed recall of the previous word list	Verbal/episodic long-term memory and retrieval
STROOP TEST		
PART 1	Read names of colors written in black ink (RED, BLUE, GREEN...)	Reading
PART 2	Name the color of bars (XXXX, XXXX, XXXX...)	Naming
PART 3	Name the color of words written in different inks (RED, BLUE, GREEN...) – Comparison to Part 1 or 2	Mental flexibility, divided attention, executive functioning

Typically, the Stroop effect is derived by comparing the correct responses and performance times of the third task to either one of the control tasks by subtracting the control task from the third task. In effect, this subtraction leaves the actual cognitive process we are interested in (the “cost” of inhibiting the response to the automatic word reading process in the third task). Thereby, the Stroop variable used for this study corresponds to the subtraction of the reaction time obtained in the interference task minus the reaction time obtained in the color naming task. The higher the value, the higher the effort needed to selectively filter out unattended information and focus on attended ones.

Wechsler Memory Scale III

The WMS-III is a neuropsychological test designed to measure different memory functions (Wechsler, 1997b). In the present study, we administered the following four WMS-III subtests: Letter-Number Sequencing, Spatial Span (forward and backward), Wordlists I, and Wordlists II, which measure, verbal, spatial, and episodic memory components, respectively. Letter-Number Sequencing and Spatial Span were used to calculate the Working Memory Index (WMI). More details on the tests can be found in **Table 3**.

Big Five Inventory

The BFI contains 44 items designed to measure an individual on five main dimensions of personality: Openness to experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (John and Srivastava, 1999). Items are rated on a five-point scale (where 1 corresponds to strongly disagree and 5 to strongly agree), and the score for each personality dimension corresponds to the average rating for the relevant items. Examples of the multiple possible choices for the item: “*I see myself as someone who...*” are: “*Is talkative,*” “*Is reserved,*” “*Is full of energy,*” and “*Can be tense.*”

Procedure and Statistical Analysis

The participants were invited to the Advanced Magnetic Imaging (AMI) laboratory for the neuroimaging session of the broad Tunteet project (coordinated by EB). There, before and after the brain scanning session, they were administered the following tests by a graduate student of psychology under supervision of a licensed and expert psychologist (TS): Stroop test, WMS-III, and WAIS-III. In another session taking place at the Biomag laboratory at Helsinki Central University Hospital, the same participants were invited for the second part of the brain scanning and personality data were collected by administering

the complete BFI. The total duration of each experimental session was around 3 h. The psychological tests, in total, did not last longer than 2 h. For the purposes of the present study we only used the results of the psychological tests.

Before testing for group differences in cognitive abilities along musicianship, we controlled that there were no significant group differences in background variables. Therefore, two regression models were performed: the first includes background variables of age, years of general education, SES, personality traits variables and years of music playing as predictors of musicianship (classification in non-musicians, amateurs, and musicians); the second model was performed for amateurs and musicians only and included music-background variables such as onset of musical training and musical practice, average of weekly hours spent in practicing and on listening to music, together with years of music playing as predictors of musicianship. By doing so, we obtained the relative contribution of each variable in predicting group differences when holding constant the others.

Age, years of general education, SES and personality traits variables were normally distributed across participants. Because years of music playing was not normally distributed, we decided to square-root the variable and use its transformation in the analyses. Results and means values for each variable of the first regression model are provided in **Table 1**, whereas the others are provided in **Table 2**. The means displayed for years of music playing are the original values (not the square-root transformed).

In order to test for group differences along cognitive abilities, we performed Multivariate analysis of variance (MANOVA) inserting musicianship as between-subjects factor and the main indices of the cognitive tests scores (Stroop, WMS-WMI, and WAIS-FSIQ) as dependent variables. Having more than two dependent variables and because they significantly correlated with each other (FSIQ-WMI: $r = 0.628$, $p < 0.001$; FSIQ-Stroop: $r = -0.337$, $p = 0.001$; WMI-Stroop: $r = -0.327$, $p = 0.001$), we opted for MANOVA. Indeed, such statistical test is able to take the relationship between dependent variables into account (Warne, 2014). Assumptions of linearity and absence of collinearity were tested before proceeding with the analysis. A separate ANOVA was then performed to examine the differences among groups along subtests of the WAIS-FSIQ, namely WAIS-VIQ and WAIS-PIQ. *Post hoc* tests with Bonferroni correction were performed for both the MANOVA and ANOVA models to avoid false positive discoveries while calculating group comparisons. The Bonferroni-adjusted alpha level for *post hoc* tests was obtained by dividing the standard alpha at 0.05 by the number of comparisons [defined as $N(N-1)/2$]. In our case, with 3 groups and 3 variables, there were 9 comparisons; thus, the alpha level was reduced at 0.0056. To be noted, the p -values reported in the Results section are Bonferroni-corrected p -values, so that a corrected $p < 0.05$ corresponds to a non-corrected $p < 0.005$ and hence is interpreted as a significant effect.

To deepen the exploration of the relationship between musical practice and cognitive abilities, and to control for the influence of potential confounding variables, we performed three backward stepwise linear regression analyses inserting background variables of age, years of education, years of music

playing and the five personality trait indices as predictors of FSIQ, WMI, and Stroop, respectively. By doing so, we would obtain the unique contribution of each of the variables, and of music practice, in predicting the variance observed in the cognitive test scores. Backward stepwise regression starts with a saturated model and gradually eliminates (stepwise) variables from the regression model in order to find the predictors that best explain variance in the dependent variable. Therefore, multiple models are generated until model fit cannot be further improved.

Because of missing data along some of the demographic variables, these models only included 60 participants from our sample (20 subjects per group). Therefore, to estimate curve fit along our whole sample we performed three further independent linear regression models by inserting years of music playing as the only predictor of, respectively, FSIQ, WMI, and Stroop. Lastly, to estimate the partial correlation of each of the cognitive measures to musical practice, FSIQ, WMI, and Stroop were included in the same model and regressed against years of music playing. Models' effect sizes are always reported as adjusted R^2 .

RESULTS

Musicians, Amateurs and Non-musicians

As compared to amateurs, musicians had spent more years and hours practicing an instrument. Musical background information for amateurs and musicians is provided in **Table 2**, along with mean, SD and the associated p , B , partial correlation values and 95% confidence interval (CI) for the group comparisons.

Despite the absence of differences in background variables, musicians performed better in all cognitive tests as compared to the other groups, as shown in the histogram in **Figure 1** and in **Table 4**. The MANOVA performed to compare participants' FSIQ, WMI, and Stroop cognitive scores exhibited a significant group effect: Pillai's Trace [$F(2,98) = 2.889$, $p = 0.01$]. The test of between-subject effects reported significant group differences in WAIS-FSIQ [$F(2,98) = 4.00$, $p = 0.021$, adjusted $R^2 = 0.057$], WMS-WMI [$F(2,98) = 4.11$, $p = 0.019$, adjusted $R^2 = 0.059$], Stroop [$F(2,98) = 6.68$, $p = 0.002$, adjusted $R^2 = 0.102$] as provided in **Table 4** (on the left side). The effect sizes of these adjusted R^2 are moderate (Cohen, 1988). A separate ANOVA model was performed to assess group differences along VIQ and PIQ and reported significant differences for the former only: VIQ [$F(2,98) = 3.46$, $p = 0.035$]; PIQ [$F(2,98) = 1.95$, $p = 0.148$]. Results are provided on the right side of **Table 4**.

Post hoc tests controlled by Bonferroni correction reported significantly higher values in favor of musicians as opposed to non-musicians for all the different tests: FSIQ ($p = 0.018$), WMI ($p = 0.018$), Stroop ($p = 0.001$), and VIQ ($p = 0.030$) as provided in **Table 4**. In turn, amateurs did not differ significantly from musicians and non-musicians in either of tests.

To deepen our understanding of the relationship between musical training and cognitive abilities, we performed stepwise backward linear regression modeling by inserting 8 demographic variables (age, years of music playing, years of general education and personality indices of neuroticism, extraversion, openness to

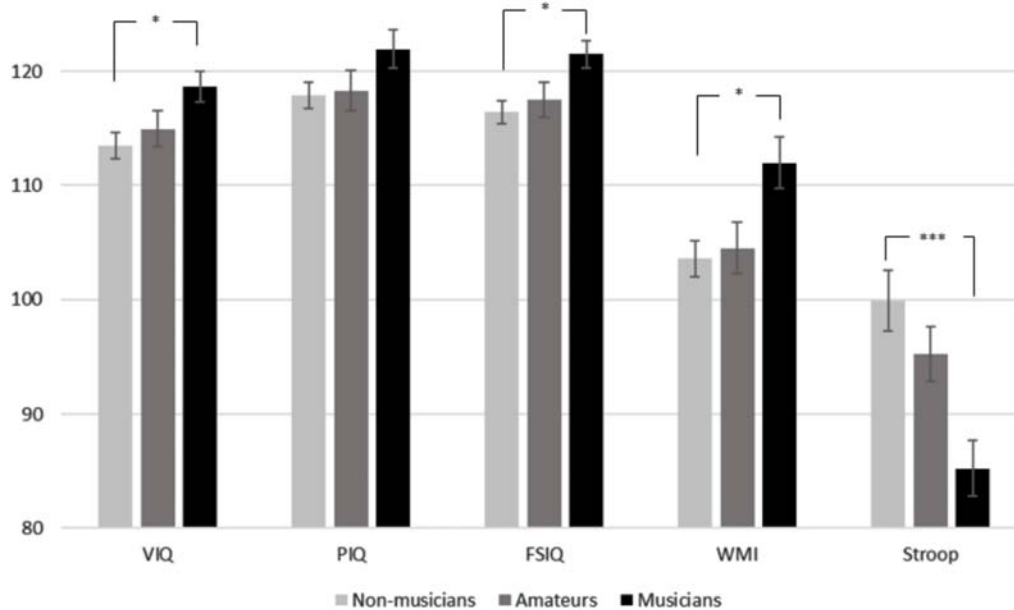


FIGURE 1 | Cognitive scores means along Wechsler Adult Intelligence Scale III (WAIS-III) subscales (in order: Verbal IQ, Performance IQ, Full Scale IQ), Working Memory Index from WMS and Stroop test reported according to musicianship. Bars show confidence intervals (CI). Musicians performed better in all tests compared to other groups. To be noted, Stroop values refer to reaction times: smaller reaction times indicate better performance. Alpha level is Bonferroni-corrected, so that $p < 0.05$ is a significant effect; * $p \leq 0.05$, ** $p \leq 0.01$, and *** $p \leq 0.001$.

experience, agreeableness and conscientiousness) as predictors of FSIQ, WMI, and Stroop (independently). Backward regressions generated in all cases 7 models in which all variables mentioned above were excluded one-by-one except for years of musical practice, which resulted, in all the cases, the only significant factor associated with FSIQ [$F(1,59) = 6.76$, $p = 0.012$, partial-correlation = 0.321, $\beta = 0.321$], WMI [$F(1,59) = 7.23$, $p = 0.009$, partial-correlation = 0.330, $\beta = 0.330$], and Stroop [$F(1,59) = 6.81$, $p = 0.012$, partial-correlation = -0.324 , $\beta = -0.324$].

This approach evidenced that (i) when holding constant the other background variables, years of music playing was the only factor significantly associated with the three cognitive measures (FSIQ, WMI, and Stroop). Besides, by excluding the other factors from the model, we found that (ii) years of music playing was the only predictor necessary to significantly explain the variance in the dependent variables.

Because of missing values in some background variables, not all of the participants were included in these regression models. When regressed independently against years of music playing, WMI and Stroop showed a significant association: WMI [$F(1,99) = 7.80$, $p = 0.006$, adjusted $R^2 = 0.064$, $\beta = 0.27$], Stroop [$F(1,99) = 8.46$, $p = 0.004$, adjusted $R^2 = 0.069$, $\beta = -0.28$], and FSIQ [$F(1,99) = 3.37$, $p = 0.069$, adjusted $R^2 = 0.023$, $\beta = 0.18$]. **Figure 2** represents the curve fit for years of music playing with WMI and Stroop: WMI showed a positive association with music playing duration, whereas Stroop showed a negative relation.

To test for the unique association of each of the cognitive variables with years of music playing, an additional model included the three cognitive variables (FSIQ, WMI, and Stroop) and regressed them against years of music playing. This

model resulted significant [$F(3,97) = 4.55$, $p = 0.005$, adjusted $R^2 = 0.096$]. When holding constant the other cognitive variables, WMI and Stroop were still significantly associated with years of musical playing, whereas FSIQ was not: WMI ($p = 0.041$, $\beta = 0.24$), Stroop ($p = 0.022$, $\beta = -0.24$), and FSIQ ($p = 0.70$, $\beta = -0.05$). Standardized coefficients, significance levels, 95% CI and partial correlation values of the regression model for the three cognitive variables (FSIQ, WMI, and Stroop) are provided in **Table 5**. Furthermore, we repeated the regression excluding the zero values for the variable Years of Music Playing. This operation resulted in exclusion of 36 non-musicians scoring less than 1 in the square-root transformed variable years of music playing, hence in a consistent decrease of statistical power. In spite of this, we could still observe tendencies for an association between cognitive measures of FSIQ, WMI, and Stroop and Years of playing: WMI [$F(1,64) = 2.90$, $p = 0.094$, adjusted $R^2 = 0.028$, $\beta = 0.208$], Stroop [$F(1,64) = 4.40$, $p = 0.04$, adjusted $R^2 = 0.05$, $\beta = -0.25$], and FSIQ [$F(1,64) = 3.55$, $p = 0.064$, adjusted $R^2 = 0.038$, $\beta = 0.23$].

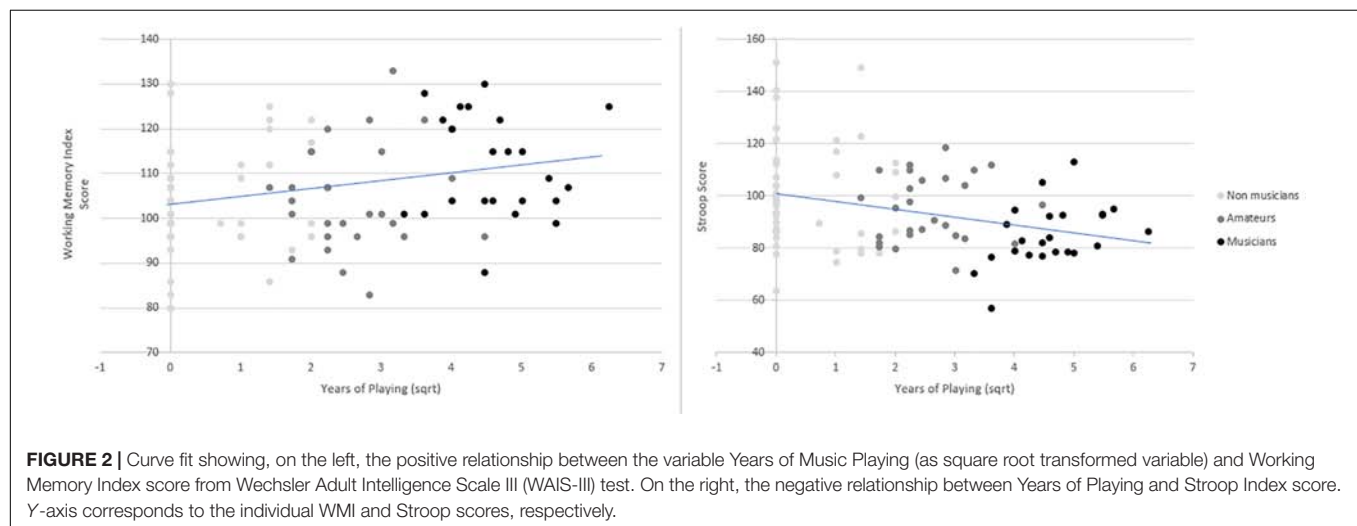
DISCUSSION

The aim of the present study was to investigate the relationship between musical training and higher-order cognitive functions. Although several studies have highlighted anatomical and functional differences between the brains of expert musicians and non-musicians, only few studies have investigated intelligence and executive functions in adults with long-term musical training controlling for possible confounding variables. Our results

TABLE 4 | Mean and statistical group comparisons along cognitive abilities.

Means of cognitive abilities per group					
	WAIS-FSIQ	WMS-WMI	Stroop	WAIS-VIQ	WAIS-PIQ
Non-musicians(NM)	116.45 ± 1.01	103.57 ± 1.61	99.91 ± 2.65	113.51 ± 1.17	117.9 ± 1.16
95% CI	±0.28	±0.44	±0.72	±0.32	±0.32
Amateurs(A)	117.52 ± 1.50	104.52 ± 2.24	95.29 ± 2.43	114.96 ± 1.56	118.33 ± 1.78
95% CI	±0.57	±0.85	±0.92	±0.59	±0.67
Musicians(M)	121.5 ± 1.22	112 ± 2.29	85.22 ± 2.45	118.67 ± 1.38	121.96 ± 1.69
95% CI	±0.50	±0.96	±1.00	±0.56	±0.69
Mean differences of cognitive abilities					
A – NM	1.07	0.95	–4.62	1.45	0.431
M – A	4.00	7.13	10.06	3.70	3.62
M – NM	5.07	8.08	–14.69	5.16	4.06
Between-subject MANOVA results			Between-subject ANOVA results		
F(2,98)	4.00	4.11	6.68	3.46	1.95
p	0.021	0.019	0.002	0.035	0.148
η ² p	0.057	0.059	0.102	0.046	0.018
ωp ²	0.56	0.058	0.101	0.046	0.018
Post hoc p (Bonferroni corrected)			Post hoc p (Bonferroni corrected)		
A – NM (p)	1.000	1.000	0.683	1.000	1.000
95% CI	–3.10 5.23	–5.71 7.61	–13.9 4.65	–3.14 6.04	–4.51 5.37
M – A (p)	0.157	0.093	0.087	0.297	0.398
95% CI	–0.96 8.97	–0.8 15.07	–21.12 0.99	–1.71 9.11	–2.20 9.45
M – NM (p)	0.018	0.018	0.001	0.030	0.172
95% CI	0.68 9.46	1.06 15.11	–24.47 –4.90	0.380 9.93	–1.08 9.20

Top: means, standard errors and mean difference for the scores of Verbal Intelligence Quotient (VIQ), Performance Intelligence Quotient (PIQ), Full-Scale Intelligence Quotient (FSIQ), Working Memory Index (WMI), and Stroop for participants grouped into non-musicians, amateurs and musicians. To be noted, Stroop values refer to reaction times: smaller reaction times indicate better performance. Bottom: on the left side, results of the Multivariate analysis of variance (MANOVA) performed to compare musicians (M), amateurs (A) and non-musicians (NM) reporting significant differences along Full Scale IQ (WAIS-FSIQ), Working Memory Index (WMS-WMI), and Stroop test. On the right side, results of a further ANOVA showing the differences between VIQ and PIQ scores. Post hoc comparisons and p-values corrected for Bonferroni correction are provided at the end of the table for both models. Significant statistical values are reported in bold.



contribute to the literature by showing that adults exposed to professional long-term musical training outperform adult non-musicians in standardized cognitive tasks designed to measure

general intelligence (g), WM and attentive abilities and that these group differences are not associated with any of the examined background variables except for duration of musical playing.

TABLE 5 | Regression model for cognitive abilities. Standardized coefficients for the regression model where Stroop, FSIQ and WMI were included and regressed against the variable “years of music playing.”

	Standardized coefficients				
	Beta	t	Sig.	95% CI	Partial correlation
(Constant)		0.537	0.593	−5.221 9.09	
Stroop	−0.241	−2.328	0.022	−0.049 −0.004	−0.230
FSIQ	−0.047	−0.386	0.700	−0.072 0.270	−0.039
WMI	0.238	2.067	0.041	0.001 0.072	0.205

Beta and t-values, as well as significance levels and 95% confidence intervals (CI) are provided for each of the variables when holding constant the others. Significant statistical values are reported in bold.

Specifically, when testing group differences with analysis of variance we found significantly higher performance in musicians compared to non-musicians in the cognitive tests' general indices of WMI, FSIQ, and Stroop and significantly greater scores. Slightly less strong effects were obtained in one subscale of the WAIS-III, assessing (VIQ). Moreover, by using regression models we noticed an association between all participants' cognitive abilities and years of music playing, which, however, did not explain all of the observed variance. To be noted, this association might not necessarily be linear because when removing the participants completely lacking any musical background, the significance threshold of the association was not reached. Nevertheless, overall these findings converge to demonstrate a positive relationship between musical training and cognitive functions. As proposed by Schellenberg (2006), this relationship could be seen as a continuum dependent on the duration and intensity of training. These findings are in line with previous research (both correlational and experimental) showing associations between musical training and intelligence measures (Chan et al., 1998; Gromko and Poorman, 1998; Cheek and Smith, 1999; Hetland, 2000; Brandler and Rammsayer, 2003; Brochard et al., 2004; Schellenberg, 2004, 2006; Swaminathan et al., 2017, 2018). For instance, a previous study with children conducted by Schellenberg (2006) reported an association between musical training and cognitive abilities (VIQ, FSIQ). Lastly, the increased VIQ we found when comparing musicians to non-musicians might be in line with previous findings, which associated music training with improvements in verbal abilities, such as reading (Anvari et al., 2002; Kraus et al., 2014) and phonology (Francois et al., 2015; see Moreno, 2009 for a review). However, differently from previous studies, we show that these effects are not affected by potential confounding variables: indeed, by controlling for age, education years, SES and personality variables, we demonstrate that the relationship between executive functions and years of music playing is statistically significant. In particular, although not explaining all of the variance, we found a positive effect for WM: the longer the musical practice, the higher the WM functions. In turn, Stroop attentive scores show a negative slope: the longer the musical practice, the shorter interference time, the bigger the attentive abilities.

Together with previous studies, our results allow us to argue that cognitive benefits associated with music practice might be evident along the lifespan. In accordance with previous evidence, we argue that ‘widespread effects of musical training on cognitive processing might occur because music lessons train attentional and executive functioning, which benefit almost all cognitive tasks’ (Hannon and Trainor, 2007). It is important to consider, though, that additional variables not considered in the present study, such as genetic factors and parental personality traits, might have had a relevant influence on the choice of our participants to engage in and persist with musical training (Mosing et al., 2014; Mosing and Ullen, 2018), as well as in the development of their cognitive abilities, as pointed out by previous studies (Butkovic et al., 2015; Corrigan and Schellenberg, 2015). Moreover, cognitive advantages might be evident particularly for those individuals who take music lessons and/or play music in addition to other academic and studying activities. Indeed, in a previous study the cognitive effects of musical training were not visible in participants who studied only music (Schellenberg, 2011). In contrast, all participants in our sample were recruited among university students and qualified professionals and reported a mean IQ higher than the average Finnish population. This long academic background might be the key difference between the present and previous studies. Because musicians differed from the other participants only regarding their musical expertise and given the positive relation between executive functions (WM, Stroop-derived attentive score) and length of musical training, our results suggest that cognitive abilities might be influenced by musical practice.

We suggest that the observed differences in cognitive performance might represent the behavioral manifestations of brain differences identified when comparing musicians with non-musicians. Indeed, neuroimaging and neurophysiological studies showed stronger and faster neural responses (especially to sounds) and enlarged neuroanatomical structures in musicians as compared with non-musicians, particularly in (pre)motor, auditory, prefrontal and visual regions (Munte et al., 2002; Gaser and Schlaug, 2003; Pantev et al., 2003; Zatorre and McGill, 2005; Hyde et al., 2009; Zuk et al., 2014; Baer et al., 2015; Bonetti et al., 2017, 2018). These modifications of brain functionality and anatomy have been associated with use-dependent regional growth of neuronal cells engaged throughout the training and their structural adaptation in response to the intense environmental demands of music practice (for reviews, see Kolb and Muhammad, 2014; Reybrouck and Brattico, 2015).

To conclude, our study highlights an association between musical training and cognitive abilities. We showed that adult participants with similar educational background but varying in their musical expertise exhibited differences in intelligence, working-memory and attentive abilities and that executive functions are significantly associated with the duration of music practice.

ETHICS STATEMENT

Coordinating Ethics Committee of the Hospital District of Helsinki and Uusimaa (approval number: 315/13/03/00/11, obtained on March the 11th, 2012).

AUTHOR CONTRIBUTIONS

EB conceived the study, prepared the ethics permission, coordinated the data collection, designed the initial statistical analyses, and fully edited the final version of the manuscript. AC conducted the final statistical analyses together with LB and wrote the initial version of the manuscript. MK contributed to the data collection, experimental design as well as coordinated the data scoring, and ensured the data quality. TS selected the psychological tests, and supervised the data collection and scoring. All authors

contributed to writing and approved the final version of the manuscript.

FUNDING

The study was supported by various funds from the University of Helsinki, Aalto University, and the Academy of Finland. The Center for Music in the Brain (MIB) is supported by the Danish National Research Foundation (DNRF 117).

ACKNOWLEDGMENTS

We wish to thank the entire team who contributed to the “Tunteet” data collection: Brigitte Bogert, Benjamin Gold, Johanna Normström, Taru Numminen-Kontti, Mikko Heimölä, Toni Auranen, Marita Kattelus, Jyrki Mäkelä, Mikko Sams, Petri Toivainen, Mari Tervaniemi, Anja Thiede, and Alessio Falco.

REFERENCES

- Alluri, V., Toivainen, P., Burunat, I., Kliuchko, M., Vuust, P., and Brattico, E. (2017). Connectivity patterns during music listening: evidence for action-based processing in musicians. *Human Brain Mapp.* 38, 2955–2970. doi: 10.1002/hbm.23565
- Anvari, S. H., Trainor, L. J., Woodside, J., and Levy, B. A. (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. *J. Exp. Child Psychol.* 83, 111–130. doi: 10.1016/S0022-0965(02)00124-8
- Baer, L. H., Park, M. T. M., Bailey, J. A., Chakravarty, M. M., Li, K. Z. H., and Penhune, V. B. (2015). Regional cerebellar volumes are related to early musical training and finger tapping performance. *NeuroImage* 109, 130–139. doi: 10.1016/j.neuroimage.2014.12.076
- Bangert, M., and Schlaug, G. (2006). Specialization of the specialized in features of external human brain morphology. *Eur. J. Neurosci.* 24, 1832–1834. doi: 10.1111/j.1460-9568.2006.05031.x
- Bonetti, L., and Costa, M. (2017). Musical mode and visual-spatial cross-modal associations in infants and adults. *Mus. Sci.* 23, 50–68. doi: 10.1177/1029864917705001
- Bonetti, L., Haumann, N. T., Brattico, E., Kliuchko, M., Vuust, P., Särkämö, T., et al. (2018). Auditory sensory memory and working memory skills: association between frontal MMN and performance scores. *Brain Res.* 1700, 86–98. doi: 10.1016/j.brainres.2018.06.034
- Bonetti, L., Haumann, N. T., Vuust, P., Kliuchko, M., and Brattico, E. (2017). Risk of depression enhances auditory Pitch discrimination in the brain as indexed by the mismatch negativity. *Clin. Neurophysiol.* 128, 1923–1936. doi: 10.1016/j.clinph.2017.07.004
- Brandler, S., and Rammsayer, T. H. (2003). Differences in mental abilities between musicians and non-musicians. *Psychol. Music.* 31, 123–138. doi: 10.1177/0305735603031002290
- Brattico, E., Pallesen, K. J., Varyagina, O., Bailey, C., Anourova, I., Järvenpää, M., et al. (2009). Neural discrimination of nonprototypical chords in music experts and laymen: an MEG study. *J. Cogn. Neurosci.* 21, 2230–2244. doi: 10.1162/jocn.2008.21144
- Brochard, R., Dufour, A., and Després, O. (2004). Effect of musical expertise on visuospatial abilities: evidence from reaction times and mental imagery. *Brain Cogn.* 54, 103–109. doi: 10.1016/S0278-2626(03)00264-1
- Burunat, I., Alluri, V., Toivainen, P., Numminen, J., and Brattico, E. (2014). Dynamics of brain activity underlying working memory for music in a naturalistic condition. *Cortex* 57, 254–269. doi: 10.1016/j.cortex.2014.04.012
- Burunat, I., Brattico, E., Hartmann, M., Vuust, P., Särkämö, T., and Toivainen, P. (2018). Musical training predicts cerebello-hippocampal coupling during music listening. *Psychomusicol. Music Mind Brain* 28, 152–163. doi: 10.1037/pmu0000215
- Burunat, I., Brattico, E., Puoliväli, T., Ristaniemi, T., Sams, M., Toivainen, P., et al. (2015). Action in perception: prominent visuo-motor functional symmetry in musicians during music listening. *PLoS One* 10:e0138238. doi: 10.1371/journal.pone.0138238
- Burunat, I., Toivainen, P., Alluri, V., Bogert, B., Ristaniemi, T., Sams, M., et al. (2016). The reliability of continuous brain responses during naturalistic listening to music. *NeuroImage* 124(Pt A), 224–231. doi: 10.1016/j.neuroimage.2015.09.005
- Butkovic, A., Ullén, F., and Mosing, M. A. (2015). Personality related traits as predictors of music practice: underlying environmental and genetic influences. *Pers. Individ. Differ.* 74, 133–138. doi: 10.1016/j.paid.2014.10.006
- Chan, A. S., Ho, Y. C., and Cheung, M. C. (1998). Music training improves verbal memory. *Nature* 396:128. doi: 10.1038/24075
- Cheek, J. M., and Smith, L. R. (1999). Music training and mathematics achievement. *Adolescence* 34, 759–761.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Cambridge, MA: Academic Press.
- Collins, A., and Koechlin, E. (2012). Reasoning, learning, and creativity: frontal lobe function and human decision-making. *PLoS Biol.* 10:e1001293. doi: 10.1371/journal.pbio.1001293
- Corrigall, K. A., and Schellenberg, E. G. (2015). Predicting who takes music lessons: parent and child characteristics. *Front. Psychol.* 6:1–8. doi: 10.3389/fpsyg.2015.00282
- Corrigall, K. A., Schellenberg, E. G., and Misura, N. M. (2013). Music training, cognition, and personality. *Front. Psychol.* 4:222. doi: 10.3389/fpsyg.2013.00222
- Diamond, A. (2013). Executive functions. *Ann. Rev. Psychol.* 64, 135–168. doi: 10.1146/annurev-psych-113011-143750
- Duncan, J., Seitz, R. J., Kolodny, J., Bor, D., Herzog, H., Ahmed, A., et al. (2000). A neural basis for general intelligence. *Science* 289, 457–460. doi: 10.1126/science.289.5478.457
- Forgeard, M., Winner, E., Norton, A., and Schlaug, G. (2008). Practicing a musical instrument in childhood is associated with enhanced verbal ability and nonverbal reasoning. *PLoS One* 3:e3566. doi: 10.1371/journal.pone.0003566
- François, C., Chobert, J., Besson, M., and Schön, D. (2013). Music training for the development of speech segmentation. *Cerebral Cortex* 23, 2038–2043. doi: 10.1093/cercor/bhs180
- Francois, C., Grau-Sánchez, J., Duarte, E., and Rodriguez-Fornells, A. (2015). Musical training as an alternative and effective method for neuro-education and neuro-rehabilitation. *Front. Psychol.* 6:1–15. doi: 10.3389/fpsyg.2015.00475

- Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., and Pantev, C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *J. Cogn. Neurosci.* 16, 1010–1021. doi: 10.1162/0898929041502706
- Garner, J. (2009). Conceptualizing the relations between executive functions and self-regulated learning. *J. Psychol.* 143, 405–426. doi: 10.3200/JRLP.143.4.405-426
- Gaser, C., and Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *J. Neurosci.* 23, 9240–9245. doi: 10.1523/jneurosci.23-27-09240.2003
- Gold, B. P., Frank, M. J., Bogert, B., and Brattico, E. (2013). Pleasurable music affects reinforcement learning according to the listener. *Front. Psychol.* 4:541. doi: 10.3389/fpsyg.2013.00541
- Gottfredson, L. S. (1997). Mainstream science on intelligence: an editorial with 52 signatories, history, and bibliography. *Intelligence* 24, 13–23. doi: 10.1016/S0160-2896(97)90011-8
- Gromko, J. E., and Poorman, A. S. (1998). The effect of music training on preschoolers' spatial-temporal task performance. *J. Res. Music Educ.* 46, 173–181. doi: 10.2307/3345621
- Guare, R., and Dawson, P. (2004). Executive skills in children and teens—Parents, teachers and clinicians can help. New York, NY: Guilford Press.
- Hannon, E. E., and Trainor, L. J. (2007). Music acquisition: effects of enculturation and formal training on development. *Trends Cogn. Sci.* 11, 466–472. doi: 10.1016/j.tics.2007.08.008
- Hansen, M., Wallentin, M., and Vuust, P. (2013). Working memory and musical competence of musicians and non-musicians. *Psychol. Music* 41, 779–793. doi: 10.1177/0305735612452186
- Helmbold, N., Rammsayer, T., and Altenmüller, E. (2005). Differences in primary mental abilities between musicians and nonmusicians. *J. Individ. Differ.* 26, 74–85. doi: 10.1027/1614-0001.26.2.74
- Herholz, S. C., Lappe, C., and Pantev, C. (2009). Looking for a pattern: an MEG study on the abstract mismatch negativity in musicians and nonmusicians. *BMC Neurosci.* 10:42. doi: 10.1186/1471-2202-10-42
- Hetland, L. (2000). Learning to make music enhances spatial reasoning. *J. Aesthet. Educ.* 34, 179–238. doi: 10.2307/3333643
- Hollingshead (1975). Four factor index of socioeconomic status. *Yale. J. Soc.* 8, 21–52.
- Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., et al. (2009). The effects of musical training on structural brain development: a longitudinal study. *Ann. N. Y. Acad. Sci.* 1169, 182–186. doi: 10.1111/j.1749-6632.2009.04852.x
- John, O. P., and Srivastava, S. (1999). The big five trait taxonomy: history, measurement, and theoretical perspectives. *Handb. Pers. Theory Res.* 2, 102–138.
- Kleber, B., Zeitouni, A. G., Friberg, A., and Zatorre, R. J. (2013). Experience-dependent modulation of feedback integration during singing: role of the right anterior insula. *J. Neurosci.* 33, 6070–6080. doi: 10.1523/jneurosci.4418-12.2013
- Kliuchko, M., Heinonen-Guzejev, M., Monacis, L., Gold, B. P., Heikkilä, K. V., Spinoza, V., et al. (2015). The association of noise sensitivity with music listening, training, and aptitude. *Noise Health* 17, 350–357. doi: 10.4103/1463-1741.165065
- Kliuchko, M., Heinonen-Guzejev, M., Vuust, P., Tervaniemi, M., Brattico, E., Moss, F., et al. (2016). A window into the brain mechanisms associated with noise sensitivity. *Sci. Rep.* 6:39236. doi: 10.1038/srep39236
- Kliuchko, M., Puoliväli, T., Heinonen-Guzejev, M., Tervaniemi, M., Toivainen, P., Sams, M., et al. (2018). Neuroanatomical substrate of noise sensitivity. *NeuroImage* 167, 309–315. doi: 10.1016/j.neuroimage.2017.11.041
- Koenke, S., Lutz, K., Wüstenberg, T., and Jäncke, L. (2004). Long-term training affects cerebellar processing in skilled keyboard players. *Neuroreport* 15, 1279–1282. doi: 10.1097/01.wnr.0000127463.10147
- Kolb, B., and Muhammad, A. (2014). Harnessing the power of neuroplasticity for intervention. *Front. Human Neurosci.* 8:1–13. doi: 10.3389/fnhum.2014.00377
- Kraus, N., and Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nat. Rev. Neurosci.* 11, 599–605. doi: 10.1038/nrn2882
- Kraus, N., Slater, J., Thompson, E. C., Hornickel, J., Strait, D. L., Nicol, T., et al. (2014). Auditory learning through active engagement with sound: biological impact of community music lessons in at-risk children. *Front. Neurosci.* 8:351. doi: 10.3389/fnins.2014.00351
- Krings, T., Töpper, R., Foltys, H., Erberich, S., Sparing, R., Willmes, K., et al. (2000). Cortical activation patterns during complex motor tasks in piano players and control subjects. a functional magnetic resonance imaging study. *Neurosci. Lett.* 278, 189–193. doi: 10.1016/S0304-3940(99)00930-1
- Moreno, S. (2009). Can music influence language and cognition? *Contemp. Music Rev.* 28, 329–345. doi: 10.1080/07494460903404410
- Mosing, M. A., Madison, G., Pedersen, N. L., Kuja-Halkola, R., and Ullen, F. (2014). Practice does not make perfect: no causal effect of music practice on music ability. *Psychol. Sci.* 25, 1795–1803. doi: 10.1177/0956797614541990
- Mosing, M. A., and Ullen, F. (2018). Genetic influences on musical specialization: a twin study on choice of instrument and music genre. *Ann. N. Y. Acad. Sci.* doi: 10.1111/nyas.13626 [Epub ahead of print].
- Münste, T. F., Altenmüller, E., and Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. *Nat. Rev. Neurosci.* 3, 473–478. doi: 10.1038/nrn843
- Pantev, C., Ross, B., Fujioka, T., Trainor, L. J., Schulte, M., and Schulz, M. (2003). Music and Learning-Induced Cortical Plasticity. *Ann. N. Y. Acad. Sci.* 999, 438–450. doi: 10.1196/annals.1284.054
- Perrotin, A., Tournelle, L., and Isingrini, M. (2008). Executive functioning and memory as potential mediators of the episodic feeling-of-knowing accuracy. *Brain Cogn.* 67, 76–87. doi: 10.1016/j.bandc.2007.11.006
- Rauscher, F. (2002). Music education: a resource for children. *Can. Music Educ.* 44, 8–12.
- Reybrouck, M., and Brattico, E. (2015). Neuroplasticity beyond sounds: neural adaptations following long-term musical aesthetic experiences. *Brain Sci.* 5, 69–91. doi: 10.3390/brainsci5010069
- Saari, P., Burunat, I., Brattico, E., and Toivainen, P. (2018). Decoding musical training from dynamic processing of musical features in the brain. *Sci. Rep.* 8:708. doi: 10.1038/s41598-018-19177-5
- Saarikivi, K., Putkinen, V., Tervaniemi, M., and Huotilainen, M. (2016). Cognitive flexibility modulates maturation and music-training-related changes in neural sound discrimination. *Eur. J. Neurosci.* 44, 1815–1825. doi: 10.1111/ejnn.13176
- Sala, G., and Gobet, F. (2017). When the music's over. does music skill transfer to children's and young adolescents' cognitive and academic skills? a meta-analysis. *Educ. Res. Rev.* 20, 55–67. doi: 10.1016/j.edurev.2016.11.005
- Salthouse, T. A., and Pink, J. E. (2008). Why is working memory related to fluid intelligence? *Psychon. Bull. Rev.* 15, 364–371. doi: 10.3758/PBR.15.2.364
- Schellenberg, E. G. (2004). Music lessons enhance IQ. *Psychol. Sci.* 15, 511–514. doi: 10.1111/j.0956-7976.2004.00711.x
- Schellenberg, E. G. (2006). Long-term positive associations between music lessons and IQ. *J. Educ. Psychol.* 98, 457–468. doi: 10.1037/0022-0663.98.2.457
- Schellenberg, E. G. (2011). Examining the association between music lessons and intelligence. *Br. J. Psychol.* 102, 283–302. doi: 10.1111/j.2044-8295.2010.02000.x
- Schellenberg, E. G., and Moreno, S. (2010). Music lessons, pitch processing, and g. *Psychol. Music* 38, 209–221. doi: 10.1177/0305735609339473
- Schellenberg, E. G., and Weiss, M. W. (2013). "Music and cognitive abilities," in *The Psychology of Music*, ed. D. Deutsch (San Diego, CA: Elsevier Academic Press).
- Schlaug, G. (2015). Musicians and music making as a model for the study of brain plasticity. *Prog. Brain Res.* 217, 37–55. doi: 10.1016/bs.pbr.2014.11.020
- Sluming, V., Brooks, J., Howard, M., Downes, J. J., and Roberts, N. (2007). Broca's area supports enhanced visuospatial cognition in orchestral musicians. *J. Neurosci.* 27, 3799–3806. doi: 10.1523/JNEUROSCI.0147-07.2007
- Strauss, E., Sherman, E. M. S., and Spreen, O. (2006). *A Compendium of Neuropsychological Tests: Administration, Norms and Commentary*. Oxford: Oxford University Press.
- Swaminathan, S., Schellenberg, E. G., and Khalil, S. (2017). Revisiting the association between music lessons and intelligence: training effects or music aptitude? *Intelligence* 62, 119–124. doi: 10.1016/j.intell.2017.03.005

- Swaminathan, S., Schellenberg, E. G., and Venkatesan, K. (2018). Explaining the association between music training and reading in adults. *J. Exp. Psychol. Learn. Mem. Cogn.* 44, 992–999. doi: 10.1037/xlm0000493
- Tervaniemi, M. (2001). Superior formation of cortical memory traces for melodic patterns in musicians. *Learn. Mem.* 8, 295–300. doi: 10.1101/lm.39501
- Tervaniemi, M. (2009). Musicians - Same or different. *Ann. N. Y. Acad. Sci.* 1169, 151–156. doi: 10.1111/j.1749-6632.2009.04591.x
- Van Zuijen, T. L., Sussman, E., Winkler, I., Näätänen, R., and Tervaniemi, M. (2005). Auditory organization of sound sequences by a temporal or numerical regularity - a mismatch negativity study comparing musicians and non-musicians. *Cogn. Brain Res.* 23, 270–276. doi: 10.1016/j.cogbrainres.2004.10.007
- Vuust, P., Pallesen, K. J., Bailey, C., Van Zuijen, T. L., Gjedde, A., Roepstorff, A., et al. (2005). To musicians, the message is in the meter: pre-attentive neuronal responses to incongruent rhythm are left-lateralized in musicians. *NeuroImage* 24, 560–564. doi: 10.1016/j.neuroimage.2004.08.039
- Warne, R. T. (2014). A primer on multivariate analysis of variance (MANOVA) for behavior scientists. *Pract. Assess., Res. Eval.* 19:17.
- Wechsler, D. (1997a). *WAIS-III Administration and Scoring Manual*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (1997b). *Wechsler Memory Scale*, 3rd Edn. San Antonio, TX: The Psychological Corporation.
- Weiss, L. G., Saklofske, D. H., Holdnack, J. A., Prifitera, A., Shaywitz, B. A., and Shaywitz, S. E. (2016). *WISC-V Assessment and Interpretation*. Cambridge, MA: Academic Press.
- Zatorre, R., and McGill, J. (2005). Music, the food of neuroscience? *Nature* 434, 312–315. doi: 10.1038/434312a
- Zuijen, T. L., van Sussman, E., Winkler, I., Näätänen, R., and Tervaniemi, M. (2004). Grouping of Sequential sounds—an event-related potential study comparing musicians and nonmusicians. *J. Cogn. Neurosci.* 16, 331–338. doi: 10.1162/089892904322984607
- Zuk, J., Benjamin, C., Kenyon, A., and Gaab, N. (2014). Behavioral and neural correlates of executive functioning in musicians and non-musicians. *PLoS One* 9:e99868. doi: 10.1371/journal.pone.0099868

Conflict of Interest Statement: 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.

Copyright © 2019 Criscuolo, Bonetti, Särkämö, Kliuchko and Brattico. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Auditory T-Complex Reveals Reduced Neural Activities in the Right Auditory Cortex in Musicians With Absolute Pitch

Masato Matsuda, Hironaka Igarashi and Kosuke Itoh*

Center for Integrated Human Brain Science, Brain Research Institute, Niigata University, Niigata, Japan

OPEN ACCESS

Edited by:

Claude Alain,
Rotman Research Institute (RRI),
Canada

Reviewed by:

Peter Schneider,
Heidelberg University, Germany
Gavin M. Bidelman,
The University of Memphis,
United States

*Correspondence:

Kosuke Itoh
ito@bri.niigata-u.ac.jp

Specialty section:

This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Neuroscience

Received: 11 March 2019

Accepted: 22 July 2019

Published: 06 August 2019

Citation:

Matsuda M, Igarashi H and Itoh K
(2019) Auditory T-Complex Reveals
Reduced Neural Activities in the Right
Auditory Cortex in Musicians With
Absolute Pitch.
Front. Neurosci. 13:809.
doi: 10.3389/fnins.2019.00809

Absolute pitch (AP) is the ability to identify the pitch names of arbitrary musical tones without being given a reference pitch. The acquisition of AP typically requires early musical training, the critical time window for which is similar to that for the acquisition of a first language. This study investigated the left–right asymmetry of the auditory cortical functions responsible for AP by focusing on the T-complex of auditory evoked potentials (AEPs), which shows morphological changes during the critical period for language acquisition. AEPs evoked by a pure-tone stimulus were recorded in high-AP musicians, low-AP musicians, and non-musicians ($n = 19$ each). A balanced non-cephalic electrode (BNE) reference was used to examine the left–right asymmetry of the N1a and N1c components of the T-complex. As a result, a left-dominant N1c was observed only in the high-AP musician group, indicating “AP negativity,” which has previously been described as an electrophysiological marker of AP. Notably, this hemispheric asymmetry was due to a diminution of the right N1c rather than enhancement of the left N1c. A left-dominant N1a was found in both musician groups, irrespective of AP. N1c and N1a exhibited no left–right asymmetry in non-musicians. Hence, music training and the acquisition of AP are both accompanied by a left-dominant hemispheric specialization of auditory cortical functions, as indexed by N1a and N1c, respectively, but the N1c asymmetry in AP possessors was due to reduced neural activities in the right hemisphere. The use of a BNE is recommended for evaluating these radially oriented components of the T-complex.

Keywords: music training, language, auditory evoked cortical potentials, brain maturation and development, balanced non-cephalic electrode

INTRODUCTION

Absolute pitch (AP) is the ability to identify pitch names of arbitrary musical tones without being given a reference pitch (Miyazaki, 1988, 1990; Takeuchi and Hulse, 1993). In addition to a possible genetic predisposition (Baharloo et al., 2000), the acquisition of AP usually requires early musical training (<7 years of age; Miyazaki, 1988; Takeuchi and Hulse, 1993; Miyazaki and Ogawa, 2006), the critical period similar to acquisition of a first language (Newport, 1990; Kuhl, 2000; Russo et al., 2003; Vitouch, 2003; Deutsch et al., 2006). Specialized neural circuitries that underpin AP are hypothesized to be established through early musical training

(Pantev et al., 1998; Ohnishi et al., 2001), the timing of which is coincident with the maturation of language-related brain functions (Krashen, 1973; Newport, 1990; Kuhl, 2000). These parallels between AP and language are consistent with the view that AP is essentially a verbal function for labeling pitches with their names (Itoh et al., 2005).

Cerebral hemispheric asymmetry is a hallmark of language, and also of AP. AP possessors show left-dominant neural responses to both speech and music stimuli (Ohnishi et al., 2001; Itoh et al., 2005; Schulze et al., 2009, 2013; Wilson et al., 2009; Oechslin et al., 2010b), whereas non-possessors typically show left-dominant neural activities for speech stimuli alone. The geometric location of electromagnetic sources of neural responses to musical pitches are asymmetrically localized in AP possessors but symmetrically localized in non-possessors (Hirata et al., 1999). Functional connectivity in the left superior temporal gyrus is enhanced in AP possessors (Loui et al., 2012). The fractional anisotropy of the superior longitudinal fasciculus, which is a major cortical fiber that is crucial for language and music functions, is left-dominant in musicians with AP, whereas it is symmetrical in musicians with relative pitch (Oechslin et al., 2010a). Volumetric studies have also shown that AP is associated with greater left-dominant asymmetry of gray matter volume of the perisylvian brain areas (Schlaug et al., 1995; Keenan et al., 2001; Ohnishi et al., 2001; Wilson et al., 2009), which was due to either a reduction of the gray matter volume in the right hemisphere (Schlaug et al., 1995; Keenan et al., 2001; Wilson et al., 2009), or its increase in the left hemisphere (Zatorre et al., 1998). However, other studies have reported right-dominant neural activities (Hirose et al., 2005; Wengenroth et al., 2014; Burkhard et al., 2019), increased right auditory cortical volume (Wengenroth et al., 2014), or higher myelination associated with greater functional connectivity in the right auditory cortex (Kim and Knösche, 2016, 2017) in AP possessors. Integrating all these findings into a coherent hypothesis remains difficult. Nevertheless, the available evidence is consistent with the view that the typical hemispheric specialization of cerebral cortical functions for auditory processing is altered in AP possessors, the details of which remain to be elucidated.

This study analyzed the left–right asymmetry of the auditory cortical functions in AP possessors by focusing on the T-complex of auditory evoked potentials (AEPs). The T-complex is a subcomponent of the auditory N1 response, which has multiple generators in the auditory and other brain regions (Näätänen and Picton, 1987). Regarding the auditory cortical generators, two categories of sources can be distinguished on their dipole orientation, tangential or radial (Scherg et al., 1989; Woods, 1995). The tangential sources are located on the superior temporal plane, and the generated electrical potentials project vertically to the fronto-central scalp: the N1b (~100 ms), or the vertex N1, is the most representative example. The radial sources are located on the lateral surface of the temporal lobe, and therefore, the electrical potentials project laterally to the temporal scalp. The T-complex, which is the focus of this study, is a radial component that consists of three peaks that are recorded over the temporal scalp: N1a (75–95 ms), Ta (100–115 ms), and N1c (130–170 ms) (Näätänen and Picton, 1987; Woods,

1995). The sources of T-complex have been estimated in the secondary or higher auditory cortex of the superior temporal gyrus (Scherg et al., 1989; Woods, 1995; Shahin et al., 2003).

The T-complex is a potentially promising index for investigating the hemispheric asymmetry of auditory cortical functions in AP possessors. First, as these potentials originate from radially oriented sources in the temporal lobe, the waves recorded over the left and right temporal scalp reflect neural activities in the left and right auditory cortices, respectively, (Knight et al., 1988). Clear separation of left and right neural activities would be difficult with the tangential components of AEP that are distributed maximally along the midline, such as the N1b. Second, and more importantly, the waveform of the T-complex shows morphological changes during the critical period of language acquisition (Pang and Taylor, 2000; Ponton et al., 2002; Tonnquist-Uhlen et al., 2003; Rinker et al., 2017), which likely reflect maturational changes in auditory cortical functions, and it is also affected by music training in childhood (Shahin et al., 2003). Hence, the functional hemispheric asymmetry for AP may manifest as an altered morphology of the T-complex.

In fact, a previously identified electrophysiological marker of AP, or “AP negativity” (Itoh et al., 2005), closely resembles the N1c component of the T-complex in terms of polarity, amplitude, latency, and scalp distribution. AP negativity is a negative AEP of approximately 1 μ V in amplitude that peaks at approximately 150 ms in latency over the posterior temporal scalp (when recorded with a linked earlobe reference). Its amplitude is greater (i.e., more negative) in the left hemisphere, in accordance with the reported left-dominant pitch processing in AP (Brancucci et al., 2009; Wilson et al., 2009). However, the original study by Itoh et al. (2005) had several issues. First, a linked earlobe reference was used, which was not optimal for evaluating the T-complex; electric potentials from radially oriented sources could have “activated” the earlobe electrodes and thus might have altered the wave morphology. Second, the number of trials for averaging was relatively small (90 trials) for evaluating the small potential of approximately 1 μ V. Furthermore, no subsequent study has yet replicated the elicitation of AP negativity in possessors of AP, to the best of our knowledge.

Therefore, we examined the morphology of the T-complex in AP possessors with a particular focus on the left–right asymmetry of the N1c component with an appropriate study design. Many trials ($n = 300$) were used to record the T-complex and evaluate its amplitudes with high reliability. To obviate the confounding effects of general music training, three groups of participants were used: musicians with high levels of AP (high-AP musicians), musicians with low-levels of AP (low-AP musicians), and non-musicians. Comparisons between the high-AP and low-AP musicians would delineate the effects of AP, whereas comparisons of the two musician groups with the non-musicians would identify the general effects of music training that are not specific to AP.

Furthermore, and quite important, a balanced non-cephalic electrode (BNE; Stephenson and Gibbs, 1951) was employed as the reference for evaluating the AEP amplitudes. Since no electrodes placed on the head (including those on the earlobes)

can be assumed to be ‘quiet’ with respect to cortical potentials, measurements of AEP amplitude are always confounded by neural activities recorded at the cephalic reference channels. This poses a serious problem when the mastoid or earlobe electrodes are used as a reference (or a part of the reference) to evaluate the T-complex, because the radial sources of the T-complex are expected to “activate” these reference channels. The BNE method ameliorates this problem by using a non-cephalic reference that is placed outside the head. Contaminations of the non-cephalic channel by electrocardiograms (ECGs) are eliminated by pairing two electrodes that are placed anteriorly and posteriorly at the base of the neck. As the ECGs recorded from these electrodes have an approximately the same magnitude but with opposite polarities (with respect to the scalp), the cardiac signals can be canceled by balancing the electrical resistance between these two non-cephalic electrodes. This is the first experiment to use the BNE reference to record electroencephalograms (EEGs) in AP possessors, to our knowledge.

MATERIALS AND METHODS

Participants

Fifty-seven healthy undergraduate students (age: 18–26 years, 15 males) participated in this study, after providing written informed consent. The participants were predominantly females, because subjects with AP were more readily available to us in this gender. This was a limitation of the study, as there is some evidence for gender differences in hemispheric lateralization related to music and AP (Luders et al., 2004; Merrett et al., 2013). Nevertheless, our results were apparently not affected by the gender of the participants (see “Results”). All participants were right-handed, as confirmed using the Edinburgh Handedness Inventory (Oldfield, 1971).

The participants were categorized into three groups according to their levels of music training and AP ability: high-AP musicians ($n = 19$, 3 males), low-AP musicians ($n = 19$, 5 males), and non-musicians ($n = 19$, 7 males). All participants had received basic-level music education as one of the curriculum requirements in Japan. Therefore, the level of music training in this study was defined in terms of the number of years in music training given by professional music teachers outside standard school education. With this definition, the average (\pm standard deviation) years in music training was 15.2 (± 2.0) in the high-AP group and 13.8 (± 3.2) in the low-AP group; the difference was not significant ($t = 1.5$, degrees of freedom (df) = 36, $P > 0.05$). The age of commencement of music training was also comparable between the high-AP musicians (4.3 ± 1.2 years old) and the low-AP musicians (4.9 ± 1.5 years old; $t = 1.4$, $df = 36$, $P > 0.05$). The non-musicians had received less than 5 years of music training, and the average was 0.9 (± 1.5) years.

The participant's AP ability was evaluated using a previously described pitch naming test (Itoh et al., 2005), in which they were instructed to identify pitch classes (e.g., C, C#, and D, without distinguishing octaves) of 60 random piano tones covering five octaves. The AP test was essentially identical to the AP test developed by Miyazaki (1988), which has been validated by

many experiments (e.g., Miyazaki, 1990; Itoh et al., 2003, 2005; Miyazaki and Ogawa, 2006; Miyazaki et al., 2012; Itoh and Nakada, 2018). Critical steps were taken to prevent the use of relative pitch. First, no reference tone was presented at any point of the test, and feedback on the accuracy of the participant's responses was not provided. Therefore, the participants had to use their own internal long-term memory to correctly identify the pitch class. Second, the sequence of the test tones was randomized to make the use of relative pitch difficult. Finally, the inter-trial interval was set relatively short. The test sounds were presented every 5 s and the duration of sounds was approximately 1 s long; therefore, the notes had to be identified within a response time window of approximately 4 s. The criterion for high-AP musicians was 90% correct or higher, and for low-AP musicians, it was 40% or lower.

This study conformed to The Code of Ethics of the World Medical Association (Declaration of Helsinki), and was conducted in accordance with the human research guidelines of the Internal Review Board of the University of Niigata.

Stimuli and Procedure

The stimulus was a sinusoidal tone of 1046.5 Hz, which corresponded to the frequency of C6 (American notation). A single, fixed frequency was used, because the use of multiple frequencies introduces pitch changes, neural responses to which might confound the results by the mechanisms of neural adaptation and/or mismatch negativity (Tervaniemi et al., 1993; Elmer et al., 2015; Rogenmoser et al., 2015; Greber et al., 2018). With a repeated presentation of the note, the participants could perceive it as the keynote of the stimulus sequence, but it required AP to identify the specific note as we did not provide any information about the stimulus. The sound had a duration of 350 ms (10-ms rise-time and 50-ms fall-time). The stimulus was presented monaurally (left-ear and right-ear) or binaurally in a random order via air-conduction insert earphones (Eartone 3A, Etymotic Research, Elk Grove Village, IL, United States) with a stimulus intensity of 65 dB SL. A total of 900 tones (300 tones for each ear condition) were presented using randomly varying stimulus onset asynchrony in the range of 1700–1900 ms (mean: 1800 ms). Only binaural AEPs were analyzed in this study, as it is more natural than monaural listening situations; we seldom receive completely lateralized monaural stimulations while listening to music. The monaural data will be analyzed in a subsequent paper.

The participants sat in a comfortable chair in a temperature-controlled and sound-attenuated room during the EEG recordings. They listened to the sounds passively while playing a video game (Nintendo DS, Nintendo Co., Ltd., Kyoto, Japan) to maintain wakefulness and prevent the participants from explicitly labeling the tones.

EEG Recording and Analysis

Twenty-two Ag electrodes were applied to the scalp of each participant according to the international 10–20 system (Jasper, 1958), and the electrodes were positioned at Fp1, Fp2, Fz, F3, F4, F7, F8, Cz, C3, C4, T3, T4, CPz, Pz, P3, P4, T5, T6, O1, O2, and the left and right earlobes. Horizontal and vertical

electro-oculograms (EOGs) were also recorded. In addition, two electrodes were placed on the right sternoclavicular junction and the tip of the seventh cervical spine for recording BNE data (Stephenson and Gibbs, 1951). All electrodes were referenced to CPz while collecting data. The EEGs and EOGs were amplified by a SynAmp amplifier (Neuroscan Labs, El Paso, TX, United States) with 16-bit resolution, a gain of 5000, and an analog–digital conversion rate of 10 kHz, band-passed between 0.05 and 2000 Hz. The electrode's impedance was kept below 5 k Ω .

After data acquisition, the EEG data were downsampled to 1 kHz, re-referenced to the BNE, segmented (from –100 to 200 ms relative to the stimulus onset), and baseline-corrected to the pre-stimulus period average. The segmented data were checked for artifacts using a threshold criterion of $\pm 100 \mu\text{V}$ using the Fp1, Fp2, F7, and F8, and horizontal and vertical EOG channels to remove ocular artifacts. The non-rejected data were averaged and time-locked to the stimulus onset to obtain AEPs for each participant, which were low-pass filtered at 50 Hz (48 dB/oct). The number of non-rejected data segments was in the range of 224–288 (mean 262) in the high-AP musicians, 223–296 (mean 267) in the low-AP musicians, and 220–300 (mean 267) in the non-musicians. Finally, the AEPs were averaged across participants to obtain grand average waveforms.

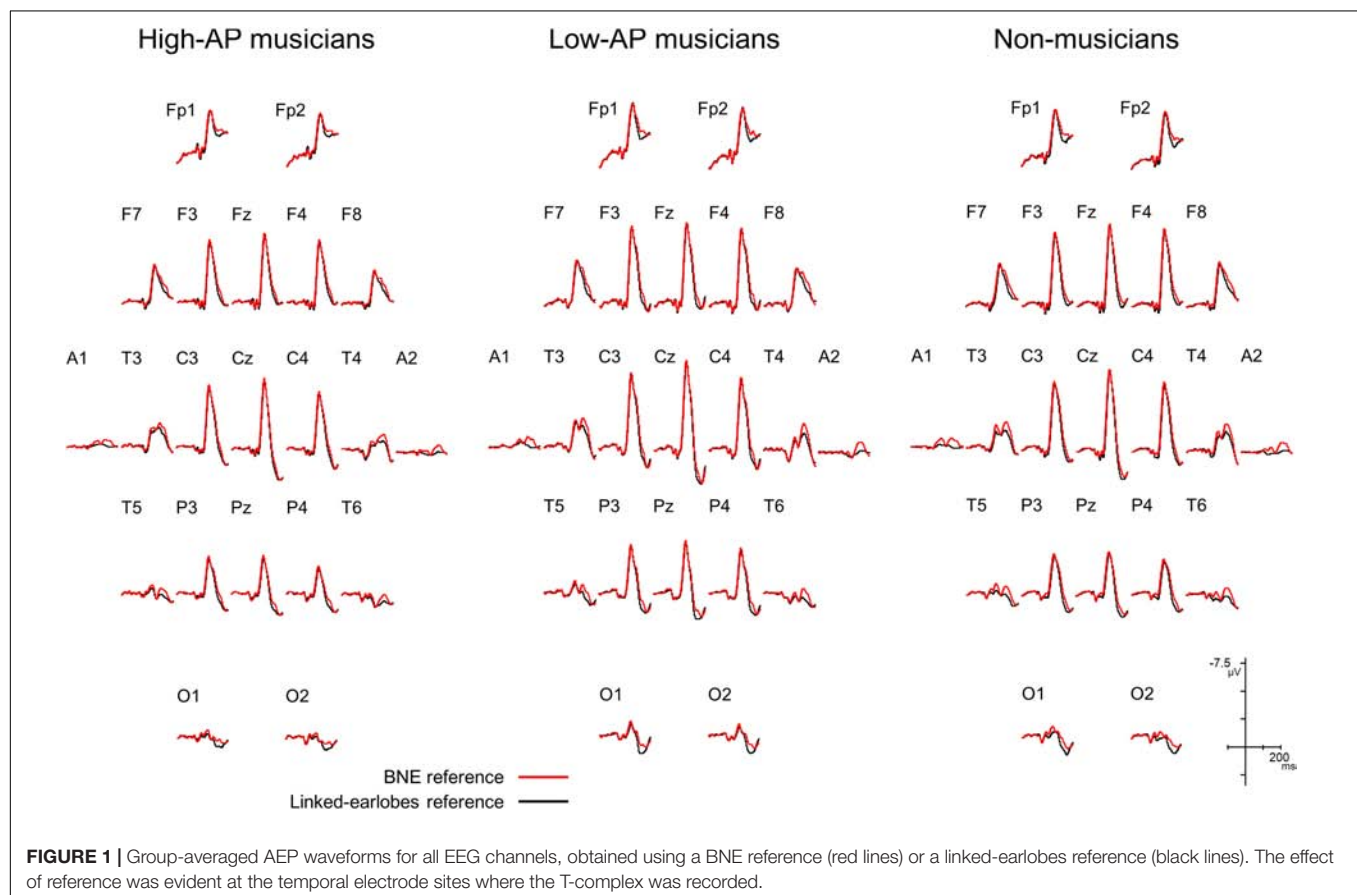
The peak amplitudes of N1a and N1c at the bilateral temporal electrode sites (T3 and T4) were analyzed. The N1b measured at the central electrode site (Cz) was also analyzed for comparison.

The Ta component was not analyzed because it was small in amplitude and difficult to measure. The grand average waveforms were used to define the time windows for N1a (70–80 ms at T3, 73–83 ms at T4), N1b (77–97 ms) and N1c (112–132 ms at T3, 119–139 ms at T4), and the peaks were defined as the most negative point in these time slots. These time windows were centered at the peak latency of the components as identified in the grand average waveforms and had a time width of 10 ms for N1a, and 20 ms for N1b and N1c.

The peak amplitudes of N1a and N1c were analyzed using a mixed-design two-way analysis of variance (ANOVA) with the group (high-AP musicians, low-AP musicians, and non-musicians) as the between-subjects factor and the hemisphere (T3 and T4) as the within-subjects factor. For the N1b amplitude, a one-way ANOVA with the group as the between-subjects factor was performed. An alpha level of $P = 0.05$ was used as the significance criterion. P -values were corrected for multiple comparisons by using the Bonferroni method wherever appropriate.

RESULTS

We first evaluated how the use of BNE reference affected the AEP waveforms. **Figure 1** compares the group-averaged AEPs for all channels, which were obtained using the BNE reference



and the linked-earlobes reference. Substantial differences in the waveforms were observed at the temporal electrode sites. This was an expected result, because the radially oriented dipoles of the T-complex would produce electrical potentials at the earlobe electrodes (Wolpaw and Wood, 1982), which was also confirmed in our BNE-referenced data. When the linked-earlobes reference was used, the potentials at the earlobes were subtracted from the AEPs in the other channels to alter the wave morphologies. Specifically, the T-complex obtained with the linked-earlobes reference was artifactually reduced in amplitude when compared to that obtained with the BNE reference.

Accordingly, we used the BNE reference to evaluate the left–right asymmetry of N1a and N1c amplitudes in high-AP musicians, low-AP musicians, and non-musicians. **Figure 2** plots the T-complex waveforms at the temporal electrode sites, where they were maximal. **Figure 3** shows the N1a and N1c amplitudes for all individual subjects. Four main findings were obtained.

First, only the high-AP musicians showed a left-dominant asymmetry in the N1c amplitude (**Figures 2, 3**). The two-way ANOVA revealed a significant group \times hemisphere interaction [$F(2,54) = 3.4$, $P = 0.042$] and follow-up one-way ANOVAs indicated a significant hemispheric asymmetry in high-AP musicians, [$F(1,54) = 6.4$, $P = 0.014$] but not in low-AP musicians [$F(1,54) = 0.3$, $P = 0.604$] or in non-musicians [$F(1,54) = 1.3$, $P = 0.264$].

Second, the left-dominant asymmetry of N1c in high-AP musicians was due to a diminution of the right N1c rather than enhancement of the left N1c (**Figure 4**). When the N1c amplitude was analyzed using a one-way ANOVA with the group as a factor, the main effect was significant at the right temporal electrode [T4, $F(2,54) = 5.1$, $P = 0.009$] but not at the left temporal electrode [T3, $F(2,54) = 0.8$, $P = 0.469$]. *Post hoc* analyses at T4 indicated that the right N1c amplitude was significantly smaller in the high-AP musicians than in the low-AP musicians ($P = 0.033$, Bonferroni corrected) and non-musicians ($P = 0.017$, Bonferroni corrected).

Third, both high-AP musicians and low-AP musicians showed a left-dominant asymmetry in the N1a amplitude (**Figures 2, 3**). The two-way ANOVA revealed a significant group \times hemisphere interaction [$F(2,54) = 3.3$, $P = 0.043$], indicating that the effect of the hemisphere varied between the groups. Follow-up one-way ANOVAs indicated a significant main effect of the hemisphere in high-AP musicians [$F(1,54) = 18.4$, $P < 0.001$] and low-AP musicians [$F(1,54) = 19.0$, $P < 0.001$] but not in non-musicians [$F(1,54) = 1.3$, $P = 0.251$].

The above analyses were conducted by treating AP as a categorical variable. Considering that AP is a graded trait (Miyazaki, 1990; Bermudez and Zatorre, 2009), we additionally performed multiple regression analyses in which the N1c amplitudes were evaluated with respect to the continuous independent variables of AP test score (AP) and years in music training (Years). As a result, the model was a significant predictor of N1c at the right temporal electrode site T4 [$F(2,35) = 7.1$, $P = 0.003$]: The variable AP contributed significantly to the model ($\beta = 0.562$, $t(35) = 3.8$, $P = 0.001$), but Years did not ($\beta = -0.123$, $t(35) = 0.8$, $P = 0.417$). Regarding the left N1c at T3, the model was not a significant predictor of its amplitude [$F(2,35) = 1.7$, $P = 0.206$], with an R^2 of 0.086: The contribution

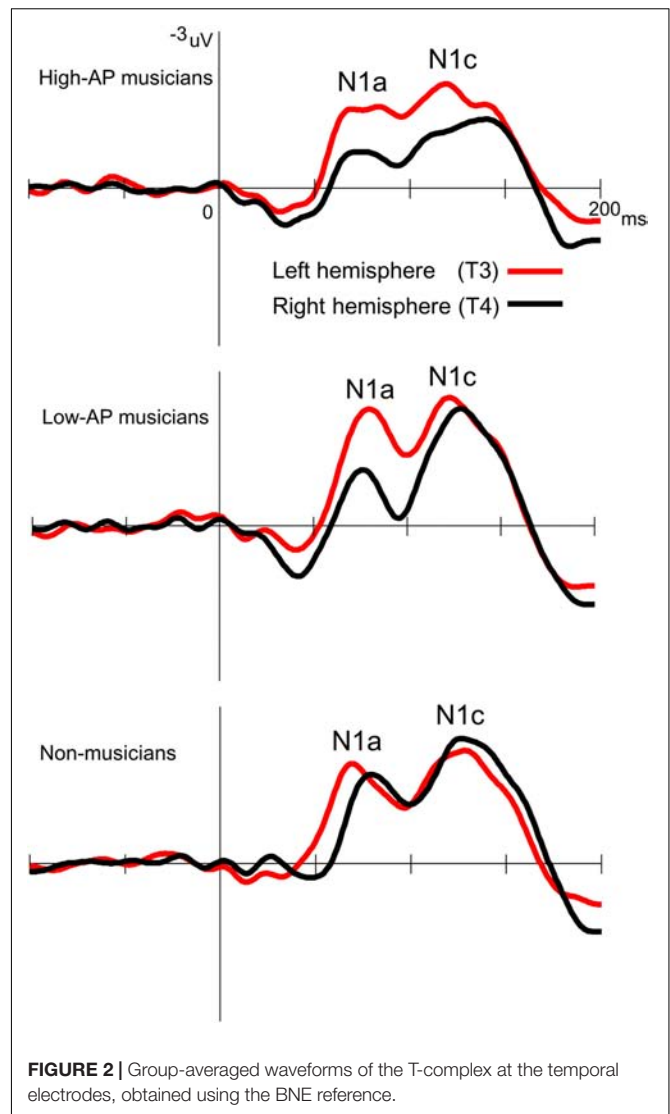


FIGURE 2 | Group-averaged waveforms of the T-complex at the temporal electrodes, obtained using the BNE reference.

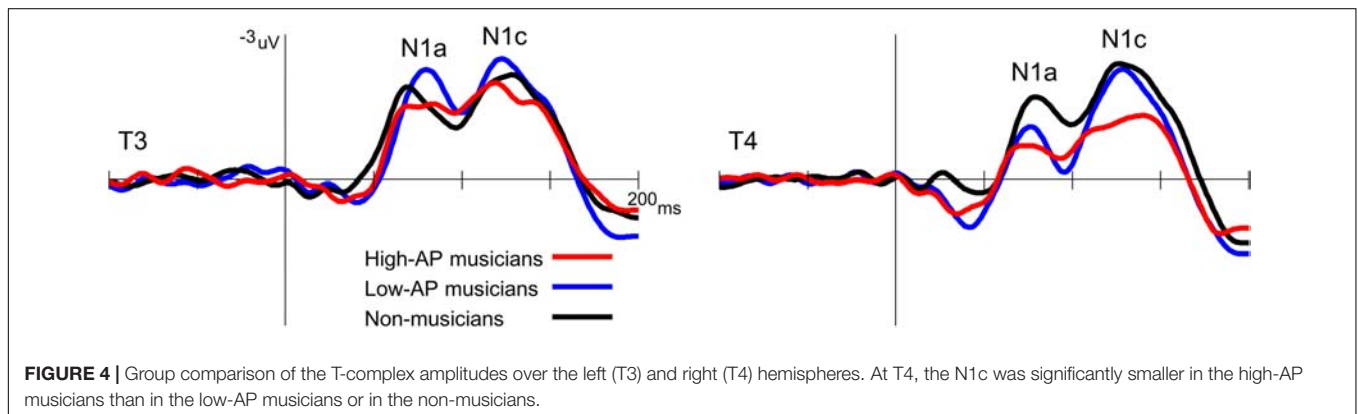
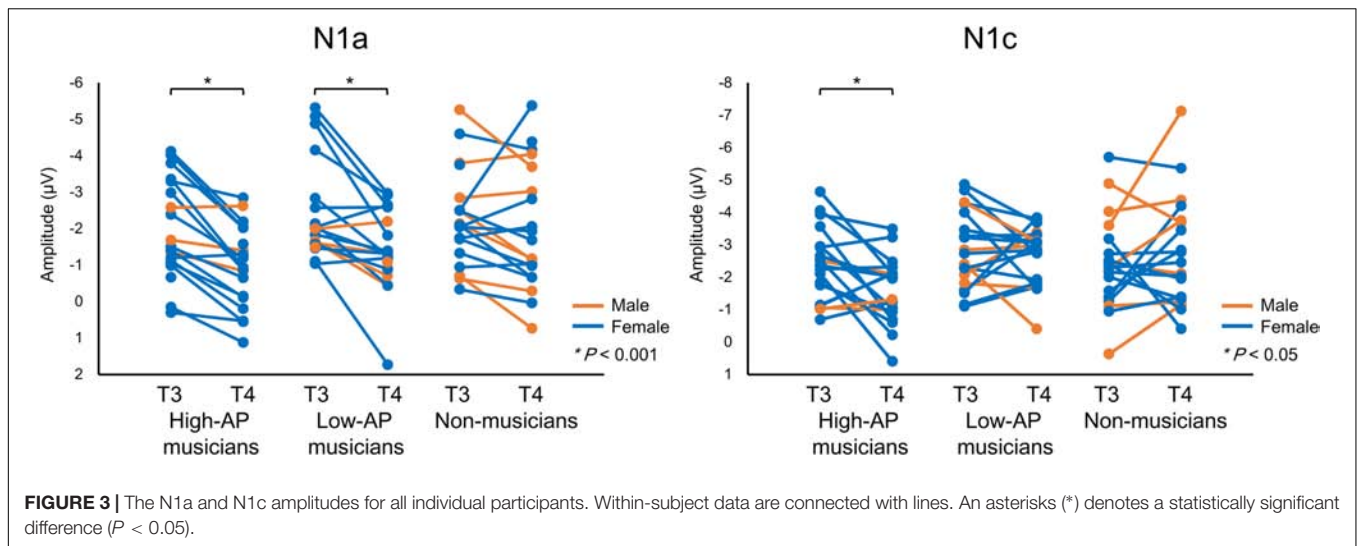
of AP was not significant ($\beta = 0.308$, $t(35) = 1.8$, $P = 0.078$), and the contribution of Years was also not significant ($\beta = -0.103$, $t(35) = 0.6$, $P = 0.549$). These results were consistent with the above ANOVA findings regarding the N1c.

No statistically significant effect was observed for the N1b amplitude. The one-way ANOVA with the group as a factor revealed that the main effect was not significant [Cz, $F(2,54) = 1.7$, $P = 0.195$].

One limitation of the experiment was that there were fewer male participants than female participants. Nevertheless, there were no apparent gender effects on the above findings, as could be appreciated in **Figure 3**.

DISCUSSION

The present study focused on the left–right asymmetry of the T-complex to investigate hemispheric asymmetry of the auditory cortical functions underpinning AP. Compared to musicians who



had low levels of AP, musicians with high levels of AP showed a greater left-dominant asymmetry of N1c amplitude. This represented AP negativity, which has previously been described as an electrophysiological marker of AP (Itoh et al., 2005). Additionally, the use of the BNE reference in this experiment revealed that the left-dominance of N1c was caused by a reduction of N1c amplitude in the right hemisphere rather than enhancement in the left hemisphere. This AP-specific effect on N1c was distinguishable from the more general effect of music training, which manifested as a left-dominant asymmetry of N1a.

The N1c response to pure-tone stimuli is typically recorded at temporal electrode sites with no left-right asymmetry or with right-dominance in a normal population (Woods, 1995; Pang and Taylor, 2000). This well-known property of N1c was confirmed in non-musicians and musicians with low levels of AP. By contrast, only musicians with high levels of AP showed a left-dominant N1c, which closely resembled AP negativity (Itoh et al., 2005) in terms of polarity (negative), amplitude (approximately 2 μ V), scalp distribution (left temporal), and latency (100–200 ms). The novel finding of this experiment was that the left-dominant asymmetry of N1c (or AP negativity) in the high-AP musician group was caused by a diminution of the right N1c, rather than enhancement of the left N1c.

A distinct methodological advantage of the current experiment is that a non-cephalic reference was used. The present recording (Figure 1) showed that the earlobe electrodes were clearly activated by the T-complex as predicted, as the sources were radially oriented in the temporal lobes (Wolpaw and Wood, 1982). Therefore, it was likely that the use of the linked-earlobes reference in Itoh et al. (2005) altered the N1c amplitudes measured on the temporal scalp electrodes. We therefore recommend the use of non-cephalic references for future recordings of AP negativity.

Notably, the right N1c was diminished in musicians with high levels of AP. The sources of N1c have been estimated in the secondary or higher auditory cortex of the superior temporal gyrus (Scherg et al., 1989; Woods, 1995; Shahin et al., 2003). Thus, although scalp AEP amplitudes are affected by many factors including source location and orientation, one possible interpretation of our finding is that the pitch stimulus activated a smaller number of right auditory cortical neurons in the high-AP musician group than in the low-AP musician group. This hypothesis is in line with previous findings that the volume of the right planum temporale is smaller in AP possessors than in non-possessors (Schlaug et al., 1995; Keenan et al., 2001; Wilson et al., 2009), although Wengenroth et al. (2014) have

reported an increased right auditory cortical volume in AP possessors, and Zatorre et al. (1998) have found an increased left planum polare volume in AP possessors. Considering that AP is typically acquired by early musical training, the functional and anatomical features of the right temporal lobe in AP possessors might be established in the course of brain maturation during which AP is acquired. In typical brain maturation, the N1c evoked by speech sounds gradually decreases in amplitude over the right hemisphere but not over the left hemisphere, commencing around the age of 7 years (Pang and Taylor, 2000). Moreover, the decrease in right N1c amplitude might be correlated with normal language development because the right N1c amplitude for speech sounds is apparently larger in children with language difficulties than in normal children (Groen et al., 2008). Because the core function of AP, or pitch labeling, is essentially verbal (Itoh et al., 2003), a common AEP component (i.e., right N1c) can reasonably be assumed to index both AP and language.

Two major hypotheses have been proposed regarding the neural mechanisms for AP: the early categorical perception hypothesis (Siegel, 1974) and the late labeling hypothesis (Levitin and Rogers, 2005). The early categorical perception hypothesis posits that the pitch labeling function of AP is subserved by some specialized neural circuitries in the auditory areas (Hirata et al., 1999; Itoh et al., 2005; Schulze et al., 2013), which is supported by anatomical findings that the auditory cortical structures in musicians with AP are organized differently from musicians without AP (Schlaug et al., 1995; Keenan et al., 2001; Wilson et al., 2009; Wengenroth et al., 2014; Kim and Knösche, 2016, 2017). By contrast, the late labeling hypothesis (also called the two-component model of AP) proposes that the labeling function of AP is subserved by late stages of cortical processing that occur outside the auditory areas (Zatorre et al., 1998; Levitin and Rogers, 2005). Specifically, the dorsolateral prefrontal cortex has been identified as the core brain structure that associates pitches with their names (Zatorre et al., 1998). This hypothesis is supported by studies that have found enhanced anatomical and functional connectivity between the auditory areas and the frontal lobe (Oechslin et al., 2010a; Elmer et al., 2015).

Previous AEP and event-related potential (ERP) experiments on this topic have yielded mixed results. An observation of the effect of AP on the late components of AEP/ERP (e.g., >200 ms), together with an absence of such effect on earlier components (e.g., <200 ms), is sometimes taken as evidence for the late labeling hypothesis (Tervaniemi et al., 1993; Rogenmoser et al., 2015). However, several AEP/ERP experiments have demonstrated AP-related differences in the early stages of auditory cortical processing (Hirata et al., 1999; Itoh et al., 2005; Wu et al., 2008; Coll et al., 2019), and our present results corroborate these findings.

In addition to the main findings regarding the neural correlates of AP, a left-dominant asymmetry of N1a was fortuitously identified as a novel electrophysiological marker of music training. Music training is known to enhance the tangentially oriented component of N1 (or its magnetic

counterpart N1m) elicited by tone stimuli, which typically has a left-dominant distribution at a peak latency of approximately 100 ms (Pantev et al., 1998; Kuriki et al., 2006; Baumann et al., 2008; Itoh et al., 2012). Our result adds to these findings by revealing that the well-known enhancement of N1 (N1m) in musicians is preceded by left-dominant neural activities at an earlier stage of auditory cortical processing at approximately 80 ms. This is in line with the findings that plastic changes due to music training occur throughout the auditory pathway, including the subcortical (Musacchia et al., 2007) and cochlear (Bidelman et al., 2016) levels of processing.

CONCLUSION

In conclusion, two main findings were obtained: (1) a left-dominant N1c (at approximately 130 ms) indexes AP, and (2) a left-dominant N1a (at approximately 80 ms) indexes music training. We conclude that the faculties for music and AP are both accompanied by a left-dominant hemispheric specialization of auditory cortical functions, but that they affect distinct stages of pitch processing in the human auditory cortex.

DATA AVAILABILITY

The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

ETHICS STATEMENT

This human study was carried out in accordance with the recommendations of the Ethical Guidelines for Medical and Health Research Involving Human Subjects (Ministry of Education, Culture, Sports, Science, and Technology; Ministry of Health, Labor, and Welfare) with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Internal Review Board of the University of Niigata.

AUTHOR CONTRIBUTIONS

MM and KI conceived and designed the study. MM and KI conducted the experiments. MM analyzed the data. All authors wrote the manuscript.

FUNDING

This work was supported by JSPS KAKENHI (19H01770, 17K00202, and 19H05309).

REFERENCES

- Baharloo, S., Service, S. K., Risch, N., Gitschier, J., and Freimer, N. B. (2000). Familial aggregation of absolute pitch. *Am. J. Hum. Genet.* 67, 755–758. doi: 10.1086/303057
- Baumann, S., Meyer, M., and Jäncke, L. (2008). Enhancement of auditory-evoked potentials in musicians reflects an influence of expertise but not selective attention. *J. Cogn. Neurosci.* 20, 2238–2249. doi: 10.5167/uzh-4931
- Bermudez, P., and Zatorre, R. J. (2009). The absolute pitch mind continues to reveal itself. *J. Biol.* 8:75. doi: 10.1186/jbiol171
- Bidelman, G. M., Nelms, C., and Bhagat, S. P. (2016). Musical experience sharpens human cochlear tuning. *Hear. Res.* 335, 40–46. doi: 10.1016/j.heares.2016.02.012
- Brancucci, A., di Nuzzo, M., and Tommasi, L. (2009). Opposite hemispheric asymmetries for pitch identification in absolute pitch and non-absolute pitch musicians. *Neuropsychologia* 47, 2937–2941. doi: 10.1016/j.neuropsychologia.2009.06.021
- Burkhard, A., Elmer, S., and Jäncke, L. (2019). Early tone categorization in absolute pitch musicians is subserved by the right-sided perisylvian brain. *Sci. Rep.* 9:1419. doi: 10.1038/s41598-018-38273-38270
- Coll, S. Y., Vuichoud, N., Grandjean, D., and James, C. E. (2019). Electrical neuroimaging of music processing in pianists with and without true absolute pitch. *Front. Neurosci.* 13:142. doi: 10.3389/fnins.2019.00142
- Deutsch, D., Henthorn, T., Marvin, E., and Xu, H. (2006). Absolute pitch among American and Chinese conservatory students: prevalence differences, and evidence for a speech-related critical period. *J. Acoust. Soc. Am.* 119, 719–722. doi: 10.1121/1.2151799
- Elmer, S., Rogenmoser, L., Kühnis, J., and Jäncke, L. (2015). Bridging the gap between perceptual and cognitive perspectives on absolute pitch. *J. Neurosci.* 35, 366–371. doi: 10.1523/JNEUROSCI.3009-14.2015
- Greber, M., Rogenmoser, L., Elmer, S., and Jäncke, L. (2018). Electrophysiological correlates of absolute pitch in a passive auditory oddball paradigm: a direct replication attempt. *eNeuro* 5:ENEURO.0333-18.2018. doi: 10.1523/ENEURO.0333-18.2018
- Groen, M. A., Alku, P., and Bishop, D. V. (2008). Lateralisation of auditory processing in Down syndrome: a study of T-complex peaks Ta and Tb. *Biol. Psychol.* 79, 148–157. doi: 10.1016/j.biopsycho.2008.04.003
- Hirata, Y., Kuriki, S., and Pantev, C. (1999). Musicians with absolute pitch show distinct neural activities in the auditory cortex. *Neuroreport* 10, 999–1002. doi: 10.1097/00001756-199904060-199904019
- Hirose, H., Kubota, M., Kimura, I., Yumoto, M., and Sakakihara, Y. (2005). Increased right auditory cortex activity in absolute pitch possessors. *Neuroreport* 16, 1775–1779. doi: 10.1097/01.wnr.0000183906.00526.51
- Itoh, K., Miyazaki, K., and Nakada, T. (2003). Ear advantage and consonance of dichotic pitch intervals in absolute-pitch possessors. *Brain Cogn.* 53, 464–471. doi: 10.1016/S0278-2626(03)00236-237
- Itoh, K., and Nakada, T. (2018). Absolute pitch is not necessary for pitch class-color synesthesia. *Conscious. Cogn.* 65, 169–181. doi: 10.1016/j.concog.2018.08.010
- Itoh, K., Okumiya-Kanke, Y., Nakayama, Y., Kwee, I. L., and Nakada, T. (2012). Effects of musical training on the early auditory cortical representation of pitch transitions as indexed by change-N1. *Eur. J. Neurosci.* 36, 3580–3592. doi: 10.1111/j.1460-9568.2012.08278.x
- Itoh, K., Suwazono, S., Arao, H., Miyazaki, K., and Nakada, T. (2005). Electrophysiological correlates of absolute pitch and relative pitch. *Cereb. Cortex* 15, 760–769. doi: 10.1093/cercor/bbh177
- Jasper, H. H. (1958). The ten-twenty electrode system of the international federation. *Electroencephalogr. Clin. Neurophysiol.* 10, 371–375.
- Keenan, J. P., Thangaraj, V., Halpern, A. R., and Schlaug, G. (2001). Absolute pitch and planum temporale. *Neuroimage* 14, 1402–1408. doi: 10.1006/nimg.2001.0925
- Kim, S. G., and Knösche, T. R. (2016). Intracortical myelination in musicians with absolute pitch: quantitative morphometry using 7-T MRI. *Hum. Brain Mapp.* 37, 3486–3501. doi: 10.1002/hbm.23254
- Kim, S. G., and Knösche, T. R. (2017). Resting state functional connectivity of the ventral auditory pathway in musicians with absolute pitch. *Hum. Brain Mapp.* 38, 3899–3916. doi: 10.1002/hbm.23637
- Knight, R. T., Scabini, D., Woods, D. L., and Clayworth, C. (1988). The effects of lesions of superior temporal gyrus and inferior parietal lobe on temporal and vertex components of the human AEP. *Electroencephalogr. Clin. Neurophysiol.* 70, 499–509. doi: 10.1016/0013-4694(88)90148-90144
- Krashen, S. (1973). Lateralization, language learning and the critical period: some new evidence. *Lang. Learn.* 23, 63–74. doi: 10.1111/j.1467-1770.1973.tb00097.x
- Kuhl, P. K. (2000). A new view of language acquisition. *Proc. Natl. Acad. Sci. U.S.A.* 97, 11850–11857. doi: 10.1073/pnas.97.22.11850
- Kuriki, S., Kanda, S., and Hirata, Y. (2006). Effects of musical experience on different components of MEG responses elicited by sequential piano-tones and chords. *J. Neurosci.* 26, 4046–4053. doi: 10.1523/JNEUROSCI.3907-05.2006
- Levitin, D. J., and Rogers, S. E. (2005). Absolute pitch: perception, coding, and controversies. *Trends Cogn. Sci.* 9, 26–33. doi: 10.1016/j.tics.2004.11.007
- Loui, P., Zamm, A., and Schlaug, G. (2012). Enhanced functional networks in absolute pitch. *Neuroimage* 63, 632–640. doi: 10.1016/j.neuroimage.2012.07.030
- Luders, E., Gaser, C., Jancke, L., and Schlaug, G. (2004). A voxel-based approach to gray matter asymmetries. *Neuroimage* 22, 656–664. doi: 10.1016/j.neuroimage.2004.01.032
- Merrett, D. L., Peretz, I., and Wilson, S. J. (2013). Moderating variables of music training-induced neuroplasticity: a review and discussion. *Front. Psychol.* 4:606. doi: 10.3389/fpsyg.2013.00606
- Miyazaki, K. (1988). Musical pitch identification by absolute pitch possessors. *Percept. Psychophys.* 44, 501–512. doi: 10.3758/BF03207484
- Miyazaki, K. (1990). The speed of musical pitch identification by absolute pitch possessors. *Music Percept.* 8, 177–188. doi: 10.2307/40285495
- Miyazaki, K., Makomaska, S., and Rakowski, A. (2012). Prevalence of absolute pitch: a comparison between Japanese and Polish music students. *J. Acoust. Soc. Am.* 132, 3484–3493. doi: 10.1121/1.4756956
- Miyazaki, K., and Ogawa, Y. (2006). Learning absolute pitch by children: a cross-sectional study. *Music Percept.* 24, 63–78. doi: 10.1525/mp.2006.24.1.63
- Musacchia, M., Sams, E., Skoe, E., and Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proc. Natl. Acad. Sci. U.S.A.* 104, 15894–15898. doi: 10.1073/pnas.0701498104
- Näätänen, R., and 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, 375–425. doi: 10.1111/j.1469-8986.1987.tb00311.x
- Newport, E. L. (1990). Maturation constraints on language learning. *Cogn. Sci.* 14, 11–28. doi: 10.1207/s15516709cog1401_2
- Oechslin, M. S., Imfeld, A., Loenneker, T., Meyer, M., and Jäncke, L. (2010a). The plasticity of the superior longitudinal fasciculus as a function of musical expertise: a diffusion tensor imaging study. *Front. Hum. Neurosci.* 3:76. doi: 10.3389/fpsyg.2010.09.076.2009
- Oechslin, M. S., Meyer, M., and Jäncke, L. (2010b). Absolute pitch—functional evidence of speech-relevant auditory acuity. *Cereb. Cortex* 20, 447–455. doi: 10.1093/cercor/bhp113
- Ohnishi, T., Matsuda, H., Asada, T., Aruga, M., Hirakata, M., Nishikawa, M., et al. (2001). Functional anatomy of musical perception in musicians. *Cereb. Cortex* 11, 754–760. doi: 10.1093/cercor/11.8.754
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-90064
- Pang, E. W., and Taylor, M. J. (2000). Tracking the development of the N1 from age 3 to adulthood: an examination of speech and non-speech stimuli. *Clin. Neurophysiol.* 111, 388–397. doi: 10.1016/S1388-2457(99)00259-X
- Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., and Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature* 392, 811–814. doi: 10.1038/33918
- Ponton, C., Eggermont, J. J., Khosla, D., Kwong, B., and Don, M. (2002). Maturation of human central auditory system activity: separating auditory evoked potentials by dipole source modeling. *Clin. Neurophysiol.* 113, 407–420. doi: 10.1016/S1388-2457(01)00733-737
- Rinker, T., Shafer, V. L., Kiefer, M., Vidal, N., and Yu, Y. H. (2017). T-complex measures in bilingual Spanish-English and Turkish-German children and monolingual peers. *PLoS One* 12:e0171992. doi: 10.1371/journal.pone.0171992
- Rogenmoser, L., Elmer, S., and Jäncke, L. (2015). Absolute pitch: evidence for early cognitive facilitation during passive listening as revealed by reduced P3a amplitudes. *J. Cogn. Neurosci.* 27, 623–637. doi: 10.1162/jocn_a_00708

- Russo, F. A., Windell, D. L., and Cuddy, L. L. (2003). Learning the “special note”: evidence for a critical period for absolute pitch acquisition. *Music Percept.* 21, 119–127. doi: 10.1525/mp.2003.21.1.119
- Scherg, M., Vajsar, J., and Picton, T. W. (1989). A source analysis of the late human auditory evoked potentials. *J. Cogn. Neurosci.* 1, 336–355. doi: 10.1162/jocn.1989.1.4.336
- Schlaug, G., Jäncke, L., Huang, Y., and Steinmetz, H. (1995). In vivo evidence of structural brain asymmetry in musicians. *Science* 267, 699–701. doi: 10.1126/science.7839149
- Schulze, K., Gaab, N., and Schlaug, G. (2009). Perceiving pitch absolutely: comparing absolute and relative pitch possessors in a pitch memory task. *BMC Neurosci.* 10:106. doi: 10.1186/1471-2202-10-106
- Schulze, K., Mueller, K., and Koelsch, S. (2013). Auditory stroop and absolute pitch: an fMRI study. *Hum. Brain Mapp.* 34, 1579–1590. doi: 10.1002/hbm.22010
- Shahin, A., Bosnyak, D. J., Trainor, L. J., and Roberts, L. E. (2003). Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *J. Neurosci.* 23, 5545–5552. doi: 10.1523/JNEUROSCI.23-13-05545.2003
- Siegel, J. A. (1974). Sensory and verbal coding strategies in subjects with absolute pitch. *J. Exp. Psychol.* 103, 37–44. doi: 10.1037/h0036844
- Stephenson, W. A., and Gibbs, F. A. (1951). A balanced non-cephalic reference electrode. *Electroencephalogr. Clin. Neurophysiol.* 3, 237–240. doi: 10.1016/0013-4694(51)90017-X
- Takeuchi, A. H., and Hulse, S. H. (1993). Absolute pitch. *Psychol. Bull.* 113, 345–361. doi: 10.1037/0033-2909.113.2.345
- Tervaniemi, M., Alho, K., Paavilainen, P., Sams, M., and Näätänen, R. (1993). Absolute pitch and event-related brain potentials. *Music Percept.* 10, 305–316. doi: 10.2307/40285572
- Tonnquist-Uhlen, I., Ponton, C. W., Eggermont, J. J., Kwong, B., and Don, M. (2003). Maturation of human central auditory system activity: the T-complex. *Clin. Neurophysiol.* 114, 685–701. doi: 10.1016/S1388-2457(03)00005-1
- Vitouch, O. (2003). Absolutist models of absolute pitch are absolutely misleading. *Music Percept.* 21, 111–117. doi: 10.1525/mp.2003.21.1.111
- Wengenroth, M., Blatow, M., Heinecke, A., Reinhardt, J., Stippich, C., Hofmann, E., et al. (2014). Increased volume and function of right auditory cortex as a marker for absolute pitch. *Cereb. Cortex* 24, 1127–1137. doi: 10.1093/cercor/bhs391
- Wilson, S. J., Lusher, D., Wan, C. Y., Dudgeon, P., and Reutens, D. C. (2009). The neurocognitive components of pitch processing: insights from absolute pitch. *Cereb. Cortex* 19, 724–732. doi: 10.1093/cercor/bhn121
- Wolpaw, J. R., and Wood, C. C. (1982). Scalp distribution of human auditory evoked potentials. I. Evaluation of reference electrode sites. *Electroencephalogr. Clin. Neurophysiol.* 54, 15–24. doi: 10.1016/0013-4694(82)90227-9
- Woods, D. L. (1995). The component structure of the N1 wave of the human auditory evoked potential. *Electroencephalogr. Clin. Neurophysiol. Suppl.* 44, 102–109.
- Wu, C., Kirk, I. J., Hamm, J. P., and Lim, V. K. (2008). The neural networks involved in pitch labeling of absolute pitch musicians. *Neuroreport* 19, 851–854. doi: 10.1097/WNR.0b013e3282ff63b1
- Zatorre, R. J., Perry, D. W., Beckett, C. A., Westbury, C. F., and Evans, A. C. (1998). Functional anatomy of musical processing in listeners with absolute pitch and relative pitch. *Proc. Natl. Acad. Sci. U.S.A.* 95, 3172–3177. doi: 10.1073/pnas.95.6.3172

Conflict of Interest Statement: 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.

Copyright © 2019 Matsuda, Igarashi and Itoh. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Comparing the Effects of Rhythm-Based Music Training and Pitch-Based Music Training on Executive Functions in Preschoolers

Ulrike Frischen*, Gudrun Schwarzer and Franziska Degé

Department of Developmental Psychology, Faculty of Psychology and Sports Science, Justus-Liebig-University, Giessen, Germany

OPEN ACCESS

Edited by:

Assal Habibi,
University of Southern California,
United States

Reviewed by:

Adam Winsler,
George Mason University,
United States
Heiner Rindermann,
Technische Universität Chemnitz,
Germany

*Correspondence:

Ulrike Frischen
ulrike.frischen@psychol.uni-
giessen.de

Received: 15 May 2019

Accepted: 02 August 2019

Published: 27 August 2019

Citation:

Frischen U, Schwarzer G and Degé F
(2019) Comparing the Effects of
Rhythm-Based Music Training and
Pitch-Based Music Training on
Executive Functions in Preschoolers.
Front. Integr. Neurosci. 13:41.
doi: 10.3389/fnint.2019.00041

Previous research has indicated the beneficial effects of music training on executive functions (EFs) in children. However, researchers have not clearly determined which component of music training produces these beneficial effects or whether different components exert different effects on EFs. In the present study, we examined the impact of rhythm-based music training compared to pitch-based music training and sports training as a control on EFs in preschoolers. Children aged between 5 and 6 years ($N = 76$) were randomly assigned to one of the three training groups and received training in small groups three times a week for 20 min in German kindergartens. Before and after training, children completed tests designed to assess inhibition, set-shifting, and visuospatial working memory. Parental education, family income, personality, and IQ served as control variables. We observed a significant training group \times time interaction for the measure of inhibition. Children from the rhythm group exhibited significant improvements in inhibition from pre- to post-tests ($d_{RM} = 0.56$), whereas children from the other groups did not. Furthermore, children from the rhythm group significantly differed from the sports control group at post-test ($d_{corr} = 0.82$). Concerning the measures of set-shifting and visuospatial working memory, the descriptive data revealed similar results; however, we did not observe significant training group \times time interactions. Based on our findings, rhythm-based music training specifically enhances inhibition in preschoolers and might affect other EFs, such as set-shifting and visuospatial working memory.

Keywords: music training, rhythm, pitch, executive functions, inhibition, preschoolers

INTRODUCTION

Playing music is one of the most challenging tasks for the human brain because it involves many cognitive processes simultaneously. Most of these processes require cognitive control, which is summarized into a concept known as *executive functions* (EFs). EFs are described as a family of top-down mental processes including goal-directed behavior, planning and problem solving, such as inhibition, set-shifting, and working memory (Diamond, 2013). Since EFs are associated with academic achievement, intelligence, health, and wellbeing, EFs should be promoted, even early in childhood (Diamond, 2013). Music training might be an effective intervention to improve EFs,

because making music, such as playing a musical instrument or singing a song, involves several EFs at the same time without focusing on a particular EF. In the last decade, some studies have reported an association between music training and EFs in children (e.g., Degé et al., 2011a). Additionally, musical training has been shown to enhance EFs in children (e.g., Jaschke et al., 2018). In preschoolers, musical training improves inhibition (e.g., Bugos and DeMarie, 2017). However, researchers have not clearly determined which specific components of music training programs are responsible for these beneficial effects on EFs. Components of musical training, which might be crucial for EFs include rhythmic entrainment (Miendlarzewska and Trost, 2014) or melodic encoding. Therefore, in the present study, we aimed to investigate whether different types of music training improved EFs such as inhibition, set-shifting, and working memory in preschoolers compared to a control sports training program. In particular, we were interested in determining whether rhythm-based music training would produce different beneficial effects on EFs than pitch-based music training.

Many cognitive processes are activated in parallel when an individual plays a musical instrument. The musician must read the music and translate the notes into movements (e.g., finger movements when playing the piano). At the same time, the individual is listening to the music and for mistakes. Additionally, the automatic responses must be inhibited when the key is changing and different chromatic signs are played (inhibition). A shift between different dynamics, rhythms, and tempos within one piece of music is often required (set-shifting). Working memory capacities are involved while remembering musical excerpts. Furthermore, the processes will become more complex when an individual is playing with others, because in addition to controlling your own playing with all of these processes, you must adapt your playing to the sound and music of the other players. For example, a musician must shift his/her attention between parts (set-shifting) and inhibit the impulse to play another part while playing the assigned part (inhibition). These processes are only examples of the variety of processes the human brain must perform while an individual plays music. However, all these processes share the requirement for voluntary cognitive control, which involves EFs. According to Diamond (2013) and Miyake et al. (2000) the three core EFs include inhibition, not acting impulsively, but prevent prepotent responses; set-shifting, switching between tasks or mental sets and working memory, monitoring and updating representations. EFs are associated with health, wealth, academic achievement and quality of life (Diamond, 2013), and studies have shown that EFs are trainable in children (for a review see Diamond and Lee, 2011) and adults (e.g., Richmond et al., 2011; Voss et al., 2011). Therefore, an opportunity to foster EFs in early childhood would be important to reduce disparities in EFs with the aim of affecting later academic achievement and well-being. As Diamond and Lee (2011) stated, repeated practice and an increasing level of difficulty is crucial for the successful training of EFs. Hence, musical training, which is associated with practicing regularly and a continuously increasing level of difficulty, might represent a perfect intervention to enhance EFs. Interestingly, and as

described above, musical training affects several EFs at the same time.

Correlational research observed positive associations between music lessons and EFs, showing that music lessons are associated with inhibition (Degé et al., 2011a; Joret et al., 2016), set-shifting, selective attention, planning (Degé et al., 2011a) and fluency (Zuk et al., 2014) in children. Regarding preschoolers, Winsler et al. (2011) observed better inhibitory control in children aged 4 years or older who were enrolled in early music classes than their age-matched peers who had never taken early music classes ($d = 0.28$). Moreover, regardless of age, children who are currently taking early music classes outperform children who are not taking early music classes ($d = 0.41$). The authors postulated that early music classes promote children's motor behaviors *via* guiding their movements with the aid of music. However, since these studies used a correlational design, they do not allow for inferences of causality. If we consider longitudinal studies of the association between music training and EFs, we also find some evidence that music training has the potential to enhance EFs in children. A study by Holochwost et al. (2017) indicated that intensive school-based music training improves performance on tasks assessing set-shifting (ds from 0.18 to 0.35), short-term-memory ($d = 0.25$), and inhibition (ds from 0.40 to 0.57) in children from grades 1 to 8. Music training occurred daily and consisted of 40 min of instrumental music classes administered in a small group and 40 min of ensemble rehearsal. In this study, the music training group was compared to an untrained control group. Thus, the authors were unable to exclude the possibility that a Hawthorne or a schooling effect influenced their results. Jaschke et al. (2018) conducted another study using a school-based instrumental music training. The authors compared the effects of weekly school-based music classes to school-based visual arts classes on 6-year-olds. The music classes comprised music theory, joint singing, and playing instruments for one or 2 h per week for 2 years. In contrast to the study performed by Holochwost et al. (2017), the lessons were held with the entire class instead of small groups. Children from the music program were divided into two different groups, one group including children who did not have prior music knowledge and another group including children who were receiving private music lessons in addition to the school curriculum. The school-based music classes enhanced some, but not all, measures of EF. Both music groups exhibited improved performance on tests of inhibition and planning and outperformed the visual arts and no-arts control groups after 2 years of training. Regarding visual working memory, the visual arts group outperformed both music groups and the no arts control group. However, as the children were not randomly assigned to the groups, the findings might have been affected by pre-existing differences between groups that were not controlled. According to Roden et al. (2012), school-based music training enhances verbal (ds from 0.65 to 1.27), but not visual memory in primary school children compared to natural science training. Children from the music group received weekly 45 min of instrumental music training conducted in a small group with a maximum of five children. In addition, they participated in singing, rhythm and pitch exercises. Another study investigating the effect of music

training on working memory showed that 2 years of an extended school-based music curriculum enhanced visual ($\eta^2 = 0.12$) and auditory working memory ($\eta^2 = 0.23$) in children aged between 9 and 11 years (Degé et al., 2011b). Similar to the study by Holochwest et al. (2017), children participating in the extended music curriculum received intensive training in five to seven music classes per week, including music theory, instrumental instruction and participating in the choir and/or the orchestra. Since the measure for visual working memory also relied on verbal components, the authors were unable to clearly determine if the music training actually affected the visual working memory or whether the training also enhanced the verbal working memory that was also involved in the task. Accordingly, researchers have not clearly determined whether music training has the potential to improve visual working memory. Another reason for this uncertainty is that the studies by Roden et al. (2012) and Degé et al. (2011b) did not randomize the participants; therefore, the results might have been influenced by pre-existing differences in children. In summary, previous studies suggest that school-based music lessons enhance EFs, such as inhibition (Holochwest et al., 2017; Jaschke et al., 2018), set-shifting (Holochwest et al., 2017), planning (Jaschke et al., 2018) and working memory (Degé et al., 2011b; Roden et al., 2012).

Regarding longitudinal studies with younger children, two studies have shown that music training enhances inhibition in preschoolers. In a study by Moreno et al. (2011), children aged 4–6 years were randomly assigned either to a group receiving computerized music listening training or to a group receiving computerized visual arts training five times a week for 45 min over a period of 20 days. Before and after training, all children completed a Go-Nogo task measuring inhibition. Children from the music group exhibited significant improvements from the pre-test to the post-test and outperformed children from the visual arts group ($\eta_p^2 = 0.12$). Therefore, the authors concluded that a short-term intense computerized hearing-based music training program improves inhibition in preschoolers. Another study compared preschoolers receiving 45 min music classes twice weekly for a period of 6 weeks in kindergarten to a control group of preschoolers receiving the same amount of LEGO training (Bugos and DeMarie, 2017). In contrast to the findings reported by Moreno et al. (2011), the music training program was a practical music training program involving tasks such as vocal development, improvisation and bimanual gross motor coordination. Children were administered the Matching Familiar Figures Test (MFFT) and the Day/Night Stroop Task before and after training to test inhibition. Based on the results, the scores of all children on both tests improved from before to after training. Moreover, the music group outperformed the LEGO group on the MFFT after training ($d = 0.99$). Hence, the authors concluded that short-term music classes enhance complex inhibition tasks in preschoolers. Since inhibition appears to be the key EF during early childhood (Wiebe et al., 2008), music training likely has no impact on other EFs, such as set-shifting and working memory, in preschoolers. However, the reported studies did not investigate the effect of music training on other EFs. Therefore, researchers have not yet clearly determined whether music training improves

inhibition alone or if it has the ability to influence other EFs in preschoolers. Taken together, the current state of knowledge indicates that music training enhances several EFs in children. In preschoolers, the reported studies showed that a short-term, computerized music listening training (Moreno et al., 2011) and a comprehensive practical music training program including vocal development, improvisation and gross motor coordination (Bugos and DeMarie, 2017) enhance inhibition. However, based on these studies, we are unable to conclude that a special component of musical training improves inhibition or whether a combination of different music components improve inhibition in preschoolers.

Most of the studies reported above used instrumental music training programs (e.g., Roden et al., 2012; Holochwest et al., 2017) or a comprehensive music training program (e.g., Bugos and DeMarie, 2017) including various components of music, such as pitch, rhythm, improvisation and bimanual coordination. The instrumental training programs consisted of general instrumental music instruction conducted in a small group (e.g., Roden et al., 2012) or in a class (Jaschke et al., 2018). In some studies, children also received lessons in ensemble playing (Holochwest et al., 2017) or music theory (Jaschke et al., 2018). The music training taught in kindergarten consisted of vocal development exercises, gross motor coordination using different instruments such as drums and xylophones, and improvisation tasks (Bugos and DeMarie, 2017). Moreno et al. (2011) did not administer practical music training in their study, but instead provided computerized hearing-based music training that comprised tasks related to pitch and rhythm. The music training programs varied in the length of lessons, ranging from 45 min per week (Roden et al., 2012) to 80 min per day (Holochwest et al., 2017), and in the duration of the training, ranging from a few weeks (e.g., Moreno et al., 2011; Bugos and DeMarie, 2017) to 2 years (e.g., Jaschke et al., 2018). In summary, substantial variety exists in the music training programs administered in current research. However, all music training programs include a mixture of components related to pitch and rhythm. Therefore, none of these studies provides information about possible differential effects of the single components on EFs. Only one study by Patscheke et al. (2018) examined the effects of a pitch-based music training compared to a rhythm-based music training on phonological awareness, but not EFs, in preschoolers. The pitch-based music training, but not the rhythm-based music training, influenced phonological awareness skills. This study proposed the hypothesis that different components of music training exert different effects on cognitive abilities in general. Therefore, single components of a music training program, such as rhythm and pitch training, might also exert different effects on EFs. Furthermore, Miendlarzewska and Trost (2014) suggested that rhythmic entrainment in particular might be one important component of a musical training program that would affect cognitive processes, such as EFs. The term *entrainment* is defined as the synchronization of at least two independent rhythmic processes (Clayton et al., 2004). In music, a simple example is clapping or tapping to the beat of a song. The ability to synchronize clapping or

tapping to a specific rhythm requires the auditory perception of the rhythm, coordination of movements and sensorimotor integration. According to a study by Krause et al. (2010), musicians appear to exhibit increased connectivity in a brain network involving the premotor cortex, posterior parietal cortex and thalamus. These areas are also associated with attentional processes and motor planning (Coull, 2004), prompting the hypothesis that rhythmic entrainment plays a key role in the effect of music training on the development of EFs (Miendlarzewska and Trost, 2014). Thus, musical rhythm training in a group setting potentially represents an appropriate intervention to promote rhythmic entrainment abilities, because the players must automatically synchronize to each other in a natural setting.

We conducted the present study to investigate the differential effects of pitch-based music training compared to a rhythm-based music training on EFs in German preschoolers. Previous studies have already reported an effect of music training on performance on different tests of inhibition (Moreno et al., 2011; Bugos and DeMarie, 2017). However, researchers have not determined whether a special aspect of music training leads to the reported effects or if different musical aspects train various EFs. Miendlarzewska and Trost (2014) proposed that rhythmic entrainment is one important factor leading to cognitive enhancement. Moreover, the study by Patscheke et al. (2018) indicated different effects of various components of a music training program, such as pitch training and rhythm training, on phonological awareness. Therefore, we decided to use the same training program as Patscheke et al. (2018) to explore possible differences in the effects of a pitch-based music training compared to a rhythm-based music training on EFs in preschoolers. In addition to the music training groups, we added a third group receiving sports training as a no-music training control group. Based on the findings reported by Miendlarzewska and Trost (2014), we predicted a greater improvement in inhibition in children from the rhythm group, because the rhythm training focused on rhythmic perception and production, which are strongly associated with rhythmic entrainment. Inhibition appears to be the most important EF during early childhood (Wiebe et al., 2008). Since previous studies showed an improvement in inhibition in preschoolers receiving music training, we were specifically interested in investigating possible differences in the effects of pitch-based music training compared to rhythm-based music training on inhibition. Nevertheless, we decided to test the effects of the training programs on the other core EFs, including set-shifting and working memory, because studies with older children suggest that music training affects several EFs (e.g., Holochwost et al., 2017). The reported studies did not examine whether music training affected other EFs or inhibition alone in preschoolers.

MATERIALS AND METHODS

Participants

At the pre-test, the sample consisted of 95 preschoolers (57 females) aged 5–6 years ($M = 5.7$ years, $SD = 0.3$ years).

Participants were recruited from different kindergartens¹ in the City of Giessen, Germany and the surrounding area. Participants were randomly assigned either to a music group receiving pitch training ($n = 33$), a music group receiving rhythm training ($n = 33$), or a control group receiving sports training ($n = 30$). We determined the inclusion criteria for our analyses to a training participation rate of at least 66%. The remaining sample included in the analyses comprised 76 children (see below for details). None of these children received music lessons or participated in a music group such as a choir or an ensemble. Some children (16%) had taken early music education for an average of $M = 14.4$ months ($SD = 7.6$ months). The socioeconomic status was assessed based on the parents' education levels and the monthly family income. The parents of most children (65%) did not have a university degree, one parent of 16% of the children had a university degree, and both parents of 11% of the children attained a university degree. Some parents did not provide information about their educational level (11.6%). The monthly family income in the sample ranged from less than 1,000€ (8%) to more than 5,000€ (3%). Most families reported a monthly income ranging from 1,000€ and 2,000€ (28%) or 2,000€ and 3,000€ (20%). Some parents (18%) did not provide details about their monthly income. We did not detect group differences in demographic variables. For additional details, please see Table 1.

Materials

Training Programs

Preschoolers were trained for 20 min three times a week for a period of 20 weeks. Trained research assistants conducted all sessions. Training programs were based on a manual. Every week, the research assistants and the supervisor of the study met to prepare and practice the training session for the following week to ensure that every training session was implemented in the same way. A typical training session on 1 day consisted of two to four different tasks lasting for a total of 20 min. The training sessions were implemented as described in the study by Patscheke et al. (2018) and based on a well-established early music education program designed by Nykrin et al. (2007).

Rhythm training focused on rhythmic exercises, including meter execution, perception, imitation and production of different rhythms. The exercises were conducted using sound gestures (e.g., clapping and stomping) and different Orff instruments (e.g., taborets, claves, and maracas). Exercises in meter execution included the synchronization of body movements to a given meter, dancing and playing rhythms with percussion instruments. Typical perception and imitation tasks consisted of imitating rhythms using rhythm language (ta-a-a-a, ta-a, ta, and titi) or percussion instruments.

Pitch training focused on discriminating sounds, intonation, sound production, and joint singing. Typical exercises for

¹In Germany, children aged between 3 and 6 years attend kindergarten. Kindergarten is not compulsory and kindergartens do not have a consistent curriculum. In the last year of kindergarten, when children are aged between 5 and 6 years, they receive some preschool classes. However, children are playing, drawing or doing handicrafts for most of the time in kindergarten.

TABLE 1 | Inferential statistics for group comparisons of control variables.

Variable	Pitch <i>M</i> (<i>SD</i>)	Rhythm <i>M</i> (<i>SD</i>)	Sports <i>M</i> (<i>SD</i>)	<i>df</i>	Statistical value	<i>p</i>
Gender	17 F/10 M	18 F/8 M	11 F/12 M	2, <i>n</i> = 76	$\chi^2 = 2.44$	0.295
Age ¹	67.44 (3.49)	68.77 (3.43)	69.43 (3.75)	2, 73	$F = 2.06$	0.134
IQ	99.93 (11.56)	99.08 (8.39)	98.13 (12.62)	2, 73	$F = 0.17$	0.846
Early music education ¹	3.36 (7.40)	1.50 (5.38)	1.14 (3.61)	2, 67	$F = 1.01$	0.368
Participation in training ²	84.19 (9.42)	82.23 (9.74)	80.00 (8.37)	2, 73	$F = 1.28$	0.285
Enjoyment of training	4.43 (0.55)	4.34 (0.95)	4.56 (0.68)	2, 59	$F = 0.49$	0.617
Parental education	0.6 (0.82)	0.39 (0.72)	0.19 (0.40)	2, 66	$F = 2.06$	0.135
Family income	3.48 (1.41)	2.60 (1.27)	2.68 (0.86)	2, 59	$F = 3.40$	0.04 ^a
Extraversion scale BFI	3.93 (0.67)	4.02 (0.69)	3.90 (0.76)	2, 69	$F = 0.32$	0.725
Agreeableness scale BFI	3.72 (0.40)	3.58 (0.52)	3.59 (0.58)	2, 69	$F = 0.58$	0.560
Conscientiousness scale BFI	3.47 (0.67)	3.41 (0.73)	3.58 (0.80)	2, 69	$F = 0.30$	0.740
Neuroticism scale BFI	2.48 (0.51)	2.53 (0.62)	2.65 (0.54)	2, 69	$F = 0.54$	0.588
Openness scale BFI	3.75 (0.61)	3.66 (0.40)	3.70 (0.46)	2, 69	$F = 0.21$	0.811

¹ In Months, ² in %, ^a post hoc analyses did not reach significance. * $p < 0.05$.

discriminating sounds were listening to different musical instruments or sounds from a CD and then naming or pointing to the instrument on a picture. Intonation was trained with the call and response method, as the teacher sang intervals or short melodies and the children subsequently repeated the intervals or melodies. Pitch discrimination was trained by listening to different tones from the mallet instrument and indicating which tones were higher or lower or by listening two different voices from a CD and indicating which was higher or lower. In pitch training, we did not use any percussion instruments or sound gestures for rhythmic accompaniment. Conversely, in rhythm training, we did not use melodic instruments or sing any melodies. Nevertheless, some overlap occurred between the training programs, because rhythm is to some extent also connected to prosodic features and song cannot be sung without using a certain rhythm. However, we reduced the overlap between trainings as much as possible.

The manual for sports training was the same as used by Patscheke et al. (2018) and comprised different exercises to practice body perception, motor skills and body coordination by supporting balance, physical strength, endurance and relaxation. Similar to the music training programs, a typical session included two to four different tasks for a total of 20 min. The children were asked to perform different exercises for body perception (e.g., exercises from yoga), balance (e.g., balancing objects on different parts of the body), motor skills (e.g., throwing balls into a box from a near or far distance) and different coordination and cooperation games in a group (e.g., walking altogether as a caterpillar). The sports training program was based on *Yoga and active games for kids* by Dunemann-Gulde (2005).

All training programs involved the same level of engagement of the children, and each trainer conducted classes in each training program to exclude trainer bias.

Measures

Control Variables

Demographic variables such as socioeconomic status (assessed based on parental education and family income) and the musical background of the children (e.g., if the child had participated

in courses for early music education) were assessed using a questionnaire. The parental education level was coded as follows: 0 = no parents holding a university degree, 1 = one parent holding a university degree, and 2 = two parents holding a university degree. Family income was assessed with a six-point scale ranging from less than 1,000€ per month to more than 5,000€ per month. Early music education was assessed in months. Furthermore, parents completed the German version of the Big Five Inventory (BFI; Rammstedt and Danner, 2017) to assess the personality of each of their participating children. The BFI is a questionnaire comprising 45 items designed to measure the Big Five factors of personality (openness, agreeableness, extraversion, conscientiousness and neuroticism). Parents were asked to rate the extent to which they would agree with specific statements (e.g., “I see my child as someone who is talkative”) on a 5-point Likert scale. As described in the study by Rammstedt and Danner (2017), we calculated the means for each of the personality factors. Reversed items were recoded. Higher scores indicate a stronger characteristic of the personality factor. The five scales of the German BFI show an internal consistency of $\alpha = 0.74$ to $\alpha = 0.86$. The test-retest reliability for the scales ranges from $r = 0.78$ to $r = 0.93$. We administered the revised version of the Culture Fair Test (CFT 1-R) created by Weiß and Osterland (2012) to assess fluid intelligence. The CFT 1-R is divided into two parts including three subtests. The first part measures figural perception and processing speed with the subtests *substitutions*, *mazes* and *similarities*. The second part includes the subtests *continuing series*, *classifications* and *matrices* to measure figural reasoning. All subtests contain 15 items that must be processed in a certain time. Research assistants explained examples of each subtest according to the manual before the subtests began. The test provides age-corrected standard values for children aged from 5.3 to 9.11 years. The test-retest-reliability is $r = 0.88$ for the first part, $r = 0.94$ for the second part and $r = 0.95$ for the whole test. We assessed children’s enjoyment of participating in the trainings to control for potential biases concerning their motivation and willingness. Every 5 weeks of the trainings, children rated on a 5-point-Likert scale how much they enjoyed participating in the training. The scale ranged from 1 = not a bit to 5 = very much and was presented to the children *via* smileys (1 = very sad smiley; 5 = very happy smiley;

3 = neutral smiley). For each child, we generated an average score of the four measurements.

Dependent Variables

The EF measures inhibition, set-shifting and working memory served as our dependent variables and were assessed before and after training. Inhibition was measured using the subtest “statue” from the NEPSY-II (Korkman et al., 2007), a developmental neuropsychological assessment for children. The test is designed to assess motor persistence and inhibition. During the test, the child was asked to stay in a particular position with his/her eyes closed for 75 s. The child was asked not to respond to any sound distractor during this time. The experimenter made different noises as distractors at certain times (e.g., after 20 s the experimenter drops a pencil onto the table). At every 5-s interval, the experimenter recorded if the child committed any error by opening his/her eyes, moving the body or responding verbally to the sound or if the child had no errors. For every 5-s interval, the child received two points if no errors were recorded, one point if two errors were recorded or zero points if more than two errors were recorded during that interval. We added the points for all time intervals as the outcome measure and compared them with age-corrected norms, which are provided for children aged from 3 to 6 years. The test-retest-reliability of this subtest is $r = 0.88$ for children aged from 5 to 6 years.

Set-shifting was assessed with the Dimensional Change Card Sort (DCCS; Zelazo, 2006). We administered both the standard and more challenging border versions. In the standard version, children were presented two target cards (a blue bunny and a red boat) and invited to sort the test cards (red bunnies and blue boats) according to one dimension. In the pre-shift phase, children were asked to sort all cards according to the dimension “color,” and were required to place a red card to the side of the red boat and a blue card to the side of the blue bunny. After 6 trials, the experimenter interrupted the game and changed the dimension to “shape,” and thus the child sorted cards with the bunny shape to the blue bunny and cards with the boat shape to the red boat (post-shift). Both the pre-shift and post-shift phases comprised 6 trials, with a total of 12 trials for the standard version. After completing the standard version of the test, we administered the more challenging extended version, which combined both dimensions from the standard version. In the extension, the child was asked to sort cards with a black frame according to the dimension “color” and cards without a frame according to the dimension “shape.” Similar to the standard version, the extended version also consisted of 12 trials. In both versions, we repeated the instructions after every trial to ensure the lowest demands on working memory as possible. In the pre-shift and post-shift phases, children were required to sort at least five of six cards correctly to pass. In the extended version, children were required to sort at least 9 of 12 cards correctly to pass. We used the scoring system suggested by Zelazo (2006): children who failed the pre-shift phase received a “0,” children who passed the pre-shift phase received a “1,” children who passed the post-shift phase but failed the extended version received a “2” and children who passed both the standard and extended versions received a “3.”

The DCCS shows an intra-class correlation (*ICC*) of 0.94 for the standard version and $ICC = 0.90$ for the extended version (Beck et al., 2011).

We measured visual-spatial working memory by administering the subtests Matrix Span Test (for visual cache) and Corsi Block Test (for inner scribe) from the Working Memory Test Battery for 5- to 12-year-old children (“Arbeitsgedächtnistestbatterie 5–12,” AGTB 5–12) described by Hasselhorn et al. (2012). According to Logie (1995), visual working memory is divided into two subcompartments: the visual cache and the inner scribe. The visual cache is responsible for storing information about shape and color, whereas the inner scribe analyses information about location and movement. Both tests consisted of a standardized computerized introduction with two practice trials and 10 test trials. The difficulty level at the beginning of the test trials depended on the child’s age. For example, 6-year-old children started at a more difficult level than 5-year-old children. Both subtests used an adaptive procedure that increased the difficulty for correct responses and decreased the difficulty for incorrect responses (described in detail below). During the test trials, children were not informed about their test performance. The AGTB provides age-corrected standard values for children aged from 5 to 12 years. The Matrix Span Test measures the memory of static visual patterns. Children were presented a 4×4 matrix composed of white and black squares on a touchscreen for 4 s. After the matrix had disappeared, children were shown a new matrix with white squares only and asked to tag the squares that were black in the previous image. If the child responded correctly in two consecutive trials, the number of the black squares increased by one (to a maximum of nine black squares). On the other hand, if the child responded incorrectly in two consecutive trials, the number of black squares was reduced by one (to a minimum of two black squares). The Corsi Block Test assessed the memory of dynamic changes in spatial locations. In this test, children were asked to memorize the path of a yellow smiley face that moved between nine white squares displayed on a gray background on the touchscreen. After the full path of the smiley face was shown, children recorded their response by touching the order of the squares in which the smiley face had moved. If the child responded correctly in two trials, the path of the smiley face was increased by one square (to a maximum of nine squares). If the child responded incorrectly in two trials, the path became one square shorter (to a minimum of two squares). The test-retest-reliabilities for both subtests are $r = 0.66$ for children aged from 5 to 8 years.

Procedure

The experiments were conducted in accordance with ethical guidelines of the ethics committee of the Faculty of Psychology and Sports Science at Giessen University (application number 2015-0001). Prior to pre-tests, parents were provided information sheets about the aims and contents of the study and informed consent was obtained for each participant. All test sessions were conducted as single sessions in a separate room in kindergarten. Trained research assistants from the

TABLE 2 | Means and standard deviations of dependent variables for treatment groups and the control group at the pre-test (T0) and post-test (T1).

Variable	Pitch <i>M (SD)</i>		Rhythm <i>M (SD)</i>		Sports <i>M (SD)</i>	
	T0	T1	T0	T1	T0	T1
Inhibition						
Statue test	11.11 (2.14)	11.81 (2.20)	10.46 (2.47)	12.17 (1.93)	11.00 (2.27)	10.86 (2.26)
Set-shifting						
DCCS	2.07 (0.55)	2.41 (0.57)	1.96 (0.66)	2.50 (0.51)	1.83 (0.39)	2.13 (0.55)
Working memory						
Matrix span	46.81 (9.12)	51.11 (9.16)	50.27 (7.50)	48.92 (10.96)	49.87 (8.19)	53.74 (11.30)
Corsi block	46.78 (8.51)	50.00 (6.58)	44.88 (7.25)	50.65 (6.14)	46.78 (9.66)	47.57 (7.62)

department, who were at all times blinded to the conditions and the hypotheses, administered test sessions. We started test sessions with the DCCS, followed by the statue test and the tests of visual working memory. The CFT 1-R was administered in an extra session on the same day after an appropriate break or on a consecutive day. After the pre-tests, children were randomly assigned to one of the music groups or to the sports group. All training programs were conducted in small groups of 5–8 children for 20 weeks. Groups participated in the designated training program three times a week for 20 min each. After the training phase was complete, post-tests of the dependent measures were immediately administered using the same method described for pre-test. When post-tests were complete, each child received a small present and a certificate as a reward for participating in the study.

RESULTS

All analyses were computed using IBM® SPSS® Statistics 25. First, we analyzed the dropout rate to ensure that the children who dropped out did not differ significantly from the remaining sample. Children with a training participation rate of less than 66% were excluded and counted as a dropout. Of the 95 children enrolled, 19 did not meet this inclusion criterion, and thus the dropout-rate in our sample was 18.1%. Major reasons for the low training participation rate of these children were illness and winter holidays. Based on our analyses, these children did not significantly differ from the included sample in terms of gender, $\chi^2_{(1)} = 0.04$, $p = 0.834$, age, $t_{(93)} = 0.31$, $p = 0.757$, parental education $t_{(82)} = -0.86$, $p = 0.392$, family income, $t_{(75)} = -1.18$, $p = 0.243$ and IQ, $t_{(93)} = -0.05$, $p = 0.960$. A power analysis using G*Power (Erdfeider et al., 1996) suggested that a total sample size of 66 participants was sufficient to ascertain small to medium effects ($f = 0.25$) in a repeated-measures within-between design (α : 0.05, power (1- β): 0.95, correlations between repeated measures: $r = 0.50$).

Preliminary Analyses

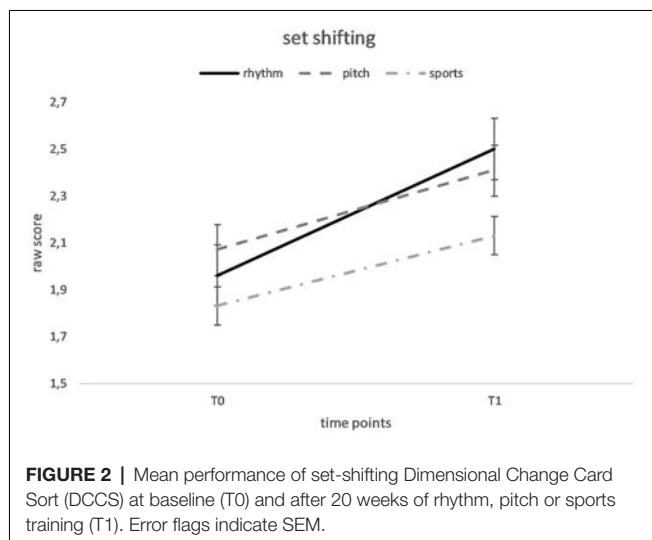
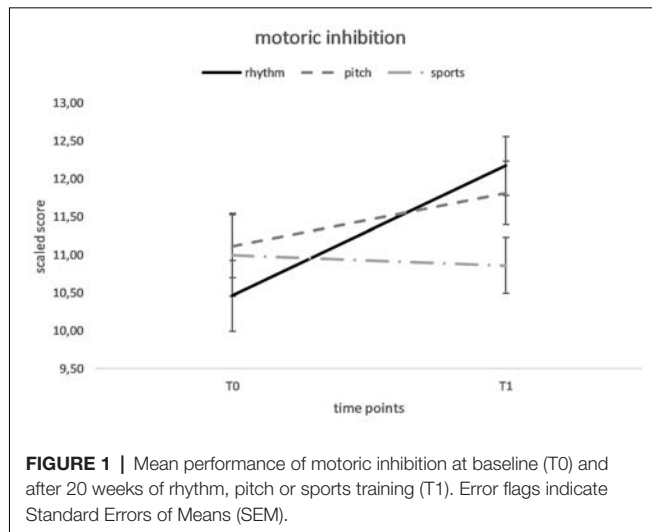
We compared differences in control variables between the training groups. As shown in Table 1, significant differences in gender, age, IQ, early music education, training participation, enjoyment of the training, parental education and personality were not observed between the training groups. A significant difference in family income was identified, indicating that

families from the pitch group had the highest income ($M = 3.48$, $SD = 1.41$), followed by the sports group ($M = 2.68$, $SD = 0.89$) and the rhythm group ($M = 2.60$, $SD = 1.27$). However, Bonferroni-adjusted *post hoc* analyses did not reveal significant ($p < 0.05$) differences between the pitch and rhythm groups [95% CI (−0.05, 1.8)], between the pitch and sports groups [95% CI (−0.14, 1.73)] or between the rhythm and sports groups [95% CI (−0.105, 0.88)]. We performed ANOVAs for the EFs measured in the pre-test to exclude the possibility of any pre-existing differences in EFs between groups. No differences in inhibition, $F_{(2,73)} = 0.45$, $p = 0.64$, set-shifting $F_{(2,73)} = 1.26$, $p = 0.29$, and working memory $F_{(2,73)} = 1.36$, $p = 0.26$ (Matrix Span), $F_{(2,73)} = 0.43$, $p = 0.65$ (Corsi Block) were observed between groups. Taken together, the preliminary analyses did not reveal significant systematic differences in control variables or dependent measures at the pre-test.

Main Analyses

We performed repeated measures ANOVAs with the training group as the between-subjects factor and time as the within-subjects factor for the tests of inhibition, set-shifting, and working memory. Mean values and standard deviations for the dependent variables are presented in Table 2.

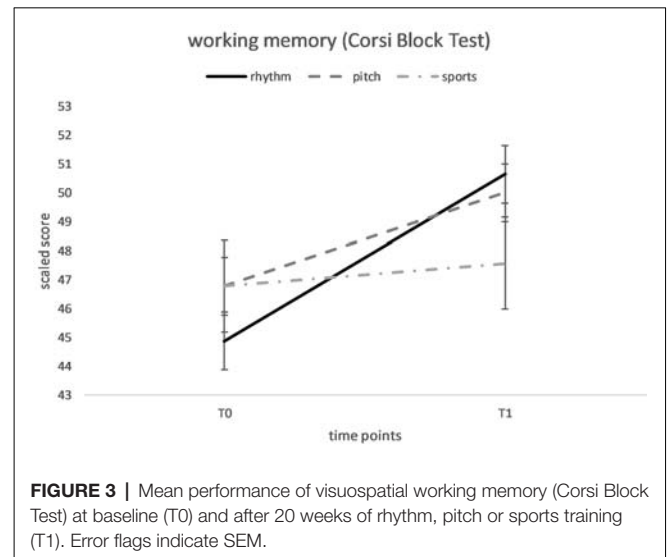
For inhibition, the repeated measures ANOVA indicated a significant main effect of time, $F_{(1,73)} = 9.03$, $p = 0.004$, $d_{RM} = 0.33$, showing an overall improvement from the pre-test ($M = 10.86$, $SD = 2.28$) to the post-test ($M = 11.64$, $SD = 2.02$), as well as a significant training group \times time interaction $F_{(2,71)} = 4.18$, $p = 0.019$, $d = 0.70$. The main effect of group was nonsignificant, $F_{(1,73)} = 0.52$, $p = 0.598$. We analyzed the training group \times time interaction by performing a 2 (time) \times 2 (training group) ANOVA and observed a significant training group \times time interaction when comparing the rhythm and sports groups, $F_{(1,42)} = 7.14$, $p = 0.011$, $d = 0.84$. We did not observe significant interactions with time when the pitch and sports groups were compared, $F_{(1,47)} = 4.28$, $p = 0.124$ or when the pitch and rhythm groups were compared, $F_{(1,49)} = 6.41$, $p = 0.12$. Paired *t*-tests showed a significant improvement from the pre-test to post-test for the rhythm group, $t_{(23)} = -3.21$, $p = 0.004$, $d_{RM} = 0.56$, whereas the results for the pitch group, $t_{(27)} = -1.94$, $p = 0.062$, and the sports group, $t_{(21)} = 0.34$, $p = 0.734$ were nonsignificant. Independent *t*-tests did not reveal differences between the rhythm and the sports group at the pre-test, $t_{(47)} = -0.27$, $p = 0.79$, but indicated a significant difference between the



rhythm and the sports group at the post-test, $t_{(42)} = 2.34, p = 0.02$, $d_{\text{corr}} = 0.82$ (see **Figure 1**).

Regarding set-shifting, a repeated measures ANOVA revealed a significant main effect of time, $F_{(1,73)} = 21.58, p < 0.001$, $d_{\text{RM}} = 0.55$, showing an overall improvement from the pre-test ($M = 1.96, SD = 0.55$) to the post-test ($M = 2.35, SD = 0.56$), as well as a significant main effect of the training group, $F_{(2,73)} = 3.23, p = 0.045, d = 0.59$. However, Bonferroni-adjusted *post hoc* group comparisons between the rhythm and sports groups [95% CI (−0.03, 0.54)], the rhythm and pitch groups [95% CI (−0.28, 0.26)] and the pitch and sports groups [95% CI (−0.02, 0.55)] were nonsignificant (all p -values > 0.05). Furthermore, the training group \times time interaction was nonsignificant, $F_{(2,73)} = 0.77, p = 0.47$ (see **Figure 2**).

Repeated measures ANOVAs for working memory did not reveal a significant main effect of time on the Matrix Span result, $F_{(1,73)} = 3.55, p = 0.064$, or a significant main effect of the training group $F_{(2,71)} = 0.88, p = 0.421$. The training group \times time interaction was also nonsignificant, $F_{(2,71)} = 2.31, p = 0.11$. For



the Corsi Block Test, we observed a significant main effect of time, $F_{(1,73)} = 10.67, p = 0.002, d_{\text{RM}} = 0.35$, showing an overall improvement from the pre-test ($M = 46.13, SD = 8.42$) to the post-test ($M = 49.49, SD = 6.80$), but a significant main effect of the training group was not detected, $F_{(2,71)} = 0.23, p = 0.795$. The training group \times time interaction was also nonsignificant, $F_{(2,71)} = 2.02, p = 0.14$ (see **Figure 3**).

DISCUSSION

We explored differences in the effects of rhythm-based music training and pitch-based music training on EFs in preschoolers. Based on our results, 20 weeks of rhythm training enhanced motor inhibition, but not pitch-based music training and sports training. The effect of the rhythm training program corroborates the hypothesis that rhythmic entrainment represents an important underlying mechanism of music training that supports cognitive function (Miendlarzewska and Trost, 2014). Moreover, since the rhythm training program required a high level of motor coordination we postulated that the motor component of the rhythm training program specifically influenced preschoolers' inhibition skills.

Our study confirms results from previous studies by showing an effect of music lessons on inhibition in school-children (Holochwest et al., 2017; Jaschke et al., 2018) and preschoolers (Bugos and DeMarie, 2017). However, most of these studies had some methodological issues, such as the lack of an active control group (Holochwest et al., 2017) or a nonrandomized sample (Jaschke et al., 2018), and thus these results might also be attributed to a general training effect or to some non-controlled influencing factors. Accordingly, only our study and the study by Bugos and DeMarie (2017) actually allow for interpretations of causality because the children were randomly assigned to the different training programs. We expand the results of the study by Bugos and DeMarie (2017) by showing that rhythm training particularly appeared to improve inhibition in preschoolers. The rhythmic synchronization while drumming

and the precise timing required while producing rhythms and rhythmical movements to music might have improved the inhibition abilities. Since all these rhythmic activities required children's motor control, it is likely that rhythm-guided motor control, in general, improves motor inhibition in children. Therefore, our findings support the idea of Winsler et al. (2011), who suggested that modulating children's movements with the aid of music may be beneficial for motor behavior and self-regulation. As shown in the present study, rhythm-guided motor control was implemented in the rhythm training program at a high level and improved motor inhibition, whereas tasks focusing on pitch with less motor involvement did not affect motor inhibition skills in preschoolers. Moreover, the sports training program, which also required a high level of motor control, did not show an effect on motoric inhibition. Thus, we can conclude that rhythm-guided motor control in particular and not motor control alone improved inhibition skills in preschoolers.

Regarding set-shifting, the results showed an overall improvement from the pre-test to the post-test, but the training programs did not exert significantly different effects, since the training group \times time interaction was nonsignificant. Hence, we were unable to generate statistically verified conclusions about a special effect of a single training program. However, the descriptive data (see **Figure 2**) appear to show greater improvements in the music groups than in the sports group, with the rhythm group attaining the greatest benefit. Therefore, our results are consistent with the findings from the study by Holochwost et al. (2017), which showed that music training enhanced set-shifting abilities in older children. However, since we did not identify a significant group by time interaction and all groups exhibited significant improvements in set-shifting from the pre-test to the post-test, further studies are required to corroborate this finding.

Regarding visuospatial working memory, we did not find any significant effect of the training programs on the Matrix Span Test, which assesses the visual cache (memory of static visual patterns), but an overall improvement was observed on the Corsi Block Test, which measures the inner scribe (memory of a dynamic series of spatial locations), from the pre-test to the post-test. Since the training group \times time interaction was nonsignificant, statistical confirmation of an effect of a single training program was not obtained. However, based on the descriptive data (see **Figure 3**), the music groups exhibited improvements in the inner scribe component, with the rhythm group showing the greatest improvement from the pre-test to the post-test, whereas the sports group recorded similar performances on the pre- and post-tests. A potential explanation is an effect of the rhythm-based music training on the inner scribe component, but due to less power and high variances, we did not observe a statistically significant training group \times time interaction. Interpreting the descriptive data, our study also confirms the findings reported by Holochwost et al. (2017) of an association between music training and visuospatial working memory in older children to some extent. The improvement on the Corsi Block Test but not the Matrix Span Test might indicate that the music training program did not improve

the memory capacity, but improved the processing of visual information. This hypothesis is consistent with the findings from studies assessing verbal memory, showing that music lessons do not enhance verbal storage itself but improve the articulatory rehearsal (Franklin et al., 2008; Degé and Schwarzer, 2017). Therefore, a similar situation might occur for the visual working memory such that the music training program did not affect the visual storage but the processing of information in the visual memory. However, as stated above, we did not find a significant group by time interaction and must interpret these results with caution.

Our results confirm the hypothesis that different components of a music training program exert different effects on cognitive abilities in general. Furthermore, the study by Patscheke et al. (2018) revealed that pitch training, but not rhythm training, affected phonological awareness. Since we administered the same program for the pitch and the rhythm training, a comparison of the results suggests that the training programs exert different effects on cognitive abilities. Consequently, the rhythmic motor aspect of a music training program might improve general cognitive abilities, such as EFs, and intonation and pitch are strongly correlated with verbal abilities such as phonological awareness.

In summary, a rhythm-based music training program enhanced inhibition in preschoolers. As mentioned above, the motor component of the rhythm training program and rhythmic synchronization might be the components of a music training program specifically affect EF. Furthermore, our results show a descriptive tendency that music training affects other EFs, such as set-shifting and visual-spatial working memory, in preschoolers as well. However, since we were unable to statistically confirm this hypothesis, further studies are required to investigate these effects. As inhibition appears to be a key factor in the development of EFs in early childhood, the effect of the rhythm-based music training program on inhibition is an important finding. Since EFs exert far-reaching effects on crucial developmental functions, such as intelligence, academic achievement, wellbeing and health (Diamond, 2013), EFs should be enhanced, even early in childhood. As the present study suggests that regular music training, particularly training focusing on rhythmic elements, improves inhibition, one recommendation is to implement elements of a rhythm-based music training program in already established EF training programs. Furthermore, the development of a music-based EF training program that can be integrated in the daily lives kindergarteners might produce promising effects. The implementation of a music-based EF training program in kindergarten might allow all children to improve their EFs in a manner independent of social contexts to ensure equal opportunities to benefit from the program. Nevertheless, further studies are needed to examine the applicability and the effectiveness of these training programs.

LIMITATIONS AND FUTURE DIRECTIONS

We explored differences in the effects of music training programs on EFs in preschoolers. The results showed an effect of rhythm-

based music training on motor inhibition. One limitation of the present study is that our music training programs had some overlapping aspects, as rhythm training cannot be completely separated from pitch training, since melodies and songs are always based on rhythm. However, the main difference between the training programs was that the rhythm training program focused on the music-guided motor component by perceiving rhythms with the whole body and producing rhythms using sound gestures, percussion instruments or rhythmical body movements. In contrast, the pitch training program concentrated on pitch and intonation without any rhythmic motor action. Regarding the sample size, the present study already exhibited an acceptable power, particularly if we consider the costs and efforts associated with the high training intensity and the organization required to work with many different kindergartens. Nevertheless, future studies with larger sample sizes should be conducted to support the generalizability of our findings. Although the power analysis revealed a sufficient sample size in the present study, a larger sample is preferred and might even reveal smaller effects that could have been overlooked in our study. As rhythm training appears to be an important factor influencing inhibition in preschoolers, an investigation of the effect of dancing on EF would also be interesting. Similar to music training, dancing relies on rhythm and meter and includes the musical component as the dancer executes movements to the rhythm of music. Furthermore, similar to our rhythm training program, dancing concentrates on music-guided motor control and rhythmic entrainment, and thus dancing might exert a similar effect on EFs.

CONCLUSION

Taken together, music training influences inhibition in preschoolers and might affect other EFs, such as set-shifting and working memory. This study is the first to investigate the different effects of distinct music training programs on several EFs in this age group. In particular, a rhythm-based music training program affects inhibition in preschoolers. In contrast to some previous studies, our results allow for causal interpretations, since we used a randomized controlled design. Therefore, the motor component of rhythm training and rhythmic entrainment represent important components

of a music training program that is designed to improve EFs. Further studies examining different age groups and larger sample sizes are required to confirm our findings. Moreover, we recommend a study examining the effect of dancing on EFs, because dancing relies on rhythmic motor movements as well. Therefore, dancing might exert a similar effect on EF. Finally, music-based rhythm training administered in a small group appears to be a suitable intervention to improve inhibition in preschoolers.

DATA AVAILABILITY

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

Human Subject Research

The studies involving human participants were reviewed and approved by Lokale-Ethikkommission Fachbereich 06 (LEK-FB06), Justus-Liebig-University of Giessen, Germany. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

FD and GS conceived the study. FD designed the study and obtained funding through a DFG grant. UF conducted the experiments and analyzed the data. UF, GS and FD wrote the manuscript.

FUNDING

This study was funded by a grant from the Deutsche Forschungsgesellschaft (DFG, No. DE2198/2-1).

ACKNOWLEDGMENTS

We thank our student assistants who were involved in implementing the training sessions and collecting data. We would like to specifically thank all children and their parents for their participation, as well as the kindergarten teachers for their support.

REFERENCES

- Beck, D. M., Schaefer, C., Pang, K., and Carlson, S. M. (2011). Executive function in preschool children: test-retest reliability. *J. Cogn. Dev.* 12, 169–193. doi: 10.1080/15248372.2011.563485
- Bugos, J. A., and DeMarie, D. (2017). The effects of a short-term music program on preschool children's executive functions. *Psychol. Music* 45, 855–867. doi: 10.1177/0305735617692666
- Clayton, M., Sager, R. D., and Will, U. (2004). In time with the music: the concept of entrainment and its significance for ethnomusicology. *ESEM CounterPoint* 1. Available online at: https://www.researchgate.net/publication/42789808_In_Time_With_the_Music_The_Concept_of_Entrainment_and_Its_Significance_for_Ethnomusicology. Accessed May 13, 2019.
- Coull, J. T. (2004). fMRI studies of temporal attention: allocating attention within, or towards, time. *Cogn. Brain Res.* 21, 216–226. doi: 10.1016/j.cogbrainres.2004.02.011
- Degé, F., Kubicek, C., and Schwarzer, G. (2011a). Music lessons and intelligence: a relation mediated by executive functions. *Music Percept. Interdiscip. J.* 29, 195–201. doi: 10.1525/mp.2011.29.2.195
- Degé, F., Wehrum, S., Stark, R., and Schwarzer, G. (2011b). The influence of two years of school music training in secondary school on visual and auditory memory. *Eur. J. Dev. Psychol.* 8, 608–623. doi: 10.1080/17405629.2011.590668

- Degé, F., and Schwarzer, G. (2017). Music lessons and verbal memory in 10- to 12-year-old children: investigating articulatory rehearsal as mechanism underlying this association. *Psychomusicol. Music Mind Brain* 27, 256–266. doi: 10.1037/pmu0000201
- Diamond, A. (2013). Executive functions. *Annu. Rev. Psychol.* 64, 135–168. doi: 10.1146/annurev-psych-113011-143750
- Diamond, A., and Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science* 333, 959–964. doi: 10.1126/science.1204529
- Dunemann-Gulde, A. (2005). *Yoga und Bewegungsspiele für Kinder. Für 4- bis 10-Jährige*. Munich: Kösel.
- Erdfelder, E., Faul, F., and Buchner, A. (1996). GPOWER: a general power analysis program. *Behav. Res. Methods Instrum. Comput.* 28, 1–11. doi: 10.3758/bf03203630
- Franklin, M. S., Moore, K. S., Yip, C. Y., Jonides, J., Rattray, K., and Moher, J. (2008). The effects of musical training on verbal memory. *Psychol. Music* 36, 353–365. doi: 10.1177/0305735607086044
- Hasselhorn, M., Schumann-Hengsteler, R., Gronauer, J., Grube, D., Mähler, C., Schmid, I., et al. (2012). *Arbeitsgedächtnistestbatterie für Kinder im Alter von 5–12 Jahren (AGTB5–12) [Working Memory Test Battery for Children from 5–12 Years of Age (AGTB5–12)]*. Göttingen: Hogrefe.
- Holochwost, S. J., Propper, C. B., Wolf, D. P., Willoughby, M. T., Fisher, K. R., Kolacz, J., et al. (2017). Music education, academic achievement, and executive functions. *Psychol. Aesthet. Creat. Arts* 11, 147–166. doi: 10.1037/aca0000112
- Jaschke, A. C., Honing, H., and Scherder, E. J. A. (2018). Longitudinal analysis of music education on executive functions in primary school children. *Front. Neurosci.* 12:103. doi: 10.3389/fnins.2018.00103
- Joret, M.-E., Germeys, F., and Gidron, Y. (2016). Cognitive inhibitory control in children following early childhood music education. *Music. Sci.* 21, 303–315. doi: 10.1177/1029864916655477
- Korkman, M., Kirk, U., and Kemp, S. (2007). *A Developmental Neuropsychological Assessment; NEPSY-II*. San Antonio, TX: Harcourt Assessment, Inc.
- Krause, V., Schnitzler, A., and Pollok, B. (2010). Functional network interactions during sensorimotor synchronization in musicians and non-musicians. *Neuroimage* 52, 245–251. doi: 10.1016/j.neuroimage.2010.03.081
- Logie, R. H. (1995). *Visuo-Spatial Working Memory*. Hove: Lawrence Erlbaum Associates Ltd.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “Frontal Lobe” tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734
- Miendlarzewska, E. A., and Trost, W. J. (2014). How musical training affects cognitive development: rhythm, reward and other modulating variables. *Front. Neurosci.* 7:279. doi: 10.3389/fnins.2013.00279
- Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., and Chau, T. (2011). Short-term music training enhances verbal intelligence and executive function. *Psychol. Sci.* 22, 1425–1433. doi: 10.1177/0956797611416999
- Nykrin, R., Grüner, M., and Widmer, M. (Eds.). (2007). *Musik und Tanz für Kinder*. Mainz: Schott.
- Patscheke, H., Degé, F., and Schwarzer, G. (2018). The effects of training in rhythm and pitch on phonological awareness in four- to six-year-old children. *Psychol. Music* doi: 10.1177/0305735618756763
- Rammstedt, B., and Danner, D. (2017). Die Facettenstruktur des big five inventory (BFI). *Diagnostica* 63, 70–84. doi: 10.1026/0012-1924/a000161
- Richmond, L. L., Morrison, A. B., Chein, J. M., and Olson, I. R. (2011). Working memory training and transfer in older adults. *Psychol. Aging* 26, 813–822. doi: 10.1037/a0023631
- Roden, I., Kreutz, G., and Bongard, S. (2012). Effects of a school-based instrumental music program on verbal and visual memory in primary school children: a longitudinal study. *Front. Psychol.* 3:572. doi: 10.3389/fpsyg.2012.00572
- Voss, M. W., Nagamatsu, L. S., Liu-Ambrose, T., and Kramer, A. F. (2011). Exercise, brain, and cognition across the life span. *J. Appl. Physiol.* 111, 1505–1513. doi: 10.1152/japplphysiol.00210.2011
- Weiß, R. H., and Osterland, J. (2012). *Grundintelligenztest Skala 1-Revision. CFT 1-R*. Göttingen: Hogrefe.
- Wiebe, S. A., Espy, K. A., and Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Dev. Psychol.* 44, 575–587. doi: 10.1037/0012-1649.44.2.575
- Winsler, A., Ducenne, L., and Koury, A. (2011). Singing one's way to self-regulation: the role of early music and movement curricula and private speech. *Early Educ. Dev.* 22, 274–304. doi: 10.1080/10409280903585739
- Zelazo, P. D. (2006). The dimensional change card sort (DCCS): a method of assessing executive function in children. *Nat. Protoc.* 1, 297–301. doi: 10.1038/nprot.2006.46
- Zuk, J., Benjamin, C., Kenyon, A., and Gaab, N. (2014). Behavioral and neural correlates of executive functioning in musicians and non-musicians. *PLoS One* 9:e99868. doi: 10.1371/journal.pone.0099868

Conflict of Interest Statement: 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.

Copyright © 2019 Frischen, Schwarzer and Degé. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Musicians Show Better Auditory and Tactile Identification of Emotions in Music

Andréanne Sharp^{1*}, Marie-Soleil Houde¹, Benoit-Antoine Bacon² and François Champoux¹

¹ École d'Orthophonie et d'Audiologie, Université de Montréal, Montreal, QC, Canada, ² Department of Psychology, Carleton University, Ottawa, ON, Canada

OPEN ACCESS

Edited by:

Assal Habibi,
University of Southern California,
United States

Reviewed by:

Patrick Bruns,
Universität Hamburg, Germany
Vesa Putkinen,
Turku PET Centre, Finland

*Correspondence:

Andréanne Sharp
andreanne.sharp@umontreal.ca;
sharp_andreanne@hotmail.com

Specialty section:

This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Psychology

Received: 23 April 2019

Accepted: 13 August 2019

Published: 28 August 2019

Citation:

Sharp A, Houde M-S, Bacon B-A
and Champoux F (2019) Musicians
Show Better Auditory and Tactile
Identification of Emotions in Music.
Front. Psychol. 10:1976.
doi: 10.3389/fpsyg.2019.01976

Musicians are better at processing sensory information and at integrating multisensory information in detection and discrimination tasks, but whether these enhanced abilities extend to more complex processes is still unknown. Emotional appeal is a crucial part of musical experience, but whether musicians can better identify emotions in music throughout different sensory modalities has yet to be determined. The goal of the present study was to investigate the auditory, tactile and audiotactile identification of emotions in musicians. Melodies expressing happiness, sadness, fear/threat, and peacefulness were played and participants had to rate each excerpt on a 10-point scale for each of the four emotions. Stimuli were presented through headphones and/or a glove with haptic audio exciters. The data suggest that musicians and control are comparable in the identification of the most basic (happiness and sadness) emotions. However, in the most difficult unisensory identification conditions (fear/threat and peacefulness), significant differences emerge between groups, suggesting that musical training enhances the identification of emotions, in both the auditory and tactile domains. These results support the hypothesis that musical training has an impact at all hierarchical levels of sensory and cognitive processing.

Keywords: emotion, music, auditory perception, tactile perception, brain plasticity

INTRODUCTION

It is well established that musical training can lead to functional and structural changes in the brain, and that these changes correlate with improved music processing as measured by pitch, timing and timbre discriminations (for a review see Kraus and Chandrasekaran, 2010). Of particular importance to the present study, a number of studies have revealed that long-term musical training promotes brain plasticity and generates reorganization in regions related to audiotactile processing (e.g., Pantev et al., 2003; Baumann et al., 2007; Zimmerman and Lahav, 2012).

At the behavioral level, it has been shown that in detections tasks, musicians react faster to auditory and tactile stimuli (Landry and Champoux, 2017) and are also better at integrating auditory and tactile information (Landry et al., 2017). In auditory frequency discrimination tasks, musicians have lower threshold compared to controls (Spiegel and Watson, 1984), and this effect appears to be correlated with years of musical expertise (Kishon-Rabin et al., 2001).

To examine whether such discrimination enhancements extended to multisensory processing, Young et al. (2017) used a two-alternative forced choice task in which participants had to determine whether a pair of stimuli were the same or different. Participant could hear the stimuli, combined or not with a corresponding tactile stimulation transmitted through a glove. The results revealed that compared to controls, musician frequency discrimination threshold was improved significantly by the addition of tactile stimulation.

Recent results from our laboratory have confirmed such frequency discrimination enhancements in the auditory and audiotactile domains and have extended the latter by demonstrating that musicians were also better at discriminating tactile-only stimuli applied to the hand (Sharp et al., 2019). Taken together, these results suggest that musical training can have an impact on sensory processing, at least in detection or discrimination tasks. Whether such enhanced abilities can extend to more complex processes remains a matter of debate.

During the last decades, the study of emotions in music has become an increasingly popular research field. It is known that the ability to identify emotion in music starts early in life and that young children base their judgments on basic psychoacoustic cues such as tempo, loudness and pitch (Adachi et al., 2004). At 3 years of age, children are sensitive to the positive and negative connotations of music but their analysis is not yet sufficiently nuanced to distinguish between more specific emotions (Kastner and Crowder, 1990). It is only around 5 years of age that children begin to discriminate happiness and sadness (Terwogt and Van Grinsven, 1991).

Around 11 years of age, children are able to identify emotions at the adult level (Hunter et al., 2011). Since the identification of emotions in music is based on psychoacoustic cues and musical features, the possibility that musical training might enhance this ability has long been surmised. Indeed it appears that musicians are more accurate than non-musicians in the identification of emotions in music (Vieillard et al., 2008). Decline due to age in the identification of emotion in music is also less marked in musicians (Castro and Lima, 2014). Emotion identification abilities in musicians have not been examined further and the capacity of musicians to better identify emotion in music throughout different sensory modalities also remains to be determined.

The present study aims at investigating the auditory, tactile and audiotactile identification of various emotions in musicians using the stimuli of Vieillard et al. (2008) and tactile stimulation technology developed by Young et al. (2017). This study will be the first to examine tactile and auditory-tactile identification of emotion abilities in musicians versus controls.

METHODS

Participants

Seventeen professional musicians (7 women, 10 men, average age = 28.9 years) and 17 matched non-musicians (8 women, 9 men, average age = 34.4 years) participated in the study. Non-musicians and musicians were matched for age, sex, handedness,

educational level, and hearing thresholds. Only participants with less than 1 year of musical training were recruited for the non-musician (control) group. The sample size of this study is justified by the restrictive criteria used for inclusion in the musicians' group. All musicians were working in the music field or studying music at the university level. The musicians specialized in piano ($n = 9$), guitar ($n = 2$), trumpet ($n = 2$), violin ($n = 1$), percussion ($n = 1$), flute ($n = 1$) and oboe ($n = 1$). They reported playing only one instrument ($n = 4$), playing two instruments ($n = 2$) or playing more than two instruments ($n = 11$). The average age of beginning to learn their first instrument was 7 years old. The average number of years of active practice of music was 20 years. Hearing thresholds were determined with an audiometer (Astera, GN Otometrics, Denmark). For both groups, pure-tone detection thresholds at octave frequencies ranging from 250 to 4000 kHz were within normal limits in both ears. The Research Committee for Health Sciences of the University of Montreal and the Center for Interdisciplinary Research in Rehabilitation of Greater Montreal approved all procedures, and each participant provided written informed consent. All experiments were performed in accordance with relevant guidelines and regulations.

Stimuli and Procedure

The stimuli used in this study were developed by Vieillard et al. (2008). They are 56 melodies produced by a digital synthesizer in piano timbre. These instrumental stimuli were composed in the tonal musical tradition to express four emotions: happiness, sadness, fear/threat and peacefulness. The stimuli vary in mode, dissonance, pitch range, tone density, rhythmic regularity, and tempo but do not vary in performance-related expressive features (e.g., vibrato or variations of articulation/phrasing). Therefore the identification of emotions was based exclusively on the compositional structure. The mean duration of each stimuli was 12.4 s. All stimuli were originally validated by Vieillard et al. (2008) and were also cross-culturally validated by Fritz et al. (2009). These stimuli have been designed to elicit specific emotions that can be universally recognized.

The battery of Vieillard et al. (2008) was selected for this experiment because the four emotions evoked by the melodies are easily recognized and discriminated. Furthermore, all stimuli were validated cross-culturally by Fritz et al. (2009), and across age groups by Lima and Castro (2011). Finally, peacefulness is the most likely stimulus in this experiment to avoid a ceiling effect in musicians which show near perfection identification for better known emotions such as happy and sad.

Each of the 56 melodies were presented in a randomized order in three stimulation conditions: auditory-only, tactile-only and auditory-tactile. There were 14 stimuli for each type of emotion. For each stimuli, participants had to rate how much the melody expressed each of the four emotions on a 10-point intensity scale ranging from 0 (absent) to 9 (present). The four scales were presented immediately after each stimulus, and always in the same order (happy/sad/scary/peaceful). Each melody was presented only once in random order during each block (auditory only, tactile only and auditory-tactile) and no feedback was given. All conditions for stimulation and emotion were randomized. For example, one participant started in the tactile condition with a

peaceful stimulus, while another started in the auditory condition with a sad stimulus. To exactly replicate the standardized task of Vieillard et al. (2008), the order of the scale presented after each stimulus was not counterbalanced.

Participants were seated in a soundproof room and stimuli were presented via headphones (TDH-39, Diatec, Canada) for the auditory-only condition, via a vibrating glove device for the tactile-only condition, and via both headphones and a vibrating glove for the auditory-tactile condition. During the tactile-only condition, white noise was presented via headphones and the participant wore earplugs. The participant had to adjust the volume during practice trials so as not to hear the vibrating glove.

The vibrating glove was a replication of the glove used by Young et al. (2017) and was equipped with six independent audio-haptic voice-coil exciters. The voice-coil transducers (TEAX14C02-8 Compact Audio Exciter) had a diameter of 14 mm and were designed to deliver vibrotactile output. The frequency range of these speakers is 300 to 20,000 Hz. Stimuli were sent via a Dayton Audio DTA3116S Class D Micro Mini Amplifier (2 × 15 W), linked via an audio cable to the software Psyscope 1.2.5 (Cohen et al., 1993) on a Mac computer.

Analysis

The percentage of accurate responses, defined as the highest rating score for a melody corresponding to the intended emotion, was calculated for each participant for each emotion. For example, given a happy melody and a rating of Happy = 7, Sad = 3,

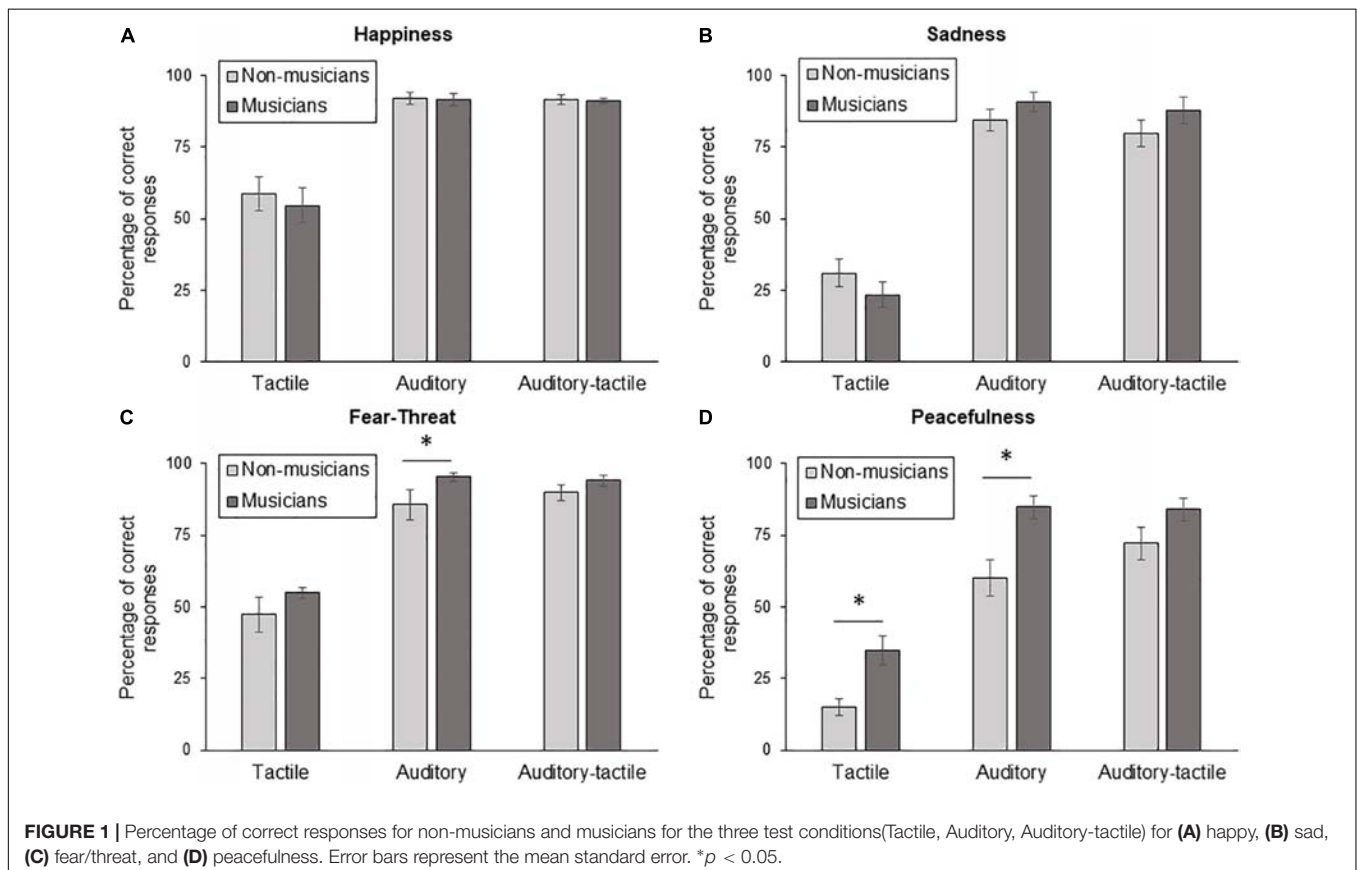
Fear = 2, Peaceful = 6, the response would be counted as correct, whereas Happy = 6, Sad = 3, Fear = 2, Peaceful = 7 would be counted as incorrect. The same rating could never be used twice for any of the melody ratings.

An ANOVA was used as an omnibus test to compare the percentage of accurate responses for stimulation conditions and emotions as within-subject factors and groups as a between-subject factor. A multivariate analysis of variance was used to compare the percentage of accurate responses between groups. To provide an estimation of multisensory benefits compared to unimodal stimulation, the increase in performance was measured by subtracting the score in the auditory only condition from the score in the auditory-tactile condition. The results provide an estimation of the contribution of tactile stimulation.

RESULTS

Figure 1 displays the percentage of accurate responses for auditory, tactile and auditory-tactile conditions for each of the emotions. An ANOVA for stimulation conditions, emotions and groups was used as an omnibus test. There was a significant difference between groups ($F(1,32) = 10.834$, $p = 0.002$). There was also a significant interaction between the condition and emotion variables ($p < 0.0001$).

The multivariate analysis of variance used to compare the percentage of accurate responses revealed a statistically



significant difference in conditions based on Group ($F(12, 21) = 2.585$, $p = 0.027$; Wilk's $\Lambda = 0.404$, partial $\eta^2 = 0.596$). **Table 1** shows that there were significant differences between groups for fear/threat auditory, peacefulness auditory and peacefulness tactile whereas no significant differences between groups were found in the other conditions.

Uncorrected t -tests revealed that for both groups, mean percentage of responses was above chance for auditory and auditory-tactile stimulation conditions for all type of emotions ($p < 0.001$). For tactile stimulation, uncorrected t -tests revealed that the mean percentage of responses was above chance for both groups for happy ($p < 0.001$) and fear/threat emotions (controls: $p = 0.002$, musicians: $p < 0.001$), but not for sad (controls: $p = 0.153$, musicians: $p = 0.747$). Finally, an uncorrected t -test showed that musicians were performing above chance for tactile stimulation for peaceful emotion ($t(16) = 2.170$, $p = 0.045$) while on the contrary, another uncorrected t -test showed that controls were performing below chance for tactile stimulation for peaceful ($t(16) = -4.629$, $p < 0.001$).

For the happiness and sadness conditions, no increases in performance were observed in the auditory-tactile compared to the auditory-only condition in either groups (mean under 0%). For sadness and peacefulness, there were increases measured for controls (Sadness: 4% and Peacefulness: 12%), but not for musicians (mean under 0%). After correcting for multiple comparisons the increase in performance between auditory and auditory-tactile stimulation was not significant for either musicians or controls (see **Table 2** for more details).

ANOVAs were used as an omnibus test to compare the number of errors between groups for each expected emotion (4). The dependent variable was the number of errors and independent variables were groups, stimulation conditions and categories of the emotion scale. The ANOVAs for happiness ($F(1,32) = 0.141$, $p = 0.710$), sadness ($F(1,32) = 0.196$, $p = 0.661$) and fear/threat ($F(1,31) = 3.061$, $p = 0.090$) revealed no differences between groups. There was a significant difference between groups for peacefulness ($F(1,32) = 10.691$, $p = 0.003$). t -Test analysis revealed differences between groups for sadness

when the expected emotion was peacefulness. This emotion had the higher rate of error for both groups. The difference between groups was the number of error, but not the type of emotion wrongly associated with peacefulness. In all conditions, both group were doing the same kind of errors for each type of emotion as shown in **Table 3**. In the auditory stimulation condition, for both groups, the emotion with which happiness and sadness was most often confused with was peacefulness. Similarly, for both groups, the emotion with which fear/threat and peacefulness were most often confused with was sadness. Results were the exact same in the auditory-tactile stimulation. In the tactile stimulation condition, for both groups, the emotion with which happiness was most often confused with was fear/threat. For all other emotions in the tactile stimulation condition, errors were distributed across the other three type of emotions. The missing values in **Table 3** are due to the fact that it was not possible to categorize some errors, because some participants were giving a 0 score to all types of emotions in the scale for a few trials.

DISCUSSION

The main objective of the present study was to investigate auditory, tactile and auditory-tactile identification of emotion in musicians versus non-musicians. A significant difference between groups was found, with musicians showing better emotion identification for fear/threat in the auditory condition and for peacefulness in both the auditory and tactile conditions. Additionally, even if the difference does not remain significant after correcting for multiple comparisons, the trend indicates a possible gain from adding tactile stimulation to the auditory stimuli in peacefulness condition for controls (12%), but not for musicians (under 0%).

The significant differences found between controls and musicians can be linked to the complexity of the emotions displayed. It is well-known that happiness and sadness are the easiest emotions to identify because they are mainly based on tempo (see Terwogt and Van Grinsven, 1991). As such it is not surprising that results revealed no difference between controls and musicians for happy (auditory, auditory-tactile and

TABLE 1 | Statistical results from the multivariate analysis of variance used to compare percentage of accurate responses between groups (auditory, tactile, auditory-tactile) for all four emotions.

	Auditory	Tactile	Auditory-Tactile
Happy	$F(1, 32) = 0.313$ $p = 0.579$ partial $\eta^2 = 0.0.010$	$F(1, 32) = 0.023$ $p = 0.881$ partial $\eta^2 = 0.0.001$	$F(1, 32) = 0.020$ $p = 0.887$ partial $\eta^2 = 0.0.001$
Sad	$F(1, 32) = 3.010$ $p = 0.092$ partial $\eta^2 = 0.0.086$	$F(1, 32) = 1.859$ $p = 0.182$ partial $\eta^2 = 0.0.055$	$F(1, 32) = 0.797$ $p = 0.379$ partial $\eta^2 = 0.0.024$
Fear/Threat	$F(1, 32) = 4.23$ $p = 0.048^*$ partial $\eta^2 = 0.117$	$F(1, 32) = 2.379$ $p = 0.133$ partial $\eta^2 = 0.0.069$	$F(1, 32) = 1.940$ $p = 0.173$ partial $\eta^2 = 0.0.057$
Peacefulness	$F(1, 32) = 15.838$ $p < 0.001^*$ partial $\eta^2 = 0.0.331$	$F(1, 32) = 8.432$ $p = 0.007^*$ partial $\eta^2 = 0.209$	$F(1, 32) = 1.531$ $p = 0.225$ partial $\eta^2 = 0.0.046$

Bold and * $p < 0.05$.

TABLE 2 | Mean percentage of increase in performance from adding tactile stimulation to auditory stimulation (auditory-tactile performance – auditory only performance).

	Group	Mean (%)	Standard error of the mean	t-Test: Auditory versus Auditory-tactile performance
Happiness	Controls	-0.50	2.04	$t(16) = 0.833$, $p = 0.417$
	Musicians	-0.39	2.15	$t(16) = 0.190$, $p = 0.851$
Sadness	Controls	-4.59	5.96	$t(16) = 2.048$, $p = 0.057$
	Musicians	-2.87	3.00	$t(16) = 0.975$, $p = 0.344$
Fear/Threat	Controls	4.24	3.51	$t(16) = -1.357$, $p = 0.193$
	Musicians	-1.20	1.52	$t(16) = 0.824$, $p = 0.422$
Peacefulness	Controls	12.19	5.73	$t(16) = 2.318$, $p = 0.034$
	Musicians	0.78	3.50	$t(16) = 0.235$, $p = 0.818$

TABLE 3 | Mean percentage of correct responses and mean percentage of errors per emotion classified by the type of emotion wrongly identified.

Auditory								
	Happiness				Sadness			
	Correct	Sadness	Fear/Threat	Peacefulness	Correct	Happiness	Fear/Threat	Peacefulness
Controls	92.02	0.84	0.84	6.72	86.98	0.42	2.10	7.14
Musicians	91.61	0.84	0.00	7.56	90.76	0.42	2.10	7.14
	Fear/Threat				Peacefulness			
	Correct	Happiness	Sadness	Peacefulness	Correct	Happiness	Sadness	Fear/Threat
Controls	84.45	7.14	10.50	1.68	61.77	19.75	23.93*	0.84
Musicians	95.38	0.00	3.36	0.42	84.87	6.30	8.86*	0.00
Auditory-tactile								
	Happiness				Sadness			
	Correct	Sadness	Fear/Threat	Peacefulness	Correct	Happiness	Fear/Threat	Peacefulness
Controls	90.23	0.00	0.00	3.78	76.47	0.42	2.94	10.50
Musicians	91.19	1.26	0.00	6.30	87.84	0.00	4.62	5.46
	Fear/Threat				Peacefulness			
	Correct	Happiness	Sadness	Peacefulness	Happiness	Sadness	Fear/Threat	Peacefulness
Controls	89.50	0.42	3.36	0.84	75.64	9.24	17.65	0.84
Musicians	94.12	0.42	4.20	0.42	84.05	6.30	8.82	0.00
Tactile								
	Happiness				Sadness			
	Correct	Sadness	Fear/Threat	Peacefulness	Correct	Happiness	Fear/Threat	Peacefulness
Controls	55.87	8.82	26.07	8.82	33.18	21.43	25.63	18.49
Musicians	54.63	6.30	19.71	16.81	23.51	15.97	25.21	31.93
	Fear/Threat				Peacefulness			
	Correct	Happiness	Sadness	Peacefulness	Correct	Happiness	Sadness	Fear/Threat
Controls	45.78	21.85	24.79	13.39	13.85	41.60	3.00	23.11
Musicians	55.03	9.24	19.33	13.03	35.71	27.31	16.39	17.65

Bold and * $p < 0.05$.

tactile) and sad (auditory and auditory-tactile) conditions as there were ceiling effects. The average performance for sad for tactile stimulation did not differ between groups, but also, did not differ from chance for both groups. A more sensitive task would be needed to determine whether musical expertise can lead to more accurate identification of these emotions via auditory, tactile and auditory-tactile stimulation. Fear/threat is a musically less straightforward emotion than happiness and sadness (Vieillard et al., 2008; Tan et al., 2017). Hence, compared to controls, musicians more accurately identified that emotion in the auditory condition. In the same vein, the most complex and ambiguous emotion displayed in the sample melodies, namely peacefulness (Vieillard et al., 2008; Tan et al., 2017), was more accurately identified by musicians than by controls in both the auditory and the tactile conditions.

Results from the auditory condition are consistent with the extensive literature demonstrating that musical training leads to brain plasticity and can improve music processing as measured by pitch, timing and timbre discriminations (for a review see Kraus and Chandrasekaran, 2010). Since the identification of emotions in music is based on psychoacoustic cues and musical features, the enhanced performance of musicians in the auditory condition was also to be expected. Furthermore, an important component of musical training is aimed at understanding and experiencing the full range of emotional meaning and expressiveness, however faint, of a musical performance (Castro and Lima, 2014). As such, it is not surprising that improved performance was only found in conditions where musicians had to identify subtle emotional qualities.

One recent study have investigated recognition of emotions in an auditory-only stimulation condition. They suggest a correlation between years of musical training and accuracy at identifying emotion in music and revealed a significant difference between groups for older musicians with respect to sad and fear emotions (Castro and Lima, 2014). It should be noted that a major limitation of this study was that the range of musical expertise of participants as measured in years was large (8–18 years), and that the average age of training onset was over 7 years of age, the known threshold beyond which music-induced structural changes and learning effects become less pronounced (for a review see Habib and Besson, 2009). As such the lesser musical expertise of their younger participants may explain why they could not find any significant differences between groups. In contrast, results from the present study were obtained with participants whose average age of learning onset was 7 years of age, and whose average number of years of active practice of music was 20.2 years. All participants were working or studying full-time in the field of music and can be considered professional musicians. In addition, the average age of the participants was 34.4 years for controls and 28.9 years for musicians, which corresponds to the younger group of Castro and Lima (2014).

The present study was the first to investigate the tactile identification of emotions in music. Results revealed that both musicians and controls were able to identify emotions via tactile stimulation only, which is in itself a new and major finding. No study to date has investigated purely tactile identification of emotion in music. The only existing study along these lines was performed by Branje et al. (2013) and suggests that multisensory stimulation can increase emotion perception in film. By using the Emoti-Chair, a device that induces vibration in the back of normal-hearing participants, they found increases in skin conductance levels when vibrotactile stimuli were added to audio/visual film content. They also observed that not only the intensity of vibration but also the frequency of the vibrotactile stimuli was playing a role in the observed reactions. The present study results are consistent with Branje et al. (2013) and further support the hypothesis that both controls and musicians are able to extract meaningful information from the frequency characteristics of a signal presented through vibrations only. Furthermore, for the emotion of peacefulness, results revealed a significant difference between musicians and controls for tactile stimulation. These results are consistent with a previous study from our laboratory, the first to demonstrate that musicians were better at discriminating frequencies via tactile stimulation applied to the hand (Sharp et al., 2019). The enhanced ability to identify peaceful emotions in music via tactile stimulation suggests that more complex processes are improved following long-term musical training. This hypothesis should be verified using other types of complex emotions that are easier to identify via tactile stimulation than peacefulness. Indeed, results in the peacefulness condition are above chance for musicians, but not for controls and the comparison of performance would be easier to interpret if both groups were above chance.

It is well-known that the frequency spectrum treated is more limited than that of the hair cells of the cochlea (1 to 1000 kHz) (Rovan and Hayward, 2000). Which musical

components is perceived though the tactile modality remains a question of debate. Some studies suggest that non-musicians can detect different musical notes via the tactile modality (Hopkins et al., 2016) and that they can discriminate timbre (Russo et al., 2012). Furthermore, low frequencies in music are important to understanding beat and can be transmitted via vibrotactile devices (Van Dyck et al., 2013; Tranchant et al., 2017). All these psychoacoustic cues are known to be transmitted via the tactile modality and are all important for emotion identification in music. Further study should investigate if other cues are used in the identification of emotion in the tactile domain or if some of these cues are more important than the others. All these studies support our results suggesting that non-musicians and musicians are able to identify emotion via tactile stimulation only.

Finally, the lack of significant difference between musicians and non-musicians in the auditory-tactile condition can be explained by the trend for controls toward exhibiting gain from tactile stimulation compared to musicians, as the latter were already too skilled in the auditory domain to benefit from tactile stimulation. Further studies should use more complex emotional stimuli to assess whether there could be a tactile gain for musicians, and investigate whether non-musicians' performance could become similar to that of musicians with training and feedback.

DATA AVAILABILITY

The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of Research Committee for Health Sciences of the University of Montreal and the Center for Interdisciplinary Research in Rehabilitation of Greater Montreal with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Research Committee for Health Sciences of the University of Montreal and the Center for Interdisciplinary Research in Rehabilitation of Greater Montreal.

AUTHOR CONTRIBUTIONS

AS and FC designed and performed the experiment. All authors wrote the manuscript, discussed the results and implications, and commented on the manuscript at all stages.

FUNDING

This work was supported by Natural Sciences and Engineering Research Council of Canada (RGPIN-2016-05211).

REFERENCES

- Adachi, M., Trehub, S. E., and Abe, J. I. (2004). Perceiving emotion in children's songs across age and culture 1. *Jpn. Psychol. Res.* 46, 322–336. doi: 10.1111/j.1468-5584.2004.00264.x
- Baumann, S., Koeneke, S., Schmidt, C. F., Meyer, M., Lutz, K., and Jancke, L. (2007). A network for audio-motor coordination in skilled pianists and non-musicians. *Brain Res.* 1161, 65–78. doi: 10.1016/j.brainres.2007.05.045
- Branje, C., Nespoil, G., Russo, F., and Fels, D. I. (2013). The effect of vibrotactile stimulation on the emotional response to horror films. *Comput. Entertain.* 11:4.
- Castro, S. L., and Lima, C. F. (2014). Age and musical expertise influence emotion recognition in music. *Music Percept. Interdiscip. J.* 32, 125–142. doi: 10.1525/mp.2014.32.2.125
- Cohen, J. D., Macwhinney, B., Flatt, M. R., and Provost, J. (1993). PsyScope: a new graphic interactive environment for designing psychology experiments. *Behav. Res. Methods Instrum. Comput.* 25, 257–271.
- Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., et al. (2009). Universal recognition of three basic emotions in music. *Curr. Biol.* 19, 573–576. doi: 10.1016/j.cub.2009.02.058
- Habib, M., and Besson, M. (2009). What do music training and musical experience teach us about brain plasticity? *Music Percept. Interdiscip. J.* 26, 279–285. doi: 10.1525/mp.2009.26.3.279
- Hopkins, C., Maté-Cid, S., Fulford, R., Seiffert, G., and Ginsborg, J. (2016). Vibrotactile presentation of musical notes to the glabrous skin for adults with normal hearing or a hearing impairment: thresholds, dynamic range and high-frequency perception. *PLoS One*, 11:e0155807. doi: 10.1371/journal.pone.0155807
- Hunter, P. G., Schellenberg, E. G., and Stalinski, S. M. (2011). Liking and identifying emotionally expressive music: age and gender differences. *J. Exp. Child Psychol.* 110, 80–93. doi: 10.1016/j.jecp.2011.04.001
- Kastner, M. P., and Crowder, R. G. (1990). Perception of the major/minor distinction: IV. Emotional connotations in young children. *Music Percept. Interdiscip. J.* 8, 189–201. doi: 10.2307/40285496
- Kishon-Rabin, L., Amir, O., Vexler, Y., and Zaltz, Y. (2001). Pitch discrimination: are professional musicians better than non-musicians? *J. Basic Clin. Physiol. Pharmacol.* 12, 125–144.
- Kraus, N., and Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nat. Rev. Neurosci.* 11, 599–605. doi: 10.1038/nrn2882
- Landry, S. P., and Champoux, F. (2017). Musicians react faster and are better multisensory integrators. *Brain Cogn.* 111, 156–162. doi: 10.1016/j.bandc.2016.12.001
- Landry, S. P., Sharp, A., Pagé, S., and Champoux, F. (2017). Temporal and spectral audiotactile interactions in musicians. *Exp. Brain Res.* 235, 525–532. doi: 10.1007/s00221-016-4813-3
- Lima, C. F., and Castro, S. L. (2011). Emotion recognition in music changes across the adult life span. *Cogn. Emot.* 25, 585–598. doi: 10.1080/02699931.2010.502449
- Pantev, C., Ross, B., Fujioka, T., Trainor, L. J., Schulte, M., and Schulz, M. (2003). Music and learning-induced cortical plasticity. *Ann. N.Y. Acad. Sci.* 999, 438–450. doi: 10.1196/annals.1284.054
- Rovan, J., and Hayward, V. (2000). "Typology of tactile sounds and their synthesis in gesture-driven computer music performance," in *Trends in Gestural Control of Music*, eds M. Wanderley, and M. Battier (Paris: IRCAM), 297–320.
- Russo, F. A., Ammirante, P., and Fels, D. I. (2012). Vibrotactile discrimination of musical timbre. *J. Exp. Psychol. Hum. Percept. Perform.* 38, 822–826. doi: 10.1037/a0029046
- Sharp, A., Houde, M. S., Maheu, M., Ibrahim, I., and Champoux, F. (2019). Improve tactile frequency discrimination in musicians. *Exp. Brain Res.* 237, 1–6. doi: 10.3758/s13414-014-0634-2
- Spiegel, M. F., and Watson, C. S. (1984). Performance on frequency-discrimination tasks by musicians and nonmusicians. *J. Acoust. Soc. Am.* 76, 1690–1695. doi: 10.1121/1.391605
- Tan, S. L., Pfordresher, P., and Harré, R. (2017). *Psychology of Music: From Sound to Significance*. Abingdon: Routledge.
- Terwogt, M. M., and Van Grinsven, F. (1991). Musical expression of moodstates. *Psychol. Music* 19, 99–109. doi: 10.1177/0305735691192001
- Tranchant, P., Shiell, M. M., Giordano, M., Nadeau, A., Peretz, I., and Zatorre, R. J. (2017). Feeling the beat: bouncing synchronization to vibrotactile music in hearing and early deaf people. *Front Neurosci.* 11:507. doi: 10.3389/fnins.2017.00507
- Van Dyck, E., Moelants, D., Demey, M., Deweppe, A., Coussement, P., and Leman, M. (2013). The impact of the bass drum on human dance movement. *Music Percept. Interdiscip. J.* 30, 349–359. doi: 10.1525/mp.2013.30.4.349
- Vieillard, S., Peretz, I., Gosselin, N., Khalfa, S., Gagnon, L., and Bouchard, B. (2008). Happy, sad, scary and peaceful musical excerpts for research on emotions. *Cogn. Emot.* 22, 720–752. doi: 10.1080/02699930701503567
- Young, G. W., Murphy, D., and Weeter, J. (2017). Haptics in music: the effects of vibrotactile stimulus in low frequency auditory difference detection tasks. *IEEE Trans. Haptics* 10, 135–139. doi: 10.1109/TOH.2016.264637
- Zimmerman, E., and Lahav, A. (2012). The multisensory brain and its ability to learn music. *Ann. N.Y. Acad. Sci.* 1252, 179–184. doi: 10.1111/j.1749-6632.2012.06455.x

Conflict of Interest Statement: 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.

Copyright © 2019 Sharp, Houde, Bacon and Champoux. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Dynamic Orchestration of Brains and Instruments During Free Guitar Improvisation

Viktor Müller^{1*} and Ulman Lindenberger^{1,2,3}

¹ Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany, ² Max Planck UCL Centre for Computational Psychiatry and Ageing Research, Berlin, Germany, ³ Max Planck UCL Centre for Computational Psychiatry and Ageing Research, London, United Kingdom

OPEN ACCESS

Edited by:

Assal Habibi,
University of Southern California,
United States

Reviewed by:

Donald Glowinski,
Université de Genève, Switzerland
Shoji Tanaka,
Sophia University, Japan

*Correspondence:

Viktor Müller
vmueller@mpib-berlin.mpg.de

Received: 15 May 2019

Accepted: 20 August 2019

Published: 04 September 2019

Citation:

Müller V and Lindenberger U (2019)
Dynamic Orchestration of Brains and
Instruments During Free Guitar
Improvisation.
Front. Integr. Neurosci. 13:50.
doi: 10.3389/fnint.2019.00050

Playing music in ensemble requires enhanced sensorimotor coordination and the non-verbal communication of musicians that need to coordinate their actions precisely with those of others. As shown in our previous studies on guitar duets, and also on a guitar quartet, intra- and inter-brain synchronization plays an essential role during such interaction. At the same time, sensorimotor coordination as an essential part of this interaction requires being in sync with the auditory signals coming from the played instruments. In this study, using acoustic recordings of guitar playing and electroencephalographic (EEG) recordings of brain activity from guitarists playing in duet, we aimed to explore whether the musicians' brain activity synchronized with instrument sounds produced during guitar playing. To do so, we established an analytical method based on phase synchronization between time-frequency transformed guitar signals and raw EEG signals. Given phase synchronization, or coupling between guitar and brain signals, we constructed so-called *extended* hyper-brain networks comprising all possible interactions between two guitars and two brains. Applying a graph-theoretical approach to these networks assessed across time, we present dynamic changes of coupling strengths or dynamic orchestration of brains and instruments during free guitar improvisation for the first time. We also show that these dynamic network topology changes are oscillatory in nature and are characterized by specific spectral peaks, indicating the temporal structure in the synchronization patterns between guitars and brains. Moreover, extended hyper-brain networks exhibit specific modular organization varying in time, and binding each time, different parts of the network into the modules, which were mostly heterogeneous (i.e., comprising signals from different instruments and brains or parts of them). This suggests that the method capturing synchronization between instruments and brains when playing music provides crucial information about the underlying mechanisms. We conclude that this method may be an indispensable tool in the investigation of social interaction, music therapy, and rehabilitation dynamics.

Keywords: intra- and inter-brain coupling, brain-instrument coupling, graph-theoretical approach, EEG hyperscanning, phase synchronization, extended hyper-brain networks, social interaction

INTRODUCTION

As noted by D'Ausilio et al. (2015): "Group-level musical coordination can be considered as a microcosm of social interaction." A recently emerging view in music neuroscience with regard to hyperscanning methods holds that playing music in groups is not only social and interactive (Keller et al., 2014; Acquadro et al., 2016; Chang et al., 2017), but that it requires strong inter-brain synchronization and specific hyper-brain network activity supporting interpersonal action coordination (Lindenberger et al., 2009; Sängers et al., 2011, 2012, 2013; Müller et al., 2013, 2018b). This hyper-brain network activity including both intra- and inter-brain synchronization is enhanced during periods of high demands on musical coordination and is accompanied by the emergence of so-called hyper-brain modules composed of nodes from two or more brains. It has also been shown that the topology of hyper-brain networks involving two (Sängers et al., 2012; Müller et al., 2013) or even four (Müller et al., 2018b) brains revealed small-world properties with high segregation and integration of brain function, and had a tendency to become more random at lower frequencies and more regular at higher frequencies. Moreover, this topology was characterized by a higher number of hub-connectors at the delta and theta frequency of brain signals during joint, as compared to solo guitar playing (Müller et al., 2013). Furthermore, different types of information flow—intra- vs. inter-modular—were found when playing guitar in quartet (Müller et al., 2018b). Nevertheless, little, if anything, is known about the interaction between brain processes or mechanisms implementing interpersonally coordinated behavior when playing music and the instruments used for music production.

In the current study, we attempt to close the conceptual gap between these important elements of musically coordinated behavior—music production and its neuronal implementation. Using acoustic recordings of guitar playing and electroencephalographic (EEG) recordings of the brain activity of guitarists playing guitar in duet, we tried to establish a method to investigate the couplings within and between all components of duet playing, i.e., the guitars and brains. All these couplings were then used to construct and to analyze a so-called *extended* hyper-brain network including two guitars and two brains and all connections within and between them. We describe different situations of connections between guitars and brains here, which together exhibit complex networks with network topology changing in time and reflecting the dynamic orchestration of brains and instruments in guitar duos.

MATERIALS AND METHODS

Participants

Two pairs of professional guitarists participated in the study. In both pairs, the lead guitarist was always the same individual. The guitarists in the duo were not known to each other. All participants were right-handed and had played the guitar professionally for more than 5 years. The study was approved by the ethics committee of Max Planck Institute for Human Development (Berlin), and therefore performed in accordance

with the ethical standards laid down in the 1964 Declaration of Helsinki. All subjects volunteered for this experiment and gave their written informed consent prior to their inclusion in the study.

Experimental Procedure and Data Acquisition

During the experiment, the guitarists sat facing each other and freely improvised in duet for about 5–6 min. Before playing, the guitarists had the opportunity to briefly discuss the theme of their improvisation. Typically, one guitarist played a single-line melody or solo, while the other accompanied with chords or in another way. They also switched roles several times during the improvisation. Participants were instructed to avoid unnecessary movement in order to reduce movement artifacts.

Acoustic and EEG measurement took place in an acoustically and electromagnetically shielded cabin. EEG was simultaneously recorded from both members of the pair using two electrode caps with 64 Ag/AgCl electrodes placed according to the international 10–10 system, with the reference electrode placed at the right mastoid. For further analysis, we used 40 EEG channels for each subject. These channels or electrodes were distributed across the entire cortex, so that the information of the remaining electrodes would be rather redundant. Separate amplifiers with separate grounds were used for each individual, optically coupled to a computer. The vertical and horizontal electrooculogram (EOG) was also recorded to control for eye blinks and eye movements. Sampling rate was 5,000 Hz. The anti-aliasing filter was set to 1,000 Hz. A notch filter was used to suppress 50 Hz noise. EEG recordings were re-referenced offline to an average of the left and right mastoids and then filtered with a band pass ranging from 0.5 to 70 Hz. Eye movement correction was accomplished by independent component analysis (Vigário, 1997). Thereafter, artifacts from head and body movements were rejected by visual inspection. Spontaneous EEG activity was resampled at 250 Hz and divided into artifact-free 10-s epochs. In all, 10 artifact-free 10-s epochs were used for coupling and network analyses.

The sounds of the guitars were recorded through two microphones (i.e., one for each guitar) on two EEG channels, simultaneously with the EEG recordings. These two sound signals were divided into corresponding 10-s epochs without resampling. In addition, video and sound were recorded using Video Recorder Software (Brain Products, Munich, Germany) synchronized with EEG data acquisition.

Data Analysis

To investigate phase synchronization or coupling between the signals, we first normalized the high-frequency auditory signals and applied an analytic Morlet wavelet transform (Figures 1A,B) to calculate the power or amplitude within the four different frequency ranges: low (50–250 Hz), middle (250–500 Hz), high (500–2,000 Hz), and whole range (50–2,000 Hz). By averaging the amplitude within these four frequency ranges, we generated low-frequency time series that varied in a frequency range comparable to the EEG time series (see Figure 1 for details). To investigate phase coupling between the given signals in a directed and frequency-resolved manner, we calculated the

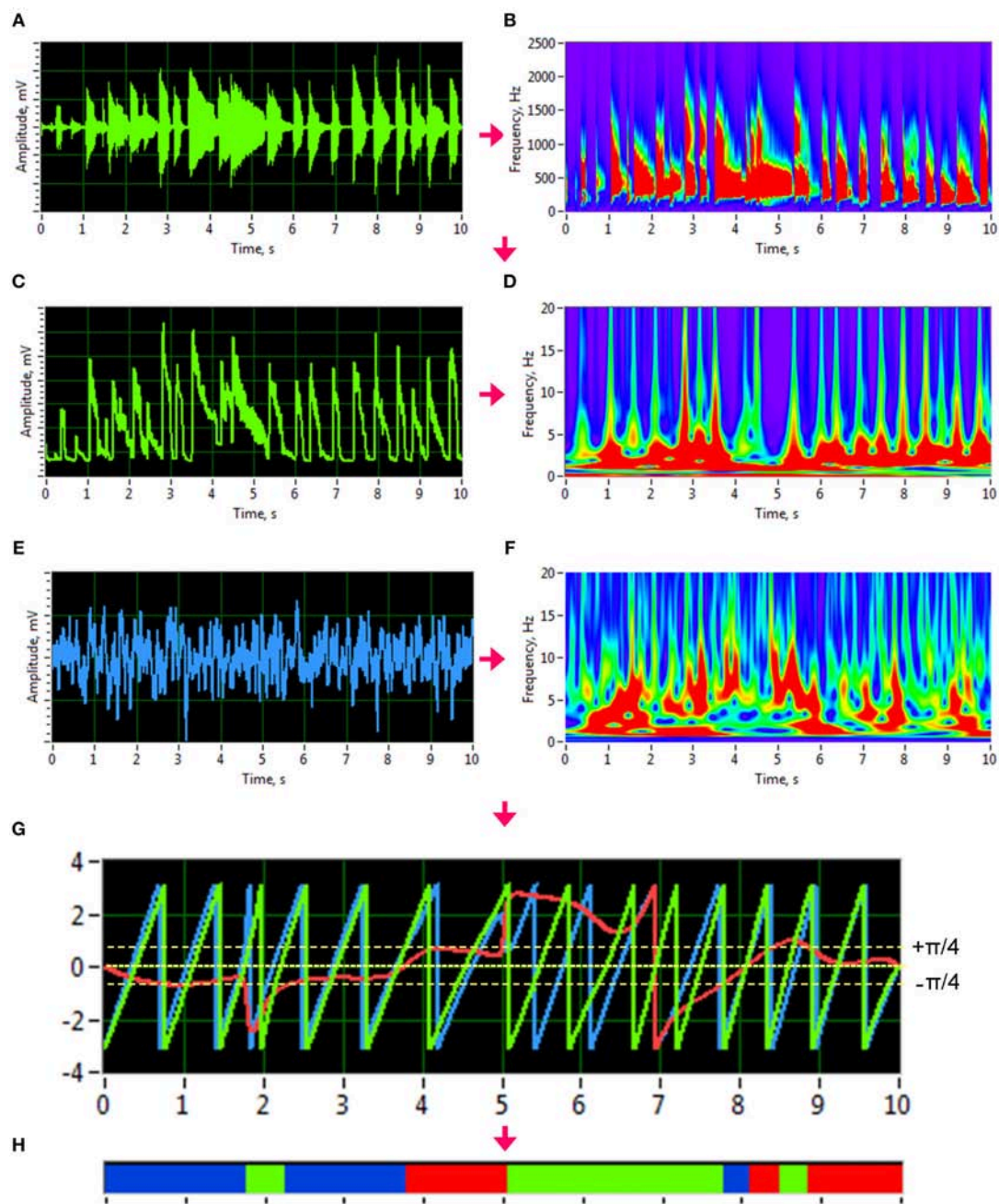


FIGURE 1 | Transformation of guitar signals and calculation of the Integrative Coupling Index (ICI). **(A)** Raw signal of the guitar recording. **(B)** Wavelet transform of the guitar recording in time-frequency domain. **(C)** New guitar signal derived from the wavelet-transformed signal by averaging the wavelet power across the frequency bins of interest. The whole range between 50 and 2,000 Hz is presented here. **(D)** Wavelet transform of the new guitar signal in the time-frequency domain used for calculation of the ICI. **(E)** Raw EEG signal. **(F)** Wavelet transform of the EEG signal in time-frequency domain. **(G)** Time course of instantaneous phases from two signals and their phase difference. Phase of the guitar signal = green curve; phase of the EEG signal = blue curve; phase difference ($\Delta\phi$) between the two signals = red curve). **(H)** Coding of the phase difference ($-p/4 < \Delta\phi < 0$: blue stripes; $0 < \Delta\phi < +p/4$: red stripes; $\Delta\phi < -p/4$ or $\Delta\phi > +p/4$: green stripes = non-synchronization).

Integrative Coupling Index (ICI) described elsewhere (Müller and Lindenberger, 2011; Müller et al., 2013). To do so, we applied an analytic complex-valued Morlet wavelet transform to both the generated auditory (Figures 1C,D) and EEG time

series (Figures 1E,F), and computed the instantaneous phases for four frequencies of interest or frequency components (FC): 1.25, 2.5, 5, and 10 Hz (FC1, FC2, FC3, and FC4, respectively). The complex mother Morlet wavelet, also called

Gabor wavelet, has a Gaussian shape around its central frequency f :

$$w(t, f) = \frac{1}{\sqrt[4]{\sigma^2 \pi}} \exp\left(-\frac{t^2}{2\sigma^2} + \frac{3}{2}\pi jft\right), j = \sqrt{-1} \quad (1)$$

where σ is the standard deviation of the Gaussian envelope of the mother wavelet. The frequency resolution of the wavelet transform was fixed at 0.125 Hz, and time resolution was fixed at 4 ms.

In order to identify the phase relations between any two channels (X and Y), the instantaneous phase difference $\Delta\Phi_{XY}(t, f)$ was computed from the wavelet coefficients for all possible electrode and transformed acoustic signal pairs:

$$\Delta\Phi_{XY}^k(t, f) = \text{mod}\left(\Phi_X^k(t, f) - \Phi_Y^k(t, f), 2\pi\right) \quad (2)$$

With instantaneous phases of the two signals across k data points in the segment: $\Phi_X^k(t, f) = \arg\{z_X^k(t, f)\}$ and $\Phi_Y^k(t, f) = \arg\{z_Y^k(t, f)\}$.

Given the estimates of the phase difference between the two signals, it is possible to ascertain how long the phase difference remains stable in defined phase angle boundaries by counting the number of points that are phase-locked within a defined time window. In accordance with the procedure described by Müller and Lindenberger (2011) (cf. Müller et al., 2013), we divided the range between $-\pi/4$ and $+\pi/4$ into two ranges and distinguished between positive and negative deviations from phase zero (**Figure 1G**). We marked negative deviations in the range between $-\pi/4$ and 0 in blue (coded with “−1”) and the positive deviations in the range between 0 and $+\pi/4$ in red (coded with “+1”). Phase difference values beyond these ranges were marked in green (coded with “0”) and represent non-synchronization (**Figure 1H**). For two channels X and Y , a blue stripe in the diagram would mean that the phase of channel Y precedes the phase of channel X , while a red stripe would mean that the phase of channel X precedes the phase of channel Y . We then counted the number of phase-locked data points, for both ranges separately. Before counting, successive points in the defined range (between $-\pi/4$ and $+\pi/4$) with a time interval shorter than a period ($T_i = 1/f_i$) of the corresponding oscillation at the given frequency f_i were discarded from the analysis. This cleaning procedure effectively eliminated instances of accidental synchronization. On the basis of this counting, we obtained four synchronization indices: (1) the *Positive Coupling Index*, *PCI*, or the relative number of phase-locked points in the positive range (between 0 and $\pi/4$); (2) the *Negative Coupling Index*, *NCI*, or the relative number of phase-locked points in the negative range (between $-\pi/4$ and 0); (3) the *Absolute Coupling Index*, *ACI*, or the relative number of phase-locked points in the positive and negative range (i.e., between $-\pi/4$ and $+\pi/4$); and (4) the *Integrative Coupling Index*, *ICI*, calculated by the formula:

$$ICI = \frac{PCI + ACI}{2 \cdot ACI} \cdot \sqrt{PCI} \quad (3)$$

The *ICI* is equal to 1 when all points are phase-locked and positive; if all phase-locked points are negative or are out of range, the *ICI* will approach 0. Thus, the *ICI* measure ranges from 0 to 1 and is asymmetric ($ICI_{XY} \neq ICI_{YX}$), indicating the relative extent of the positive shift in phase difference between two signals. We restrict the description of our study results to the *ICI* measure, which is the most informative index due to its directionality. For dynamic representation of coupling within the 10-s segments, we calculated phase coupling using moving time windows of 2,000 ms width and 80 ms time delay. Overall, within a segment of 10-s duration, coupling measures across 101 time windows were collected by this shifting procedure. The Matlab code to calculate the *ICI* measure from the phase difference between the two signals can be found in the **Supplementary Material**.

Network Construction and Calculation of Strengths

For construction of the *extended* hyper-brain network, we calculated the *ICI* between all electrode pairs within and between the brains as well as within and between the guitars using the four different frequency components corresponding to the four ranges described above (low, middle, high, and whole range). In addition, we calculated the *ICI* between the brains and the guitars. Given all these couplings, we finally constructed an extended hyper-brain network comprising 88 nodes ($40 + 40 + 4 + 4$) and 7,656 edges (all possible couplings between the nodes) for each FC (1.25, 2.5, 5, and 10 Hz) and each time window. **Figure 2** shows an example of an extended hyper-brain network in the form of the connectivity matrix (**Figure 2A**) and connectivity maps (**Figure 2B**).

In order to determine the network properties, we set the cost level (ratio of the number of actual connections divided by the maximum possible number of connections in the network) to 25%, which makes it possible to investigate sparse networks with the same number of edges at different FCs and time windows. The connectivity threshold was always higher than the significance level determined by the surrogate data procedure (i.e., networks at this cost or sparsity level always included significant connections). This allowed more accurate examination of the network topology in the different duos and playing conditions. Surrogate data were created in two ways: (1) by random permutations of the time series under consideration, and (2) by phase permutation of the given time series. The phase permutation procedure involved: (a) computing the amplitude and phase spectrum of a real signal using a Fourier transformation; (b) phase shuffling, whereby the phase values of the original spectrum are used in random order and the sorted values of the surrogate sequence are replaced by the corresponding sorted values of the reference sequence; and (c) inverse Fourier transformation back to the time domain. In this way, the real and the surrogate data retain the same power spectrum but a different time course due to phase shuffling. For both surrogate data procedures, 10,000 permutations were used.

As *ICI* is a directed weighted measure, we obtained the nodes' in- and out-strengths, with the in-strength defined as the sum of weights of all incoming connections (w_{ji}), $S_{in} = \sum_{j \in N} w_{ji}$, and the

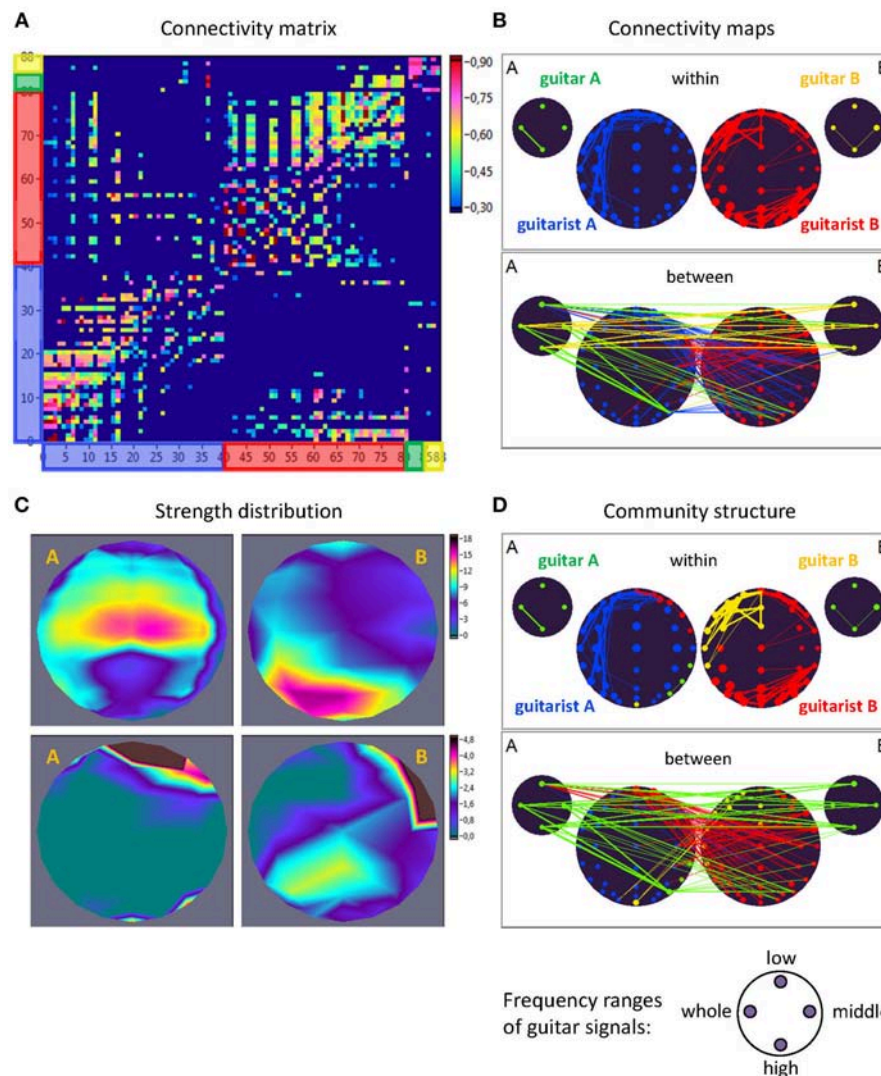


FIGURE 2 | Extended hyper-brain network. **(A)** Connectivity matrix. The network consisting of 88 nodes and 7,656 edges in total includes 40 nodes (electrodes) of guitarist A's brain, 40 nodes of guitarist B's brain, four nodes of guitar A, and 4 nodes of guitar B. The four nodes in the guitars display four different frequency ranges used during signal transformation. The connection strength in the matrix is displayed by colors ranging from dark blue (threshold connectivity values) to dark red (high connectivity values). **(B)** Connectivity maps. The upper panel represents the connectivity within the guitars and brains, and the lower panel represents the connectivity between the guitars and brains. The strength of the nodes (sum of all out-going connections) is coded by circle size, and the strength of edges is coded by the width of the line. The different parts of the network are color-coded: guitar A, green; guitar B, yellow; guitarist A's brain, blue; guitarist B's brain, red. **(C)** Topological distribution of the strengths. The upper two maps represent the topological distribution of the out-strengths within the brains of the two guitarists, and the lower ones display the topological distribution of the out-strengths going from guitarist A's brain to guitarist B's brain (left) and vice versa (right). **(D)** Community structure of the network. The connectivity maps are presented as in **(B)** but nodes and corresponding edges are color-coded according to module affiliation. Note that most of the modules are heterogeneous, comprising different guitars and brains, with exception of the blue module, which is located within the brain of guitarist A.

out-strength as the sum of weights of all outgoing connections (w_{ij}), $S_{out} = \sum_{j \in N} w_{ij}$. For representation of network dynamics, we used out-strengths (S_{out}) that were first determined for each node separately, and then grouped and summed for: (a) the out-strength going from each node of the guitar (A and B, separately) to the both brains of the guitarists, (b) the out-strength going from guitar A to guitar B and vice versa, (c) the coupling within the brains for each guitarist (A and B) separately, (d) the

coupling between the brains with the out-strength going from guitarist A's brain to guitarist B's brain and vice versa, (e) the hyper-brain network comprising electrodes or nodes from two guitarists' brains.

As shown in **Figure 2**, the out-strengths are visualized in two different ways. First, in the connectivity maps (**Figure 2B**), the strengths are coded by the size of the nodes. Second, we present the topological distribution of strength, as displayed in **Figure 2C**.

Modularity Analysis and Modular Organization of the Extended Hyper-Brain Network

To further investigate the modular organization of the networks, community structures and the modularity index (Q) were determined. For this calculation, we used the modularity optimization method for directed graphs that is implemented in the Brain Connectivity Toolbox (Rubinov and Sporns, 2010). The optimal community structure is a subdivision of the network into non-overlapping groups of nodes or communities in a way that maximizes the number of within-module edges and minimizes the number of between-module edges. Q is a statistic that quantifies the degree to which the network may be subdivided into such clearly delineated groups or modules. For directed networks, this is given by the formula (Leicht and Newman, 2008):

$$Q = \frac{1}{l} \sum_{i,j \in N} \left[a_{ij} - \frac{k_i^{\text{in}} k_j^{\text{out}}}{l} \right] \cdot \delta_{m_i, m_j}, \quad (4)$$

where $l = \sum_{i,j} a_{ij}$ is the number of edges in the graph, and a_{ij} is defined to be 1 if there is an edge from j to i , and 0 otherwise, k_i^{in} and k_i^{out} are the in- and out-degrees of the node i , and δ_{m_i, m_j} is the Kronecker delta, where $\delta_{m_i, m_j} = 1$ if $m_i = m_j$, and 0 otherwise. High modularity values indicate strong separation of the nodes into modules. For this analysis, the extended hyper-brain network including all the nodes of the instruments and brains was used. Due to the fact that within-module edges are maximized by this partition procedure, the module or network community comprises those nodes with the strongest connections, which can represent different parts of the brains or instruments. One can assume that these modules or communities must have a specific functional meaning (Müller et al., 2018b). The community structures are presented as connectivity maps, where different modules are coded by color (see **Figure 2D** for an example).

RESULTS

Here, we exemplarily present data on the coupling between the instrument sounds and the brains of two guitarists when improvising freely in duets. **Figures 3, 4** display the traces of guitars A and B (panel A), dynamic changes of coupling strengths going from guitar A (panel B) and guitar B (panel C) to the brains of both guitarists for each of the four frequency components of the guitar signals, dynamic changes of coupling strengths going from guitar A to the guitar B and vice versa (panel D), dynamic changes of coupling strengths within the brains of each of the two guitarists (panel E), dynamic changes of coupling strengths going from the guitarist A's brain to the guitarist B's brain and vice versa (panel F), and brain connectivity maps (left column) of the coupling within (upper map) and between (lower map) the brains, with extended hyper-brain community structures (middle column) as well as the topological distribution of coupling strengths (right column)

also within (upper maps) and between (lower maps) the brains of the two guitarists at the three representative time points (panel G). **Figures 3, 4** display the coupling strengths of the 10-s segment for FC1 (i.e., 1.25 Hz) in the two guitar duos, respectively. Results of coupling strengths for the other FCs (i.e., 2.5, 5, and 10 Hz) of these musical sequences of the two duos can be found in **Supplementary Figures 1–6**. The auditory representation of the 10-s segments analyzed here and the corresponding visualization of coupling strengths in real time can also be found in **Supplementary Movies 1, 2**. Please note that the soundtracks used for the coupling analyses described here were reconstructed from microphone records stored on the EEG computer with a sampling rate of 5,000 Hz. Notwithstanding this low sampling frequency, the auditory signals return the guitar tones well (**Supplementary Movies 1, 2**).

Coupling Dynamics During a 10-s Segment in Duo 1

As can be seen in **Figure 3**, during the first time period (3.0 s) the communication between the guitars was realized through the connections coming from guitar A, which also sent strong connections to the right-temporal regions of guitarist A's brain. The brain of guitarist A also communicated with both guitars, while guitarist B's brain was connected to both guitars as well as guitarist A's brain. The strengths within the brain of guitarist A were predominantly strongest at parieto-occipital and frontal regions, and at parieto-occipital and left-temporal regions within the brain of guitarist B. The strongest between-brain strengths predominantly came from right-temporal regions of guitarist B's brain. Modularity analysis of the extended hyper-brain network including both guitars and both brains revealed two modules: one module (blue) comprised guitarist A's brain (except for two left fronto-temporal nodes) and both guitars (except for the low-frequency node of the guitar B), the second module (red) correspondingly comprised guitarist B's brain together with two nodes from guitarist A's brain and one node from guitar B, mentioned above. During the next time period of 6.0 s, both guitars and both brains were synchronized with each other to some extent. Guitar A exhibited bidirectional connections with guitarist B's brain, while guitar B communicated with both brains, but especially guitarist B's. The within-brain connection strengths were asymmetric in the two brains with right-temporal (guitarist A) and left-temporal (guitarist B) location predominance, also in frontal and parieto-occipital regions. The between-brain connections were rather scarce and predominantly located in right-temporal regions in both brains. The modularity analysis also partitioned the extended hyper-brain network into two modules, with a blue module comprising guitarist A's brain (except for four right fronto-temporal nodes) and three nodes of guitar B, and a red module comprising all remaining nodes (see **Figure 3** for details). Finally, at the end of the 10-s segment (9.5 s), both brains were under the strong influence of both guitars, and they also strongly communicated with their own guitars and each other. Within the brains, the strongest strengths of guitarist A were located predominantly in the parieto-occipital regions, while

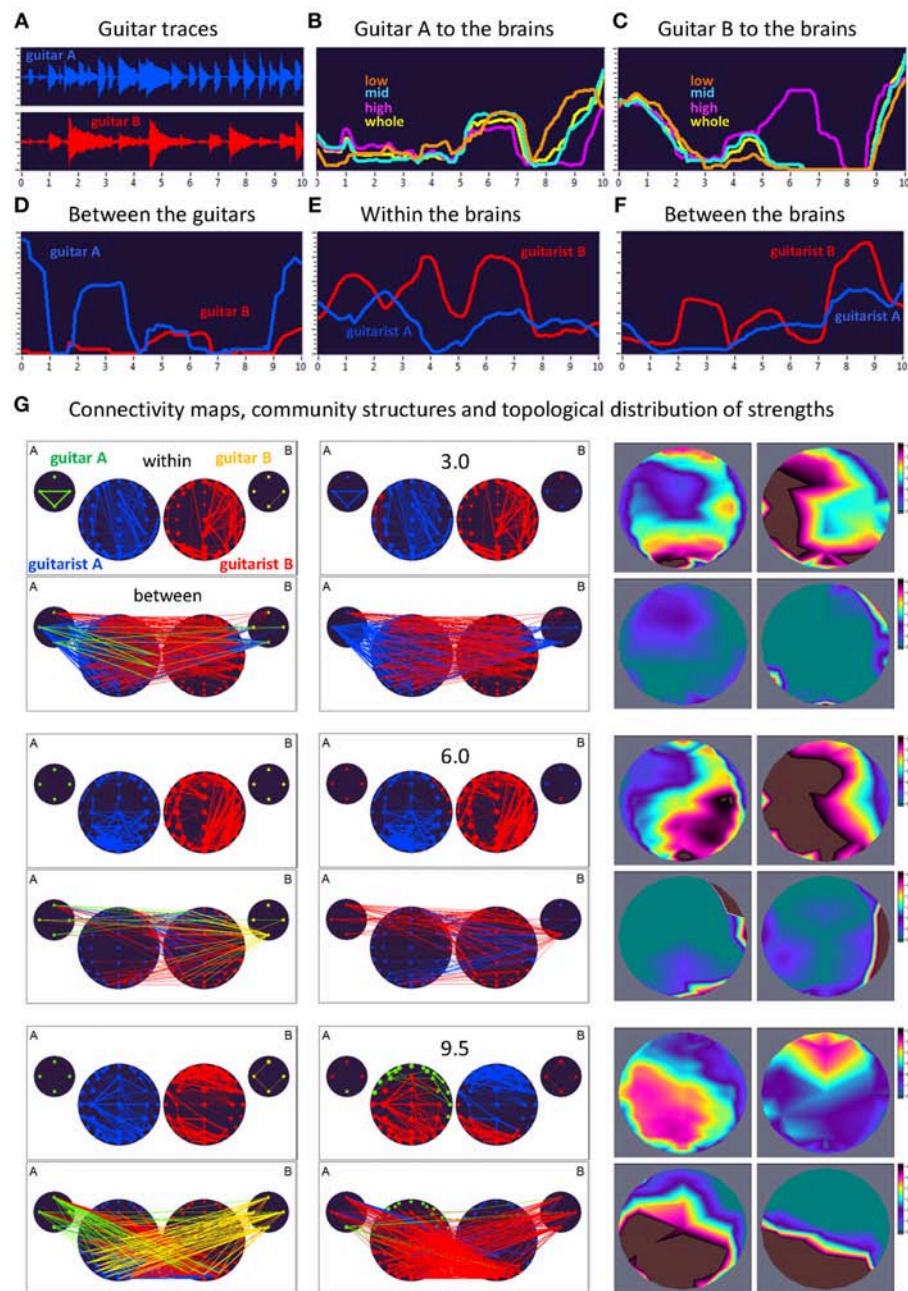


FIGURE 3 | Dynamic changes of strengths during a 10-s improvisation period for FC1 (1.25 Hz) in duo 1. **(A)** Guitar traces obtained by microphone recording: guitar A, blue; guitar B, red. **(B)** Dynamic changes of coupling strengths going from guitar A to the brains of both guitarists for each of the four frequency ranges of the guitar signal, which are indicated by color: low range, brown; middle range, cyan; high range, purple; whole range, yellow. **(C)** Dynamic changes of coupling strengths going from guitar A to the brains of both guitarists for each of the four frequency ranges of the guitar signal, which are indicated by the same colors as in **(B)**. **(D)** Dynamic changes of coupling strengths going from guitar A to the guitar B (blue) and vice versa (red). **(E)** Dynamic changes of coupling strengths within the brains of each of the two guitarists. **(F)** Dynamic changes of coupling strengths going from guitarist A's brain to guitarist B's brain (blue) and vice versa (red). **(G)** Brain connectivity maps, community structures, and topological distribution of coupling strengths. See **Figure 2** for explanations.

those of guitarist B were strongest in the fronto-central regions. The strongest between-brain connection strengths were densely located in parieto-occipital and left-temporal regions in both brains. This time modularity analysis revealed three modules.

The biggest (red) comprised three nodes from guitar A, all nodes from guitar B, central and parieto-occipital nodes from guitarist A's brain, and occipito-parietal and left-temporal nodes from guitarist B's brain. The second module (blue) is restricted to

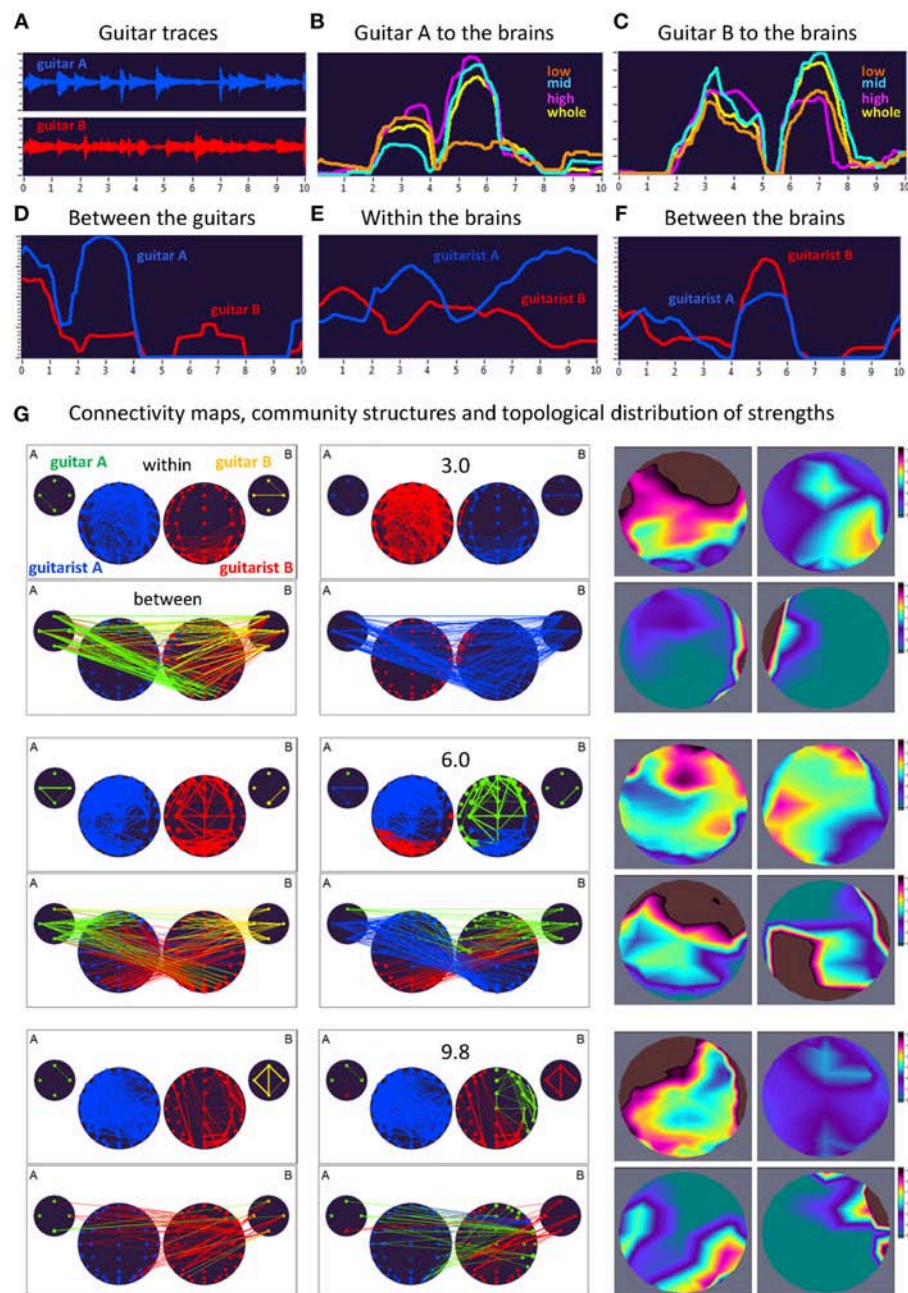


FIGURE 4 | Dynamic changes of strengths during a 10-s improvisation period for FC1 (1.25 Hz) in duo 2. **(A)** Guitar traces obtained by microphone recording: guitar A, blue; guitar B, red. **(B)** Dynamic changes of coupling strengths going from guitar A to the brains of both guitarists for each of the four frequency ranges of the guitar signal, which are indicated by color: low range, brown; middle range, cyan; high range, purple; whole range, yellow. **(C)** Dynamic changes of coupling strengths going from guitar B to the brains of both guitarists for each of the four frequency ranges of the guitar signal, which are indicated by the same colors as in (B). **(D)** Dynamic changes of coupling strengths going from guitar A to guitar B (blue) and vice versa (red). **(E)** Dynamic changes of coupling strengths within the brains of each of the two guitarists. **(F)** Dynamic changes of coupling strengths going from guitarist A's brain to guitarist B's brain (blue) and vice versa (red). **(G)** Brain connectivity maps, community structures, and topological distribution of coupling strengths. See Figure 2 for explanations.

guitarist B and comprised the remaining nodes from guitarist B's brain. The third module (green) comprised the fronto-temporal nodes from the guitarist A's brain, and also one node from guitar A.

Coupling Dynamics During a 10-s Segment in Duo 2

As can be seen in Figure 4, during the time period of 3.0 s, both guitars showed strong connections to both brains (especially

to guitarist B's brain) and to each other. These connections were mostly bidirectional going from and to the brains/guitars. The within-brain connections were strongest in guitarist A's brain, with strongest strengths predominantly located in fronto-temporal regions, while those in guitarist B's brain were located predominantly right-parietally. The between-brain strengths were strongest at the right and left temporal regions, for guitarist A and B, respectively. Modularity analysis of the extended hyper-brain network revealed two modules: one module (blue) comprised guitarist B's brain (except for three left fronto-temporal nodes) and both guitars, the second module (red) correspondingly comprised guitarist A's brain together with three nodes from guitarist B's brain, mentioned above. During the next time period of 6.0 s, both guitars also showed strong connections to both brains and to each other, with links going from guitar A to both brains and links from guitar B predominantly to guitarist B's brain. These connections were mostly unidirectional, going from guitars to brains (especially, from guitar B), indicating strong influence coming from guitars. The within-brain connections are strongest in both guitarists at the fronto-temporal regions, while the between-brain strengths are densely located at frontal regions in guitarist A's brain and at occipital and right-temporal regions in guitarist B's brain. The modularity analysis partitioned the extended hyper-brain network into three modules, with two modules (blue and red) comprising both guitarists' brains (with the blue module additionally comprising three nodes from guitar A), and a third module (green) comprising the remaining node from guitar A, all nodes from guitar B, and frontal, central as well as right-temporal nodes from guitarist B's brain (see **Figure 4** for details). Finally, at the end of the 10-s segment (9.8 s), both guitars and both brains were synchronized with each other to some extent. Within the brains, the strengths of guitarist A were strongest predominantly in the frontal and temporal regions, while those of guitarist B were considerably reduced and occupy the frontal and occipital regions. The between-brain connection strengths were strongest in the right parieto-occipital regions in guitarist A's brain and in right-temporal and frontal regions in guitarist B's brain. The modularity analysis revealed three modules: one module (blue) comprised all nodes from guitarist A's brain and one fronto-temporal node from guitarist B's brain, the second module (red) comprised the left part and some right parieto-occipital nodes of guitarist B's brain, together with one node from guitar A and all nodes from guitar B, and the third module (green) comprised the remaining part of guitarist B's brain and three remaining nodes from guitar A.

The inspection of other frequency components in both segments (see **Supplementary Figures 1–6**) showed (i) different temporal dynamic changes of network strengths for these frequency components, and (ii) that the networks within the same time periods exhibit different connectivity and community structures. All this indicates a complex interplay between different frequency components and underlying networks of the guitarists and instruments when improvising freely. Furthermore, it can be seen that dynamic changes of strengths themselves are oscillatory in nature and can be considered as second-order oscillations (Müller et al., 2018b).

Power Spectral Density (PSD) of Second-Order Oscillations

Further, we investigated the PSD of the second-order oscillations emerging through dynamic changes of coupling strengths between the instruments and brains as shown previously. To do so, we calculated the PSD across all 10 trials and 4 frequency components within the two duos for all the coupling strengths presented before. The PSD values were then averaged for the four frequency components separately. As can be seen in **Figure 5**, there are several PSD peaks in the frequency range between 0 and 1.5 Hz. Closer inspection of spectral peaks and the PSD course indicates that these peaks or frequency components can be divided into harmonics with a fixed frequency ratio (e.g., 1:2, 1:3, 1:4, etc.). This suggests that the coupling strengths between the instruments and the brains have a specific temporal structure, which apparently facilitate the free guitar improvisation and the underlying musical performance.

DISCUSSION

The primary objective of this study was to investigate the dynamic orchestration of brains and instruments during free guitar improvisation based on phase synchronization and extended hyper-brain network architecture including all the coupling types within and between the musicians' brains and guitars. The main findings are that such extended hyper-brain networks when playing guitar in duet (1) exhibit different temporal dynamic changes of network strengths, which are oscillatory in nature and have specific temporal structure for different EEG and instrument frequency components, and (2) feature different connectivity and community structures combining different brain regions and frequency components of instrument sounds, which share different, mostly heterogeneous modules (i.e., comprising different brains and instruments). It should be noted that the instrument's sound is a result of the musician's behavior, which is based on sensorimotor synchronization and action. At the same time, this sound influences the behavior of musicians through auditory sensory pathways and is in this sense an actor. In our view, music improvisation and interaction can be understood only when considering both bidirectional influences.

Previously, we have shown that both coupling strengths and community structures change their patterns depending on the oscillation frequency and musical situation (Müller et al., 2013, 2018b). We also pointed out that coupling strengths oscillate and exhibit so-called second-order oscillations, and that "the most important characteristic of the hyperbrain network organization is the existence of so-called hyperbrain modules sharing electrodes from two, three, or even four brains and characterized by strong connections or information flow within the modules and weak connections or information flow between the modules" (Müller et al., 2018b, p. 207). Here, we expand this observation to *extended* hyper-brain networks including instruments and their players' brains. It has been shown that modules or community structures can comprise one or two instruments as well as one or two brains, or different

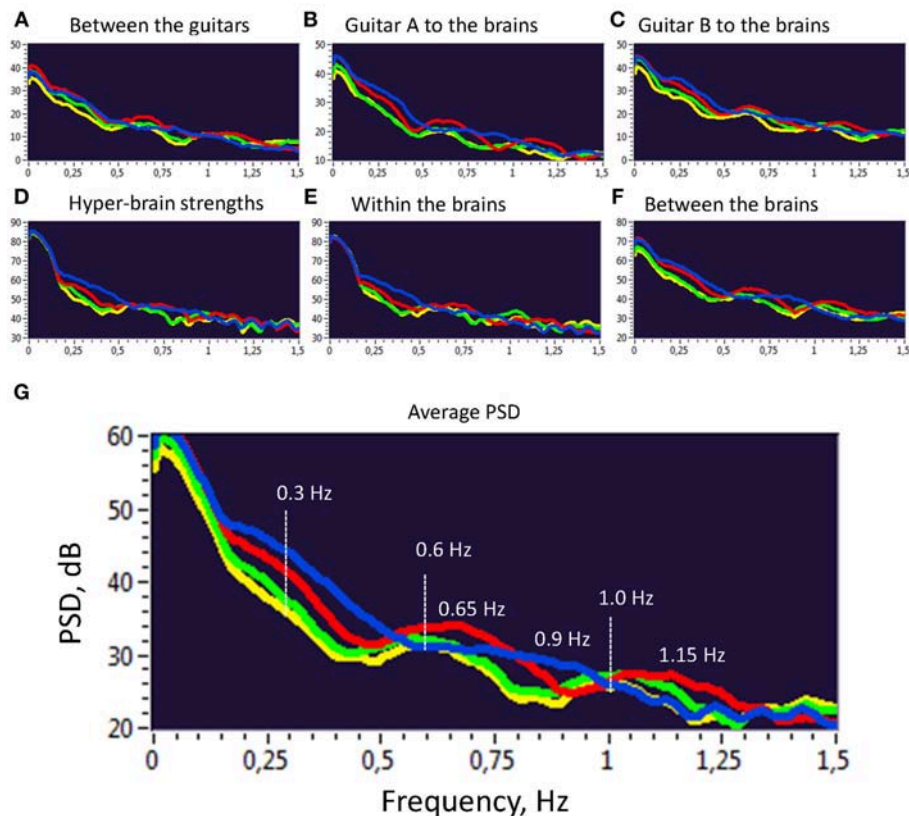


FIGURE 5 | Power Spectral Density (PSD) of second order oscillations represented by strength dynamics. **(A)** PSD of coupling strengths between two guitars. **(B)** PSD of coupling strengths going from guitar A to the brains of both guitarists. **(C)** PSD of coupling strengths going from guitar B to the brains of both guitarists. **(D)** PSD of the hyper-brain coupling strengths. **(E)** PSD of the within-brain coupling strengths. **(F)** PSD of the between-brain coupling strengths. **(G)** Average PSD of all coupling strengths presented in A-F. For representation, PSD was averaged across 10 trials and two duos for the four FCs: FC1 (1.25 Hz), blue; FC2 (2.5 Hz), red; FC3 (5 Hz), green; and FC4 (10 Hz), yellow.

parts of them. Such attuned modular organization of extended hyper-brain networks provides efficient information flow within and between the brains and instruments (through behavioral control), and apparently supports the free improvisation when playing guitar in duet. The fact that modularity structure and coupling strengths dynamically change during playing indicates that brain-brain and brain-instrument interactions are never controlled by the same brain regions but rather alter regional dominances in accordance with the current musical situation and/or expectations. Even if it can be assumed that sensorimotor brain regions play an important role in music production, it is to be expected that fronto-parietal brain regions, which have to do with touch, theory of mind or intentions of others, together with the visual, auditory, and somatosensory cortices, are important teammates in music improvisation (Zatorre et al., 2007; Keller et al., 2014). It has been suggested that adaptive timing, motor sequencing, special organization of movements as well as anticipatory and attentional processes that enable rhythmic joint action and interaction are supported by distributed networks of cortical and subcortical brain regions (Zatorre et al., 2007; Keller et al., 2014). How all these functions and underlying cortical

regions or brain structures are related to modular organization of extended hyper-brain networks must remain to be seen.

Recently, it has been reported that vocalizing patterns of singers are coupled to their respiratory and cardiac oscillations during choir singing (Müller et al., 2018a, 2019). Furthermore, the coupling between an instrument (piano assessed by MIDI tone onsets) and the brains of pianists performing a musical duet was investigated by using amplitude envelope correlations between EEG and MIDI signals (Zamm et al., 2018). In contrast, the method described here is based on the phase coupling between frequency components of amplitude variation in acoustic signals measured directly from the guitar and frequency components of raw EEG signals. In our view, this method provides more options to investigate the coupling between the instruments and the brains and also offers more information about behavioral and brain synchrony, especially when integrated into *extended* hyper-brain networks. We also used a directed coupling measure, which indicates the direction of the phase difference shift between two signals or refers to the preceding phase of one of the two signals. The preceding phase of one signal related to another can be understood two

fold: (i) one signal influences the other or (ii) anticipation is at work. Both these processes are obviously very important during music improvisation (Biasutti and Frezza, 2009; Pecenka and Keller, 2011; Badino et al., 2014). In sum, this method provides high flexibility and accuracy in investigating different synchronization patterns between different instruments and musicians' brains when playing music in groups or assemblies.

In a number of studies, it has been shown that hyperscanning as a neuroimaging technique to investigate dynamic social interaction nowadays plays a crucial role in understanding the neural and physiological mechanisms of interacting or collective behavior (Dumas et al., 2011, 2014; Sängers et al., 2011; Keller et al., 2014; Acquadro et al., 2016; Müller et al., 2018a,b, 2019). Music performance has been considered a powerful catalyzer for social interaction (Keller et al., 2014; Acquadro et al., 2016). Extended hyper-brain networks including instrument-brain coupling would provide further insights into the interplay of music performance and underlying brain-body interactions. Moreover, it has been suggested that hyperscanning has great potential for music therapy (Hunt, 2015). Fachner et al. (2019) measured dual-EEG of an experienced therapist ("Guide") and client ("Traveler") in a real music therapy session, which was combined with audiovisual recordings. They identified and quantitatively investigated therapeutically important moments of interest (MOI) and no-interest (MONI). The authors suggested that combining dual-EEG (hyperscanning) with detailed audiovisual and qualitative data can provide pivotal information for further research into music therapy (Fachner et al., 2019). There is no doubt that registration of instrument-brain coupling would further improve this interesting and therapeutically significant approach.

In conclusion, this study shows that data acquisition and analysis methods for simultaneous EEG and instrument sound recordings from multiple persons are important for discovery of extended hyper-brain synchrony during interpersonal interactions. Synchronization patterns during free guitar improvisation assessed in terms of phase alignment for instrument-instrument, brain-brain, and instrument-brain interactions seem to reflect the complex interplay of different functions and underlying temporal dynamics of interpersonal coordination. These functions may include prediction of changed group behavior and executive control during group

interaction. Future research needs to explore how these different functions assessed on the neuronal, behavioral, and group levels interact with each other and constitute mechanisms that establish and sustain interbrain oscillatory couplings in communication, voluntary action coordination, and social cognitive development. This method opens interesting perspectives for research on music interaction and may be an indispensable tool in the investigation of social interaction, music therapy, and rehabilitation dynamics.

DATA AVAILABILITY

The datasets for this study will not be made publicly available because restrictions included in the consent statement that the participants of the study signed only allow the present data to be used for research purposes within the Max Planck Institute for Human Development in Berlin.

AUTHOR CONTRIBUTIONS

VM and UL designed the study and discussed the results, wrote the article, and read and approved the final version of the manuscript. VM acquired and analyzed the data.

FUNDING

This research was supported by the Max Planck Society.

ACKNOWLEDGMENTS

We thank Thomas Holzhausen and the other guitarists for participation in the study. Technical assistance of Nadine Pecenka, Fanny Schuster, Carolin Schneidratus, Marlen Griefing, Veneta Mircheva, and Zurab Schera during data acquisition is greatly appreciated. The authors thank Julia Delius for language assistance and Dionysios Perdakis for providing the code to calculate the ICI measure.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnint.2019.00050/full#supplementary-material>

REFERENCES

- Acquadro, M. A. S., Congedo, M., and De Ridder, D. (2016). Music performance as an experimental approach to hyperscanning studies. *Front. Hum. Neurosci.* 10:242. doi: 10.3389/fnhum.2016.00242
- Badino, L., D'Ausilio, A., Glowinski, D., Camurri, A., and Fadiga, L. (2014). Sensorimotor communication in professional quartets. *Neuropsychologia* 55, 98–104. doi: 10.1016/j.neuropsychologia.2013.11.012
- Biasutti, M., and Frezza, L. (2009). Dimensions of music improvisation. *Creat. Res. J.* 21, 232–242. doi: 10.1080/10400410902861240
- Chang, A., Livingstone, S. R., Bosnyak, D. J., and Trainor, L. J. (2017). Body sway reflects leadership in joint music performance. *Proc. Natl. Acad. Sci. U.S.A.* 114, E4134–E4141. doi: 10.1073/pnas.1617657114
- D'Ausilio, A., Novembre, G., Fadiga, L., and Keller, P. E. (2015). What can music tell us about social interaction? *Trends Cogn. Sci.* 19, 111–114. doi: 10.1016/j.tics.2015.01.005
- Dumas, G., de Guzman, G. C., Tognoli, E., and Kelso, J. A. S. (2014). The human dynamic clamp as a paradigm for social interaction. *Proc. Natl. Acad. Sci. U.S.A.* 111, E3726–E3735. doi: 10.1073/pnas.1407486111
- Dumas, G., Lachat, F., Martinerie, J., Nadel, J., and George, N. (2011). From social behaviour to brain synchronization: review and perspectives in hyperscanning. *IRBM* 32, 48–53. doi: 10.1016/j.irbm.2011.01.002
- Fachner, J. C., Maidhof, C., Grocke, D., Pedersen, I. N., Trondalen, G., Tucek, G., et al. (2019). '... Telling me not to worry...' Hyperscanning and neural dynamics of emotion processing during Guided Imagery in Music. *Front. Psychol.* 10:1561. doi: 10.3389/fpsyg.2019.01561

- Hunt, A. M. (2015). Boundaries and potentials of traditional and alternative neuroscience research methods in music therapy research. *Front. Hum. Neurosci.* 9:342. doi: 10.3389/fnhum.2015.00342
- Keller, P. E., Novembre, G., and Hove, M. J. (2014). Rhythm in joint action: psychological and neurophysiological mechanisms for real-time interpersonal coordination. *Philos. Trans. R. Soc. B Biol. Sci.* 369:20130394. doi: 10.1098/rstb.2013.0394
- Leicht, E., and Newman, M. (2008). Community structure in directed networks. *Phys. Rev. Lett.* 100:118703. doi: 10.1103/PhysRevLett.100.118703
- Lindenberger, U., Li, S.-C., Gruber, W., and Müller, V. (2009). Brains swinging in concert: cortical phase synchronization while playing guitar. *BMC Neurosci.* 10:22. doi: 10.1186/1471-2202-10-22
- Müller, V., Delius, J. A. M., and Lindenberger, U. (2018a). Complex networks emerging during choir singing. *Ann. N. Y. Acad. Sci.* 1431, 85–101. doi: 10.1111/nyas.13940
- Müller, V., Delius, J. A. M., and Lindenberger, U. (2019). Hyper-frequency network topology changes during choral singing. *Front. Physiol.* 10:207. doi: 10.3389/fphys.2019.00207
- Müller, V., and Lindenberger, U. (2011). Cardiac and respiratory patterns synchronize between persons during choir singing. *PLoS ONE* 6:e24893. doi: 10.1371/journal.pone.0024893
- Müller, V., Sängler, J., and Lindenberger, U. (2013). Intra- and inter-brain synchronization during musical improvisation on the guitar. *PLoS ONE* 8:e73852. doi: 10.1371/journal.pone.0073852
- Müller, V., Sängler, J., and Lindenberger, U. (2018b). Hyperbrain network properties of guitarists playing in quartet. *Ann. N. Y. Acad. Sci.* 1423, 198–210. doi: 10.1111/nyas.13656
- Pecenkova, N., and Keller, P. E. (2011). The role of temporal prediction abilities in interpersonal sensorimotor synchronization. *Exp. Brain Res.* 211, 505–515. doi: 10.1007/s00221-011-2616-0
- Rubinov, M., and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52, 1059–1069. doi: 10.1016/j.neuroimage.2009.10.003
- Sängler, J., Lindenberger, U., and Müller, V. (2011). Interactive brains, social minds. *Commun. Integr. Biol.* 4, 655–663. doi: 10.4161/cib.17934
- Sängler, J., Müller, V., and Lindenberger, U. (2012). Intra- and interbrain synchronization and network properties when playing guitar in duets. *Front. Hum. Neurosci.* 6:312. doi: 10.3389/fnhum.2012.00312
- Sängler, J., Müller, V., and Lindenberger, U. (2013). Directionality in hyperbrain networks discriminates between leaders and followers in guitar duets. *Front. Hum. Neurosci.* 7:234. doi: 10.3389/fnhum.2013.00234
- Vigário, R. N. (1997). Extraction of ocular artefacts from EEG using independent component analysis. *Electroencephalogr. Clin. Neurophysiol.* 103, 395–404. doi: 10.1016/S0013-4694(97)00042-8
- Zamm, A., Debener, S., Bauer, A. R., Bleichner, M. G., Demos, A. P., and Palmer, C. (2018). Amplitude envelope correlations measure synchronous cortical oscillations in performing musicians. *Ann. N. Y. Acad. Sci.* 1423, 251–263. doi: 10.1111/nyas.13738
- Zatorre, R. J., Chen, J. L., and Penhune, V. B. (2007). When the brain plays music: auditory-motor interactions in music perception and production. *Nat. Rev. Neurosci.* 8, 547–558. doi: 10.1038/nrn2152

Conflict of Interest Statement: 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.

Copyright © 2019 Müller and Lindenberger. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Absolute Pitch and Musical Expertise Modulate Neuro-Electric and Behavioral Responses in an Auditory Stroop Paradigm

Vivek V. Sharma^{1,2}, Michael Thaut¹, Frank Russo^{1,3} and Claude Alain^{1,2,4*}

¹ Music and Health Research Collaboratory, University of Toronto, Toronto, ON, Canada, ² Rotman Research Institute, Baycrest Centre, Toronto, ON, Canada, ³ Department of Psychology, Ryerson University, Toronto, ON, Canada, ⁴ Department of Psychology, University of Toronto, Toronto, ON, Canada

OPEN ACCESS

Edited by:

Marc Schönwiesner,
Leipzig University, Germany

Reviewed by:

Mickael L. D. Deroche,
Concordia University, Canada
Stefanie Andrea Hutka,
University of Toronto, Canada

*Correspondence:

Claude Alain
calain@research.baycrest.org

Specialty section:

This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Neuroscience

Received: 27 May 2019

Accepted: 20 August 2019

Published: 06 September 2019

Citation:

Sharma VV, Thaut M, Russo F
and Alain C (2019) Absolute Pitch and
Musical Expertise Modulate
Neuro-Electric and Behavioral
Responses in an Auditory Stroop
Paradigm. *Front. Neurosci.* 13:932.
doi: 10.3389/fnins.2019.00932

Musicians have considerable experience naming pitch-classes with verbal (e.g., Doh, Ré, and Mi) and semiotic tags (e.g., musical notation). On the one end of the spectrum, musicians can identify the pitch of a piano tone or quality of a chord without a reference tone [i.e., absolute pitch (AP) or relative pitch], which suggests strong associations between the perceived pitch information and verbal labels. Here, we examined the strength of this association using auditory versions of the Stroop task while neuro-electric brain activity was measured using high-density electroencephalography. In separate blocks of trials, participants were presented with congruent or incongruent auditory words from English language (standard auditory Stroop), Romanic solemnization, or German key lexicons (the latter two versions require some knowledge of music notation). We hypothesized that musically trained groups would show greater Stroop interference effects when presented with incongruent musical notations than non-musicians. Analyses of behavioral data revealed small or even non-existent congruency effects in musicians for solfège and keycodes versions of the Stroop task. This finding was unexpected and appears inconsistent with the hypothesis that musical training and AP are associated with high strength response level associations between a perceived pitch and verbal label. The analyses of event-related potentials revealed three temporally distinct modulations associated with conflict processing. All three modulations were larger in the auditory word Stroop than in the other two versions of the Stroop task. Only AP musicians showed significant congruity effects around 450 and 750 ms post-stimulus when stimuli were presented as Germanic keycodes (i.e., C or G). This finding suggests that AP possessors may process alpha-numeric encodings as word forms with a semantic value, unlike their RP possessing counterparts and non-musically trained individuals. However, the strength of musical conditional associations may not exceed that of standard language in speech.

Keywords: music, EEG, conflict resolution, event-related potentials, absolute pitch

INTRODUCTION

Musicians have considerable experience naming pitch-classes with verbal and semiotic tags. A small subset of musicians possesses linguistic tags for decoding pitch-classes into verbal codes or symbols without having to resort to an external reference pitch. This ability to name pitches without a referent is called absolute pitch (AP). The phenomenon of AP is of particular interest because it links measurable sensory inputs to neural representations (Zattore, 2003). AP also demonstrates highly stable category judgments using verbal codes that the general population does not seem to have for isolated sound objects (Levitin and Rogers, 2005). Most musicians do not have AP and instead require an external pitch-class to be presented and maintained in working memory as a reference to compare a subsequent pitch-class in order to label it. This more common ability is referred to as relative pitch (RP), which is the ability to label the intervals between pitches. Thus, while AP musicians can name pitch categories without any external referent, RP musicians can name pitch categories once a reference pitch is provided.

Absolute pitch is thought to be “automatic” whereas almost no RP users report automatic responses when identifying a pitch-class on the basis of knowledge of a preceding one, possibly relying on focused attention and working memory (Zatorre et al., 1998). AP possessors can accurately name a pitch-class within less than a second of hearing it (Miyazaki, 1990). Evidence also shows that AP is likely acquired during a critical period of 5–7 years of age (Russo et al., 2003; Miyazaki and Ogawa, 2006). There is also a negative correlation between mean response time (RT) and accuracy across participants in AP groups (Takeuchi and Hulse, 1993). Musicians with RP are faster at naming the intervals between two pitch-class categories, most likely because AP possessors are thought to name individual elements first then extrapolate distance, adding processing stages to their interval judgments (Miyazaki, 1992). In contrast, RP can be developed by most individuals at any age and can become quite rapid and accurate with training and rehearsal (Miyazaki, 1989). Musicians with RP cannot associate abstract entities such as verbal codes to individual pitches consistently over a long-term but only to intervals or chord structures. AP possessors are quicker and most accurate when responding to piano-tones, the instrument of acquisition for most musicians with AP – some can only label piano-tones (Miyazaki, 1989). However, AP possessors have more difficulties naming individual pitch when presented with “black-key” pitch classes or simple tones (Takeuchi and Hulse, 1993). Interestingly, musicians with AP may also be susceptible to imperceptibility of “key signature” – the macroscopic organizational schema for all pitches in a piece of music, where a single pitch is perceived as the 1st element of the musical scale sequence (Terhardt and Seewann, 1983). Baharloo et al. (1998) delineated four categories of AP, where certain categories had high accuracy and low RTs to sinetone tests, while other categories did not, which informed the level of automaticity expected in individual AP possessors. Currently, it is thought that non-AP requires working memory to keep reference tones and schema of tonal intervals in echoic memory

and consciousness while AP rely more heavily on categorization networks during pitch naming (Schulze et al., 2009). Thus, RP uses working memory and conscious control while AP seems to be automatic and working memory free. In this sense, much of RP ability is schematic while AP may employ more stimulus driven responses.

In the present study, we used auditory versions of the Stroop task to explore how pitch may be internally represented in musicians with RP and AP. In the original version of the Stroop task, participants are presented with a colored word, which could either be congruent (e.g., the word RED in red ink), incongruent (e.g., the word RED in blue ink), or neutral (e.g., word printed in black or a series of the letter “X” printed in colors). Participants are slower and make more errors for incongruent than congruent or neutral stimuli (MacLeod, 1991). The Stroop task has been used to study conflict resolution in children (Prevor and Diamond, 2005) and older adults (West and Alain, 2000) as well as in neurological (Rafal and Henik, 1994) and psychiatric (McNeely et al., 2003) populations. In the auditory versions of the Stroop task, participants are presented with spoken words that could either be congruent (e.g., the word Low in low pitch) or incongruent (e.g., the word Low in high pitch) (Hamers and Lambert, 1972). As with the visual Stroop task, participants are slower and make more errors for incongruent than congruent stimuli (Itoh et al., 2005; Donohue et al., 2012). If musicians process musical annotations similarly to how the general population process words, then we should observe interference effects when they are presented with incongruent musical annotations. However, it is unclear how musicians might process incongruence between semantic meaning and linguistic information during category recognition of musical pitch-classes and notations. Musical experience may allow for behavioral flexibility in an auditory Stroop paradigm involving pitch information. For example, musicians with AP may attenuate or bypass conflicting linguistic information during pitch perception, biasing their attention toward their own internal words. Thus, developing associative bonds between language and pitch may generate reflexive, rapid, and accurate encoding of sensory features irrespective of conflicting verbal codes. That is, audible musical percepts with strongly associated codewords may be less susceptible to interference than associated verbal codes outside of the domain of music. This type of transfer effect has been discussed in the context of musical expertise by some researchers, who suggest that scalp recording of event-related potentials (ERPs) may be effective to disentangle neural time-course events to study these effects (Besson et al., 2011).

Earlier ERP studies of the Stroop effect have identified several neural correlates thought to play an important role in information conflict resolution. For example, Duncan-Johnson and Kopell (1981) found that the P300 response at midline scalp locations (i.e., Fz, Cz, and Pz) was not significantly modulated by semantic incongruence or neutrality of color and word and yet the classic Stroop stimuli did indeed elicit a typical Stroop effect, which suggested that based on their resulting RTs, the neural signature of Stroop interference should more likely rely on activity later than the P300 response. This was also supported by other investigators (Ilan and Polich, 1999; Atkinson et al., 2003).

Rebai et al. (1997) demonstrated that a late negative N400 wave is evoked when participants are asked to internally verbalize responses of incongruent color-word conditions from recordings of the Fz, Cz, and Oz plus left and right parietal electrodes. West and Alain (1999) noted four spatially and temporally distinct modulations during the visual Stroop task, which included a phasic polar positivity that peaked at about 500 ms after stimulus onset over lateral fronto-polar scalp areas, a fronto-central slow wave that began at about 500 ms after the stimuli and persisted over the remainder of trial, a left-parietal modulation peaking at 522 ms and a left temporo-parietal positivity beginning at about 600 ms after stimulus onset. Ergen et al. (2014) showed that the N450 wave was larger for incongruent stimuli in comparison to congruent stimuli and maximal over frontal electrodes. These investigators also found a slow late potential at around 600 ms after stimuli onset over the parietal electrodes of their extended 10/20 system (Ergen et al., 2014).

In the present study, non-musicians (NM) and musicians with AP or RP completed three different versions of an auditory Stroop task while we measured neuro-electric brain activity using high-density electroencephalography. In the auditory word Stroop task, participants were presented with sung English words (i.e., low or high) in one of two pitches corresponding semantically to the words. They were told to ignore the word and press a button indicating the pitch of the stimulus. In the solfège version of the auditory Stroop task, participants were instructed to do the same as the word Stroop task. However, they were presented with the English solemnizations corresponding to the two pitches (i.e., Doh or Soh). Finally, in the keycodes task, participants were presented with phonemic equivalents of key notation that corresponded to the two pitches (i.e., “C” and “G”) and instructed to do the same as the other two tasks. For the word Stroop task, we anticipated that musicians would be faster and more accurate than NMs, but that the Stroop interference effects should be similar in all three groups. For the solfège and keycodes versions of the Stroop task, we predicted greater Stroop interference in musicians with RP and AP due to an over-learned semantic response and prolonged exposure to congruencies from pitch naming training. Prior research using auditory word Stroop (Itoh et al., 2005; Donohue et al., 2012) found analogous neuro-electric modulations to those found in the visual domain in terms of increased negativity for incongruent compared to congruent conditions over midline central scalp area between 400 and 600 ms after stimulus onset (West and Alain, 1999). We anticipate a similar modulation in NMs but a possible attenuation of such responses in musically trained individuals for pitch naming-based Stroop effects that use non-musical standard English words. Finally, we hypothesize that musicians with RP and AP would show enhanced negativity for incongruent trials in the solfège and keycodes tasks because of strong internal pitch representations and prolonged exposure to pitch naming with these lexicons. Understanding the strength of pitch information to verbal code mappings may demonstrate that creating abstract representations of perceptual features may allow confirmation of what is heard with greater certainty. This may help explain how words tag onto object features in order for categories to emerge from the environment.

MATERIALS AND METHODS

Participants

Sixty participants were recruited for this study. One NM was excluded, one musician with AP and one musician with RP were excluded because of excessive eye movements during EEG recording. The final sample comprised 19 NM (mean age = 24.9 years, range = 20–39, $SD = 4.3$, five males) with <4 years exposure to any music training. It also included 19 musicians who possess AP (mean age = 25.1 years, range = 19–41, $SD = 5.9$, five males) and 19 who possess RP (mean age = 26.4 years, range = 21–39, $SD = 4.0$, 5 males) all of whom reported >7 years of formal musical training. There were no significant differences between the groups for age and years of education.

Screening Test

Stimuli and Apparatus

Stimuli consisted of 24 pure tones (44.1 kHz sampling rate). They were generated using the digital audio editing software, ProTools LE (Version 7.3.1, Daly City, CA, United States). All stimuli were 1000-ms in duration, including 5-ms onset and 30-ms linear offset ramps. Pitch-classes were equally spaced, ranging from C3 to B5 inclusively and tuned to standard A4 = 440 Hz piano tunings in equal temperament. The stimuli were presented binaurally in an order that was individually randomized for each participant.

A computer interface presented the stimuli and collected response data from the participants using customized software in MATLAB by Mathworks (Cambridge, MA, United States, 2015b). Participants were tested in a quiet room in front of the computer interface, which was equipped with a mouse that allowed the participants to classify the tones using 12 circular response regions. Each of these regions displayed a complete set of chromatic pitch-class labels for AP and diatonic for RP, including enharmonic equivalents. The response-regions were placed in a circular shape, forming a clock of pitch-class labels. A small circle was displayed around the center of this wheel for the participants to click within before hearing a tone. Category response regions were positioned such that each response was equidistant from the pointing cursor at the onset of every tone.

For RP musicians, the stimuli and apparatus were the same, except the pitch C4 was always presented preceding the pitch meant to be named. Because RP is considered to be more effortful and less acute than AP, black-key pitches were not included nor were their pitch-class labels presented on screen, which decreased the response set to ease working memory load. The RP test consisted of naming 12 stimuli ranging based on their distance from C4.

Screening Procedure

We used the screening procedure developed by Bermudez and Zatorre (2009). The screen tests employed the same computer and digital audio system as the experiment. For the purposes of this study, the criterion for AP was defined as the ability to score $\geq 85\%$ or an absolute mean deviation of ≤ 1 under these screening test conditions. RP was defined as the ability to score $\geq 80\%$ and

participants with mean RT that were three standard deviations from the group mean were eliminated. A set of structured questionnaires and follow-up interviews were then administered after testing to index musical experience and health-related data.

Experimental Test

Stimuli

The auditory stimuli were seven sung monosyllabic phonemes: /low/, /high/, /do/, /so/, /see/, /jee/, and /baw/. A professional vocalist sung the phonemes, which were recorded at a sample rate of 44,100 Hz using a Shure SM58 microphone and ProTools LE (Version 7.3.1, Daly City, CA, United States). The fundamental frequency of each tone was transformed digitally to be tuned to 261.63 Hz for the low pitch and 392 Hz for the high pitch using Adobe Audition 1.5 (San Jose, 2004). Sinusoid tones were generated with the same frequencies as the sung speech tones (i.e., C4 = 261.63 Hz and G4 = 392 Hz). The /baw/ syllable was recorded in both pitches and used as a neutral phoneme for all lexicons. Average root mean square power of the tones was then normalized ($M = -16.13$ dBs, $SD = 1.72$ dBs). The stimuli were edited to have a 500-ms duration including 5-ms rise/fall time. All phonemes were standardized to the Carnegie Mellon University Pronouncing Dictionary. Stimuli were presented binaurally at an intensity of 75 decibels sound pressure level (SPL) through insert earphones model ER-3A by Etymotic Research (Elk Grove Village, 1985).

Procedure

Participants sat in a comfortable chair in an acoustically and electrically shielded room. The stimuli were presented using the computer software Presentation 16.5 by Neurobehavioral Systems Inc. Trials consisted of 500-ms fixation cross, followed by 500-ms auditory stimulation. The fixation cross was maintained on screen throughout the stimulus, and for 500-ms after a response was made, which initiated the next trial. The inter-trial interval was 1000-ms. The fixation cross was a white cross with a 58 point font size and located in the center of a black screen. Each block of trials comprised 100 congruent, 100 incongruent, and 67 neutral trials, which were presented in completely random order within each block. Participants completed three blocks of trials with breaks offered between blocks. They responded using a computer keyboard and were instructed to press the left arrow key for “Low” pitch and the right arrow key for “High” pitch, regardless of sung word. The time between the onset of a tone and a key press was recorded as the RT for that trial. RTs that were greater than 1000-ms or less than 200-ms were excluded from further analyses.

EEG Acquisition and Preprocessing

Electroencephalogram (EEG) was recorded from 66 scalp electrodes using a BioSemi Active Two acquisition system (BioSemi V.O.F., Amsterdam, Netherlands). The electrode montage was set according to the BioSemi electrode cap based on the 10/20 system and included a common mode sense active electrode and driven right leg passive electrode serving as ground. Ten additional electrodes were placed below the hair line (both mastoid, both pre-auricular points, outer canthus of

each eye, inferior orbit of each eye, two facial electrodes) to monitor eye movements and to cover the whole scalp evenly. The latter is important because we used an average reference (i.e., the average of all scalp EEG channels as the reference for each EEG channel) for ERP analyses. Neuro-electric activity was digitized continuously at a rate of 512 Hz with a bandpass of DC-100 Hz, and stored for offline analysis. All off-line analyses were performed using Brain Electrical Source Analysis software (BESA, version 6.1; MEGIS GmbH, Gräfenberg, Germany).

The continuous EEG data were first digitally filtered with 0.3 Hz high-pass (forward, 6 dB/octave) and 40 Hz low-pass filters (zero phase, 24 dB/octave). For each participant, a set of ocular movements was identified from the continuous EEG recording and then used to generate spatial components that best account for eye movements. The spatial topographies were then subtracted from the continuous EEG to correct for lateral and vertical eye movements as well as for eye-blinks. The analysis epoch consisted of 200 ms of pre-stimulus activity and 1000 ms of post-stimulus activity time-locked to the stimulus onset. After correcting for eye movements, data for each participant were then scanned for artifacts; epochs including deflections exceeding 120 μ V were marked and excluded from the analysis. The remaining epochs were averaged according to electrode position, stimulus type (i.e., neutral, congruent, incongruent), experimental condition (i.e., auditory word Stroop, Solfège, and Keycodes), and response accuracy (correct vs. incorrect response). Each average was baseline-corrected with respect to the pre-stimulus interval (i.e., mean amplitude over the 200 ms prior to stimulus onset).

The ERP analyses involved only correct trials and focused on mean amplitude for three pre-defined time windows and electrode clusters motivated by prior auditory (i.e., Itoh et al., 2005; Donohue et al., 2012) and visual (West and Alain, 1999) Stroop studies. The first window (150–250 ms) comprised of C1, Cz, C2, FC1, FCz, and FC2 electrodes over the fronto-central scalp area to measure the P200 response. The N450 (325–525 ms) was quantified using the same electrodes. A late positive component (LPC) modulation (600–900 ms) was quantified over the P1, P3, Pz, P2, P4, PO3, POz, and PO4 electrode sites.

RESULTS

In a preliminary analysis we examined whether RTs and error rates differed for the two neutral stimuli (i.e., phonemes and sinetones). For RTs, the ANOVA with Neutral Type and Task as within subject factors and Group as between-subject factors did not reveal significant differences in RT as a function of Neutral Type [$F(1,54) = 3.16$, $p = 0.081$], nor did this factor interact with Musicianship [$F(2,54) = 2.23$, $p = 0.118$] or Task ($F < 1$). The three way interaction between Neutral Type, Task, and Musicianship was not significant [$F(4,108) = 1.04$, $p = 0.391$]. For error rates, the ANOVA with Neutral Type and Task as within-subject factors and Group as between-subject factors did not reveal significant differences in error as a function of Neutral Type [$F(1,54) = 2.96$, $p = 0.091$], nor did this factor interact with Musicianship [$F(2,54) = 1.18$, $p = 0.313$] or Task ($F < 1$).

The three way interaction between Neutral Type, Task, and Musicianship was not significant [$F(4,108) = 1.18, p = 0.322$]. Thus, the RTs and error rates for neutral phonemes and sinetones were averaged together into a single neutral condition.

Musicianship and Congruency Effect on Response Time

The group mean RTs for each task and condition are shown in **Figure 1**. A mixed design ANOVA on RTs with Musicianship (NM, RP, and AP) as a between-subject factor and Task (Words, Solfège, and Keycodes) and Congruence (Congruent, Incongruent, and Neutral) as within-subject factors and *post hoc* comparisons used the Bonferroni procedure. The ANOVA yielded a three-way interaction between Musicianship, Task, and Condition [$F(8,216) = 2.81, p < 0.01, \eta_p^2 = 0.094$]. To better understand this interaction, we examined the effect of musicianship and condition in each task separately.

For the Words task, we anticipated that musicians would be faster than non-musicians given the mounting evidence that musical training enhanced executive function and cognitive control (Moussard et al., 2016; Alain et al., 2019). The ANOVA on RTs revealed a significant group \times stimulus type interaction [$F(4,108) = 5.56, p < 0.001, \eta_p^2 = 0.171$], implying differentials of interference effects between the AP, RP, and NM groups. Nonetheless, it should be noted that the main effect of musicianship was significant [$F(2,54) = 20.97, p < 0.001, \eta_p^2 = 0.437$]. *Post hoc* comparisons show that there was no difference between AP and RP ($p = 1$) though both groups were significantly faster than NMs ($p < 0.001$ for both differences) due to an overall RT advantage for this task, despite congruency effects. As for the interaction between congruence and musicianship, NMs were significantly slower for incongruent trials, intermediate for neutral, and fastest for congruent trials (all pairwise comparisons were significant, $p < 0.01$ in all cases). For musicians with AP, the congruent condition was significantly faster than incongruent ($p < 0.001$) but not neutral ($p = 1$), and the RTs to neutral words were significantly faster than those generated for incongruent stimuli ($p < 0.01$). Finally, musicians with RP showed slower RTs to incongruent than both congruent ($p < 0.001$) and neutral ($p < 0.001$) stimuli, but no difference was seen between congruent and neutral.

For the Solfège task, we expected NMs to show little interference since they do not associate verbal codes with pitch representation using this lexicon, while musicians should show interference, particularly AP who have learned to associate pitch representations with labels automatically. Surprisingly, the interaction between musicianship and congruency was not significant. This suggests comparable interference effect in AP, RP, and NM groups. Nonetheless, the main effect of musicianship was significant [$F(2,54) = 21.05, p < 0.001, \eta_p^2 = 0.438$], with AP and RP musicians not being significantly different in RT and both musicians groups being faster than NMs ($p < 0.001$ for both differences). There was also a main effect of congruence

[$F(2,54) = 6.68, p < 0.01, \eta_p^2 = 0.110$]. Pairwise comparison shows that the congruent condition was significantly faster than both incongruent ($p < 0.001$) and neutral ($p = 0.029$), while no difference was found between incongruent and neutral.

For the Keycodes task, we expected NMs to also show little to no interference unlike musicians. The interaction between congruence and musicianship was significant [$F(4,108) = 2.97, p = 0.023, \eta_p^2 = 0.099$]. Again, the main effect of musicianship was significant [$F(2,54) = 21.04, p = 0.001, \eta_p^2 = 0.438$] in a similar manner as the other two tasks. The main effect of congruence was not significant. Surprisingly, the NM group was significantly faster for the incongruent condition when compared to the congruent condition ($p = 0.012$).

Musicianship and Congruency Effect on Error Rates

The group mean error rates for each task and condition are shown in **Figure 2**. A mixed design ANOVA with the same factors as the RTs was performed for error rates. Similar to RTs, the three-way interaction between Musicianship, Task, and Congruency was significant [$F(8,216) = 9.16, p < 0.001, \eta_p^2 = 0.253$]. To better understand this interaction, we examined the effect of musicianship and condition in each task separately.

For English words, the interaction between musicianship and condition was significant [$F(4,108) = 11.31, p < 0.001, \eta_p^2 = 0.295$]. NMs were more accurate for congruent than incongruent or neutral stimuli ($p < 0.001$, in both cases). They were also more accurate for neutral than incongruent stimuli ($p < 0.001$). Musicians with RP were more accurate for congruent than incongruent stimuli ($p < 0.01$). They were also more accurate for neutral than incongruent stimuli ($p < 0.01$). There was no difference between the neutral and congruent words condition. Musicians with AP did not show significant differences in accuracy for congruent, incongruent, and neutral stimuli for this task.

For the Solfège task, there was a significant interaction between musicianship and condition [$F(4,108) = 6.36, p < 0.001, \eta_p^2 = 0.191$]. NMs were more accurate for congruent than incongruent or neutral stimuli ($p < 0.001$ in both cases). They were also more accurate for neutral than incongruent stimuli ($p = 0.024$). Musicians with RP and AP showed similar error rates in this task for congruent, incongruent, and neutral stimuli with no significant differences.

For the keycodes task, there was a significant musicianship by condition interaction [$F(4,108) = 12.89, p < 0.001, \eta_p^2 = 0.323$]. NMs were more accurate for incongruent than congruent ($p < 0.01$). They were also more accurate for incongruent and congruent than neutral stimuli ($p < 0.001$ in both cases). Musicians with RP and AP showed comparable error rates in the keycodes task with no significant accuracy differences between congruent, incongruent, and neutral stimuli.

Effects of Musicianship on Neural Correlates of Stroop Effect

Figure 3 shows group mean ERPs elicited as a function of task and congruency. In all three groups, stimuli generated an N1 and

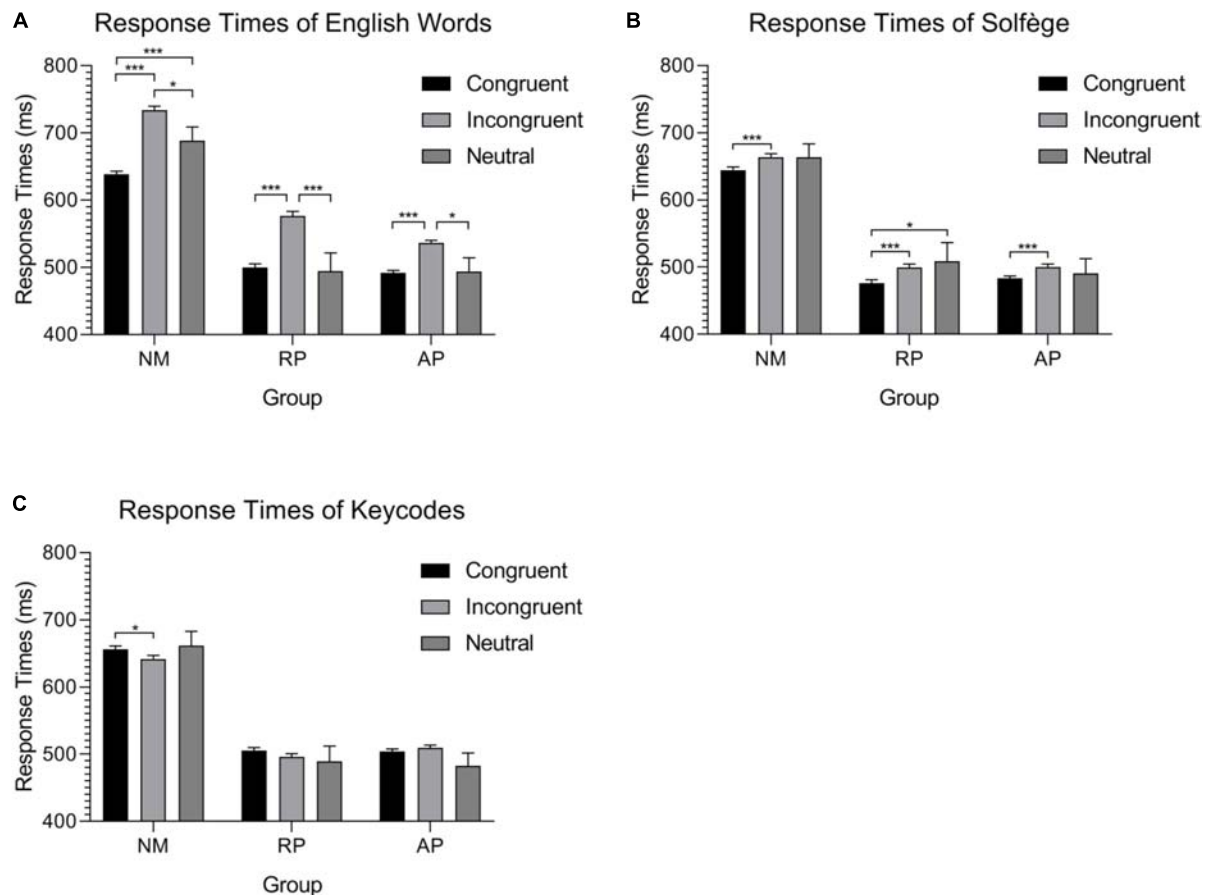


FIGURE 1 | Group mean response time during the three auditory Stroop tasks. **(A)** Musicians and non-musicians process neutral stimuli differently during the English words task. **(B)** Significant differences between congruent and incongruent stimuli in the solfège task show a possible facilitation effect of the lexicon but still demonstrate interference. **(C)** A possible reverse behavioral Stroop effect for alpha-numeric codes is found for isolated alpha-numeric codes from color-words in NMs, though the effect is marginal. No other differences are found for this task. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

P2 deflection at fronto-central sites, which inverted in polarity at inferior parietal sites and mastoids, consistent with generators in superior temporal gyrus. These sensory evoked responses were followed by a broad negativity that peaked at about 525 ms after sound onset at central scalp sites. This negativity was followed by a positive wave at parietal sites (Figure 4). The analyses of ERP amplitude focus on congruent and incongruent trials. The ERPs generated by the neutral trials were excluded because the neutral stimuli had distinctive acoustic features making it difficult to dissociate congruency effects from acoustically driven factors. Three different spatio-temporal clusters were examined: Central P200, Central N450, and Parietal LPC (Figures 5–7).

Central P200

Figures 5–7 show the group mean amplitude as a function of task and congruency. The three-way interaction between musicianship, task, and congruency was significant [$F(4,108) = 3.00$, $p = 0.045$, $\eta_p^2 = 0.085$]. To further elucidate this interaction, the tasks were analyzed separately in a similar manner as the behavioral results. The main effect of musicianship was not significant.

For English words, the interaction between musicianship and group was not significant. However, there was a main effect of congruency [$F(1,54) = 42.93$, $p < 0.001$, $\eta_p^2 = 0.443$], where the congruent condition was significantly more positive than the incongruent condition ($p = 0.017$ for NM, $p < 0.001$ for RP and AP). For the solfège and keycodes tasks, there were no significant main effects or interactions between group and congruence.

Central N450

Figures 5–7 display the differences in amplitudes for the N450 modulation. The three-way interaction between group, task, and congruency trended toward significance [$F(4,108) = 2.12$, $p = 0.084$], whereas the two-way interaction between task and congruency was significant [$F(1,54) = 8.59$, $p < 0.001$, $\eta_p^2 = 0.141$]. To further elucidate the three-way interaction, the tasks were analyzed separately. The ANOVA also yielded a main effect of congruency [$F(1,54) = 20.29$, $p < 0.001$, $\eta_p^2 = 0.273$], with greater negativity in ERPs for incongruent than congruent stimuli ($p < 0.001$) across the tasks. A significant main effect of musicianship was also found [$F(2,54) = 5.18$, $p < 0.01$, $\eta_p^2 = 0.161$]. Pairwise comparisons showed larger ERP amplitude

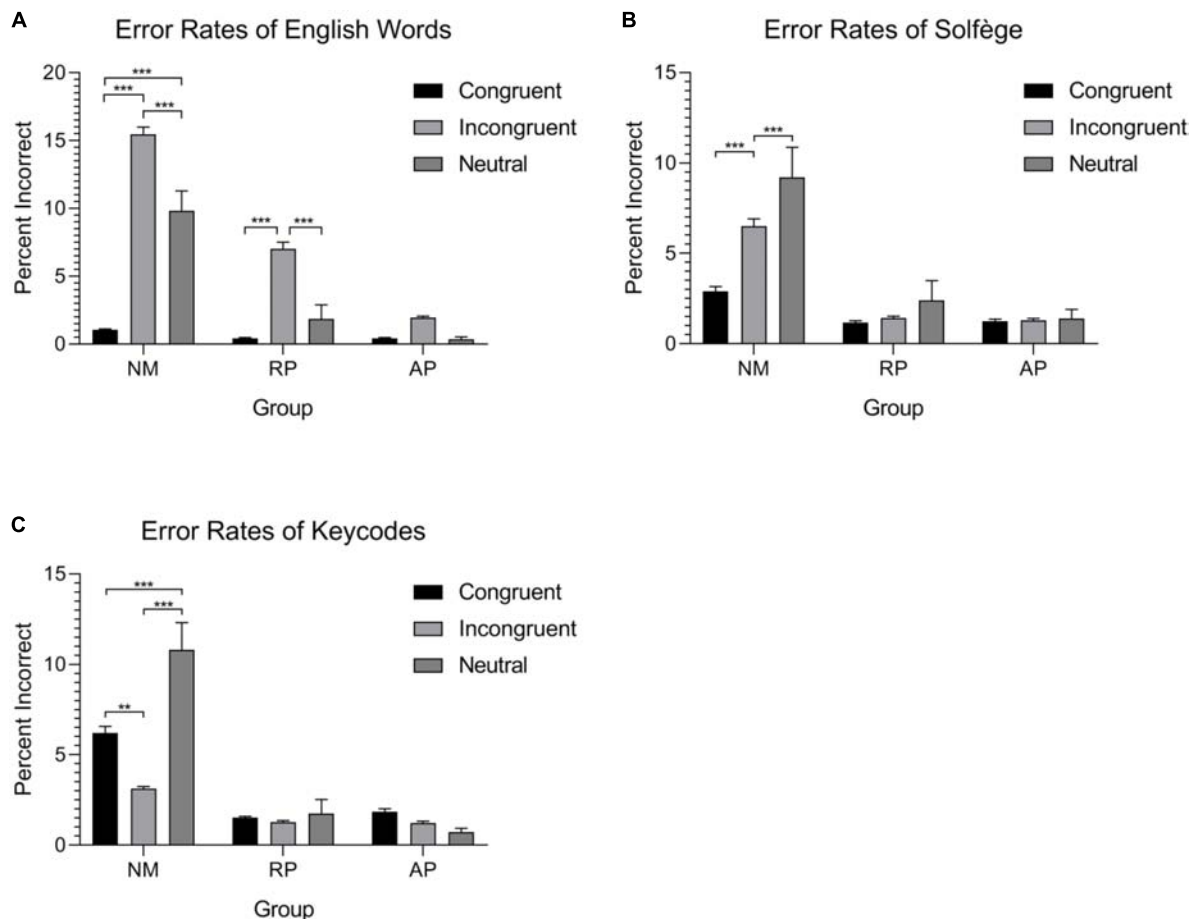


FIGURE 2 | Error-rates for the three tasks and three groups. Both musician groups have significantly lower errors overall in comparison to the non-musician group. **(A)** AP possessors show very little errors for standard word form Stroop task. **(B)** Musicians show no differences in errors for the solfège lexicon. **(C)** Only non-musicians show considerable error rate differences within the conditions and a reverse Stroop effect between congruent and incongruent stimuli. ** $p < 0.01$ and *** $p < 0.001$.

in NMs compared to AP ($p = 0.048$) and RP ($p = 0.010$) across the tasks and congruencies.

For the English word Stroop task, the group by congruency interaction was not significant [$F(2,54) = 0.371$, $p = 0.692$, $\eta_p^2 = 0.014$]. In all three groups, the incongruent stimuli generated significantly greater negative ERP amplitude than the congruent stimuli [$F(1,54) = 22.77$, $p < 0.001$, $\eta_p^2 = 0.297$]. The main effect of musicianship trended toward significance [$F(2,54) = 3.11$, $p = 0.053$, $\eta_p^2 = 0.103$].

For solfège, the group by congruency interaction was not significant [$F(2,54) = 1.60$, $p = 0.212$, $\eta_p^2 = 0.056$]. However, there was a main effect of musicianship [$F(2,54) = 6.85$, $p < 0.01$, $\eta_p^2 = 0.202$]. Pairwise comparisons revealed larger ERP amplitude in NMs than musicians with RP ($p < 0.01$) or AP ($p = 0.021$). The ERP amplitude was comparable for the RP and AP groups and not significantly different.

In the keycodes task, there was a significant interaction between musicianship and congruency [$F(2,54) = 3.67$, $p = 0.032$, $\eta_p^2 = 0.120$]. Musicians with AP showed a significant congruency effect (i.e., N450, $p < 0.01$),

while no differences in ERP amplitude were found between congruent and incongruent stimuli in NM and musicians with RP.

Parietal LPC (600–900)

Figures 5–7 show the differences in ERP waveform amplitudes for the parietal LPC. The three-way interaction between musicianship, task, and congruency was significant [$F(4,108) = 2.67$, $p = 0.036$, $\eta_p^2 = 0.090$]. Once again, to further elucidate this interaction, the tasks were analyzed separately. There was also a main effect of task [$F(2,108) = 2.67$, $p = 0.036$, $\eta_p^2 = 0.090$] and a main effect of musicianship [$F(2,54) = 6.15$, $p < 0.01$, $\eta_p^2 = 0.186$].

For the English words, the group by congruency interaction was not significant [$F(2,54) = 1.64$, $p = 0.204$, $\eta_p^2 = 0.057$]. Overall, the ERP amplitude was smaller for congruent than incongruent stimuli [$F(1,54) = 45.53$, $p < 0.001$, $\eta_p^2 = 0.457$]. The main effect of musicianship was significant [$F(2,54) = 7.36$, $p < 0.01$, $\eta_p^2 = 0.214$]. Pairwise comparisons reveal that the musicians with AP generated more negative amplitude than NM group

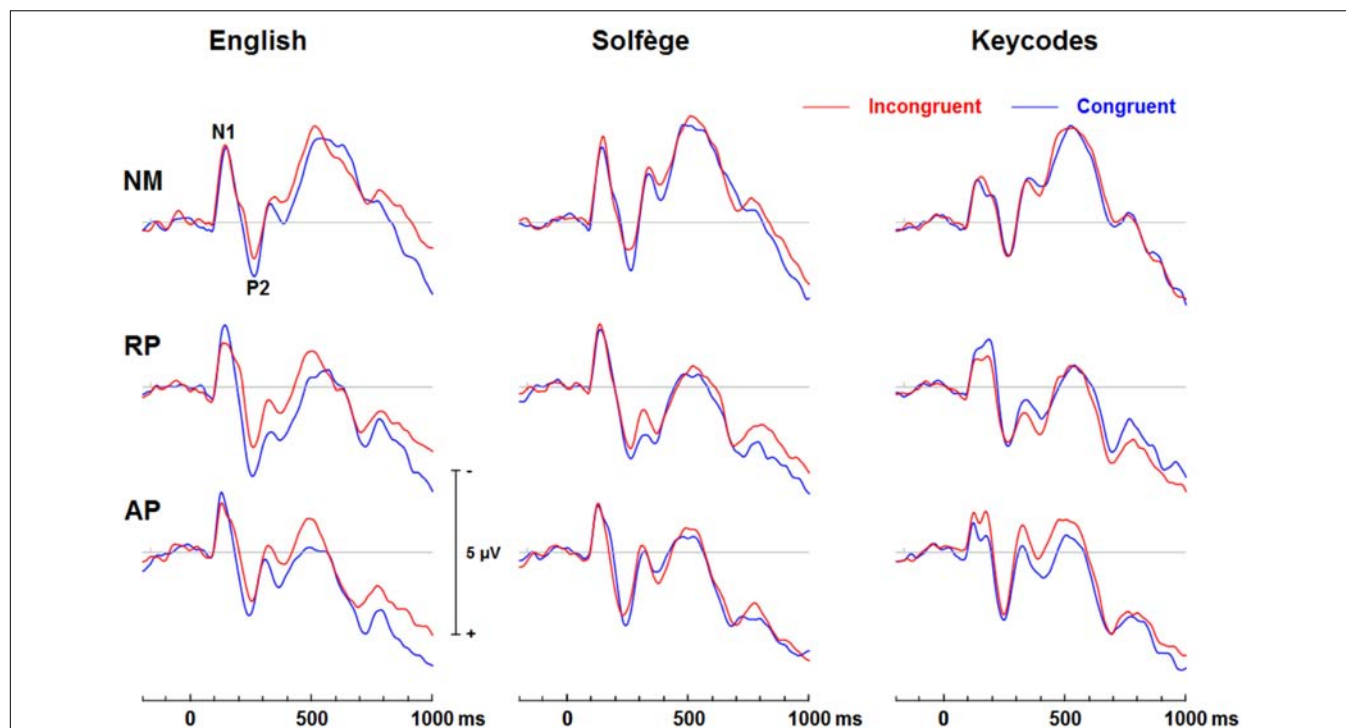


FIGURE 3 | Event-related potential waveforms of congruent and incongruent conditions for each task at the FCz site. A clear N1 and P2 neural electric modulation is visible, along with an expected N450. English words show a congruency effect for all groups for the N450 but only AP shows this effect for keycodes.

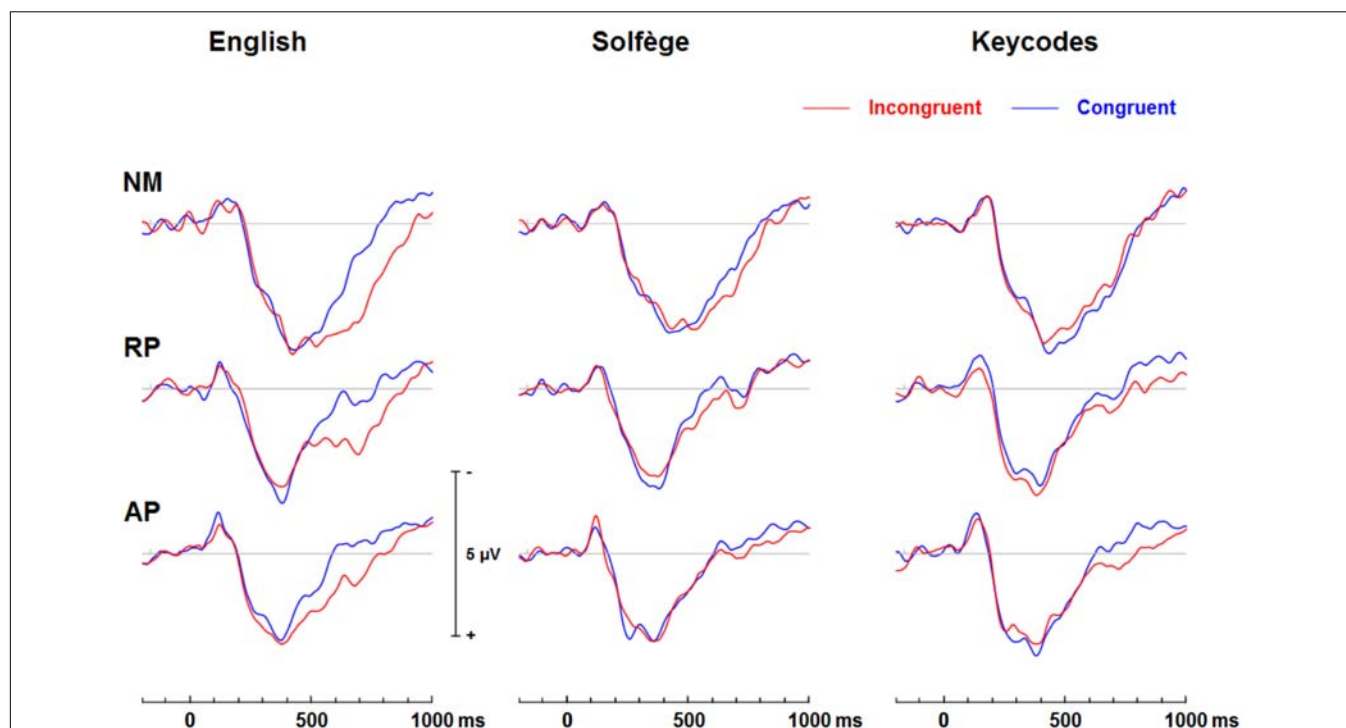
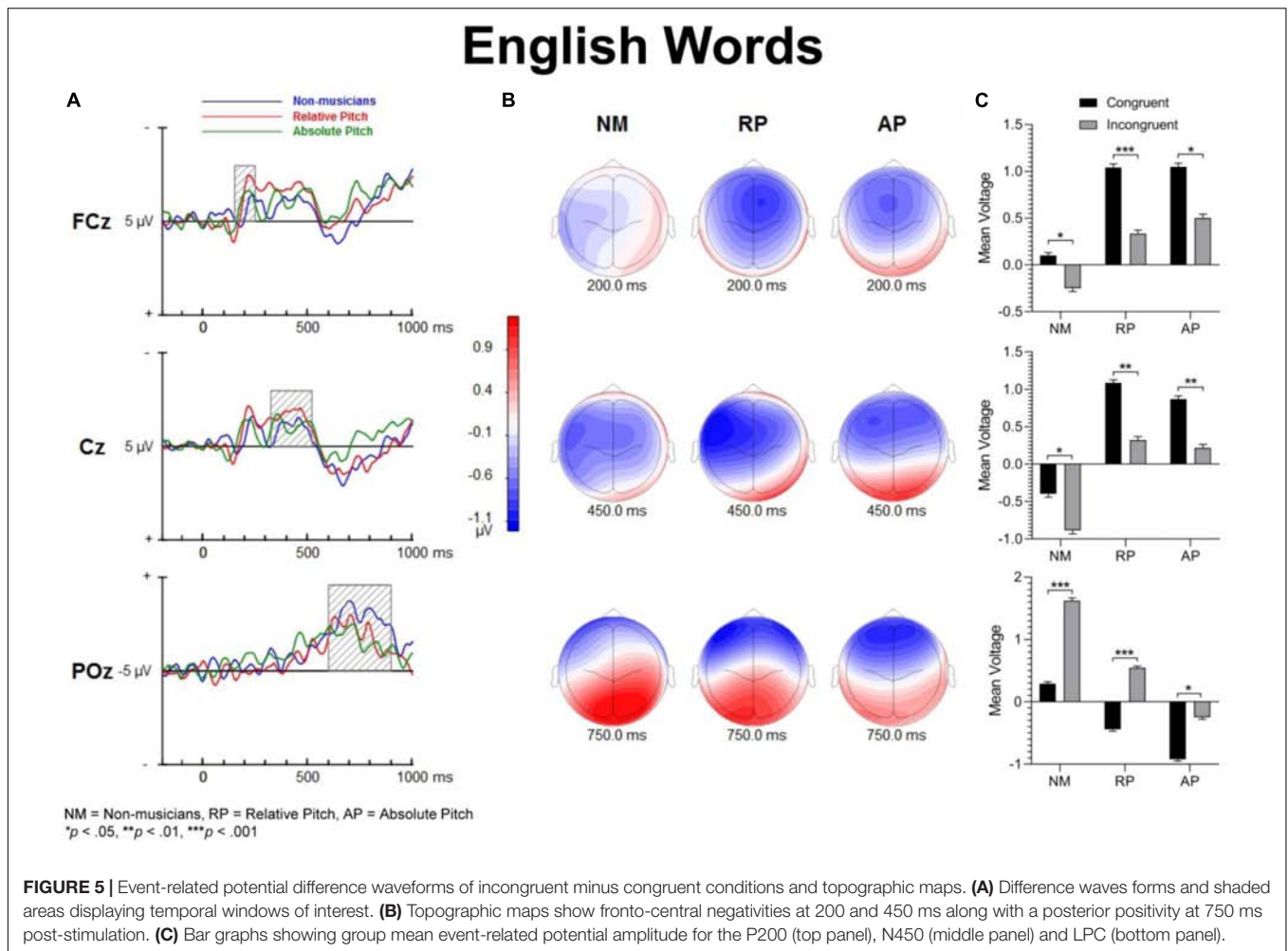


FIGURE 4 | Event-related potential waveforms at the midline parietal (Pz) electrode site showing a posterior late positive complex (LPC) wave that significantly differs in amplitude at late stages of processing for congruent and incongruent word form conditions, which may reflect a post-trial motor response awareness. English words elicit significant congruence effects but only AP shows significant LPC differences for keycodes.



($p < 0.01$). There was no difference in ERP amplitude between the NM and RP group ($p = 0.074$).

In the solfège task, the group by congruency interaction was not significant [$F(2,54) = 1.86$, $p = 0.166$, $\eta_p^2 = 0.064$]. The ERP was significantly more positive in amplitude for congruent than incongruent stimuli [$F(1,54) = 6.21$, $p = 0.016$, $\eta_p^2 = 0.103$]. There was a main effect of musicianship [$F(2,54) = 4.66$, $p = 0.014$, $\eta_p^2 = 0.147$]. Pairwise comparisons showed greater positive ERP amplitude in NMs than musicians with AP ($p = 0.013$) but not RP ($p = 0.086$). There was no difference in LPC amplitude for the solfège task between AP and RP groups.

For keycodes LPC, there was a significant interaction between musicianship and congruency [$F(2,54) = 3.41$, $p = 0.040$, $\eta_p^2 = 0.112$]. In musicians with AP, the LPC amplitude was significantly more positive for incongruent than congruent stimuli ($p < 0.01$) while no such differences in ERP amplitude were observed in musicians with RP or in the NM group. Overall, the ERP amplitude was smaller for congruent than incongruent stimuli, [$F(1,54) = 5.74$, $p = 0.022$, $\eta_p^2 = 0.094$]. The main effect of musicianship was also significant [$F(2,54) = 5.44$, $p < 0.01$, $\eta_p^2 = 0.168$], with more positivity of ERP amplitude in NMs than musicians with AP ($p = 0.012$) or RP ($p = 0.024$).

There was no significant difference in ERP amplitude between musicians with AP or RP.

DISCUSSION

In the present study, we used three different auditory versions of the Stroop task to assess the strength of association between the perceived pitch-classes and verbal labels in musicians with RP and AP. We hypothesized that musically trained groups would show greater Stroop interference effects proportionately when presented with incongruent musical notations than non-musicians. We also anticipated that musicians may perform better in absolute terms than non-musicians given the mounting evidence suggesting that musical training enhances executive functions and attentional control (Moreno et al., 2014).

Behavioral Data: Response Times and Error Rates

Overall, participants were slower and made more errors for incongruent than congruent stimuli. This congruency effect was modulated by lexicon, being stronger for word than for solfège or

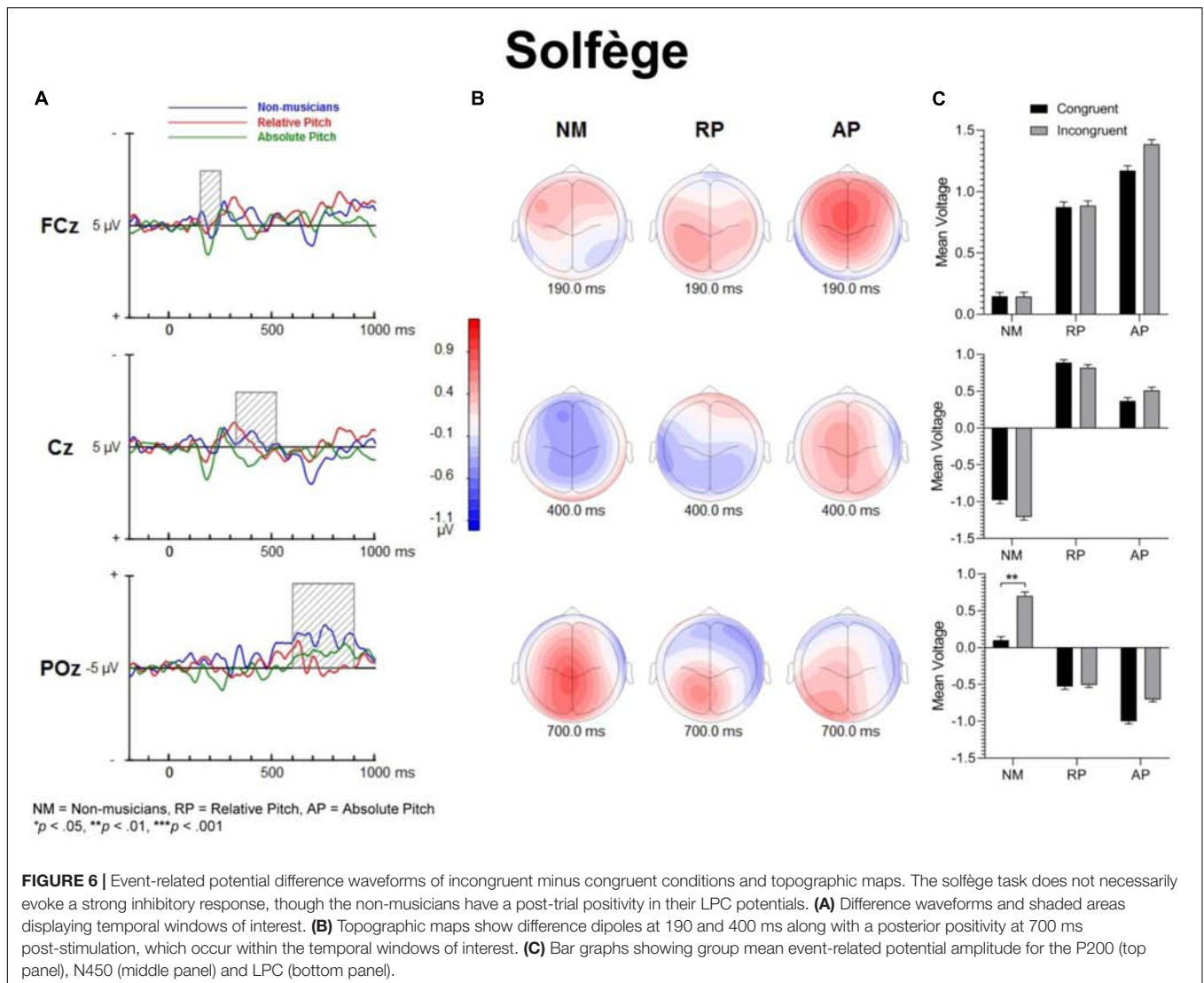
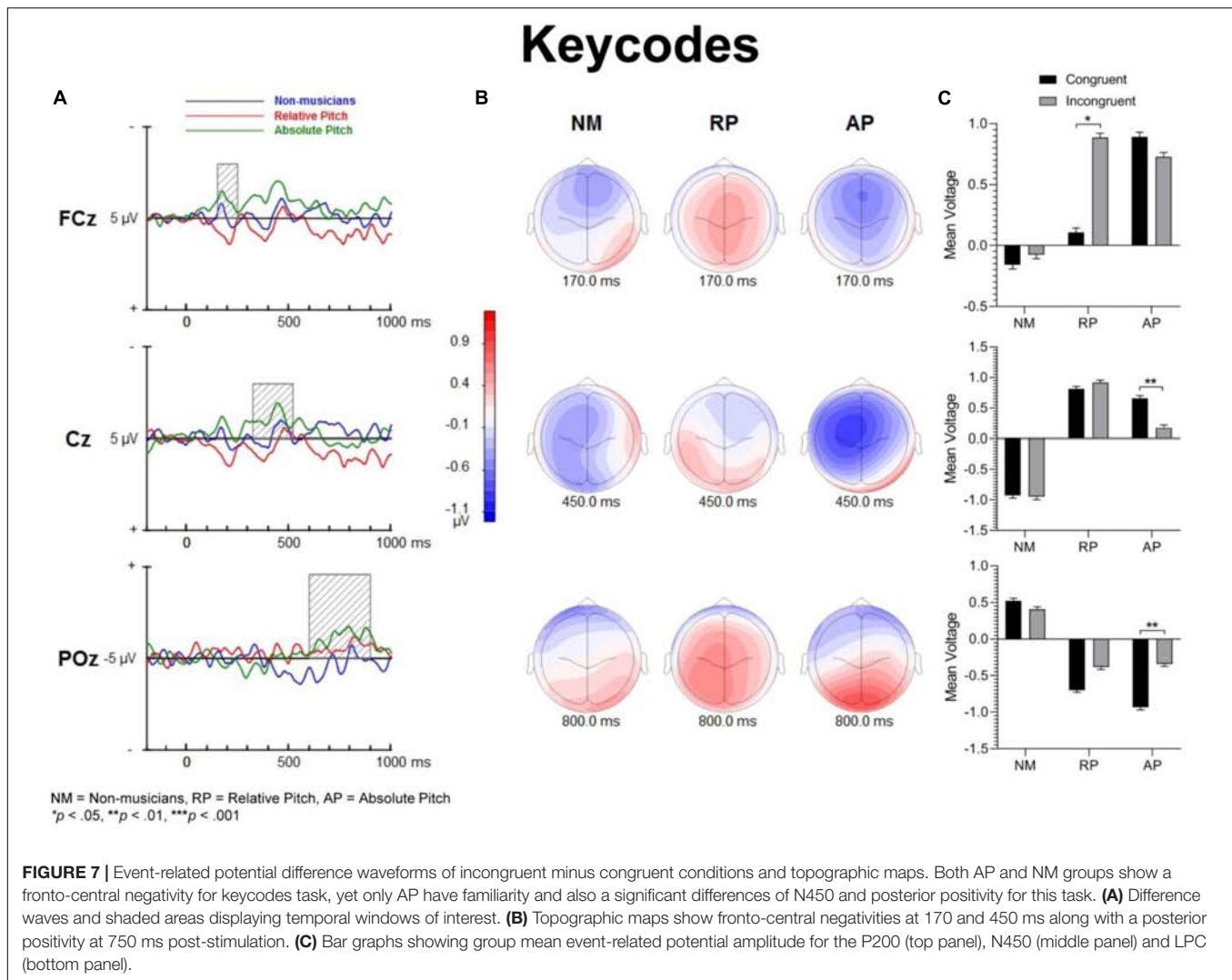


FIGURE 6 | Event-related potential difference waveforms of incongruent minus congruent conditions and topographic maps. The solfège task does not necessarily evoke a strong inhibitory response, though the non-musicians have a post-trial positivity in their LPC potentials. **(A)** Difference waveforms and shaded areas displaying temporal windows of interest. **(B)** Topographic maps show difference dipoles at 190 and 400 ms along with a posterior positivity at 700 ms post-stimulation, which occur within the temporal windows of interest. **(C)** Bar graphs showing group mean event-related potential amplitude for the P200 (top panel), N450 (middle panel) and LPC (bottom panel).

Germanic keycodes stimuli. Overall, musicians with RP and AP were faster and more accurate than non-musicians. This pattern of results is consistent with many studies showing that musicians often outperform non-musicians in a wide range of auditory tasks involving executive functions and attentional control (Moussard et al., 2016; Alain et al., 2018). However, the analyses of RT data show small or even non-existent congruency effects in musicians for solfège and keycodes versions of the Stroop task. This finding was unexpected and appears inconsistent with the hypothesis that musical training and AP are associated with a strong association between a perceived pitch and a verbal label. One possible explanation for these small and no congruency effects in the solfège and keycodes tasks would be that the musicians were successful in suppressing task-irrelevant information. The pattern of RTs during the word Stroop task is consistent with this possibility, with AP showing the smallest congruency effect, even in terms of accuracy as will be discussed. Further, the reverse Stroop effect findings for keycodes in NMs may occur

due to a lack of long-term learning, which makes affirmative association processing or evaluation of unfamiliar label to pitch mappings slower and less accurate than simply ignoring the unfamiliar verbal information and responding to pitch directly in the incongruent keycodes condition.

The analyses of error rate are also consistent with a flexibility of task-relevant attentional control account, with greater accuracy on incongruent trials across the tasks among musicians with RP and AP than non-musicians. Indeed, musicians with AP had virtually no errors in all three versions of the Stroop task, suggesting that they were able to successfully suppress task-irrelevant information and/or competing responses. While musician groups were not immune to incongruence, the findings seem consistent with several studies showing superior performance in musicians compared to non-musicians in a wide range of tasks involving auditory discrimination, executive functions, and attentional control (e.g., Kaiser and Lutzenberger, 2003; Marques et al., 2007; Putkinen et al., 2012; Alain et al.,



2018). At the same time, it should be noted that interference effects for words were observed in AP RTs rather than error rate. Interestingly, neutral word conditions were responded to more rapidly and accurately than incongruent stimuli but less rapidly than congruent stimuli by NMs only, while musician RTs did not differ between congruent and neutral English words. This may imply a RT facilitation effect in NMs for congruent trials that has reached fluency in AP and RP, allowing the musicians to name neutral trials with the same fluency NMs seem to name learned percepts such as colors in the color-word Stroop task (Stroop, 1935).

Neural Correlates: ERP Responses to Congruent and Incongruent Stimuli

The processing of congruent and incongruent stimuli was associated with three temporally and spatially distinct neural signatures. The first modulation peak of interest was directionally positive, occurred at about 200 ms after sound onset, and was larger over midline central and fronto-central areas. In a

prior spoken words version of the “low/high” auditory Stroop task (Donohue et al., 2012), a similar neural response was observed around 200 ms. This neural response was found to be significantly attenuated for incongruent compared to congruent sung words but not musical annotations, suggesting that conflicting information in musical domain maybe less salient than in standard language. In the present study, our behavioral data showed reduced congruency effects in the solfège and keycodes tasks. At the same time, the amplitude of the early modulation at 200 ms, or P200, remained comparably unaffected by congruence across both of these tasks. This appears inconsistent with current models positing a strong conditional association between musical notation and pitch, unlike language, which is consistent with the behavioral findings. If both AP and RP musicians were processing musical notation at the same processing strength as words, then one would expect to observe larger and regular congruency effects at that latency for both groups across all tasks as was observed for words across all groups. However, despite a lack of distinct congruency effects on P200 in the music tasks and lack of statistical significance

in group main effects, the overall amplitude was nevertheless slightly larger in musicians across tasks and conditions, which seems consistent with studies showing that this modulation may be an important physiological correlate for longer timescale auditory learning (Reinke et al., 2003; Ross and Tremblay, 2009; Alain et al., 2015). The somewhat greater P200 positivity in the RP and AP groups may also imply an advantage in early processing, which might allow the stimulus conflict to be more easily responded to. Certainly, the idea that musical mappings do not produce strong response conflict but musicianship may still allow for greater efficiency in stimulus conflict detection seems to be consistent with both the behavioral and merely observed P200 results. Perhaps, early sensory stream information processing of direction of change or frequency detection from prolonged familiarity to pitch naming causes P200 to be minutely greater in musicians such that the statistically significant congruence effects on P200 are measurably greater in musicians for at least words, which do indeed show evidence of strong associational bonding to pitch information.

The modulation peaking at about 200 ms was followed by another one that peaked at about 450 ms after sound onset, which was characterized by an increased negativity elicited by the incongruent relative to congruent stimuli. This modulation, referred to as the N450, had a broader and more sustained time course than the early P200, and peaked over the left central and fronto-central scalp area. This long lasting negativity could reflect both congruence detection as well as conflict resolution as it remained until a response was made (see also West and Alain, 1999). The N450 amplitude varied as a function of task, being greater for the English word Stroop task than the solfège or keycodes Stroop tasks. This may be due to greater cognitive load related to the semantic effects of the standard language Stroop tasks and greater flexibility between verbal code and musical pitch in terms of associative processing, particularly in the more common RP ability. The Stroop interference effect also decreased when participants were presented with less familiar material such as in the solfège and keycodes tasks. That is, Romanic solemnization or Germanic key lexicons may not possess as strong stimulus to response mappings as the English language (standard auditory Stroop), even for AP. In other words, there is less interference of response in general for pitch naming with musical lexicons than there is for the English language lexicon. Thus, it may also be easier to suppress abstract musical lexicons than concrete referential pitch labels such as low or high. The latter is likely to have a stronger stimulus to response mapping than more abstract musical notations despite prolonged exposure to musical lexicons. However, it may be possible that the significant differences in N450 and also LPC from congruency effects in the keycodes task suggest that only AP seem to experience stimulus conflict though not necessarily response conflict as measured with RTs and accuracy. At the same time, the N450 for both musician groups was significantly lower in amplitude than the NMs, especially in the auditory word Stroop task. This finding may reflect modified thresholds for information conflict detection in the auditory domain for musically trained populations.

The LPC is thought to index conceptual level of representation (West and Alain, 1999). In the present study, its amplitude was modulated by task and musical training, with greater difference between congruent and incongruent stimuli during the auditory word Stroop than musical annotations. Given that the LPC often occurred at the time of or even after a button-press response was made, we reasoned that the conceptual level of representation is a post-response or decision evaluation process, described in previous studies as a semantic re-activation of the lingual item preceding a response (Ergen et al., 2014). The effects of task on this modulation provide some support for such an interpretation because the conceptual level of representation differed between the three tasks, with a stronger representation for words than solfège or keycodes. Interestingly, the modulation was observed in the keycodes task for AP musicians with differences between congruent and incongruent stimuli, while no such modulation was seen in NMs or musicians with RP. This provides converging evidence supporting the notion that AP might yield a conceptual level post-response process for alpha-numeric encodings, which in turn may be modulated by incongruence when there is conflict in information. AP possessors were also significantly different in LPC amplitudes for English words, which may mean they process word forms and alpha-numeric musical encodings similarly but with less response conflict for the musical codes, contrary to our hypotheses. The interaction between musicianship, task, and congruency for the LPC also seems to imply less processing power requirements for the musical tasks, which may show a transfer effect to non-musical tasks in musicians, though both the exogenous and LPC modulations may demonstrate differential use of cognitive resources and possibly behavioral strategy for each group.

CONCLUSION

Music can be viewed as another form of language or complex communication medium. Those proficient in music (i.e., musicians) likely develop a lexicon of musical annotations that may share similarities with speech representation. That is, a familiar pitch and musical annotation may automatically trigger a verbal label with a specific meaning. If this is the case, then musicians should experience difficulty in processing incongruent musical stimuli. The present study shows that musicians, even those with AP, exhibited little interference effect at the behavioral level in the two musical versions of the Stroop task. However, the analyses of ERP data show a significant difference between congruent and incongruent in the Germanic keycodes for AP, unlike in the RP or NM groups for this task. This finding suggests that AP possessors, at least at the neural level, might process phonemic equivalents of alpha-numeric encodings of pitch information with a similar associative processing as word forms. Yet, while this processing appears to have semantic properties it is not as powerful as that of standard language. Thus, keycodes appear to have unique processing in music for their mappings to pitch classes in AP. We hypothesize that greater executive function and attentional control associated with playing a musical instrument

may help musicians process conflicting information. For this reason, the detection of stimulus conflict may generate attenuated response conflict in musicians compared to NMs, which would help to explain certain discrepancies between the neural and behavioral data in particular parts of our study, along with the more rapid and accurate musician responses. Future research is needed to better understand the apparent contradiction between behavioral and electrophysiological measures in AP. At the same time, the lack of P200 congruency effects in the musical tasks should also be examined by increasing task difficulty and attentional load. Finally, cross-modal transfer of musical training on semantic information may be studied by applying conceptual level processing to domains outside of pitch perception, such as spatialization.

DATA AVAILABILITY

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Human Research and Ethics Committee at

the Baycrest Centre, Toronto, ON, Canada. The participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

VS, MT, FR, and CA designed the experiments and interpreted the results of experiments. VS performed the experiments and collected the data. CA and VS analyzed the data. VS and CA drafted the manuscript. All authors edited, revised, and approved the final version of the manuscript.

FUNDING

This work was supported by the Natural Sciences and Engineering Research Council of Canada to CA under Grant (RGPIN-2016-05523).

ACKNOWLEDGMENTS

We would like to thank all the participants for their involvement in this study.

REFERENCES

- Alain, C., Khatamian, Y., He, Y., Lee, Y., Moreno, S., Leung, A., et al. (2018). Different neural activities support auditory working memory in musicians and bilinguals. *Ann. N. Y. Acad. Sci.* 1423, 435–446. doi: 10.1111/nyas.13717
- Alain, C., Moussard, A., Singer, J., Lee, Y., Bidelman, G. M., and Moreno, S. (2019). Music and visual art training modulate brain activity in older adults. *Front. Neurosci.* 13:182. doi: 10.3389/fnins.2019.00182
- Alain, C., Zhu, K. D., He, Y., and Ross, B. (2015). Sleep-dependent neuroplastic changes during auditory perceptual learning. *Neurobiol. Learn. Mem.* 118, 133–142. doi: 10.1016/j.nlm.2014.12.001
- Atkinson, C. M., Drysdale, K. A., and Fulham, W. R. (2003). Event-related potentials to stroop and reverse stroop stimuli. *Int. J. Psychophysiol.* 47, 1–21. doi: 10.1016/s0167-8760(02)00038-7
- Baharloo, S., Johnston, P. A., Service, S. K., Gitschier, J., and Freimer, N. B. (1998). Absolute pitch: an approach for identification of genetic and nongenetic components. *Am. J. Hum. Genet.* 62, 224–231. doi: 10.1086/301704
- Bermudez, P., and Zatorre, R. J. (2009). A distribution of absolute pitch as revealed by computerized testing. *Music Percept.* 27, 89–101. doi: 10.1525/mp.2009.27.2.89
- Besson, M., Chobert, J., and Marie, C. (2011). Transfer of training between music and speech: common processing, attention, and memory. *Front. Psychol.* 2:94. doi: 10.3389/fpsyg.2011.00094
- Donohue, S. E., Liotti, M., Perez, R., and Woldorff, M. G. (2012). Is conflict monitoring supramodal? spatiotemporal dynamics of cognitive control processes in an auditory stroop task. *Cogn. Affect. Behav. Neurosci.* 12, 1–15. doi: 10.3758/s13415-011-0060-z
- Duncan-Johnson, C. C., and Kopell, B. S. (1981). The stroop effect: brain potentials localize the source of interference. *Science* 214, 938–940. doi: 10.1126/science.7302571
- Ergen, M., Saban, S., Kirmizi-Alsan, E., Uslu, A., Keskin-Ergen, Y., and Demiralp, T. (2014). Time-frequency analysis of the event-related potentials associated with the stroop test. *Int. J. Psychophysiol.* 94, 463–472. doi: 10.1016/j.ijpsycho.2014.08.177
- Hamers, J. E., and Lambert, W. E. (1972). Bilingual interdependencies in auditory perception. *J. Verbal Learning Verbal Behav.* 11, 303–310. doi: 10.1016/s0022-5371(72)80091-4
- Ilan, A. B., and Polich, J. (1999). P300 and response time from a manual stroop task. *Clin. Neurophysiol.* 110, 367–373. doi: 10.1016/s0168-5597(98)00053-7
- Itoh, K., Suwazono, S., Arao, H., Miyazaki, K., and Nakada, T. (2005). Electrophysiological correlates of absolute pitch and relative pitch. *Cereb. Cortex* 15, 760–769. doi: 10.1093/cercor/bhh177
- Kaiser, J., and Lutzenberger, W. (2003). Induced gamma-band activity and human brain function. *Neuroscientist* 9, 475–484. doi: 10.1177/1073858403259137
- Levitin, D. J., and Rogers, S. E. (2005). Absolute pitch: perception, coding, and controversies. *Trends Cogn. Sci.* 9, 26–33. doi: 10.1016/j.tics.2004.11.007
- MacLeod, C. M. (1991). Half a century of research on the stroop effect: an integrative review. *Psychol. Bull.* 109, 163–203. doi: 10.1037/0033-2909.109.2.163
- Marques, C., Moreno, S., Castro, S. L., and Besson, M. (2007). Musicians detect pitch violation in a foreign language better than non-musicians: behavioral and electrophysiological evidence. *J. Cogn. Neurosci.* 19, 1453–1463. doi: 10.1162/jocn.2007.19.9.1453
- McNeely, H. E., West, R., Christensen, B. K., and Alain, C. (2003). Neurophysiological evidence for disturbances of conflict processing in patients with schizophrenia. *J. Abnorm. Psychol.* 112, 679–688. doi: 10.1037/0021-843x.112.4.679
- Miyazaki, K. (1989). Absolute pitch identification: effects of timbre and pitch region. *Music Percept.* 7, 1–14. doi: 10.2307/40285445
- Miyazaki, K. (1990). The speed of musical pitch identification by absolute-pitch possessors. *Music Percept.* 8, 177–188. doi: 10.2307/40285495
- Miyazaki, K. (1992). Perception of musical intervals by absolute pitch possessors. *Music Percept.* 9, 413–426. doi: 10.2307/40285562
- Miyazaki, K., and Ogawa, Y. (2006). Learning absolute pitch by children: a cross-sectional study. *Music Percept.* 24, 63–78. doi: 10.1525/mp.2006.24.1.63
- Moreno, S., Wodniecka, Z., Tays, W., Alain, C., and Bialystok, E. (2014). Inhibitory control in bilinguals and musicians: event related potential (ERP) evidence for experience-specific effects. *PLoS One* 9:e94169. doi: 10.1371/journal.pone.0094169

- Moussard, A., Bermudez, P., Alain, C., Tays, W., and Moreno, S. (2016). Life-long music practice and executive control in older adults: an event-related potential study. *Brain Res.* 1642, 146–153. doi: 10.1016/j.brainres.2016.03.028
- Prevor, M. B., and Diamond, A. (2005). Color-object interference in young children: a stroop effect in children 3 1/2–6 1/2 years old. *Cogn. Dev.* 20, 256–278. doi: 10.1016/j.cogdev.2005.04.001
- Putkinen, V., Tervaniemi, M., and Houtilainan, M. (2012). Informal musical activities are linked to auditory discrimination and attention in 2–3-year-old children: an event-related potential study. *Eur. J. Neurosci.* 37, 654–661. doi: 10.1111/ejn.12049
- Rafal, R., and Henik, A. (1994). “The neurology of inhibition: integrating controlled and automatic processes,” in *Inhibitory Processes in Attention, Memory, and Language*, eds D. Dagenbach and T. H. Carr (San Diego, CA: Academic Press), 1–51.
- Rebai, M., Bernard, C., and Lannou, J. (1997). The stroop’s test evokes a negative brain potential, the N400. *Int. J. Neurosci.* 91, 85–94. doi: 10.3109/00207459708986367
- Reinke, K., He, Y., Wang, C., and Alain, C. (2003). Perceptual learning modulates sensory evoked response during vowel segregation. *Cogn. Brain Res.* 17, 781–791. doi: 10.1016/s0926-6410(03)00202-7
- Ross, B., and Tremblay, K. (2009). Stimulus experience modifies auditory neuromagnetic responses in young and older listeners. *Hear. Res.* 248, 48–59. doi: 10.1016/j.heares.2008.11.012
- Russo, F. A., Windell, D. L., and Cuddy, L. L. (2003). Learning the “special note”: evidence for a critical period for absolute pitch acquisition. *Music Percept.* 21, 119–128.
- Schulze, K., Gaab, N., and Schlaug, G. (2009). Perceiving pitch absolutely: comparing absolute and relative pitch possessors in a pitch memory task. *BMC Neurosci.* 10:106. doi: 10.1186/1471-2202-10-106
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662. doi: 10.1037/h0054651
- Takeuchi, A. H., and Hulse, S. H. (1993). Absolute pitch. *Psychol. Bull.* 113, 345–361.
- Terhardt, E., and Seewann, M. (1983). Aural key identification and its relationship to absolute pitch. *Music Percept.* 1, 63–83. doi: 10.2307/40285250
- West, R., and Alain, C. (1999). Event-related neural activity associated with the stroop task. *Cogn. Brain Res.* 8, 157–164. doi: 10.1016/s0926-6410(99)00017-8
- West, R., and Alain, C. (2000). Age-related decline in inhibitory control contributes to the increased stroop effect observed in older adults. *Psychophysiology* 37, 179–189. doi: 10.1017/s0048577200981460
- Zatorre, R. J., Perry, D. W., Beckett, C. A., Westbury, C. F., and Evans, A. C. (1998). Functional anatomy of musical processing in listeners with absolute and relative pitch. *Proc. Natl. Acad. Sci. U.S.A.* 95, 3172–3177. doi: 10.1073/pnas.95.6.3172
- Zattore, R. J. (2003). Absolute pitch: a model for understanding the influence of genes and development on neural and cognitive function. *Nat. Neurosci.* 6, 692–695. doi: 10.1038/nn1085

Conflict of Interest Statement: 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.

Copyright © 2019 Sharma, Thaut, Russo and Alain. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Uncertain Emotion Discrimination Differences Between Musicians and Non-musicians Is Determined by Fine Structure Association: Hilbert Transform Psychophysics

Francis A. M. Manno^{1,2*}, Raul R. Cruces³, Condon Lau^{2*} and Fernando A. Barrios^{3*}

¹ School of Biomedical Engineering, Faculty of Engineering, University of Sydney, Sydney, NSW, Australia, ² Department of Physics, City University of Hong Kong, Hong Kong, China, ³ Instituto de Neurobiología, Universidad Nacional Autónoma de México, Querétaro, Mexico

OPEN ACCESS

Edited by:

Paul J. Colombo,
Tulane University, United States

Reviewed by:

Bruno Gingras,
University of Vienna, Austria
Graham Frederick Welch,
UCL Institute of Education,
United Kingdom

*Correspondence:

Francis A. M. Manno
Francis.Manno@Sydney.edu.au
Condon Lau
condon.lau@cityu.edu.hk
Fernando A. Barrios
fbarrios@unam.mx

Specialty section:

This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Neuroscience

Received: 17 June 2019

Accepted: 13 August 2019

Published: 18 September 2019

Citation:

Manno FAM, Cruces RR, Lau C
and Barrios FA (2019) Uncertain
Emotion Discrimination Differences
Between Musicians and
Non-musicians Is Determined by Fine
Structure Association: Hilbert
Transform Psychophysics.
Front. Neurosci. 13:902.
doi: 10.3389/fnins.2019.00902

Humans perceive musical sound as a complex phenomenon, which is known to induce an emotional response. The cues used to perceive emotion in music have not been unequivocally elucidated. Here, we sought to identify the attributes of sound that confer an emotion to music and determine if professional musicians have different musical emotion perception than non-musicians. The objective was to determine which sound cues are used to resolve emotional signals. Happy or sad classical music excerpts modified in fine structure or envelope conveying different degrees of emotional certainty were presented. Certainty was determined by identification of the emotional characteristic presented during a forced-choice discrimination task. Participants were categorized as good or poor performers ($n = 32$, age 21.16 ± 2.59 SD) and in a separate group as musicians in the first or last year of music education at a conservatory ($n = 32$, age 21.97 ± 2.42). We found that temporal fine structure information is essential for correct emotional identification. Non-musicians used less fine structure information to discriminate emotion in music compared with musicians. The present psychophysical experiments revealed what cues are used to resolve emotional signals and how they differ between non-musicians and musically educated individuals.

Keywords: emotion, psychophysics, modulation, fine structure, envelope, frequency, amplitude

INTRODUCTION

The process of resolving emotions has been described as the optimization of an economic choice (Seymour and Dolan, 2008). Experimentally, emotions are selected due to their perceived certainty and robust differentiation (Russell, 1980; Barrett, 1998; LeDoux, 2000; Phan et al., 2002; Koelsch, 2014). Few studies have analyzed uncertain emotions or the psychoacoustic cues that endow sound with emotion. Here, we sought to identify the sound attributes involved in musical emotion discrimination and to determine if non-musicians and musicians perceive emotional sound cues differently. We were interested in emotional uncertainty, where the perception of sound attributes is difficult to distinguish.

Cues in Emotion

Music is transmitted through temporal fine structure (TFS) and envelope (ENV) modulations, which are the perceived changes of the sound in amplitude and frequency, respectively. However, the exact contributions of TFS and ENV used to resolve emotion in music are unknown. Although the influence of pitch on emotional content of speech is well known (Lieberman and Michaels, 1962; Wildgruber et al., 2005), its contribution to music is less clear (Scherer, 1995; Coutinho and Dikken, 2013). Most studies have concentrated on emotion in speech (e.g., since very early on: Fairbanks and Pronovost, 1938; Fairbanks, 1940; Lieberman et al., 1964) or the cues (TFS or ENV) that confer intelligibility to speech (Licklider, 1946; Lieberman and Michaels, 1962; Williams and Stevens, 1972; Frick, 1985; Drullman et al., 1994, 1996; Drullman, 1995; Shannon et al., 1995; Leinonen et al., 1997). Using a Hilbert transform, which allows the signal to be deconstructed and reconstructed into its individual frequency and amplitude time components (King, 2009), researchers found that ENV was most important for speech reception, whereas TFS was most important for pitch perception (Smith et al., 2002). Several follow-up studies have corroborated the importance of ENV for speech intelligibility, in addition to the importance of TFS features (Zeng et al., 2004, 2005; Davidson et al., 2009; Fogerty, 2011; Swaminathan and Heinz, 2012; Apoux et al., 2013; Shamma and Lorenzi, 2013; Moon et al., 2014; Swaminathan et al., 2014). Unfortunately, the sound properties (ENV and TFS) that confer emotion to music have been less studied than the sound properties conveying emotion in speech, i.e., prosody (Hodges, 2010; Coutinho and Dikken, 2013). Therefore, a primary aim for the present experiment was to discern which and how the attributes in sound endow music with emotion.

Differing Ability to Discriminate Emotions in Auditory Cues: Emotional Resolvability Differences

Musicians have an enhanced auditory perception for several acoustic features, such as the ability to learn lexical tones (Wong et al., 2007), enhanced audiovisual processing (Musacchia et al., 2007), better speech-in-noise perception (Başkent and Gaudrain, 2016; Coffey et al., 2017), better pitch discrimination thresholds (Micheyl et al., 2006), and superior frequency discrimination (Liang et al., 2016; Mandikal Vasuki et al., 2016; Madsen et al., 2017) compared with non-musicians. Musical training and experience shape linguistic patterns (Wong et al., 2007) and enhance speech-in-noise discrimination (Parbery-Clark et al., 2009b), altering brainstem and cortical responses to musical and non-musical acoustic features (Parbery-Clark et al., 2009a; Kraus and Chandrasekaran, 2010; Strait et al., 2010). Musicians possess different auditory perceptual abilities than non-musicians; hence, a musician's ability to discriminate emotion in sound when linked to TFS or ENV changes should also differ from their non-musician colleagues.

Speech and Music Relations

There are several dimensions, models, and psychoacoustic features which are used to categorize emotion in music (Russell,

1980; Schubert, 2004, 2013; Gabrielsson and Lindstrom, 2010; Eerola and Vuoskoski, 2011; Eerola, 2012). Psychophysical studies (i.e., frequency and amplitude features) of emotion make up less of the literature than other aspects (i.e., anxiety, arousal, etc., Hodges, 2010). For example, Gingras et al. (2014) found that intensity and arousal ratings in music-induced emotion were largely unaffected by amplitude normalization, suggesting that additional acoustic features besides intensity could account for the variance in subjective arousal ratings. Further, some of the same neurobiological mechanisms underlying emotion in music also subserve the emotion in speech (Fitch, 2006; Juslin and Västfjäll, 2008; Kotz et al., 2018). This is important as several studies give clues as to how uncertain emotion in music might be perceived. For example, Shannon et al. (1995) found speech recognition primarily utilized temporal cues with a few spectral channels. Further to that stated above, Smith et al. (2002) found ENV is most important for speech reception, and TFS most important for pitch perception. Several follow-up studies have corroborated the importance of ENV for speech intelligibility up to a certain number of bands including aspects of TFS (Zeng et al., 2004, 2005; Davidson et al., 2009; Fogerty, 2011; Swaminathan and Heinz, 2012; Apoux et al., 2013; Shamma and Lorenzi, 2013; Moon et al., 2014; Swaminathan et al., 2014). Although the emotion carried in speech is a similar, but different auditory perception than music, resolving emotion in music will aid our understanding of identifying emotions in sound.

Present Experimental Aims

Our aim was to conduct a psychoacoustic experiment to investigate certain and uncertain emotion in musical sound. We sought to determine which cues are used to resolve emotional signals (Pfeifer, 1988). Moreover, we studied the differences between musicians and non-musicians. To accomplish these aims, we decomposed happy and sad musical stimuli in TFS or ENV using a Hilbert transform (Smith et al., 2002). The process yielded stimuli with increasing decomposition in TFS and ENV, and then we explored the different degrees of emotional certainty they conveyed. Certainty was defined as the ability to identify the decomposed stimuli based on their unaltered forms. Happy and sad stimuli with varying degrees of decomposition were presented in a forced-choice discrimination task. *First*, we expected that decomposing TFS or ENV information essential to determining emotionality in sound would reveal which cue was more important for emotion discrimination. *Second*, we expected that segregating participants by their identification with the original excerpt into good and poor performers, based on the reported classification (Peretz et al., 1998, 2001; Gosselin et al., 2007; Khalfa et al., 2008; Hopyan et al., 2016), would result in different emotional resolvability curves. *Third*, we expected that assessing musicians in their first year of study in the conservatory compared to those in their last year would reveal differences in emotional resolvability based on their musical education. *Lastly*, comparing non-musicians to musicians would reveal differences in emotional resolvability based on musical experience. Our main aim was to understand the cues used to resolve emotional signals.

MATERIALS AND METHODS

The experiment included both non-musicians and musicians who were studying music at a conservatory as participants. Participants with musical experience were from a musical conservatory located in Querétaro, Mexico. All data, sound files, and scripts are available at www.fmanno.com, the Open Science Framework (Manno et al., 2018, 2019) and GitHub¹. **Supplementary Data Sheet S1** contains extended analyses and **Supplementary Data Sheet S2** contains the original data. The research protocol was approved by the Internal Review Board of the Instituto de Neurobiología, Universidad Nacional Autónoma de México in accordance with the Declaration of Helsinki, 2013. Informed consent (verbal and written) prior to undertaking the experiment was granted and abided by as set forth in the Ethical Principles of the Acoustical Society of America for Research Involving Human and Non-human Animals in Research and Publishing and Presentations.

Study Participants

Participants were recruited from a local university (Natural Sciences and Musical Conservatory), and final participants were randomly selected from approximately 320 individuals (with the sample not differing from the population group). All individuals were native Spanish speakers. The present study included 64 individuals divided equally into 8 groups (**Table 1**). Non-musician participants were selected and classified as poor and good ($n = 32$, age 21.17 ± 2.63 SD), based on their performance in the Montreal Emotional Identification Task using greatest mean spread between the groups as the separation metric (Peretz et al., 1998, 2001; Gosselin et al., 2007, 2011; Khalfa et al., 2008; Hopyan et al., 2016). Poor and good participants did not have musical training or education, nor did this cohort play instruments. Poor and good participants were a separate sample from musicians. Musicians were separated based on musical education as first-year (low education) and last-year (high education) students at

the conservatory ($n = 32$, age 21.97 ± 2.42). Musically educated participants were recruited from the wind and string sections of the conservatory. All volunteers were free of contraindications for psychoacoustic testing. Prior to undertaking the emotional resolvability experiment, participants had audiometric testing to confirm hearing within normal limits.

Audiometric testing consisted of the examiner presenting to the participants a series of pure tones from 400 to 8,000 Hz, in addition to linear sweeps, log sweeps, and white noise in the same frequency range. The sound pressure level (SPL) decibels (dB) was modulated from 20 to 60 SPL dB. The participants included in the study confirmed hearing the series of audiometric presentations. Participants who did not confirm hearing these tones were excluded.

Acoustic Stimuli

The 32 original acoustic stimuli classified as sad or happy were taken from a previous study [Montreal Emotional Identification Task²; **Supplementary Table S1** (Peretz et al., 1998, 2001; Gosselin et al., 2007, 2011; Khalfa et al., 2008; Hopyan et al., 2016)]. Half of the stimuli in the repertoire evoked a sense of happiness (major mode with a median tempo of 138 beats per min, bpm), and the other half evoked a sense of sadness (minor mode with a median tempo of 53 bpm) (Peretz et al., 2001). The altered excerpts had the same frequency and amplitude values in terms of pitch and duration as was found in the original stimulus. The sound files of classical music were processed in MATLAB to curtail length to 3-s (in order to present the entire battery of stimuli within an hour period), restricted in frequency/amplitude range for presentation (to reduce noise in the Musical Instrument Digital Interface (MIDI) file), and analyzed spectrally for differences in TFS or ENV (see **Figure 1**). Original and altered stimuli were presented with MATLAB (Statistics Toolbox Release 2012b, The MathWorks Inc., Natick, MA) using the Psychophysics Toolbox extension³ (Manno et al.,

²www.brams.umontreal.ca/peretz

³<http://psychtoolbox.org>

¹https://github.com/rcruces/2019_uncertain_emotion_discrimination

TABLE 1 | Participant data.

Sex	Non-musicians				Musicians			
	Good performer		Poor performer		First year music		Last year music	
	Male	Female	Male	Female	Male	Female	Male	Female
Number	$n = 8$	$n = 8$	$n = 8$	$n = 8$	$n = 8$	$n = 8$	$n = 8$	$n = 8$
Age (years)	20.16 ± 2.11	22.64 ± 3.58	21.11 ± 2.27	20.74 ± 2.11	20.89 ± 1.74	20.56 ± 1.17	23.16 ± 2.01	23.27 ± 3.34
Range (years)	18–25	9–29	19–26	19–25	19–24	19–23	20–26	20–30
Music education	None	None	None	None	2 ± 0	2 ± 0	8.38 ± 2.26	7.12 ± 2.1
Correct identification	$97.66 \pm 2.21\%$	$95.31 \pm 2.36\%$	$80.08 \pm 2.86\%$	$80.86 \pm 4.24\%$	$93.36 \pm 2.00\%$	92.58 ± 5.52	$89.45 \pm 3.31\%$	$91.02 \pm 4.24\%$
Total ($n = 32$)	31.25 ± 0.66	30.50 ± 0.71	25.62 ± 0.86	25.88 ± 1.27	29.88 ± 0.64	29.62 ± 1.65	28.62 ± 0.99	29.12 ± 1.27

Group (good, poor, first year music education or last year music education) by sex (male and female). Years of music education \pm SD years. Correct identification – indicating the percent identification based on the original classification of the Montreal Emotional Identification Task. Total – indicating the total number of stimuli presented with the column representing the number out of the total.

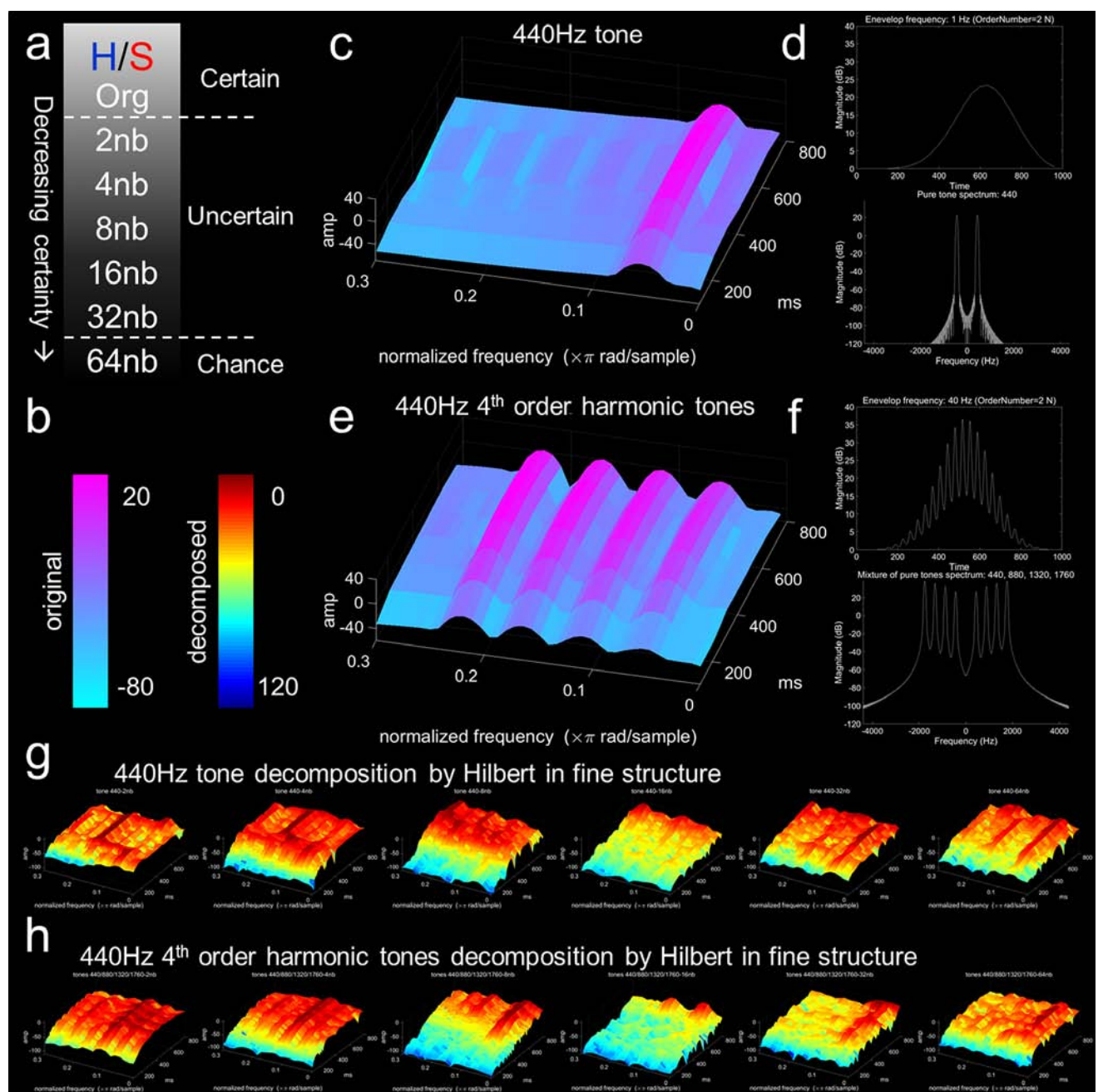


FIGURE 1 | Experimental design and example of the sound decomposition. **(a)** Happy (H)/sad (S) stimuli with decreasing emotional certainty by decomposition. Original excerpts were taken from the Montreal Emotional Identification Task (Peretz et al., 1998; Peretz et al., 2001; Gosselin et al., 2007; Khalfa et al., 2008; Hopyan et al., 2016). Decompositions are based on perceptible categorization as happy or sad, which pertain to certain (org – original) and increasingly uncertain (2nb, 4nb, 8nb, 16nb, 32nb, 64nb decompositions). **(b)** Color bar for coding spectrogram in original (cool color bar) and decomposed (jet color bar) stimuli (simple stimuli example decomposition). Color bar represents normalized power/frequency (dB/Hz) or amplitude (sound pressure level dB) by fine structure components (Hz). All spectrogram plots contain magnitude (dB) on z-axis, normalized frequency ($\times \pi$ rad/sample) on x-axis, and time in milliseconds (ms) on y-axis. **(c)** Single 440 Hz tone. **(d)** Upper panel, single tone with extracted envelope at 1 Hz (order 2N Butterworth filter) with magnitude on the y-axis and time on the x-axis. Lower panel, fast Fourier transform plot of single tone 440 Hz with magnitude on y-axis and frequency (hertz, cycles per second) on x-axis. **(e)** Complex 440 Hz 4th order harmonics with 440, 880, 1320, and 1760 Hz components. **(f)** Upper panel, complex tone with extracted envelope at 40 Hz (order 2N Butterworth filter) with magnitude on the y-axis and time on the x-axis. Lower panel, fast Fourier transform plot of complex 440 Hz 4th order harmonic tones with magnitude (dB) on y-axis and frequency hertz on x-axis. **(g)** Hilbert decomposition of single 440 Hz tone fine structure with complex 4th order harmonics envelope. **(h)** Hilbert decomposition of complex 4th order harmonics fine structure with single 440 Hz tone envelope. Progression from left to right for both panels **(g,h)** represent the Hilbert transformation for this simple example with 2nb, 4nb, 8nb, 16nb, 32nb, and 64nb decompositions. The simple example demonstrates how a complex acoustic stimulus that is categorized as emotional can be decomposed.

2018, 2019). Sound level was adjusted before psychoacoustic testing as described above.

Acoustic Stimuli Decomposition

Happy and sad stimuli were decomposed by a Hilbert transform in order to derive the altered excerpts (Smith et al., 2002; Moon et al., 2014). The decomposition process associated the acoustic aspects of emotion (happy or sad) with TFS or ENV. The ENV is represented as the magnitude of the Hilbert transform and TFS is the phase (Smith et al., 2002; King, 2009; Moon et al., 2014). Here, we created band-decomposed hybrid stimuli as mixtures of the happy and sad sounds by equal bandwidth steps. Cutting frequencies to 80, 260, 600, 1240, 2420, 4650, and 8820 Hz created six bands of decomposition: 2nb, 4nb, 8nb, 16nb, 32nb, and 64nb. Here, “nb” means number band decomposition as in the original description (Smith et al., 2002). An increase in band decomposition results in decreasing emotional resolvability for the original stimuli (Smith et al., 2002). The entire set of 224 decomposed stimuli and a script demonstrating the Hilbert transform process (Hilbert Explanation) can be found at <https://osf.io/8ws7a> (Manno et al., 2019). In our case, the Hilbert decomposition resulted in hybrid acoustic sounds and emotional resolvability was effectively tied to TFS or ENV in equally spaced decreasing bandwidths (Smith et al., 2002; Moon et al., 2014). For signal decomposition in the present study, the Hilbert transform $y(t)$ in the time domain is related to real function $x(t)$ by the analytic signal $A(t) = x(t) + iy(t)$, with $i = (-1)^{1/2}$. The Hilbert ENV is the magnitude of the analytic signal $|A(t)| = ((s_r(t))^2 + (s_i(t))^2)^{1/2}$ and the Hilbert TFS is $\cos \varphi(t)$, where $\varphi(t) = \arctan(s_r(t)/s_i(t))$ is the phase of the analytic signal (Smith et al., 2002; King, 2009). If the real (r) part pertains to cosine of the frequency contained within the signal and the imaginary (i) part pertains to sine of the frequency contained within the signal, the magnitude of the amplitude is related by the value of the cosine and sine of the signal (King, 2009). The decomposition process has been previously elaborated on for various signal-processing purposes (Oswald, 1956; Smith et al., 2002; Binder et al., 2004; King, 2009; Moon et al., 2014).

Hybrid Stimuli Example

Recombined hybrid stimuli with differing combinations of ENV and TFS from either emotion category were presented in a happy/sad descending two-interval forced-choice discrimination task (Figures 1a,b). For visualization of the Hilbert process, we created two stimuli, one pure tone (440 Hz, Figure 1c), and a harmonic series of the pure tone (440, 880, 1320, and 1760 Hz, Figure 1e). Both stimuli underwent TFS and ENV extraction by the Hilbert transformation and decomposed into hybrid stimuli for each band of decomposition. Figure 1g presents the Hilbert transformation of the pure tone in TFS combined with the harmonic series stimuli in ENV. Figure 1h presents the Hilbert transformation of the harmonic series stimuli in TFS combined with the pure tone in ENV. From left to right for Figures 1g,h, band decomposition increases from 2nb to 64nb. The pure tone and the harmonics series were given an amplitude double their predecessor (t), starting

with 10 dB SPL (decibels sound pressure level), a phase $(\pi/2)/t$ change from its lower harmonic, and separate durations. For the representation, spectrogram plots contained magnitude (dB) on the z -axis, normalized frequency ($\times \pi$ rad/sample) on the x -axis, and time in milliseconds (ms) on the y -axis. Spectrograms and acoustic stimuli were normalized across power/frequency (dB/Hz) and amplitude (dB SPL) by fine structure components (Hz).

Happy/Sad Forced-Choice Discrimination Task

A forced-choice discrimination task was conducted where participants were required to respond to the acoustic stimuli indicating if they perceived them as happy or sad. The entire repertoire of original and band-decomposed hybrid stimuli was utilized for the present experiments (see Supplementary Music Files at <https://osf.io/8ws7a>). The participants were presented with all the stimuli and asked to classify them as happy or sad (Smith et al., 2002; Moon et al., 2014). If an original sound was happy, decompositions were categorized as happy and the identification was deemed correct for happy and incorrect for sad. Participants were handed a sheet consisting of 32 rows and 7 columns (see Supplementary Music Matrix at <https://osf.io/8ws7a>). The columns in the Supplementary Music Matrix are organized as original stimuli presentation, followed from left to right by stimuli decompositions (2nb, 4nb, etc.). The trial was organized by random presentation followed by increasing decompositions. Decompositions in TFS were presented starting with their original unaltered form and continuing through 2nb to 64nb decompositions (Figure 1a).

Statistical Analysis

Average Discrimination Curves by TFS

Average response calculations were derived for TFS for happy and sad musical stimuli. The average discrimination curves were percent identification of the response based on original categorization (found in TFS for all original stimuli). Curves were analyzed independently for non-musician poor and good performers separated by male and female, in addition to first-year (low education) and last-year (high education) musicians separated by male and female.

Group Differences

Discrimination curves between groups were tested for significance with an ANOVA and follow-up t -test. The discrimination normalized ratio for identification of the stimuli as happy or sad was calculated by determining the percent identification of happy or sad over its opposite stimuli discrimination. Measures of discriminability (D'), and the corrected non-parametric measure of discriminability (A'), were utilized for determining differences in emotional resolvability (Stanislaw and Todorov, 1999; Verde et al., 2006). These measures provide an estimation of signal from noise and determination of the threshold response for the acoustic emotional stimuli (Green, 1960; Swets, 1961, 1986; Swets and Sewall, 1961; Swets et al., 1961). The interest in determining a threshold response for an acoustic emotional stimulus is to

ascertain when an emotional stimulus becomes non-emotional (i.e., threshold). Here, we group averaged identifications; however, individual values were calculated by stimuli.

Normalized Benefit of TFS and ENV

The difference between TFS or ENV emotion discrimination was examined as the ratio between the perception of one emotion (using TFS or ENV) versus the perception of the other emotion (using TFS or ENV). The ratio was calculated by the normalized benefit of TFS or ENV to the emotion discrimination curve (originally calculated for the visual contribution to speech in noise; Sumbly and Pollack, 1954; Meister et al., 2016). The adopted formula was: TFS benefit = $(\text{TFS} - \text{ENV}) / (1 - \text{ENV})$ or ENV benefit = $(\text{ENV} - \text{TFS}) / (1 - \text{TFS})$, to compare both the TFS and ENV contributions, respectively, on the scale of +1 to -1. If the percent difference, normalized by each contribution result, was a positive value, this represented benefit to perception. Conversely, a negative value indicated lack of contribution to perception.

Discriminant Analysis

A canonical discriminant analysis was used to determine the weight of our variables (nb0, nb2, nb4, etc.), which best separated our different groups (poor performer, good performer, first year musician, last year musician). The two generalized canonical discriminant analyses (one for happy and one for sad) were computed using the multivariate linear model $\text{Group} \sim \text{nb0} + \text{nb2} + \text{nb4} + \text{nb8} + \text{nb16} + \text{nb32} + \text{nb64}$ to obtain the weights associated with each variable. This model represents a transformation of the original variables in the space of maximal difference for the group.

RESULTS

Poor and Good Performers

Group Comparison Differences for Poor and Good Performers

Percent identification for poor ($df = 6, 18, F = 30.76, p < 0.0001$) and good ($df = 6, 18, F = 79.04, p < 0.0001$) performers was significantly different by decomposition, indicating a decrease in certainty (Figures 2a,b). Both poor and good performers used TFS for emotion identification, and discrimination performance decreased with increasing band (Supplementary Figures S1a–d). Supplementary Figures S1a, S2b show the average male poor and good identification curves, and Supplementary Figures S1c,d shows the average female poor and good identification curves. For visualization, happy and sad discriminability was represented by identification accuracy with the unaltered excerpt in 20th percentile bands (Figures S2c,d).

Fine Structure and Envelope Identification Differences for Happy or Sad Emotion for Poor and Good Performers

There were no apparent differences for poor performers' TFS- or ENV-based identification of happy or sad emotion. For example, good performers did not show differences for happy TFS ($df = 1, 6, F = 2.545, p = 0.1617$), sad TFS ($df = 1, 6, F = 1.494, p = 0.2674$),

happy ENV ($df = 1, 6, F = 1.897, p = 0.2176$), and sad ENV ($df = 1, 6, F = 2.885, p = 0.1403$). For good performers, happy in TFS ($df = 1, 6, F = 7.749, p = 0.0318$) and sad in ENV ($df = 1, 6, F = 7.591, p = 0.0331$) were different between male and females. Sad in TFS was significantly different between poor and good performers ($df = 6, 30, F = 3.773, p = 0.0091$), but happy in TFS did not reach significance ($df = 6, 30, F = 1.912, p = 0.1218$). Happy in ENV was significantly different between poor and good performers ($df = 1, 6, F = 3.630, p = 0.0110$), but sad in ENV did not reach significance ($df = 1, 6, F = 1.881, p = 0.1273$).

Discrimination Curves for Poor and Good Performers

Non-musician good and poor performers were significantly different for uncertain emotion discrimination ($df = 15, 90, F = 1.814, p = 0.0445$). Good and poor performers used TFS to discriminate happy and sad uncertain emotions differently, by approximately $4.01 \pm 3.33\%$ SD and $9.20 \pm 6.82\%$ SD, respectively, depending on the increasing uncertainty of stimuli (Supplementary Figures S1a–d and Discrimination curves). Males and females, irrespective of poor or good performers, used TFS to discriminate happy and sad uncertain emotions differently, by approximately $2.67 \pm 1.62\%$ SD and $2.06 \pm 1.67\%$ SD, respectively, depending on the increasing uncertainty of stimuli (Supplementary Figures S1a–d, Discrimination curves).

Average Discriminability A' and Benefit of TFS and ENV for Poor and Good Performers

The corrected non-parametric measure of discriminability (A') was used for determining differences in emotional resolvability, and no significant difference between poor performers sad ($p = 0.5957$) and happy ($p = 0.6612$) was found. Similarly, no difference between good performers sad ($p = 0.6712$) and happy ($p = 0.6644$) was found. Figure 3a demonstrates grouped A' for poor and good performers by uncertain emotion. Figure 3c depicts the averaged normalized benefit of TFS or ENV to the emotion discrimination curve for poor and good performers. The TFS was beneficial for emotion discriminability for poor and good performers for sad and happy stimuli. Poor performers had 0.2928 and 0.2377 benefit for happy and sad, respectively. Good performers had 0.3295 and 0.3355 benefit for happy and sad, respectively. The ENV was negatively beneficial for emotion discrimination for all stimuli for poor or good performers.

First and Last Year Musicians

Group Comparison Differences for First- and Last-Year Musicians

Percent identification for first-year musicians ($df = 6, 18, F = 71.69, p < 0.0001$) and last-year musicians ($df = 6, 18, F = 45.37, p < 0.0001$) were significantly different by decomposition, indicating decrease in certainty (Figures 2a,b). Both first-year and last-year musicians using TFS for emotion identification and discrimination decreased with increasing band (Supplementary Figures S2a–f). First-year ($df = 6, 18, F = 28.29, p < 0.0001$) and last-year ($df = 6, 18, F = 29.04, p < 0.0001$) musicians resolved happy and sad emotion differently. Supplementary Figures S2a,d shows the average first- or last-year musician discrimination curves for

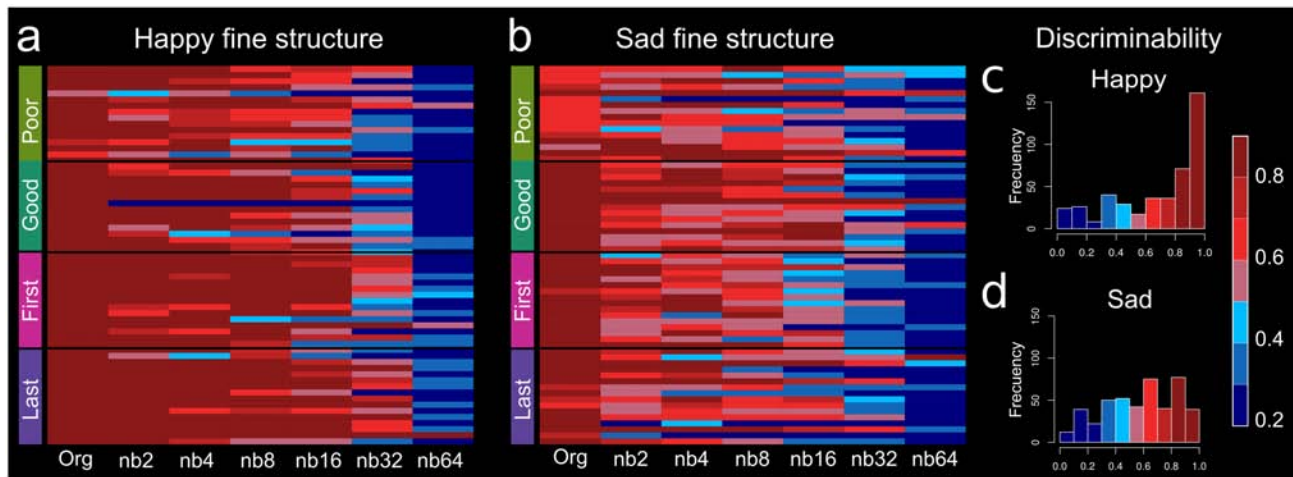


FIGURE 2 | Heatmap for accuracy of response by group and uncertainty. Discrimination heatmap for accuracy concerning uncertain (a) happy fine structure accuracy and (b) sad fine structure accuracy. The darker red color represents greater discriminability of emotion and similarity with the original un-altered excerpt. On the x-axis, band number is represented with Org, unaltered original excerpt, to nb64. Discriminability is represented by identification accuracy with the unaltered excerpt. (c) Happy discriminability and (d) sad discriminability. The color bar represents discriminability 20th percentile bands.

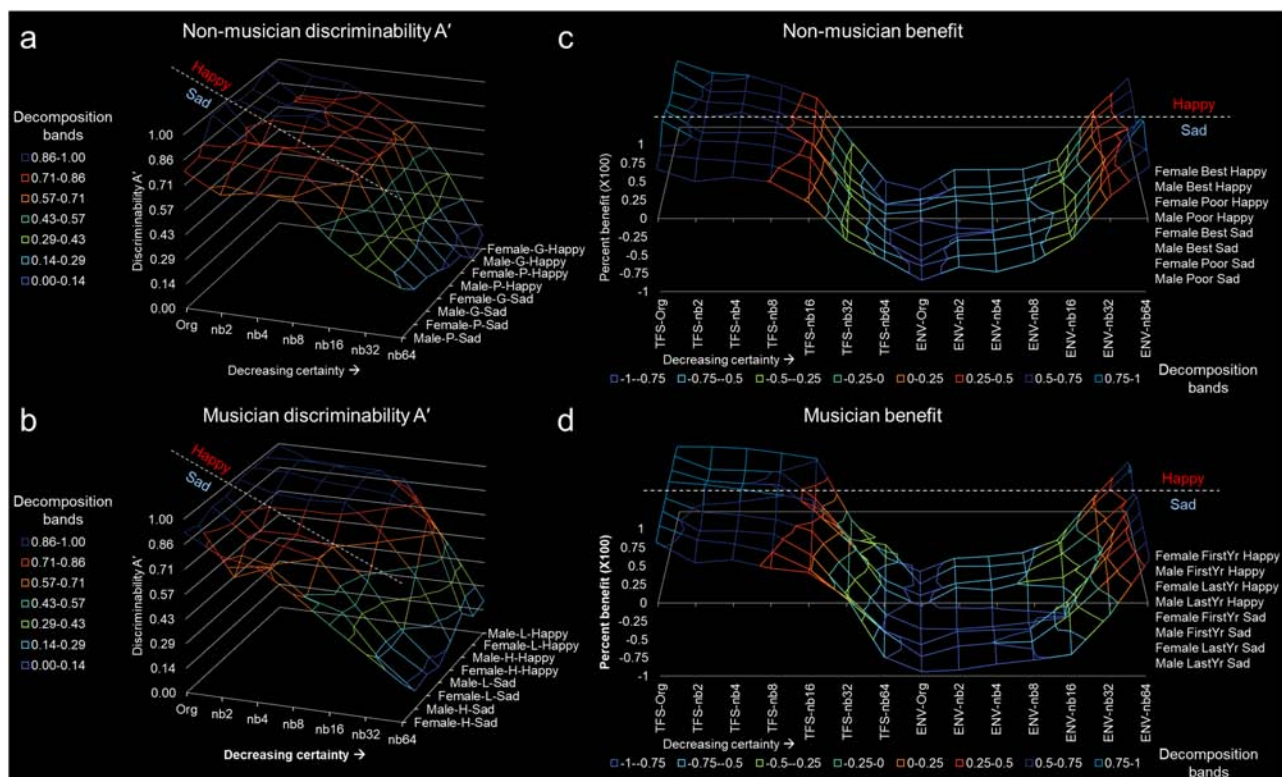


FIGURE 3 | Discriminability A' and TFS/ENV benefit to stimuli perception of uncertain emotion. (a) Non-musician discriminability A' and (b) musician discriminability A'. (c) Non-musician benefit and (d) musician benefit. For all figures, the x-axis represents band number from Org to nb64. The band decompositions were associated in ranges to represent the discriminability. The z-axis represents sex (male or female), type of group (G – good, P – poor, L – first-year musician, H – last-year musician), and emotion type (happy or sad). Here, groups are connected to visualize the trend and pattern more readily. For example panels (a,b), all groups interpret stimuli with greater certainty (i.e., greater discriminability A' shows more blue coloring) for happy more than sad. The dashed line represents the grouping of happy and sad. Note that happy is fuller than sad, indicating high A'. For benefit panels (c,d), note that for TFS there is a fuller line graph occupying more area, indicating that the majority of individuals use fine structure to discriminate uncertain emotion. Note the subtle difference between happy and sad, represented by the dashed line. Musicians benefit from TFS more than non-musicians.

happy and sad stimuli. **Supplementary Figures S2b,c** shows male and female first-year musician discrimination curves and **Supplementary Figures S2e,f** shows male and female last-year musician averaged discrimination curves.

Fine Structure and Envelope Identification Differences for Happy or Sad Emotion for First- and Last-Year Musicians

There were no apparent identification differences for first-year musicians related to TFS happy ($df = 1, 6, F = 3.364, p = 0.1163$), TFS sad ($df = 1, 6, F = 0.1967, p = 0.6729$) or ENV happy ($df = 1, 6, F = 0.1047, p = 0.7573$). Interestingly, first-year musicians used ENV sad significantly more for emotional resolvability ($df = 1, 6, F = 6.058, p = 0.0490$). Last-year musicians were significantly different for emotional resolvability for TFS happy ($df = 1, 6, F = 14.88, p = 0.0084$), TFS sad ($df = 1, 6, F = 11.91, p = 0.0136$), ENV happy ($df = 1, 6, F = 13.66, p = 0.0101$), and ENV sad ($df = 1, 6, F = 14.49, p = 0.0089$). There were significant differences in emotional resolvability between first- and last-year musicians for TFS happy ($df = 6, 18, F = 7.585, p = 0.0017$), TFS sad ($df = 6, 18, F = 4.574, p = 0.0150$), ENV happy ($df = 6, 18, F = 4.966, p = 0.0110$), and ENV sad ($df = 6, 18, F = 7.816, p = 0.0015$).

Discrimination Curves for First- and Last-Year Musicians

When comparing first- and last-year musicians, we found a significant difference in discriminating uncertain emotion ($df = 15, 90, F = 4.377, p < 0.0001$). First- and last-year musicians used TFS to discriminate happy and sad uncertain emotion differently, by approximately $2.51 \pm 1.68\%$ SD and $3.90 \pm 2.30\%$ SD, respectively, depending on the stimuli uncertainty (**Supplementary Figures S1a–f, S2a–f** – Discrimination curves). Males and females, irrespective of first/last year musical education used TFS to discriminate happy and sad uncertain emotion differently, by approximately $12.10 \pm 4.08\%$ SD and $4.24 \pm 3.16\%$ SD, respectively, depending on the increasing uncertainty of stimuli (**Supplementary Figures S2a–f** – Discrimination curves).

Average Discriminability A' and Benefit of TFS and ENV for First- and Last-Year Musicians

The corrected non-parametric measure of discriminability (A'), used for determining differences in emotional resolvability, found a significant difference between first-year and last-musicians (**Figure 3b**). First-year musicians discriminability A' for sad was 0.5794 and last-year musicians discriminability A' for sad was 0.5632. However, first- and last-year musician discriminability A' for happy was 0.7725 and 0.7805, respectively. This represents a discriminability A' difference of 24.99% for first-year and 27.83% for last-year musicians. **Figure 3b** demonstrates discriminability A' for first-year and last-year musicians by uncertain emotion. Note that the blue portion of the curve for sad is missing, indicating a lack of discriminability. **Figure 3d** depicts the averaged normalized benefit of TFS or ENV to the emotion discrimination curve for musicians.

The TFS was beneficial for emotion discriminability for first or last year musicians for sad and happy stimuli. First year musicians had 0.4813 and 0.3013 benefit for happy and sad, respectively. Last-year musicians had 0.4652 and 0.3094 benefit for happy and sad, respectively. The ENV was negatively beneficial for emotion discrimination for all stimuli for musically educated participants.

Poor and Good Performers Compared With First- and Last-Year Musicians Group Comparison Differences and Discriminant Analysis for Non-musicians and Musicians

Identification of happy differed significantly for non-musicians and musicians ($df = 9, 54, F = 15.68, p < 0.0001$) and by emotional resolvability ($df = 6, 54, F = 315.5, p < 0.0001$; **Figure 4a**). Identification of sad differed significantly between non-musicians and musicians ($df = 9, 54, F = 3.526, p = 0.0017$) and by emotional resolvability ($df = 6, 54, F = 112.1, p < 0.0001$; **Figure 4b**). **Figure 4** demonstrates the separation between the different groups.

Group separation and differences were based on emotional resolvability, and **Figure 4** shows the spread of separation weighted by the band decomposition as result of the canonical discriminant analysis. For happy, the greatest discriminant variables between groups were the original stimuli ($Can1 = 0.6264$) and nb64 ($Can1 = -0.7641$). For sad, the greatest discriminant variables were the original stimuli ($Can1 = -1.011$) and nb64 ($Can1 = 0.4193$). For both sad and happy, the greatest discriminant variables were the original stimuli and the hybrid stimulus nb64, indicating that the original and most uncertain stimuli were most discriminable between our groups (**Figure 4**).

Average Discriminability A' of TFS and ENV for Non-musicians and Musicians

The differences in discriminability (A') between non-musicians and musicians (**Figures 3a,b**) for happy were statistically significant ($df = 1.518, 9.109, F = 8.796, p = 0.0101$), as were differences in A' as a function of emotional resolvability ($df = 6, 42, F = 191.7, p < 0.0001$). Differences in A' between non-musicians and musicians (**Figures 3a,b**) for sad were statistically significant ($df = 2.934, 17.61, F = 5.086, p = 0.0107$), as were differences in A' as a function of emotional resolvability ($df = 6, 42, F = 156.1, p < 0.0001$).

Benefit of TFS and ENV for Non-musicians and Musicians

The averaged normalized benefit of happy TFS to the emotion discrimination curve (**Figures 3c,d**) was statistically different between non-musicians and musicians ($df = 2.383, 14.30, F = 6.922, p = 0.0060$), as was the difference in benefit of happy TFS to emotional resolvability ($df = 6, 42, F = 298.5, p < 0.0001$). The averaged normalized benefit of sad TFS to the emotion discrimination curve (**Figures 3c,d**) was not statistically different between non-musicians and musicians ($df = 2.643, 15.86, F = 1.826, p = 0.1870$), although the contribution to emotional resolvability was statistically

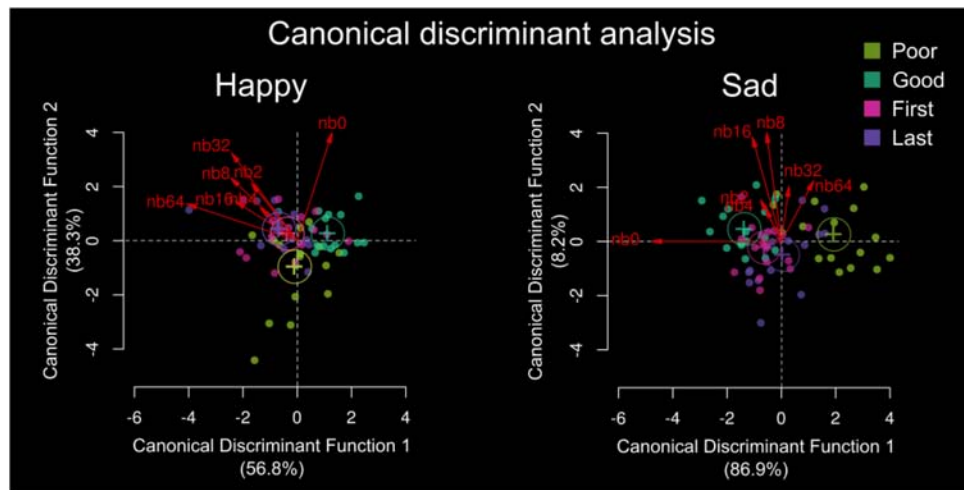


FIGURE 4 | Canonical discriminant analysis for (a) happy and (b) sad uncertain emotion discrimination. Group is sorted by poor performer, good performer, first year music education and last year music education with band number from Org (nb0) to nb64 in red. The plot shows the canonical scores for the groups. Scores are indicated by points and the canonical structure coefficients are indicated by vectors from the origin. Standardized beta coefficients are given for each variable in each discriminant (canonical) function, and the larger the standardized coefficient, the greater the contribution of the respective variable to the discrimination between groups. Here, the discriminant function coefficients denote the unique contribution of each variable to the discriminant function, while the structure coefficients denote the simple correlations between the variables and the functions. For happy, the greatest standardized beta coefficients for org was Can1 = 0.6264, and for nb64 was Can1 = -0.7641. For sad, the greatest standardized beta coefficients for org were Can1 = -1.011, and for nb64 was Can1 = 0.4193.

significant ($df = 6, 42, F = 333.0, p < 0.0001$). The averaged normalized benefit of happy ENV to the emotion discrimination curve (**Figures 3c,d**) was not statistically different between non-musicians and musicians ($df = 2.638, 15.83, F = 1.708, p = 0.2087$), although the contribution to emotional resolvability was statistically significant ($df = 6, 42, F = 342.1, p < 0.0001$). The averaged normalized benefit of sad ENV to the emotion discrimination curve (**Figures 3c,d**) was statistically different between non-musicians and musicians ($df = 2.909, 17.45, F = 10.05, p = 0.0005$), as was the difference in benefit of sad ENV to emotional resolvability ($df = 6, 42, F = 450.0, p < 0.0001$).

DISCUSSION

The objective of the present experiment was to investigate certain and uncertain emotion in musical sounds, and determine if non-musicians and musicians resolve emotion differently. Here, stimuli that varied in emotional certainty were presented in a happy/sad interval forced-choice discrimination psychophysical task. There are three results of considerable interest: *First*, TFS information was essential to discriminating emotion in sound. *Second*, different emotional resolvability curves were found to depend on whether participants were poor or good performers and on year of musical education. *Third*, non-musicians used less TFS and had reduced emotional resolvability curves compared to musicians. The aim of the present experiment was to understand the cues used to resolve emotional signals at threshold and how they differ between non-musicians and musicians.

Resolving Emotion Using Psychoacoustic Cues

Emotion in sound is transmitted through TFS and ENV modulations. In a groundbreaking study, Lieberman and Michaels (1962) found that pitch aided in the identification of emotional content by 44% while amplitude cues added only 3% more. The present study found TFS cues essential and beneficial to discriminating emotion in musical excerpts, whether individuals were good performers, poor performers, or had musical training. The results indicate that TFS cues are required for resolving emotion in sound and individuals differ in their perceptive ability to discriminate these cues. Furthermore, happy emotion was discriminated with higher accuracy than sad emotion for all groups (**Figures 2a, 3 and Supplementary Figures S1, S2**). This is most likely due to individuals using major mode and the fast tempo of tones for discriminating emotion in sound, which are prominent in happy stimuli (Peretz et al., 2001; Hopyan et al., 2016). However, how these cues determine specific emotions and are perceived by individuals is not completely understood. For example, individuals differ in their tendency to report the co-occurrence of discrete emotions of the same valence (Barrett, 1998). Individuals vary in the extent to which they distinguish between like-valence discrete emotions or did not distinguish between like-valence emotions when reporting on their subjective experience (Barrett, 1998). The results indicate that individuals are reporting several affective states together, or it may indicate they are not distinguishing between distinct emotional states. The aforementioned manuscript (Barrett, 1998) bolstered support for both the theory of discrete emotion, where individuals label emotions based on determining a subjective level of arousal,

and the dimensional theory of emotion, where individuals focus on the subjective emotional experience dimensionalized by valence, arousal, and intensity of the affective state (Russell, 1980; Barrett, 1998). The present study found emotional resolvability changed as a function of altering the TFS content of the musical excerpt, revealing that an essential cue to discriminating emotion is fine structure. Recently, a study investigating the similarities/dissimilarities of emotion in music and speech prosody found that the psychoacoustic features implicated were loudness, tempo and speech rate, melodic and prosodic contour, spectral centroid, and sharpness (Coutinho and Dibben, 2013). In contrast, the features distinct to music and speech were spectral flux and roughness, respectively. Here, the authors indicated that emotional cues in sound are encoded as psychoacoustic spatiotemporal patterns, which for music and speech rely heavily on their “shared acoustic profiles” (Coutinho and Dibben, 2013). We encourage research into determining what constitutes an emotion from non-emotion sound (Pfeifer, 1988; Barrett, 1998; LeDoux, 2000; Phan et al., 2002; Wildgruber et al., 2005; Koelsch, 2014), which will enable a more thorough classification of the neurobiology of emotion. Future studies should further explore the psychoacoustic foundations of emotion.

Musicians Compared With Non-musicians

Musicians discriminate emotion differently, likely due to their unique education. For example, in the speech-in-noise and hearing-in-noise test, musicians perform significantly better than non-musicians, derived in part from musicians’ enhanced working memory and frequency discrimination (Parbery-Clark et al., 2009a,b). Musicians in the present study discriminated emotion in sound differently than non-musicians by using more TFS through each of the nb decompositions. Within the group of musicians, last-year musically educated individuals discriminated happy or sad excerpts somewhat differently than first-year musically educated individuals. Although the greatest difference was in male and female musicians, the difference between musicians and non-musicians reveals the most, as musicians benefited more from TFS components. For example, musicians discriminating happy or sad excerpts utilized more TFS irrespective of whether individuals were in the first or last year of their music education. Recent discrimination tasks bolster these results. In a study where participants were tasked to detect frequency changes in quiet and noisy conditions, the acoustic change complex, a type of late auditory evoked potential, showed a larger P2’ amplitude in musicians than in non-musicians (Liang et al., 2016). Moreover, in a task where target speech and competing speech were presented with either their natural F0 contours or on a monotone F0, and with F0 difference between the target and masker systematically varied, F0 discrimination was significantly better for musicians (Madsen et al., 2017). Most of these frequency discrimination tasks indicate that musicians have an enhanced ability to perceive or discriminate TFS or fine structure components. Future studies should expand the range and variety of emotion discrimination paradigms, to explore differences between musicians and non-musicians.

Study Limitations and Future Directions

The present study concentrated on analyzing two emotions elicited by classical music excerpts. Constraining the variety of emotion to a dichotomous task is artificial, but aides in discerning how the basic components of sound cue emotion. Future studies should analyze the diverse emotional repertoire that exists in humans. The present study analyzed non-musicians and individuals with musical education. We chose these groups based on previous literature (Micheyl et al., 2006; Musacchia et al., 2007; Wong et al., 2007; Parbery-Clark et al., 2009a,b; Kraus and Chandrasekaran, 2010; Strait et al., 2010; Başkent and Gaudrain, 2016; Liang et al., 2016; Mandikal Vasuki et al., 2016; Coffey et al., 2017; Madsen et al., 2017), suggesting music training would influence acoustic perception in the emotional resolvability task. This manuscript found that years in music education significantly affected emotional resolvability ($F_{15,90} = 4.377$, $p < 0.0001$), with advanced musicians using more fine structure to discriminate happy uncertain emotion by $2.51 \pm 1.68\%$. Future studies could analyze different musicians (piano versus string) to determine if emotional resolvability differences are associated with the type of instrument training. The present manuscript assessed males and females separately, as it is known that gender differences in the perception of non-target emotions (incorrect) are greater for men than women (Fischer et al., 2018). Further, our entire cohort of subjects ($n = 64$) was derived from a Spanish speaking population. In this regard, identifying emotion is easier for listening individuals with similar cultural and language backgrounds (Waaramaa and Leisiö, 2013) and a second language is known to interfere with emotion recognition from speech prosody (Bhatara et al., 2016). The repertoire of music used in the present study was classical, which was familiar to all participants; therefore, we believe the effect due to cultural background should be minimal. Future studies could further assess these confounding variables to determine their affect on uncertain musical emotion.

DATA AVAILABILITY

Data and sound files used in this work can be downloaded in an anonymized format from the Open Science Framework: Manno, Francis A. M. 2018. “Music Psychophysics.” OSF. November 20. <https://osf.io/8ws7a>.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Internal Review Board of the Instituto de Neurobiología, Universidad Nacional Autónoma de México. The participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

FM designed the experiments, interviewed participants, executed experiments, analyzed the data, and wrote the manuscript.

FM and RC assisted in experimental design and analysis, and revised the document. CL and FB supervised the experiments, assisted with data curation and analysis, and assisted in writing the manuscript.

FUNDING

We thank the Consejo Nacional de Ciencia y Tecnología (CONACyT) for the funding received via the grant CB255462 to FB.

ACKNOWLEDGMENTS

The authors are grateful to L. González-Santos, M.Sc., Dr. S. Hernández-Cortés Manno, and Chair Professor S. H. Cheng for support, and Z. Gracia-Tabuenca, Jorge Gámez, and Prof. F. King for their comments, coding, statistical expertise, and mathematical assistance. The authors thank J. G. Norris for editing the manuscript. The authors are also grateful to the students of the Facultad de Bellas Artes and the Facultad de Ciencias Naturales of the Universidad Autónoma de Querétaro for their participation in the study.

REFERENCES

- Apoux, F., Yoho, S. E., Youngdahl, C. L., and Healy, E. W. (2013). Role and relative contribution of temporal envelope and fine structure cues in sentence recognition by normal-hearing listeners. *J. Acoust. Soc. Am.* 134, 2205–2212. doi: 10.1121/1.4816413
- Barrett, L. F. (1998). Discrete emotions or dimensions? The role of valence focus and arousal focus. *Cogn. Emot.* 12, 579–599. doi: 10.1080/026999398379574
- Bağkent, D., and Gaudrain, E. (2016). Musician advantage for speech-on-speech perception. *J. Acoust. Soc. Am.* 139, EL51–EL56. doi: 10.1121/1.4942628
- Bhatara, A., Laukka, P., Boll-Avetisyan, N., Granjon, L., Anger Elfenbein, H., and Bänziger, T. (2016). Second language ability and emotional prosody perception. *PLoS One* 11:e0156855. doi: 10.1371/journal.pone.0156855
- Binder, J. R., Liebenthal, E., Possing, E. T., Medler, D. A., and Ward, B. D. (2004). Neural correlates of sensory and decision processes in auditory object identification. *Nat. Neurosci.* 7, 295–301. doi: 10.1038/nn1198
- Coffey, E. B. J., Mogilever, N. B., and Zatorre, R. J. (2017). Speech-in-noise perception in musicians: a review. *Hear. Res.* 352, 49–69. doi: 10.1016/j.heares.2017.02.006
- Coutinho, E., and Dikken, N. (2013). Psychoacoustic cues to emotion in speech prosody and music. *Cogn. Emot.* 27, 658–684. doi: 10.1080/02699931.2012.732559
- Davidson, S. A., Gilkey, R. H., Colburn, H. S., and Carney, L. H. (2009). Diotic and dichotic detection with reproducible chimeric stimuli. *J. Acoust. Soc. Am.* 126, 1889–1905. doi: 10.1121/1.3203996
- Drullman, R. (1995). Temporal envelope and fine structure cues for speech intelligibility. *J. Acoust. Soc. Am.* 97, 585–592. doi: 10.1121/1.413112
- Drullman, R., Festen, J. M., and Houtgast, T. (1996). Effect of temporal modulation reduction on spectral contrasts in speech. *J. Acoust. Soc. Am.* 99(4 Pt 1), 2358–2364. doi: 10.1121/1.415423
- Drullman, R., Festen, J. M., and Plomp, R. (1994). Effect of temporal envelope smearing on speech reception. *J. Acoust. Soc. Am.* 95, 1053–1064. doi: 10.1121/1.408467

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2019.00902/full#supplementary-material>

FIGURE S1 | Discrimination curves and discrimination normalized ratio indices for non-musicians. Discrimination curves of percent identification of happy or sad stimuli. (a) Average male poor performers, (b) average male good performers, (c) average female poor performers, and (d) average female good performers. Sad curve represented with blue line and happy curve represented with red line. The discrimination normalized ratio of stimuli identification of happy or sad stimuli. (e) Male poor performers, (f) male good performers, (g) female poor performers, and (h) female good performers. Sad discrimination ratios are represented in red and happy discrimination ratios are represented in blue.

FIGURE S2 | Discrimination curves and discrimination normalized ratio indices for musicians in their first or last year of study. Discrimination curves of percent identification of happy or sad stimuli. (a) Average first year (low) and (d) average last year music education (high). Male (b) and female (c) first year and male (e) and female (f) last year music education. The discrimination normalized ratio of stimuli identification of happy or sad stimuli (g) through (j). Sad curve represented with blue line and happy curve represented with red line. (g) Male first year and (h) male last year. (i) Female first year and (j) female last year.

TABLE S1 | Montreal emotional identification task acoustic stimuli.

DATA SHEET S1 | Contains extended analyses.

DATA SHEET S2 | Contains the original data used for all analyses in a csv file.

- Eerola, T. (2012). Modeling listeners' emotional response to music. *Top. Cogn. Sci.* 4, 607–624. doi: 10.1111/j.1756-8765.2012.01188.x
- Eerola, T., and Vuoskoski, J. K. (2011). A comparison of the discrete and dimensional models of emotion in music. *Psychol. Music* 39, 18–49. doi: 10.1093/scan/nsv032
- Fairbanks, G. (1940). Recent experimental investigations of vocal pitch in speech. *J. Acoust. Soc. Am.* 11, 457–466. doi: 10.1121/1.3589249
- Fairbanks, G., and Pronovost, W. (1938). Vocal pitch during simulated emotion. *Science* 88, 382–383. doi: 10.1126/science.88.2286.382
- Fischer, A. H., Kret, M. E., and Broekens, J. (2018). Gender differences in emotion perception and self-reported emotional intelligence: a test of the emotion sensitivity hypothesis. *PLoS One* 13:e0190712. doi: 10.1371/journal.pone.0190712
- Fitch, W. T. (2006). The biology and evolution of music: a comparative perspective. *Cognition* 100, 173–215. doi: 10.1016/j.cognition.2005.11.009
- Fogerty, D. (2011). Perceptual weighting of individual and concurrent cues for sentence intelligibility: frequency, envelope, and fine structure. *J. Acoust. Soc. Am.* 129, 977–988. doi: 10.1121/1.3531954
- Frick, R. W. (1985). Communicating emotion: the role of prosodic features. *Psychol. Bull.* 97, 412–429. doi: 10.1037/0033-2909.97.3.412
- Gabrielsson, A., and Lindström, E. (2010). "The role of structure in musical expression of emotions," in *Handbook of Music and Emotion: Theory, Research, Applications*, Chap. 14, eds P. N. Juslin and S. Sloboda (Oxford: Oxford University Press).
- Gingras, B., Marin, M. M., and Fitch, W. T. (2014). Beyond intensity: spectral features effectively predict music-induced subjective arousal. *Q. J. Exp. Psychol.* 67, 1428–1446. doi: 10.1080/17470218.2013.863954
- Gosselin, N., Peretz, I., Hasboun, D., Baulac, M., and Samson, S. (2011). Impaired recognition of musical emotions and facial expressions following anteromedial temporal lobe excision. *Cortex* 47, 1116–1125. doi: 10.1016/j.cortex.2011.05.012
- Gosselin, N., Peretz, I., Johnsen, E., and Adolphs, R. (2007). Amygdala damage impairs emotion recognition from music. *Neuropsychologia* 45, 236–244. doi: 10.1016/j.neuropsychologia.2006.07.012

- Green, D. M. (1960). Psychoacoustics and detection theory. *J. Acoust. Soc. Am.* 32, 1189–1203. doi: 10.1121/1.1907882
- Hodges, D. A. (2010). “Psychophysical measures,” in *Handbook of Music and Emotion: Theory, Research, Applications*, Chap. 11, eds P. N. Juslin and S. Sloboda (Oxford: Oxford University Press).
- Hoppyan, T., Manno, F. A., Papsin, B. C., and Gordon, K. A. (2016). Sad and happy emotion discrimination in music by children with cochlear implants. *Child Neuropsychol.* 22, 366–380. doi: 10.1080/09297049.2014.992400
- Juslin, P. N., and Västfjäll, D. (2008). Emotional responses to music: the need to consider underlying mechanisms. *Behav. Brain Sci.* 31, 559–575. doi: 10.1017/S0140525X08005293
- Khalifa, S., Guye, M., Peretz, I., Chapon, F., Girard, N., Chauvel, P., et al. (2008). Evidence of lateralized anteromedial temporal structures involvement in musical emotion processing. *Neuropsychologia* 46, 2485–2493. doi: 10.1016/j.neuropsychologia.2008.04.009
- King, F. W. (2009). *Hilbert Transforms*. Cambridge: Cambridge University Press.
- Koelsch, S. (2014). Brain correlates of music-evoked emotions. *Nat. Rev. Neurosci.* 15, 170–180. doi: 10.1038/nrn3666
- Kotz, S. A., Ravignani, A., and Fitch, W. T. (2018). The evolution of rhythm processing. *Trends Cogn. Sci.* 22, 896–910. doi: 10.1016/j.tics.2018.08.002
- Kraus, N., and Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nat. Rev. Neurosci.* 11, 599–605. doi: 10.1038/nrn2882
- LeDoux, J. E. (2000). Emotion circuits in the brain. *Annu. Rev. Neurosci.* 23, 155–184. doi: 10.1146/annurev.neuro.23.1.155
- Leinonen, L., Hiltunen, T., Linnankoski, I., and Laakso, M.-L. (1997). Expression of emotional–motivational connotations with a one-word utterance. *J. Acoust. Soc. Am.* 102, 1853–1863. doi: 10.1121/1.420109
- Liang, C., Earl, B., Thompson, I., Whitaker, K., Cahn, S., Xiang, J., et al. (2016). Musicians are better than non-musicians in frequency change detection: behavioral and electrophysiological evidence. *Front. Neurosci.* 10:464. doi: 10.3389/fnins.2016.00464
- Licklider, J. C. (1946). Effects of amplitude distortion upon the intelligibility of speech. *J. Acoust. Soc. Am.* 18, 429–434. doi: 10.1121/1.1916383
- Lieberman, P., and Michaels, S. B. (1962). Some aspects of fundamental frequency and envelope amplitude as related to the emotional content of speech. *J. Acoust. Soc. Am.* 34, 922–927. doi: 10.1121/1.1918222
- Lieberman, P., Michaels, S. B., and Soron, H. I. (1964). Emotional effects of fundamental-frequency transformations. *J. Acoust. Soc. Am.* 36, 1048–1048. doi: 10.1121/1.2143375
- Madsen, S. M. K., Whiteford, K. L., and Oxenham, A. J. (2017). Musicians do not benefit from differences in fundamental frequency when listening to speech in competing speech backgrounds. *Sci. Rep.* 7:12624.
- Mandikal Vasuki, P. R., Sharma, M., Demuth, K., and Arciuli, J. (2016). Musicians' edge: a comparison of auditory processing, cognitive abilities and statistical learning. *Hear. Res.* 342, 112–123. doi: 10.1016/j.heares.2016.10.008
- Manno, F. A. M., Rodríguez-Cruces, R., Lau, C., Barrios, F. A. (2018). *Music Psychophysics*. Available at: <https://osf.io/8ws7a> (accessed November 20, 2018).
- Manno, F. A. M., Fernandez-Ruiz, J., Manno, S. H. C., Cheng, S. H., Lau, C., and Barrios, F. A. (2019). Sparse sampling of silence type I errors with an emphasis on primary auditory cortex. *Front. Neurosci.* 13:156. doi: 10.3389/fnins.2019.00516
- Meister, H., Fuersen, K., Schreitmüller, S., and Walger, M. (2016). Effect of acoustic fine structure cues on the recognition of auditory-only and audiovisual speech. *J. Acoust. Soc. Am.* 139:3116. doi: 10.1121/1.4953022
- Michéyl, C., Delhommeau, K., Perrot, X., and Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hear. Res.* 219, 36–47. doi: 10.1016/j.heares.2006.05.004
- Moon, I. J., Won, J. H., Park, M. H., Ives, D. T., Nie, K., Heinz, M. G., et al. (2014). Optimal combination of neural temporal envelope and fine structure cues to explain speech identification in background noise. *J. Neurosci.* 34, 12145–12154. doi: 10.1523/JNEUROSCI.1025-14.2014
- Musacchia, G., Sams, M., Skoe, E., and Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proc. Natl. Acad. Sci. U.S.A.* 104, 15894–15898. doi: 10.1073/pnas.0701498104
- Oswald, J. (1956). The theory of analytic band-limited signals applied to carrier systems. *IRE Trans. Circuit Theory* 3, 244–251. doi: 10.1109/tct.1956.1086333
- Parbery-Clark, A., Skoe, E., and Kraus, N. (2009a). Musical experience limits the degradative effects of background noise on the neural processing of sound. *J. Neurosci.* 29, 14100–14107. doi: 10.1523/JNEUROSCI.3256-09.2009
- Parbery-Clark, A., Skoe, E., Lam, C., and Kraus, N. (2009b). Musician enhancement for speech-in-noise. *Ear. Hear.* 30, 653–661. doi: 10.1097/AUD.0b013e3181b412e9
- Peretz, I., Blood, A. J., Penhune, V., and Zatorre, R. (2001). Cortical deafness to dissonance. *Brain* 124(Pt 5), 928–940. doi: 10.1093/brain/124.5.928
- Peretz, I., Gagnon, L., and Bouchard, B. (1998). Music and emotion: perceptual determinants, immediacy, and isolation after brain damage. *Cognition* 68, 111–141. doi: 10.1016/s0010-0277(98)00043-2
- Pfeifer, R. (1988). “Artificial intelligence models of emotion,” in *Cognitive Perspectives on Emotion and Motivation*, Vol. 44, eds V. Hamilton, G. H. Bower, and N. H. Frijda (Dordrecht: Springer), 287–320. doi: 10.1007/978-94-009-2792-6_12
- Phan, K. L., Wager, T., Taylor, S. F., and Liberzon, I. (2002). Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage* 16, 331–348. doi: 10.1006/nimg.2002.1087
- Russell, J. A. (1980). A circumplex model of affect. *J. Pers. Soc. Psychol.* 39, 1161–1178. doi: 10.1037/h0077714
- Scherer, K. R. (1995). Expression of emotion in voice and music. *J. Voice.* 9, 235–248. doi: 10.1016/s0892-1997(05)80231-0
- Schubert, E. (2004). Modeling perceived emotion with continuous musical features. *Music Perception* 21, 561–585. doi: 10.1525/mp.2004.21.4.561
- Schubert, E. (2013). Emotion felt by the listener and expressed by the music: literature review and theoretical perspectives. *Front. Psychol.* 17:837. doi: 10.3389/fpsyg.2013.00837
- Seymour, B., and Dolan, R. (2008). Emotion, decision making, and the amygdala. *Neuron* 58, 662–671. doi: 10.1016/j.neuron.2008.05.020
- Shamma, S., and Lorenzi, C. (2013). On the balance of envelope and temporal fine structure in the encoding of speech in the early auditory system. *J. Acoust. Soc. Am.* 133, 2818–2833. doi: 10.1121/1.4795783
- Shannon, R. V., Zeng, F. G., Kamath, V., Wyganski, J., and Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science* 270, 303–304. doi: 10.1126/science.270.5234.303
- Smith, Z. M., Delgutte, B., and Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature* 416, 87–90. doi: 10.1038/416087a
- Stanislaw, H., and Todorov, N. (1999). Calculation of signal detection theory measures. *Behav. Res. Methods Instrum. Comput.* 31, 137–149. doi: 10.3758/bf03207704
- Strait, D. L., Kraus, N., Parbery-Clark, A., and Ashley, R. (2010). Musical experience shapes top-down auditory mechanisms: evidence from masking and auditory attention performance. *Hear. Res.* 261, 22–29. doi: 10.1016/j.heares.2009.12.021
- Sumbly, W. H., and Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *J. Acoust. Soc. Am.* 26:212. doi: 10.1121/1.1907309
- Swaminathan, J., and Heinz, M. G. (2012). Psychophysiological analyses demonstrate the importance of neural envelope coding for speech perception in noise. *J. Neurosci.* 32, 1747–1756. doi: 10.1523/JNEUROSCI.4493-11.2012
- Swaminathan, J., Reed, C. M., Desolge, J. G., Braidia, L. D., and Delhorne, L. A. (2014). Consonant identification using temporal fine structure and recovered envelope cues. *J. Acoust. Soc. Am.* 135, 2078–2090. doi: 10.1121/1.4865920
- Swets, J. A. (1961). Is there a sensory threshold? *Science* 134, 168–177.
- Swets, J. A. (1986). Indices of discrimination or diagnostic accuracy: their ROCs and implied models. *Psychol. Bull.* 99, 100–117. doi: 10.1037/0033-2909.99.1.100
- Swets, J. A., and Sewall, S. T. (1961). Stimulus vs response uncertainty in recognition. *J. Acoust. Soc. Am.* 33, 1586–1592. doi: 10.1121/1.1908507
- Swets, J. A., Tanner, W. P., and Birdsall, T. G. (1961). Decision processes in perception. *Psychol. Rev.* 68, 301–340.
- Verde, M. F., Macmillan, N. A., and Rotello, C. M. (2006). Measures of sensitivity based on a single hit rate and false alarm rate: the accuracy, precision, and robustness of d' , A_z , and A' . *Percept. Psychophys.* 68, 643–654. doi: 10.3758/bf03208765
- Waaramaa, T., and Leisö, T. (2013). Perception of emotionally loaded vocal expressions and its connection to responses to music. A cross-cultural

- investigation: Estonia, Finland, Sweden, Russia, and the USA. *Front. Psychol.* 4:344. doi: 10.3389/fpsyg.2013.00344
- Wildgruber, D., Riecker, A., Hertrich, I., Erb, M., Grodd, W., Ethofer, T., et al. (2005). Identification of emotional intonation evaluated by fMRI. *Neuroimage* 24, 1233–1241. doi: 10.1016/j.neuroimage.2004.10.034
- Williams, C. E., and Stevens, K. N. (1972). Emotions and speech: some acoustical correlates. *J. Acoust. Soc. Am.* 52, 1238–1250. doi: 10.1121/1.1913238
- Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., and Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat. Neurosci.* 10, 420–422. doi: 10.1038/nn1872
- Zeng, F. G., Nie, K., Liu, S., Stickney, G., Del Rio, E., Kong, Y. Y., et al. (2004). On the dichotomy in auditory perception between temporal envelope and fine structure cues. *J. Acoust. Soc. Am.* 116, 1351–1354. doi: 10.1121/1.1777938
- Zeng, F. G., Nie, K., Stickney, G. S., Kong, Y. Y., Vongphoe, M., Bhargave, A., et al. (2005). Speech recognition with amplitude and frequency modulations. *Proc. Natl. Acad. Sci. U.S.A.* 102, 2293–2298. doi: 10.1073/pnas.0406460102
- Conflict of Interest Statement:** 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.

Copyright © 2019 Manno, Cruces, Lau and Barrios. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Auditory Categorization of Man-Made Sounds Versus Natural Sounds by Means of MEG Functional Brain Connectivity

Vasiliki Salvari¹, Evangelos Paraskevopoulos^{1,2}, Nikolas Chalas², Kilian Müller¹, Andreas Wollbrink¹, Christian Dobel³, Daniela Korth³ and Christo Pantev^{1*}

¹ Institute for Biomagnetism and Biosignalanalysis, University of Münster, Münster, Germany, ² School of Medicine, Faculty of Health Sciences, Aristotle University of Thessaloniki, Thessaloniki, Greece, ³ Department of Otorhinolaryngology, Friedrich-Schiller University of Jena, Jena, Germany

OPEN ACCESS

Edited by:

Claude Alain,
Rotman Research Institute (RRI),
Canada

Reviewed by:

Jochen Kaiser,
Goethe University Frankfurt, Germany
Fernando R. Nodal,
University of Oxford, United Kingdom

*Correspondence:

Christo Pantev
pantev@uni-muenster.de

Specialty section:

This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Neuroscience

Received: 10 June 2019

Accepted: 19 September 2019

Published: 04 October 2019

Citation:

Salvari V, Paraskevopoulos E, Chalas N, Müller K, Wollbrink A, Dobel C, Korth D and Pantev C (2019) Auditory Categorization of Man-Made Sounds Versus Natural Sounds by Means of MEG Functional Brain Connectivity. *Front. Neurosci.* 13:1052. doi: 10.3389/fnins.2019.01052

Previous neuroimaging studies have shown that sounds can be discriminated due to living-related or man-made-related characteristics and involve different brain regions. However, these studies have mainly provided source space analyses, which offer simple maps of activated brain regions but do not explain how regions of a distributed system are functionally organized under a specific task. In the present study, we aimed to further examine the functional connectivity of the auditory processing pathway across different categories of non-speech sounds in healthy adults, by means of MEG. Our analyses demonstrated significant activation and interconnection differences between living and man-made object sounds, in the prefrontal areas, anterior-superior temporal gyrus (aSTG), posterior cingulate cortex (PCC), and supramarginal gyrus (SMG), occurring within 80–120 ms post-stimulus interval. Current findings replicated previous ones, in that other regions beyond the auditory cortex are involved during auditory processing. According to the functional connectivity analysis, differential brain networks across the categories exist, which proposes that sound category discrimination processing relies on distinct cortical networks, a notion that has been strongly argued in the literature also in relation to the visual system.

Keywords: auditory perception, functional connectivity, magnetoencephalography, sound discrimination, man-made sound, natural sound

INTRODUCTION

Current knowledge on neuronal networks underlying auditory perception remains fragmentary, despite the fact that audition has been extensively studied (Zatorre et al., 2002). The basic network properties that have been suggested for the auditory modality resemble the structure of the visual one, dividing the information processing pathways in dorsal and ventral networks, corresponding to the processing of information to “where” and “what” respectively (Romanski et al., 2000; Kubovy and Van Valkenburg, 2001; Arnott et al., 2004; Huddleston et al., 2008; Asplund et al., 2010). Following the “what” pathway, the physical characteristics of the sound stimulus are initially encoded in the primary and secondary auditory cortex, along with their associative areas, prior to their integration into a more abstract representation (Griffiths and Warren, 2004; Bregman, 2017). Within this pathway, the processing of auditory information seems

to be performed in sound category specific channels (Caramazza and Mahon, 2003). In this line, suggestions have been raised to propose functional specificity for processing different types of sounds (Belin et al., 2000; Zatorre and Belin, 2001; Patterson et al., 2002; Lewis et al., 2004; Zatorre et al., 2004; Hunter et al., 2010).

Attempts to dissociate the processing of different sound categories at the cortical level have been made in brain-lesion case studies (Clarke et al., 2000; Taniwaki et al., 2000; Mendez, 2001; Steinke et al., 2001). Cases like auditory agnosia, represent the impaired ability to recognize sounds, when peripheral hearing is intact. However, this impairment does not necessarily apply to all sound categories; it may rather be category-specific depending on the brain damage. For instance, a patient with focal damage in the right fronto-parietal area was able to identify environmental sounds and name musical instruments, but could not recognize music (Steinke et al., 2001). On the other hand, left temporal lesion or left fronto-temporal ischemia have caused agnosia restricted to environmental sounds (Clarke et al., 2000). It should be emphasized though, that single-case studies of brain-lesioned patients are very heterogeneous and therefore they cannot provide a detailed model of cortical sound processing.

Several functional neuroimaging studies with healthy participants indicate bilateral auditory cortex activation for speech sounds (Belin et al., 2000; Zatorre and Belin, 2001) and right-lateralized activation for non-speech sounds during sound discrimination tasks (for review see Tervaniemi and Hugdahl, 2003). Other regions of the brain, such as the inferior frontal cortex have been also reported to be involved, indicating a network that involves further cognitive functions, beyond the auditory ones (for review see Price, 2012). So far, the majority of current studies are mainly focused in the differential processing of speech versus non-speech sound categories, though we still have poor knowledge about the differential processing within non-speech sound category. The existing studies have shown that sound category discrimination depends more likely on its associated manipulative characteristics (Lewis et al., 2004, 2005; Murray et al., 2006; De Lucia et al., 2009, 2012). For instance, in the context of playing a guitar, we listen to the sound while we perceive motor and visual actions of the guitar playing. Thus, at a higher cognitive level information from all sensory modalities that receive input from a stimulus are integrated in order to construct the percept of the sound. Similar model (Heekeren et al., 2008) has been proposed in the past for the visual and somatosensory system, indicating a multisensory integration already in low level early stages of cognitive processing.

Evidence on the functional organization of auditory perception shows that sounds can be categorized into “living” and “man-made” stimuli (Lewis et al., 2005), suggesting differential brain activation. In particular, a man-made object in comparison to a sound of an animal, might require a top-down mechanism which integrates semantic and multisensory features associated more with action. Similarly, Murray et al. (2006) demonstrated by means of EEG, that “man-made” sounds display stronger brain activation in the auditory “what” pathway compared to the “living” objects, and that regions of the right hemisphere and premotor cortices were mainly involved. Other differentiations have been also reported within the category of man-made objects.

The main idea is that daily used object sounds, such as the phone ringing, might trigger more response of action than a typical tone of a musical instrument (in non-musicians) and, thus stronger brain activation will be elicited (De Lucia et al., 2009). Interestingly, EEG studies focusing on the temporal dynamics have shown that the category discrimination process occurs around the N1 component, already 70 ms after the stimulus onset (Murray et al., 2006; De Lucia et al., 2012).

Nevertheless, the way distinct networks operate across different categories of sounds is still poorly understood. Although, the aforementioned studies have given some insights of when and where this differentiation appears, the source space analyses offer simple maps of activated brain regions, rather than indicating how these regions of a distributed system are functionally connected to execute a specific task. Up to date, the investigation of complex networks has been developed methodologically, given the opportunity to study cortical reorganization underpinning associated cognitive processes (Rogers et al., 2007; Bullmore and Sporns, 2009). Therefore, in the current study we aimed to further investigate the functional connectivity of the auditory processing pathway across different non-speech sound categories. The cortical responses of three different categories of sounds were compared. Namely, the *Musical* and the *Artificial* category (sounds of daily used/heard objects), representing the man-made-objects sound categories and the *Natural* category (mainly animal vocalizations). According to our knowledge, this is the first study to investigate functional connectivity across different non-speech sound categories by means of magnetoencephalography (MEG), which has high spatial resolution and excellent temporal resolution. Taking into consideration the existing literature about living versus man-made-related sounds, we expected that the *Musical* and the *Artificial* sounds would demonstrate stronger cortical responses, that would involve motor related regions and significant interconnections among these regions in comparison to the *Natural* sounds. Further, it should be noted that although the *Musical* and the *Artificial* sounds belong to the man-made category, evidence for differential activation between these groups has been previously reported (De Lucia et al., 2009) as a function of daily use. The N1 auditory evoked field was *a priori* set as the time interval of interest based on previous electrophysiological findings that report early responses in the sound category discrimination processing (Murray et al., 2006; De Lucia et al., 2009, 2012).

MATERIALS AND METHODS

Subjects

The current study was conducted with a sample of 20 young adults (mean age = 27.19, SD = 5.59, 8 males). They were recruited from the pool of subjects of our institute among those who had normal hearing, according to a clinical audiometric evaluation. All subjects were right handed, according to the Edinburgh Handedness Inventory (Oldfield, 1971). The participants were informed about the aim of the study and the ones willing to participate were provided with a consent

form that ensured the confidentiality of their identity. The study was performed according to the Declaration of Helsinki and approved by the ethics committee of the Medical faculty of the University of Münster.

Stimuli

The stimuli consisted of three different categories of sounds: *Natural*, *Musical*, and *Artificial*. The *Natural* and *Artificial* sounds were recordings obtained from online sound databases (Free Sounds Effects¹; SoundBible²; ZapSplat³). The *Musical* sounds were obtained from “McGill University Master samples” sound bank that have been created for perceptual research related to the psychology of music. The Audacity software⁴ was used to resample all sounds at 44,100 Hz and to implement onset/offset linear slopes of 20 ms. Then, the mono sounds were converted into stereo sounds. By means of the WavePad Sound Editor⁵ they were normalized by –10 dB RMS based on the Average Loudness normalization method.

The stimulus paradigm was performed via Presentation software (Version 18.0, Neurobehavioral systems, Inc., Berkeley, CA, United States)⁶. It consisted of two blocks with a short break in between. Each block included the presentation of the three different categories of sounds that were pseudo-randomly presented across blocks and across subjects, whereas the sounds of each category were presented always in the same order: A = *Artificial*, M = *Musical* and N = *Natural*; Block1: A1-A2-...-An-M1-M2-...-Mn-N1-N2-...-Nn; Block2: M1-M2-...-Mn-N1-N2-...-Nn-A1-A2-...-An. Each block contained 144 stimuli, 48 for each category, that makes a total of 288 stimuli for the whole experiment; 48 (sounds per category) × 3 (categories) × 2 (blocks). Each block contained 144 stimuli, 48 for each category, that makes a total of 288 stimuli for the whole experiment; 48 (sounds per category) × 3 (categories) × 2 (blocks). Each stimulus lasted for 1 s with a randomized Inter-Stimulus Interval (ISI) between 0.7 and 1.3 s, in order to avoid expectancy and rhythmicity. The *Natural* sounds contained sounds of living objects. The *Musical* sounds contained notes of different musical instruments, whereas the *Artificial* sounds were daily object-like sounds. Examples of the sound files used in the study can be found in the **Supplementary Material**.

MEG Recordings

Participants were examined in a magnetically shielded and acoustically quiet room by means of 275 channel whole-head system (OMEGA 275, CTF, VSM Medtech Ltd., Vancouver, BC, Canada). Data were continuously recorded with a sampling frequency of 600 Hz resulting in an off-line cut-off frequency of 150 Hz. Participants were seated in an upright position and their head was stabilized with cotton pads inside the MEG helmet. A silent movie was presented on a projector screen

mounted on the MEG system gantry, placed according to participants' best view angle, in order to keep them staying vigilant during the experiment; as been applied in previous auditory experiments (Pantev et al., 2004; Okamoto et al., 2008; Paraskevopoulos et al., 2018). After passing electro-static transducers the auditory stimuli were delivered via silicon tubes of 60 cm length and an inner diameter of 5 mm ending with a silicon earpiece fitted individually to each subject's ear. Prior to the stimulation, an audiological hearing threshold determination test with 5 dB accuracy on 1 kHz frequency, was conducted. Stimulus sound pressure levels were set to 60 dB SL above the individual hearing threshold. The whole experiment lasted for approximately 30 min.

MRI Protocol

A T1-weighted MR image was performed for all participants, in a 3 Tesla scanner (Gyrosan Intera T30, Philips), in order to obtain the individuals' Finite Element Model (FEM) of the head. The files gave images of 400 one layer-slices with 0.5 mm thickness in the sagittal plane (TR = 7.33.64 ms, TE = 3.31 ms). The matrix size of each slice was 512 × 512 with voxel size of 0.5 × 0.58 × 0.58 mm³. To ensure the reliability of investigation on brain structure within and across subjects, we used SPM12 (Statistical Parametric mapping)⁷ to regulate for intensity inhomogeneity (Ganzetti et al., 2016) and therefore, the images were resliced to isotropic voxels of 2 × 2 × 2 mm.

MEG Data Analysis

The analysis of the MEG data was run according to a previously developed analysis applied for functional connectivity networks under different auditory paradigms (Paraskevopoulos et al., 2015, 2018). The Brain Electrical Source Analysis software (BESA MRI, version 2.0, Megis Software, Heidelberg, Germany) was used to compute the individual's head model by segmenting four different head tissues (scalp, skull, CSF, and brain), based on the FEM. The four-layer FEM model gives more precise results as compared to other models, since it includes the CSF (Ramon et al., 2004; Wendel et al., 2008), which is a highly conductive layer and important for MEG source reconstruction (Wolters et al., 2006). The MEG sensors were co-registered and adjusted to the individuals' structural MRI via the nasion, and the left and the right entries of the ear-canals as landmarks. By means of 3D spline interpolation the MRIs were transformed to ACPC (anterior-posterior cingulate) and to Talairach space. A predefined option for conductivity values (c.f. Wolters et al., 2006) was set for the skin compartment to 0.33 S/m, for the skull to 0.0042 S/m, for the CSF to 0.79 S/m and for the brain tissue to 0.33 S/m.

For the pre-processing of the MEG data, the BESA research software (version 6.0, Megis Software, Heidelberg, Germany) was used. For artifact rejection, an automated electrocardiogram (ECG) and eye blinks artifact detection and correction provided by BESA (Ille et al., 2002) was applied. Data were filtered off-line with a 50 Hz notch filter, zero-phase low-pass filter of 45 Hz and zero-phase high-pass filter of 0.5 Hz. The data were divided

¹ www.freesounds.org

² www.soundbible.com

³ www.zapsplat.com

⁴ www.audacityteam.org

⁵ www.nch.com.au

⁶ www.neurobs.com

⁷ www.fil.ion.ucl.ac.uk

into epochs of 1000 ms post- and 500 ms pre-stimulus onset. A baseline correction based on a 100 ms pre-stimulus interval was applied. During the averaging of the stimulus related epochs, trials having amplitudes larger than 3 pT and data exceeding the 15% of rejected trials, were excluded from the analysis. The two measurement blocks were then averaged for each participant in order to improve the signal-to-noise ratio.

For the current density reconstruction, we used a time window around the N1 major component of the slow auditory evoked field (Pantev et al., 1993; c.f. **Figure 2**), which according to the global field power of the grand average data was between 80 ms and 120 ms after stimulus onset, including the rising slope of the N1. Low Resolution Electromagnetic Tomography (LORETA) provided by BESA was applied for the source reconstruction, for each subject and each category of sounds as it provides smooth distribution of sources as inverse solution (Pascual-Marqui et al., 1994). It is based on the weighted minimum norm method (Grech et al., 2008) and it does not rely on an *a priori* determination of activated sources.

Statistical Analysis

For the statistical analysis of the LORETA reconstruction, we used the SPM12 running on Matlab software (R2016b version; MathWorks Inc., Natick, MA, United States). An explicit mask was set, to include results only for the gray matter, thus decreasing the search volume. One-way ANOVA analysis was run with the three different categories of sounds (*Natural*, *Artificial*, and *Musical*) as within-subjects factor. F- and t-contrast matrix-tables were then designed (based on the general linear model) to test for statistical differences across the three categories and between-categories, respectively. For multiple comparisons control, the Family Wise Error (FWE) was implemented.

Connectivity Analysis

In order to examine the cortical network across the significant sources derived from the SPM12 analysis, we further implemented a connectivity analysis. Having defined the activated regions in source space via the above described analysis, we employed an equivalent current dipole model by setting one dipole to the peak of each significant cluster derived by the F-contrast. This resulted into five equivalent current dipoles in total. Due to the fact that SPM expresses coordinates based on standardized brains by the Montreal Neurological Institute (MNI coordinates), the coordinates were transferred to Talairach space to fit the brain coordinates of BESA software, where the dipole model was run. For the conversion the “NMI2TAL” applet of the Yale BioImage Suite Package was used (sprout022.sprout.yale.edu), which is based on the Lacadie et al.’s (2008) mapping coordinates. The orientation of the dipoles was fitted based on the individuals FEM volume conductor, whereas the coordinates were fixed across all subjects and conditions, as defined above. The results contained five source waveforms that corresponded to each dipole including the 80–120 ms interval.

The HERMES toolbox (Niso et al., 2013) of Matlab was used to construct a 5×5 adjacency matrix for each subject and each condition based on the Mutual Information (MI) algorithm,

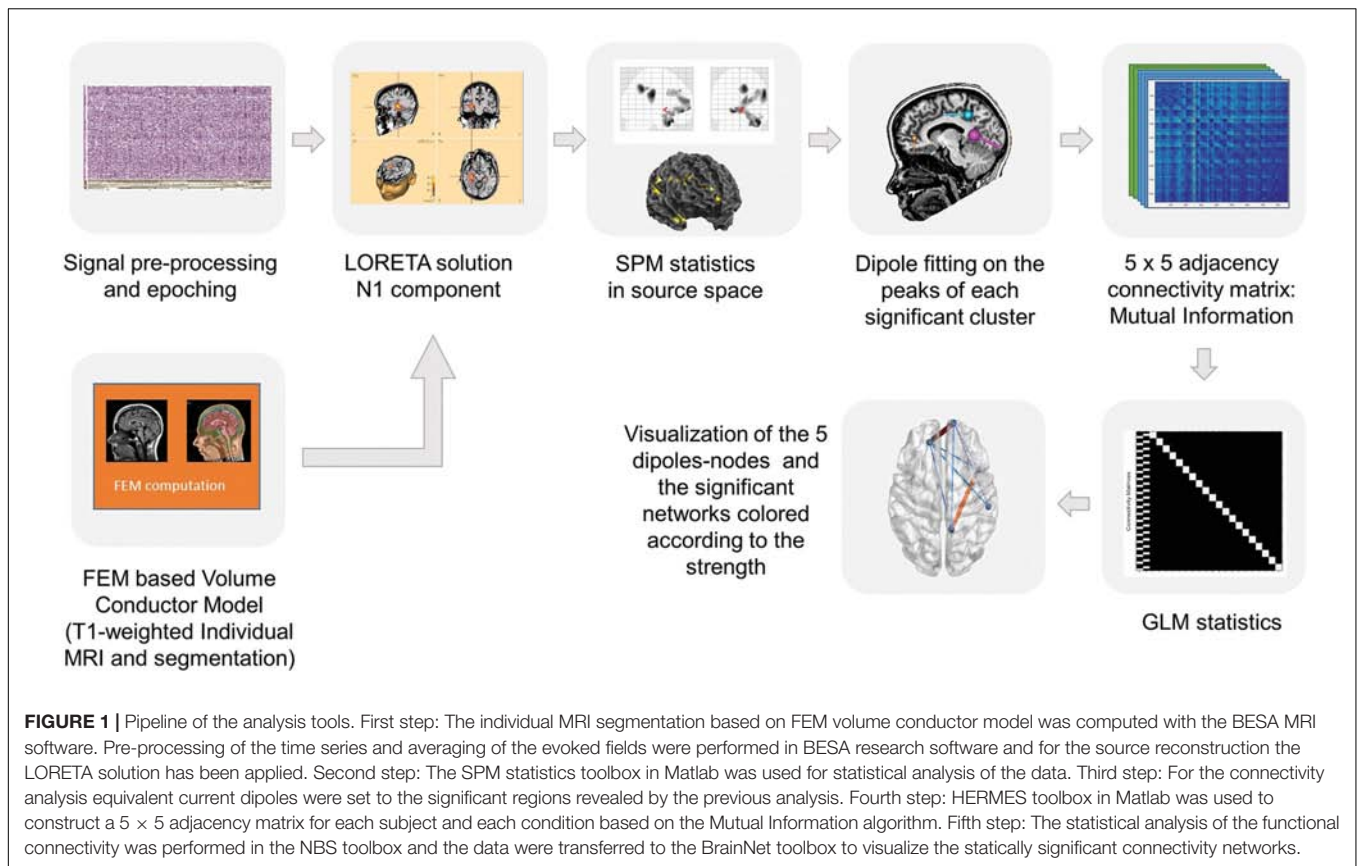
which measures the mutual dependence between variables and it detects correlations of random variables with non-linear dependence measure (Zeng, 2015). The results were then transferred to the Network Based Statistic Toolbox (NBS; Zalesky et al., 2010) to examine statistically significant connections. One way within-subjects ANOVA was run with the three conditions as the within-subjects factor. The NBS method was set for multiple correction at the significant level of $p > 0.05$ (see **Figure 1** for analysis pipeline tools). This resulted in a functional connectivity graph with nodes and edges representing the significant activated regions and their significant interconnection, respectively.

RESULTS

Source Space

The N1 auditory evoked field was *a priori* set as the time window of interest. The root-mean square time series of the grand average across subjects was computed in sensor space to depict the time window in ms around the *a priori* set N1 auditory field maximum. The **Figure 2** illustrates the mean of the root-mean square values of each sound category as well as the maximum and minimum values of their confidence intervals. The 80–120 ms interval was determined for the following source reconstructions where we performed F- and t- contrast statistics by means of the SPM12 software. **Figure 3** and **Table 1** illustrate the significant clusters obtained by the *Musical* \neq *Artificial* \neq *Natural* contrast. The biggest in size cluster involved parts of the right and left frontal cortex, as well as, parts of the temporal lobe. The peak of the current cluster was located in the anterior part of the right temporal cortex, in the most dorsal area of the superior temporal gyrus (STG) ($x = 43$, $y = 14$, $z = -29$; $F(1, 20) = 13.1$, cluster size = 3128, $p < 0.001$ FWE corrected at cluster level). A second cluster was located in the right inferior parietal lobe, with the peak in the right supramarginal gyrus (SMG) (coordinates: $x = 56$, $y = -27$, $z = 27$; $F(1, 20) = 12.80$, cluster size = 513, $p < 0.001$ FWE corrected at cluster level). The third cluster was located in the posterior cingulate cortex (PCC; overlapping with cluster two in the figure), which involved the posterior cingulate gyrus and medial part of the parietal lobe (coordinates: $x = 10$, $y = -52$, $z = 39$, $F(1, 20) = 12.42$, cluster size = 650, $p < 0.001$ FWE corrected at cluster level).

For the between groups comparisons, the results revealed significant differences for both *Musical* $>$ *Natural* and *Artificial* $>$ *Natural* comparisons as demonstrated in **Figure 4** and **Table 2**, depicting the coordinates and the source mapping results for both comparisons. In more detail, two significant clusters were obtained for *Musical* $>$ *Natural* sounds. The biggest cluster involved regions of left and right frontal cortex, as well as, temporal cortex and SMG. The peak of the current cluster was located at the SMG ($x = 56$, $y = -27$, $z = 27$, $t(20) = 5.05$, cluster size: 109960, $p > 0.001$ FWE corrected at cluster level). The second cluster was located in the PCC (coordinates: $x = 8$, $y = -54$, $z = 40$, $t(20) = 4.61$, cluster size = 3142, $p < 0.001$ FWE corrected at cluster level). For the *Artificial* $>$ *Natural* contrast we found two clusters revealing significant differential activation; one significant peak was located at the PCC (coordinates: $x = 10$,



$y = -37$, $z = 39$, $t(20) = 4.09$, cluster size = 675, $p = 0.001$ FWE corrected at cluster level) and the other one at the medial prefrontal cortex mPFC (coordinates: $x = 13$, $y = 39$, $z = 18$, $t(20) = 4.08$, cluster size = 1007, $p < 0.001$ FWE corrected at cluster level).

Connectivity Results

For the connectivity analysis, five equivalent current dipoles were set to the peaks of the significantly differential activated areas as derived from the F-contrast. These are the PCC, the aSTG, the SMG and the bilateral prefrontal cortex. Three connectivity analyses were conducted based on the significant interaction found in the source space analysis, $p < 0.05$, NBS corrected (Network-Based Statistics; Zalesky et al., 2010).

For the *Musical* ≠ *Artificial* ≠ *Natural* contrast, the results revealed connections of all the nodes having six edges in total. As **Figure 5** demonstrates, edges between PCC and right anterior superior temporal gyrus (aSTG), as well as, between left ventral mPFC and right medial dorsal prefrontal cortex had the strongest activation (as indicated by the F -value). The left mPFC yielded significant differential connections with all the nodes located in the right hemisphere and it was the only one connecting with the SMG.

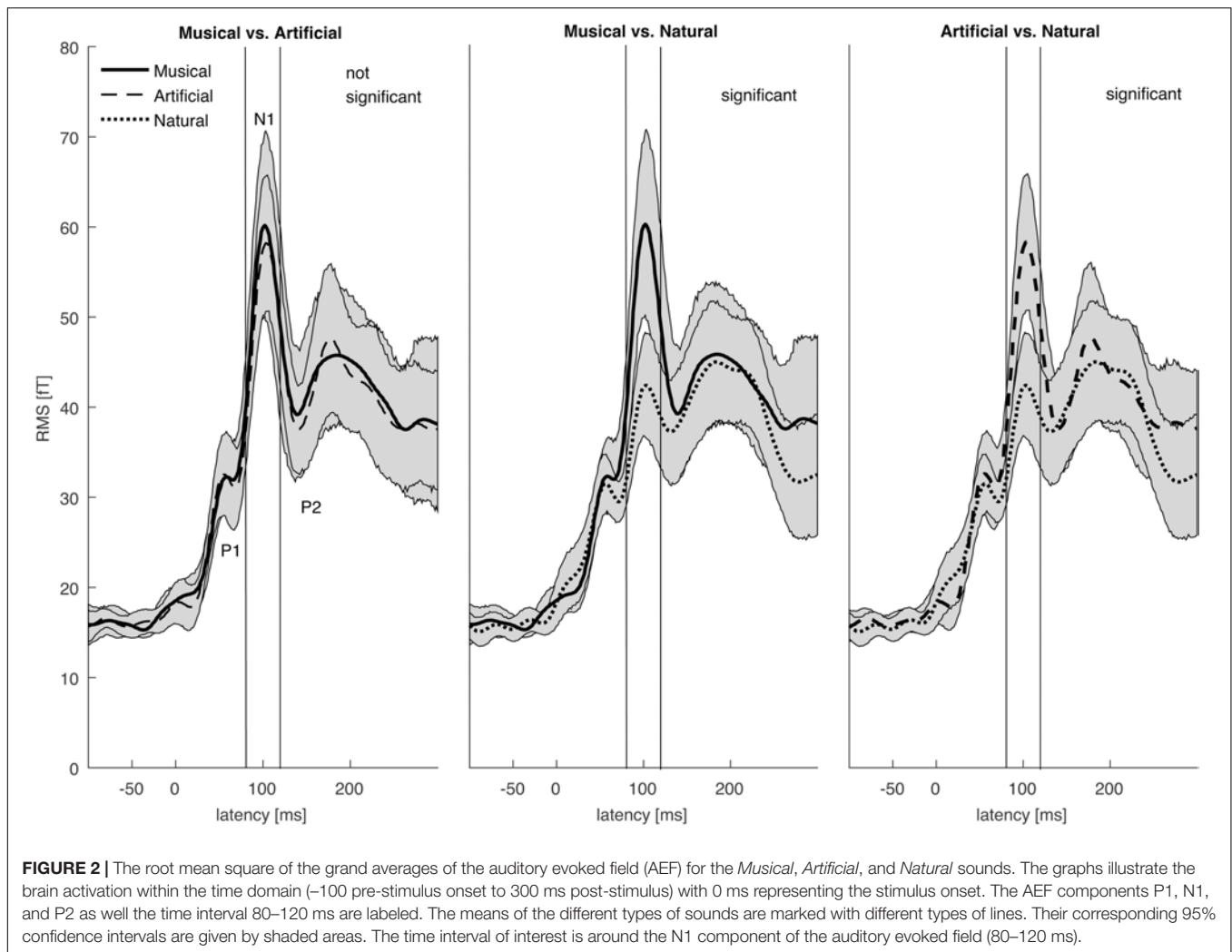
In the *Music-versus-Natural* comparison a significant differential network of 10 edges was yielded with all the nodes in the network being interconnected (c.f. **Figure 6**). According to the t -value, increased functional connectivity was demonstrated

between the right mPFC and the PCC, as well as, between the PCC and the aSTG. With smaller t -value, the left mPFC yielded significant interconnections with the right PFC, the aSTG, the SMG and the PCC. Weaker interconnections were obtained between the SMG and the aSTG, the PCC, as well as the left mPFC.

With regard to *Artificial-versus-Natural* comparisons (c.f. **Figure 6**), the results showed significant differential interconnections among all the nodes having nine edges in total. In detail, the edge-connection of the prefrontal inter-hemisphere and the connection between the aSTG and PCC nodes, were the most pronounced according to the t -value. Slighter in strength edges were revealed for the rest of interconnections, but no significant interconnection was found between PCC and SMG.

DISCUSSION

The current study examined neural responses in the processing of different categories of sounds, by means of MEG. Further, functional connectivity analysis based on the MI algorithm was used to depict the differential connectome of the regions involved in the discrimination of *Natural*, *Musical*, and *Artificial* sound-category. The obtained results demonstrated statistically significant differences across the different sound conditions in the superior temporal cortex, the posterior cingulate, the inferior parietal and the bilateral prefrontal cortex, between 80 and



120 ms after stimulus onset. Direct comparisons across the different categories, showed that both *Musical* and *Artificial* sounds demonstrated statistically significant differences and more enhanced brain activation in source space and connectivity analysis, when compared to *Natural* sounds category.

Our finding that regions of parietal cortex showed significant category modulation of activation, as derived from the contrast analysis, it is consistent with several neuroimaging studies investigating the organization of cortical auditory perception (Lewis et al., 2005; Murray et al., 2006; Staeren et al., 2009). In our study, we found two regions in the parietal lobe to be involved in the differential processing of the assessed sound categories, namely, the PCC and the right inferior parietal cortex (IPC). The peak of the IPC cluster was located in the SMG, which is part of the somatosensory cortex (Buccino et al., 2001). In fact, it is suggested to be involved in action representation and specifically in the mental representation of movement (Chong et al., 2008; Tunik et al., 2008). Within the framework of this interpretation, any event or stimuli containing actions would engage motor related regions even in the absence of tactile stimulation (Nilsson et al., 2000). Additional evidence associates

the SMG with the mirror neuron system, a system involved in imitating and identifying the actions of other persons (Carlson, 2012). On this basis, the right SMG has been suggested to be especially involved in motor planning, even when the actions are just observed and not necessarily executed (Chong et al., 2008). It seems that in the mental representation, information of how a tool is manipulated, visualized or how the sound is produced, are retrieved and integrated.

Apart from the parietal cortex, our connectivity results further stressed the involvement of the prefrontal cortex and the STG in the processing of different sound categories. Significant functional connectivity between prefrontal cortex and parietal cortex has been demonstrated in the past, suggesting that this network is part of the working memory network, linking perception and higher order cognition. This is rather multi-modal and regards both auditory and visual object processing (Husain et al., 2004; Husain and Horwitz, 2006). With respect to the STG, most of neuroimaging studies have shown the significant involvement of this region in the auditory processing. However, according to the meta-analysis of Arnott et al. (2004; including 38 studies of different neuroimaging techniques) it

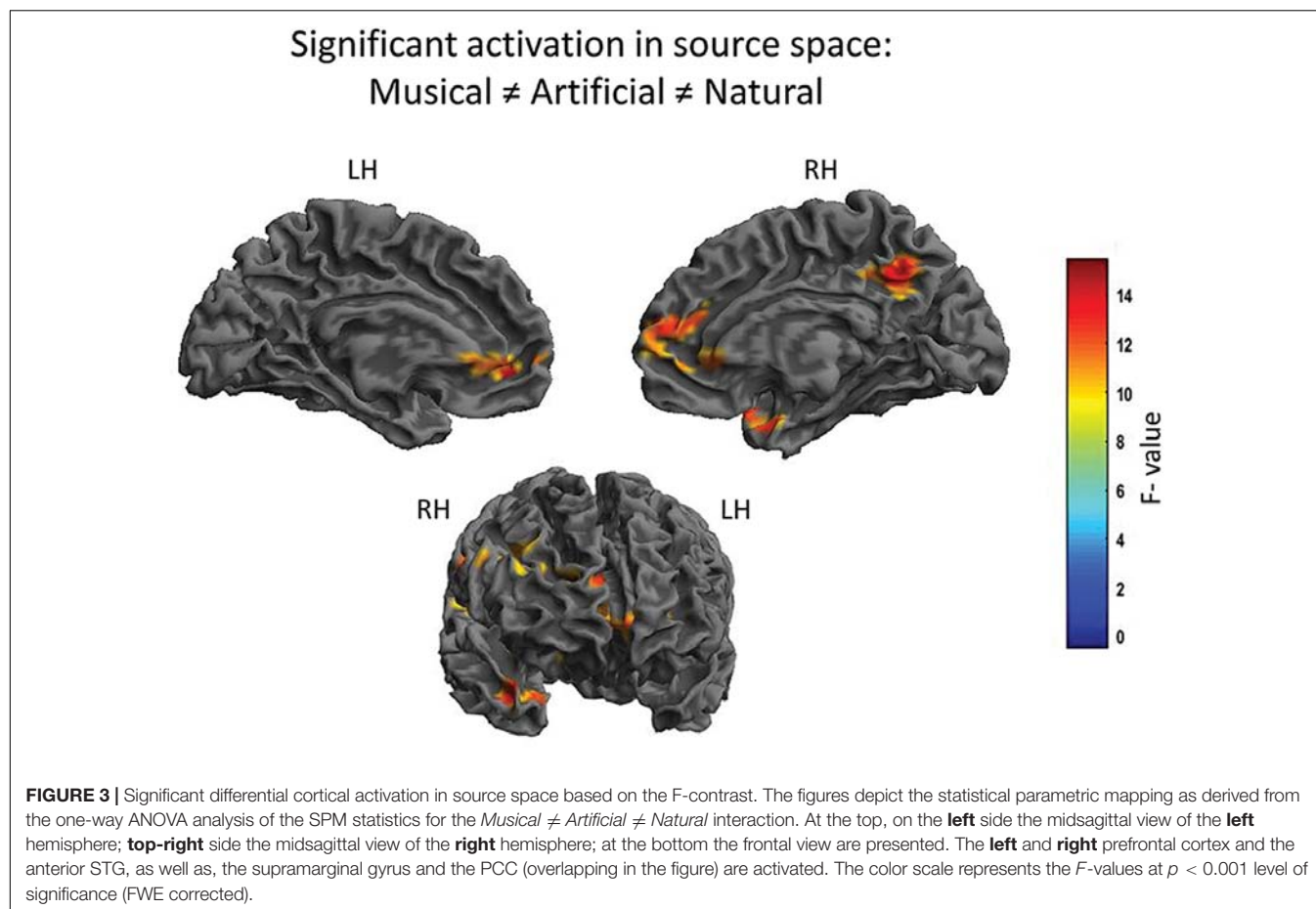


TABLE 1 | Significant clusters based on the F-contrast.

Location of cluster	MNI coordinates			Cluster size in voxel	<i>F</i> (1, 20)	P FWE-corrected
	X	Y	Z			
Right aSTC	43	14	-29	3128	13.01	<0.001
Right SMG	56	-27	27	513	3.86	<0.001
PCC	10	-52	39	650	3.81	<0.001

The locations of clusters with corresponding MNI coordinates of significant different cortical regions as derived by the source space analysis for the *Musical \neq Artificial \neq Natural* contrast are given. Cluster size in voxel, *F*- and *p*-values are also depicted, FEW corrected at $p < 0.001$ level of significance.

seems that the anterior area of the STG, as demonstrated here, is associated to greater extend with the discrimination of sounds. In line with our connectivity results, the STG region has been found to interact significantly with the prefrontal cortex during category discrimination task (Husain and Horwitz, 2006). The fact that our connectivity pattern was lateralized on the right hemisphere is consistent with the suggestion that this hemisphere is more strongly activated, when non-speech or non-living stimuli are represented (see Zatorre et al., 2002; Tervaniemi and Hugdahl, 2003). This derives not only from brain lesion studies, but also from neuroimaging studies with healthy participants (Belin and Zatorre, 2003; Murray et al., 2006).

Another possible interpretation of our results is that the processing of Musical and Artificial sounds involves more the

default mode network (DMN). The PCC area, as derived by our results, has been strongly suggested to have a central role in the DMN (for review see Leech and Sharp, 2013) with the inferior parietal lobe, prefrontal cortex and temporal cortex (however, more medial structures) as the main nodes for intrinsic connections (Raichle et al., 2001). Activation of DMN has been mainly observed under task-free conditions when attention to external stimuli is not required (for review see Buckner et al., 2008). However, increased activity during cognitive processing which requires internal attention is also found, such as memory retrieval and planning (Spreng, 2012). From this view, the processing of man-made objects sounds might require a more top-down processing of information than the living-object sounds do. Moreover, the PCC has been highly associated with

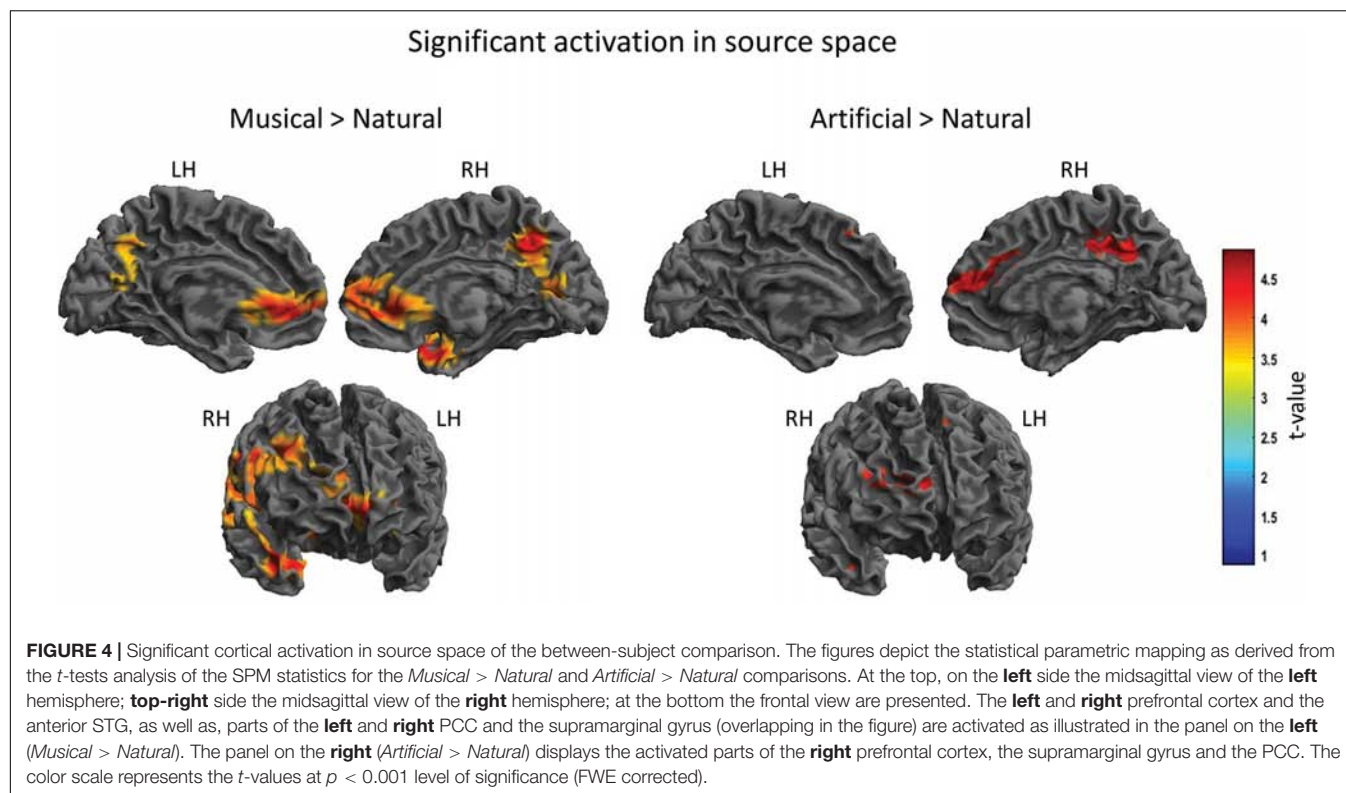


TABLE 2 | Significant clusters of the between-subject comparisons.

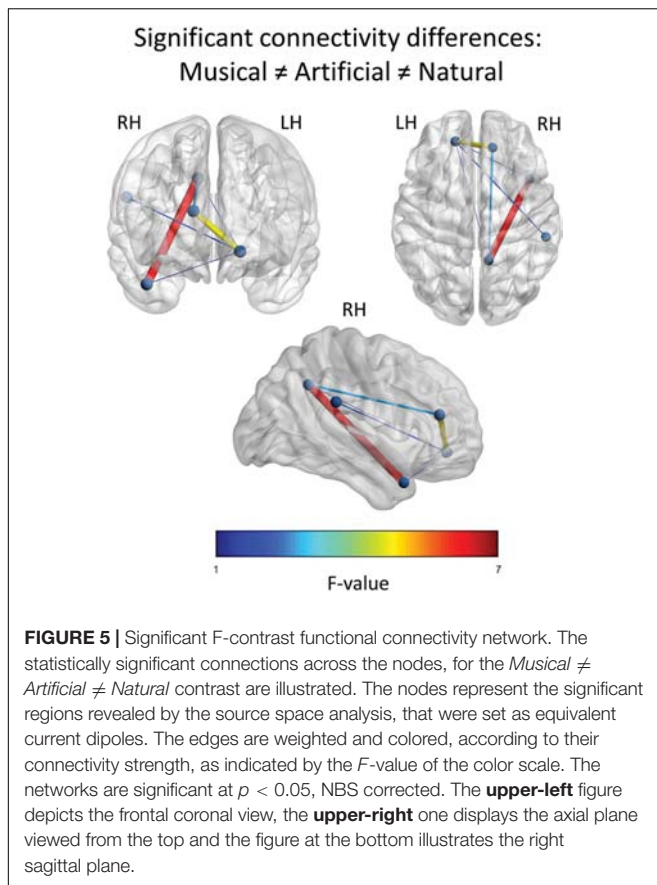
t-Contrast	Location of cluster	MNI coordinates			Cluster size in voxel	<i>t</i> (20)	P FWE-corrected
		<i>x</i>	<i>y</i>	<i>z</i>			
<i>Musical > Natural</i>	Right SMG	56	-27	27	109960	5.05	<0.001
	PCC	8	-54	40	3142	4.61	<0.001
<i>Artificial > Natural</i>	Right mPFC	13	39	18	1007	4.08	<0.001
	PCC	10	-37	39	675	4.09	0.001

The location of clusters with the corresponding MNI coordinates of the significant different cortical regions as derived by the source space analysis for the *Musical-versus-Natural* and *Artificial-versus-Natural* *t*-tests are shown. Cluster size, *t*- and *p*-values are also depicted, FEW corrected at $p < 0.001$ level of significance.

emotional processing and arousal state with higher responses and large-scale functional connectivity under arousal state and decreased on sleeping state (Leech and Sharp, 2013). As such, stimulation with Musical and Artificial sounds might evoke high-arousal emotions in relation to natural sounds. Nevertheless, the lack of a relevant behavioral measurements does not allow for further interpretation on this assumption and it is beyond the scope of this paper. In the future, a behavioral test assessing the levels of emotional arousal to different sounds would give us some insights on whether this factor affects significantly the differential brain activation.

Our main hypothesis was that the man-made related sounds would differentiate from the *Natural* sounds due to the mental representation of motor characteristics, that was indeed more pronounced. We further aimed to investigate whether statistical differences also apply to the comparison between *Musical* and

Artificial as it has been previously reported; based on the assumption that the sounds of daily-used objects might trigger more the response of action compare to the musical instruments (in non-musicians). Our results did not yield any significant differences in source space analysis, however, in the connectivity analysis, even though the *Musical-versus-Natural* and *Artificial-versus-Natural* demonstrated similar connectivity patterns, the strength of the interconnections (as given by the thickness of the edge) was different. For instance, in the *Artificial-versus-Natural* network, the inter-hemispheric prefrontal edge and the STG-PCC edge demonstrated stronger interconnections among the rest, whereas the *Musical-versus-Natural* network yielded a more distributed intensity across the interconnections with smaller values. A previous study (De Lucia et al., 2009) has shown differential discrimination between musical and tool sounds only after 300 ms stimulus onset. On this basis, it might be that the

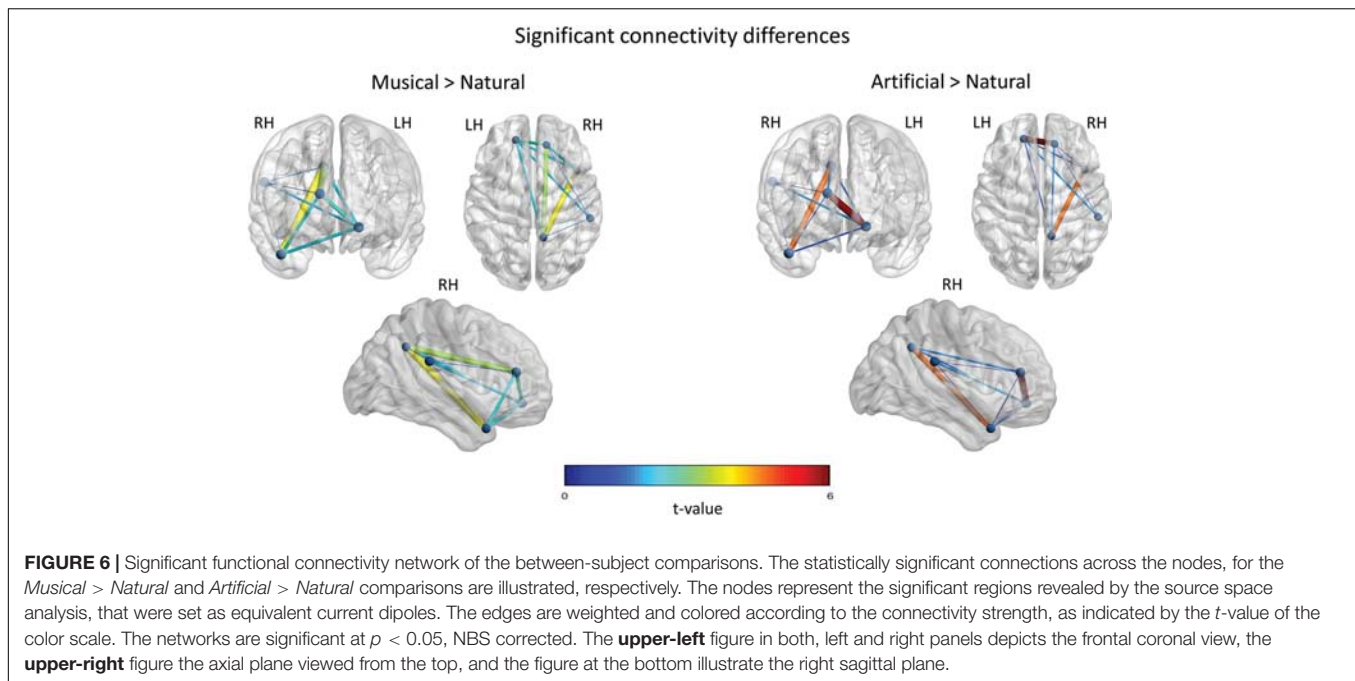


discrimination of broader categories of sounds, such as between living and man-made sounds already occur on early responses, though discrimination of subcategories could be followed by later responses. It would be interesting in the future to investigate later time intervals in which higher-level conceptual processes might be needed for the discrimination of man-made subcategories. It should be mentioned though that the *Artificial* category in our study, contained sounds that are generally man-made objects, however, some of them were less manipulative (e.g., ambulance siren). The absence of a stricter categorization based on a behavioral assessment that would divide the objects based on familiarity and the frequency of use, might have limited our interpretation regarding the significant differences between sound categories. The examination of the cortical regions in response to sound familiarity, recognizability and attention might give insights on the role that the “what” auditory neuronal pathway has in sound processing. Previous studies have shown the importance of sound novelty, task demands and attention in sound discrimination (Levy et al., 2001, 2003). However, others argue that the discrimination of living and man-made objects is present independently of behavioral proficiency, and hence familiarity (De Lucia et al., 2012), when correctly categorized sounds were compared with incorrectly categorized sounds as well as independently of consciousness (Cossy et al., 2014) when comatose patients were examined. This is still on debate in the literature and should be addressed in future studies.

A possible limitation of the current study might be that the categorization of complex sounds could be confounded by the physical differences of stimuli (Belin et al., 2000; Staeren et al., 2009), something that by purpose was not controlled here, in order to avoid any distortion in the quality of the sounds. However, natural sounds differ from human-made and more synthetic sounds by nature which we cannot manipulate. According to Theunissen and Elie’s (2014) review, the natural sounds consist of statistical properties that follow a power law relationship and are optimally encoded in the ascending auditory processing system in contrast to the sounds with more random and flat envelope power spectra (such as the sounds in the *Artificial* category). Nevertheless, similar categorization of sounds has been administered in the past and our results are in consistence (Murray et al., 2006).

The results of our study suggest that the dissociation between living and man-made objects, is based on distinct neuronal processing. However, the reason of this phenomenon is still questioned. From another perspective, it might be that sounds have been conceptually specialized in the processing of different categories of sounds due to evolutionary adaptation; a theory that has been strongly argued for the visual system, as well (Caramazza and Mahon, 2003). According to that, distinct cortical pathways corresponding to different categories of sounds, have been evolved analogous to the environmental sounds, which humans have experienced over the years. This would be in agreement with the previous mentioned review that suggests optimally neural encoding over the natural sounds compare to the sounds of human-made machines (Theunissen and Elie, 2014). From this perspective, listening to the sounds of nature would elicit weaker activation of brain in comparison to the more “modern” tool sounds (e.g., phone ringing), since we are evolutionary more adapted to the natural sounds and this in turn would require less cognitive effort (Buckner et al., 2008). This would in turn give some explanation on the abovementioned findings, where the man-made sounds showed stronger brain activation compare to the “living” objects. Furthermore, the functional connectivity maps showed that within the man-made category, there might be “key” connections for *Artificial-versus-Natural* relative to *Musical-versus-Natural*, even though the nodes remained the same for both comparisons. This could also explain cases of semantic impairments, where patients are impaired in a very specific category, while the remaining categories within the same domain are spared (McCarthy and Warrington, 2016; Muhammed et al., 2018). Therefore, it seems very probable that the perception of auditory objects relies on a large-scale distributed system, which follows distinct neuronal pathways, dissociated on the basis of the weight that each node has in the network. Based on our findings, this assumption cannot be clearly answered, though it gives rise for upcoming examination. A connectivity analysis involving also the common activated regions might provide a better picture to this assumption, however, due to the fact that this requires a different analysis, it is recommended for future study.

In general, the fact that the processing of auditory stimuli engages regions beyond the auditory cortex, such as anterior



temporal and frontal lobe (Maeder et al., 2001; Alain et al., 2008), is well documented. Similar to our cortical network analysis results, recent neuroimaging studies suggest that a network of fronto-temporo-parietal regions contributes to semantic processing (for review see Thompson-Schill, 2003; Patterson et al., 2002). This network has been proposed to be associated with the perception of both auditory and visual object identification (Goll et al., 2010; Hunter et al., 2010; Brefczynski-Lewis and Lewis, 2017). However, what is not well documented yet is how these brain responses are functionally connected. Our connectivity study underpins the connectivity of brain regions within sound discrimination. In order to obtain a better understanding of brain auditory categorization, it is not sufficient to investigate only the activated regions in isolation, but rather to understand how these regions interact. In this aspect we believe that our results are valuable for better understanding of the human brain in sound discrimination.

CONCLUSION

The present study demonstrated an enhanced brain network of man-made related sounds (*Musical* and *Artificial*) when compared to *Natural* sounds. So far the literature has provided simple brain activation maps. We additionally showed how these differentially activated brain regions are functionally connected and linked to the respective cognitive processes. We replicated previous findings supporting the engagement of other modalities beyond the auditory, to be involved in the processing of sound stimuli, as soon as this reaches the level of object representation. This in turn seems to be based on semantic categorization of the stimulus, following distinct neuronal pathways for living versus man-made objects. In addition to previous studies that

investigated only the cortical activation to different sound categories, we demonstrated significant differences in the functional connectivity between the cortical sources involved in the processing of the different sound categories.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Medical faculty of the University of Münster. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

VS contributed to the data recruitment, data analysis, and wrote the first draft of the manuscript. EP and NC involved in the statistical analysis. KM and AW involved in the informatic support. CD, DK, and CP contributed to the manuscript revision. All the authors read and approved the submitted version of the manuscript.

FUNDING

This work was supported by the Deutsche Forschungsgemeinschaft (DFG; PA 392/16-1, DO-711/10-1).

ACKNOWLEDGMENTS

The authors would like to thank the subjects for their cooperation and the technicians of the Institute for Biomagnetism and Biosignalanalysis for supporting the data acquisition.

REFERENCES

- Alain, C., McDonald, K. L., Kovacevic, N., and McIntosh, A. R. (2008). Spatiotemporal analysis of auditory “what” and “where” working memory. *Cereb. Cortex* 19, 305–314. doi: 10.1093/cercor/bhn082
- Arnott, S. R., Binns, M. A., Grady, C. L., and Alain, C. (2004). Assessing the auditory dual pathway model in humans. *Neuroimage* 22, 401–408. doi: 10.1016/j.neuroimage.2004.01.014
- Asplund, C. L., Todd, J. J., Snyder, A. P., and Marois, R. (2010). A central role for the lateral prefrontal cortex in goal-directed and stimulus-driven attention. *Nat. Neurosci.* 13, 507–512. doi: 10.1038/nn.2509
- Belin, P., and Zatorre, R. J. (2003). Adaptation to speaker's voice in right anterior temporal lobe. *Neuroreport* 14, 2105–2109. doi: 10.1097/00001756-200311140-00019
- Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., and Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature* 403, 309–312. doi: 10.1038/35002078
- Brefczynski-Lewis, J. A., and Lewis, J. W. (2017). Auditory object perception: a neurobiological model and prospective review. *Neuropsychologia* 105, 223–242. doi: 10.1016/j.neuropsychologia.2017.04.034
- Bregman, A. S. (2017). “Asking the “what for” question in auditory perception,” in *Perceptual Organization*, eds M. Kubovy, and J. Pomerantz, (Hillsdale, NJ: Routledge), 99–118. doi: 10.4324/9781315512372-4
- Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur. J. Neurosci.* 13, 400–404.
- Buckner, R. L., Andrews-Hanna, J. R., and Schacter, D. L. (2008). The brain's default network. *Ann. N. Y. Acad. Sci.* 1124, 1–38. doi: 10.1196/annals.1440.011
- Bullmore, E., and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 186–198. doi: 10.1038/nrn2575
- Caramazza, A., and Mahon, B. Z. (2003). The organization of conceptual knowledge: the evidence from category-specific semantic deficits. *Trends Cogn. Sci.* 7, 354–361. doi: 10.1016/s1364-6613(03)00159-1
- Carlson, N. R. (2012). *Physiology of Behavior*, 11th Edn. London: Pearson.
- Chong, T. T. J., Cunningham, R., Williams, M. A., Kanwisher, N., and Mattingley, J. B. (2008). fMRI adaptation reveals mirror neurons in human inferior parietal cortex. *Curr. Biol.* 18, 1576–1580. doi: 10.1016/j.cub.2008.08.068
- Clarke, S., Bellmann, A., Meuli, R. A., Assal, G., and Steck, A. J. (2000). Auditory agnosia and auditory spatial deficits following left hemispheric lesions: evidence for distinct processing pathways. *Neuropsychologia* 38, 797–807. doi: 10.1016/s0028-3932(99)00141-4
- Cossy, N., Tzovara, A., Simonin, A., Rossetti, A., and De Lucia, M. (2014). Robust discrimination between EEG responses to categories of environmental sounds in early coma. *Front. Psychol.* 5:155. doi: 10.3389/fpsyg.2014.00155
- De Lucia, M., Camen, C., Clarke, S., and Murray, M. M. (2009). The role of actions in auditory object discrimination. *Neuroimage* 48, 475–485. doi: 10.1016/j.neuroimage.2009.06.041
- De Lucia, M., Tzovara, A., Bernasconi, F., Spierer, L., and Murray, M. M. (2012). Auditory perceptual decision-making based on semantic categorization of environmental sounds. *Neuroimage* 60, 1704–1715. doi: 10.1016/j.neuroimage.2012.01.131
- Ganzetti, M., Wenderoth, N., and Mantini, D. (2016). Intensity inhomogeneity correction of structural MR images: a data-driven approach to define input algorithm parameters. *Front. Neuroinform.* 10:10. doi: 10.3389/fninf.2016.00010
- Goll, J. C., Crutch, S. J., and Warren, J. D. (2010). Central auditory disorders: toward a neuropsychology of auditory objects. *Curr. Opin. Neurol.* 23, 617–627. doi: 10.1097/WCO.0b013e32834027f6
- Grech, R., Cassar, T., Muscat, J., Camilleri, K. P., Fabri, S. G., Zervakis, M., et al. (2008). Review on solving the inverse problem in EEG source analysis. *J. Neuroeng. Rehabil.* 5, 25. doi: 10.1186/1743-0003-5-25
- Griffiths, T. D., and Warren, J. D. (2004). What is an auditory object? *Nat. Rev. Neurosci.* 5, 887–892. doi: 10.1038/nrn1538
- Heekeren, H. R., Marrett, S., and Ungerleider, L. G. (2008). The neural systems that mediate human perceptual decision making. *Nat. Rev. Neurosci.* 9, 467–479. doi: 10.1038/nrn2374
- Huddleston, W. E., Lewis, J. W., Phinney, R. E., and DeYoe, E. A. (2008). Auditory and visual attention-based apparent motion share functional parallels. *Percept. Psychophys.* 70, 1207–1216. doi: 10.3758/PP.70.7.1207
- Hunter, M. D., Eickhoff, S. B., Pheasant, R. J., Douglas, M. J., Watts, G. R., Farrow, T. F., et al. (2010). The state of tranquility: subjective perception is shaped by contextual modulation of auditory connectivity. *Neuroimage* 53, 611–618. doi: 10.1016/j.neuroimage.2010.06.053
- Husain, F. T., and Horwitz, B. (2006). Experimental-neuromodeling framework for understanding auditory object processing: integrating data across multiple scales. *J. Physiol. Paris* 100, 133–141. doi: 10.1016/j.jphysparis.2006.09.006
- Husain, F. T., Tagamets, M. A., Fromm, S. J., Braun, A. R., and Horwitz, B. (2004). Relating neuronal dynamics for auditory object processing to neuroimaging activity: a computational modeling and an fMRI study. *Neuroimage* 21, 1701–1720. doi: 10.1016/j.neuroimage.2003.11.012
- Ille, N., Berg, P., and Scherg, M. (2002). Artifact correction of the ongoing EEG using spatial filters based on artifact and brain signal topographies. *J. Clin. Neurophysiol.* 19, 113–124.
- Kubovy, M., and Van Valkenburg, D. (2001). Auditory and visual objects. *Cognition* 80, 97–126. doi: 10.1016/s0010-0277(00)00155-4
- Lacadie, C. M., Fulbright, R. K., Rajeevan, N., Constable, R. T., and Papademetris, X. (2008). More accurate Talairach coordinates for neuroimaging using non-linear registration. *Neuroimage* 42, 717–725. doi: 10.1016/j.neuroimage.2008.04.240
- Leech, R., and Sharp, D. J. (2013). The role of the posterior cingulate cortex in cognition and disease. *Brain* 137, 12–32. doi: 10.1093/brain/awt162
- Levy, D. A., Granot, R., and Bentin, S. (2001). Processing specificity for human voice stimuli: electrophysiological evidence. *Neuroreport* 12, 2653–2657. doi: 10.1097/00001756-200108280-00013
- Levy, D. A., Granot, R., and Bentin, S. (2003). Neural sensitivity to human voices: ERP evidence of task and attentional influences. *Psychophysiology* 40, 291–305. doi: 10.1111/1469-8986.00031
- Lewis, J. W., Brefczynski, J. A., Phinney, R. E., Janik, J. J., and DeYoe, E. A. (2005). Distinct cortical pathways for processing tool versus animal sounds. *J. Neurosci.* 25, 5148–5158. doi: 10.1523/jneurosci.0419-05.2005
- Lewis, J. W., Wightman, F. L., Brefczynski, J. A., Phinney, R. E., Binder, J. R., and DeYoe, E. A. (2004). Human brain regions involved in recognizing environmental sounds. *Cereb. Cortex* 14, 1008–1021. doi: 10.1093/cercor/bbh061
- Maeder, P. P., Meuli, R. A., Adriani, M., Bellmann, A., Fornari, E., Thiran, J. P., et al. (2001). Distinct pathways involved in sound recognition and localization: a human fMRI study. *Neuroimage* 14, 802–816. doi: 10.1006/nimg.2001.0888
- McCarthy, R. A., and Warrington, E. K. (2016). Past, present, and prospects: reflections 40 years on from the selective impairment of semantic memory (Warrington, 1975). *Q. J. Exp. Psychol.* 69, 1941–1968. doi: 10.1080/17470218.2014.980280
- Mendez, M. F. (2001). Generalized auditory agnosia with spared music recognition in a left-hander. Analysis of a case with a right temporal stroke. *Cortex* 37, 139–150. doi: 10.1016/s0010-9452(08)70563-x
- Muhammed, L., Hardy, C. J., Russell, L. L., Marshall, C. R., Clark, C. N., Bond, R. L., et al. (2018). Agnosia for bird calls. *Neuropsychologia* 113, 61–67. doi: 10.1016/j.neuropsychologia.2018.03.024

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2019.01052/full#supplementary-material>

- Murray, M. M., Camen, C., Andino, S. L. G., Bovet, P., and Clarke, S. (2006). Rapid brain discrimination of sounds of objects. *J. Neurosci.* 26, 1293–1302. doi: 10.1523/jneurosci.4511-05.2006
- Nilsson, L. G., Nyberg, L., Klingberg, T., Åberg, C., Persson, J., and Roland, P. E. (2000). Activity in motor areas while remembering action events. *Neuroreport* 11, 2199–2201. doi: 10.1097/00001756-200007140-00027
- Niso, G., Bruña, R., Pereda, E., Gutiérrez, R., Bajo, R., Maestú, F., et al. (2013). HERMES: towards an integrated toolbox to characterize functional and effective brain connectivity. *Neuroinformatics* 11, 405–434. doi: 10.1007/s12021-013-9186-1
- Okamoto, H., Stracke, H., and Pantev, C. (2008). Neural interactions within and beyond the critical band elicited by two simultaneously presented narrow band noises: a magnetoencephalographic study. *Neuroscience* 151, 913–920. doi: 10.1016/j.neuroscience.2007.11.038
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4
- Pantev, C., Elbert, T., Makeig, S., Hampson, S., Eulitz, C., and Hoke, M. (1993). Relationship of transient and steady-state auditory evoked fields. *Electroencephalogr. Clin. Neurophysiol.* 88, 389–396. doi: 10.1016/0168-5597(93)90015-h
- Pantev, C., Okamoto, H., Ross, B., Stoll, W., Ciurlia-Guy, E., Kakigi, R., et al. (2004). Lateral inhibition and habituation of the human auditory cortex. *Eur. J. Neurosci.* 19, 2337–2344. doi: 10.1111/j.0953-816x.2004.03296.x
- Paraskevopoulos, E., Chalas, N., Kartsidis, P., Wollbrink, A., and Bamidis, P. (2018). Statistical learning of multisensory regularities is enhanced in musicians: an MEG study. *NeuroImage* 175, 150–160. doi: 10.1016/j.neuroimage.2018.04.002
- Paraskevopoulos, E., Kraneburg, A., Herholz, S. C., Bamidis, P. D., and Pantev, C. (2015). Musical expertise is related to altered functional connectivity during audiovisual integration. *Proc. Natl. Acad. Sci.* 112, 12522–12527. doi: 10.1073/pnas.1510662112
- Pascual-Marqui, R. D., Michel, C. M., and Lehmann, D. (1994). Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *Int. J. Psychophysiol.* 18, 49–65. doi: 10.1016/0167-8760(84)90014-x
- Patterson, K., Nestor, P. J., and Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat. Rev. Neurosci.* 8, 976–987. doi: 10.1038/nrn2277
- Patterson, R. D., Uppenkamp, S., Johnsrude, I. S., and Griffiths, T. D. (2002). The processing of temporal pitch and melody information in auditory cortex. *Neuron* 36, 767–776.
- Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage* 62, 816–847. doi: 10.1016/j.neuroimage.2012.04.062
- Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., and Shulman, G. L. (2001). A default mode of brain function. *Proc. Natl. Acad. Sci.* 98, 676–682.
- Ramon, C., Schimpf, P., Haueisen, J., Holmes, M., and Ishimaru, A. (2004). Role of soft bone, CSF and gray matter in EEG simulations. *Brain Topogr.* 16, 245–248. doi: 10.1023/b:brat.0000032859.68959.76
- Rogers, B. P., Morgan, V. L., Newton, A. T., and Gore, J. C. (2007). Assessing functional connectivity in the human brain by fMRI. *Magn. Reson. Imaging* 25, 1347–1357. doi: 10.1016/j.mri.2007.03.007
- Romanski, L. M., Tian, B., Fritz, J. B., Mishkin, M., Goldman-Rakic, P. S., and Rauschecker, J. P. (2000). Reply to “What, where and how in auditory cortex”. *Nat. Neurosci.* 3:966. doi: 10.1038/79892
- Spreng, R. N. (2012). The fallacy of a “task-negative” network. *Front. Psychol.* 3:145. doi: 10.3389/fpsyg.2012.00145
- Staeren, N., Renvall, H., De Martino, F., Goebel, R., and Formisano, E. (2009). Sound categories are represented as distributed patterns in the human auditory cortex. *Curr. Biol.* 19, 498–502. doi: 10.1016/j.cub.2009.01.066
- Steinke, W. R., Cuddy, L. L., and Jakobson, L. S. (2001). Dissociations among functional subsystems governing melody recognition after right-hemisphere damage. *Cogn. Neuropsychol.* 18, 411–437. doi: 10.1080/02643290125702
- Taniwaki, T., Tagawa, K., Sato, F., and Iino, K. (2000). Auditory agnosia restricted to environmental sounds following cortical deafness and generalized auditory agnosia. *Clin. Neurol. Neurosurg.* 102, 156–162. doi: 10.1016/s0303-8467(00)00090-1
- Tervaniemi, M., and Hugdahl, K. (2003). Lateralization of auditory-cortex functions. *Brain Res. Rev.* 43, 231–246. doi: 10.1016/j.brainresrev.2003.08.004
- Theunissen, F. E., and Elie, J. E. (2014). Neural processing of natural sounds. *Nat. Rev. Neurosci.* 15, 355–366. doi: 10.1038/nrn3731
- Thompson-Schill, S. L. (2003). Neuroimaging studies of semantic memory: inferring “how” from “where”. *Neuropsychologia* 41, 280–292. doi: 10.1016/s0028-3932(02)00161-6
- Tunik, E., Lo, O. Y., and Adamovich, S. V. (2008). Transcranial magnetic stimulation to the frontal operculum and supramarginal gyrus disrupts planning of outcome-based hand-object interactions. *J. Neurosci.* 28, 14422–14427. doi: 10.1523/jneurosci.4734-08.2008
- Wendel, K., Narra, N. G., Hannula, M., Kauppinen, P., and Malmivuo, J. (2008). The influence of CSF on EEG sensitivity distributions of multilayered head models. *IEEE Trans. Biomed. Eng.* 55, 1454–1456. doi: 10.1109/TBME.2007.912427
- Wolters, C. H., Anwander, A., Tricoche, X., Weinstein, D., Koch, M. A., and Macleod, R. S. (2006). Influence of tissue conductivity anisotropy on EEG/MEG field and return current computation in a realistic head model: a simulation and visualization study using high-resolution finite element modeling. *NeuroImage* 30, 813–826. doi: 10.1016/j.neuroimage.2005.10.014
- Zalesky, A., Fornito, A., and Bullmore, E. T. (2010). Network-based statistic: identifying differences in brain networks. *Neuroimage* 53, 1197–1207. doi: 10.1016/j.neuroimage.2010.06.041
- Zatorre, R. J., and Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* 11, 946–953. doi: 10.1093/cercor/11.10.946
- Zatorre, R. J., Belin, P., and Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends Cogn. Sci.* 6, 37–46. doi: 10.1016/s1364-6613(00)01816-7
- Zatorre, R. J., Bouffard, M., and Belin, P. (2004). Sensitivity to auditory object features in human temporal neocortex. *J. Neurosci.* 24, 3637–3642. doi: 10.1523/jneurosci.5458-03.2004
- Zeng, K. (2015). Physical layer key generation in wireless networks: challenges and opportunities. *IEEE Commun. Mag.* 53, 33–39. doi: 10.1109/mcom.2015.7120014

Conflict of 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.

Copyright © 2019 Salvari, Paraskevopoulos, Chalas, Müller, Wollbrink, Dobel, Korth and Pantev. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Effects of Music Training on Inhibitory Control and Associated Neural Networks in School-Aged Children: A Longitudinal Study

Sarah L. Hennessy¹, Matthew E. Sachs¹, Beatriz Ilari^{1,2} and Assal Habibi^{1*}

¹ Brain and Creativity Institute, University of Southern California, Los Angeles, CA, United States, ² Thornton School of Music, University of Southern California, Los Angeles, CA, United States

OPEN ACCESS

Edited by:

Virginia Penhune,
Concordia University, Canada

Reviewed by:

Vesa Putkinen,
Turku PET Centre, Finland
Lutz Jäncke,
University of Zurich, Switzerland

*Correspondence:

Assal Habibi
ahabibi@usc.edu

Specialty section:

This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Neuroscience

Received: 06 July 2019

Accepted: 24 September 2019

Published: 16 October 2019

Citation:

Hennessy SL, Sachs ME, Ilari B
and Habibi A (2019) Effects of Music
Training on Inhibitory Control
and Associated Neural Networks
in School-Aged Children:
A Longitudinal Study.
Front. Neurosci. 13:1080.
doi: 10.3389/fnins.2019.01080

Inhibitory control, the ability to suppress an immediate dominant response, has been shown to predict academic and career success, socioemotional wellbeing, wealth, and physical health. Learning to play a musical instrument engages various sensorimotor processes and draws on cognitive capacities including inhibition and task switching. While music training has been shown to benefit cognitive and language skills, its impact on inhibitory control remains inconclusive. As part of an ongoing 5-year longitudinal study, we investigated the effects of music training on the development of inhibitory control and its neural underpinnings with a population of children (starting at age 6) from underserved communities. Children involved in music were compared with children involved in sports and children not involved in any systematic after-school program. Inhibition was measured using a delayed gratification, flanker, and Color-Word Stroop task, which was performed both inside and outside of an MRI scanner. We established that there were no pre-existing differences in cognitive capacities among the groups at the onset. In the delayed gratification task, beginning after 3 years of training, children with music training chose a larger, delayed reward in place of a smaller, immediate reward compared to the control group. In the flanker task, children in the music group, significantly improved their accuracy after 3 and 4 years of training, whereas such improvement in the sport and control group did not reach significance. There were no differences among the groups on behavioral measures of Color-Word Stroop task at any time point. As for differences in brain function, we have previously reported that after 2 years, children with music training showed significantly greater bilateral activation in the pre-SMA/SMA, ACC, IFG, and insula during the Color-Word Stroop task compared to the control group, but not compared to the sports group (Sachs et al., 2017). However, after 4 years, we report here that differences in brain activity related to the Color-Word Stroop task between musicians and the other groups is only observed in the right IFG. The results suggest that systematic extracurricular programs, particularly music-based training, may accelerate development of inhibitory control and related brain networks earlier in childhood.

Keywords: music, inhibitory control, executive function, longitudinal research, music training, inhibition, neuroplasticity

INTRODUCTION

Executive functions are broadly defined as top-down processes related to goal acquisition and attention that primarily recruit the brain's prefrontal network (Miller and Cohen, 2001). Inhibitory control, a sub-construct of executive function (Miyake et al., 2000), refers to the ability to suppress a primary response. It is correlated with the activation in the dorsolateral prefrontal cortex (DLPFC), anterior cingulate cortex (ACC), supplementary and pre-supplementary motor cortex (SMA/pre-SMA), insula, and inferior frontal gyrus (IFG) (Niendam et al., 2012). Inhibitory control has been shown to be predictive of academic success (Mischel et al., 1989; Alloway et al., 2005; Duckworth and Seligman, 2005; Kirby et al., 2005), career success (Bailey, 2007), positive socioemotional wellbeing (Hughes and Dunn, 1998; Lengua, 2003; Eisenberg et al., 2005), wealth (Moffitt et al., 2011), reduced substance abuse risk and incarceration (Moffitt et al., 2011), and physical health (Seeyave et al., 2009; Miller et al., 2011).

The evidence supporting the association between inhibitory control and positive life outcomes has encouraged the development of educational programs aimed at improving these skills in childhood. Several studies observed enhancement of inhibitory control after short-term exercise programs among children who were overweight (Davis et al., 2011), who were diagnosed with attention deficit hyperactivity disorder (Chang et al., 2012), as well as among typically developing, non-overweight children (Chen et al., 2014). Similar improvements have been reported in children involved in martial arts (Lakes and Hoyt, 2004), mindfulness training (Flook et al., 2010), and classroom-based programs such as Tools of Mind (Diamond et al., 2007). In addition to behavioral improvements, these interventions produced training-related activation increases in regions of the cognitive control network changes in the cognitive control network in both children (Davis et al., 2011; Voss et al., 2011) and adults (Allen et al., 2012; Berkman et al., 2014). Others cite neural activation decreases associated with increased efficiency of inhibitory control after an intervention (Chaddock-Heyman et al., 2013; Nishiguchi et al., 2015).

Recently, there has been increasing interest in the impact of music interventions on developing inhibition control. Playing a musical instrument requires a variety of functions, including coordination of fine motor skills and different streams of auditory input (Zatorre et al., 2007), the rapid adjustment of one's motor behavior in response to mistakes (Jentzsch et al., 2014). To meet the technical demands of playing their instrument, musicians must continue to play as they inhibit attention away from one hand to focus on a different movement in the other hand. Like any practice-dependent activity or acquiring any skill, playing a musical instrument requires focused attention and self-discipline, prioritizing practice over other, more instantly gratifying, activities. This is particularly important in a group setting, where there are many social distractors, such as side conversations or other playing a piece incorrectly, that musicians must ignore for the benefit of the performance. For these reasons, musicians have been often cited as a model from which to investigate neuroplasticity (Münte et al., 2002).

In spite of this, contributions of musical abilities on inhibition in adults are not clear. While some studies have found that professional adult musicians show faster reaction times on a Color Stroop Task than amateur adult musicians (Travis et al., 2011) and untrained individuals (Bialystok and Depape, 2009), and faster reaction times than untrained individuals on a visual Simon task (Bialystok and Depape, 2009), Zuk et al. (2014) did not find differences in response times between adult musicians and untrained individuals on a Color Stroop Task, and Slevc et al. (2016) found music training to be unrelated to scores in an Auditory Stroop and Simon Arrows task.

Studies with children are similarly inconclusive. In a Simon task, children aged 10–11 who were musically trained showed reduced reaction time differences between congruent and incongruent trials when compared to children without music training (Joret et al., 2017). When asked to name shapes and arrows in an opposite manner (e.g., 'up' when presented with a 'down' arrow), musically trained children aged 3–9 performed significantly faster than musically un-trained children (Saarikivi et al., 2016). In another study, Holochwost et al. (2017) reported improvements in go-no-go task accuracy in 7–13-year-old children participating in an *El Sistema* orchestral program after 2 years and again after 3 years of training. Children aged 5–7 involved in school music education had greater increases in scores on a go-no-go task over two and a half years when compared to a visual arts and passive control group (Jaschke et al., 2018). Similarly, 4–6-year-old children involved in a music listening program performed better than those in an arts training program on go-no-go accuracy at post-test (Moreno et al., 2011). In an fMRI investigation of cognitive control networks, Zuk et al. (2014), demonstrated that musically trained children aged 9–12 had greater activation in the pre-SMA/SMA and right VLFC during a set shifting task as compared to non-musician children. Our group recently reported that, children aged 8–9 with 4 years of music training, as compared to children without any systematic training, had greater BOLD signal during incongruent trials of a color-word Stroop task in inhibition-related neural regions, including the pre-SMA/SMA, precentral sulcus, ACC and IFG (Sachs et al., 2017). However, no significant differences were observed between musically trained children and children involved in sports training. Additionally, neither our group (Sachs et al., 2017), nor Zuk et al. (2014) observed behavioral differences between groups, and others have reported no significant benefits of music training on inhibition (Schellenberg, 2011), leaving the findings inconclusive.

Several factors possibly contribute to the lack of consistency in these findings, including the absence of an active comparison group in some studies, and a great degree of variation in the music training programs studied. For example, Moreno et al. (2011) used a computerized curriculum with a focus on listening rather than learning to play a musical instrument, which involves sensory-motor learning and participation in a group setting (e.g., Holochwost et al., 2017). Even among studies that agree on the definition of "music training" as instrumental instruction, many differ on the length of training required to label participants as

“musicians” or “non-musicians.” In Schellenberg (2011) study, child musicians were described as having at least 2 years of music lessons outside of school; whereas, children in Zuk et al. (2014) study had been playing for an average of 5 years, while Degé and Schwarzer (2011) defined music training as a continuous variable. These differences are of particular importance due to the rapid development of inhibitory control during early to middle childhood, in which one or 2 years of regular participation in a music program could provide significant improvements in skill performance. Differences in individual and group music training between studies may also affect comparability between studies, since the social aspect of music making requires more integration of cognitive and social functions than solo playing (Keller, 2008). Playing in an ensemble where one is required to follow a conductor’s command, to attend to other members of the ensemble, and to adjust behaviors in response to other players, could improve inhibition skills at a faster rate than individual musical practice.

Finally, a limitation of current research in the field is the difficulty in recruiting a study sample that reflects society’s growing socio-economic status (SES) and cultural diversity. While a few groups (e.g., Holochwost et al., 2017) have investigated underrepresented populations, many studies are conducted with participants from what Henrich et al. (2010) called WEIRD societies (white, English speaking, industrial, and from rich and democratic countries), and may be generalizable only to a narrow portion of society at large. To draw broader conclusions about the impact of music training throughout society, more research involving participants from diverse backgrounds is needed.

In the present study, we investigate the effects of music training on inhibitory control using behavioral and neuroimaging methods, aiming to extend the previously reported findings (Sachs et al., 2017). Over the course of 4 years, we compare children involved in music training with children involved either in sports training or no systematic enrichment training. By conducting this investigation longitudinally, and implementing a pre-training baseline assessment, we attempt to differentiate effects of training from pre-existing biological contributions. To our knowledge, this is the first longitudinal study using neuroimaging to specifically assess inhibitory control in children involved in music training. Additionally, by comparing children involved in music training to children involved in sports training, we assess whether any effects observed in measures of inhibition are related to music training specifically or are associated with any type of extra-curricular activity that is socially engaging and motivating.

We hypothesize: (a) that children involved in music training will show greater improvements on behavioral measurements of inhibition than will children involved in no systematic training and children involved in sports training, evidenced by reduced reaction time and improved accuracy; (b) that, during an fMRI inhibition task, children involved in music training as compared to children with no systematic training will show greater activation of brain regions associated with inhibitory control, including the IFG, SMA/pre-SMA, ACC, and insula, continuing the pattern reported after 2 years of training (Sachs et al., 2017).

MATERIALS AND METHODS

Participants

Data for this report were collected as part of an ongoing longitudinal study investigating the effects of music training on child brain, cognitive, and socioemotional development (Habibi et al., 2014). Eighty-eight participants (36 female, mean age = 6.81 years, $SD = 0.69$) were recruited from community music and sports programs, and public elementary schools in the Greater Los Angeles Area. Participants were from three groups: 28 children (11 female) who had enrolled and were about to begin participation in the Youth Orchestra of Los Angeles at the Heart of Los Angeles program (hereafter called “music group”), 29 children (12 female) who had enrolled and were about to begin participation in community based soccer or swimming training formed the first comparison group (hereafter called “sports group”), 31 children (13 female) recruited from public elementary schools in the same Los Angeles Area who, at the time of recruitment, were not engaged in an organized and systematic after-school music or sports programs formed the second comparison group (hereafter called “control group”).

The Music Training Program

The Youth Orchestra of Los Angeles at Heart of Los Angeles (YOLA at HOLA) is inspired by the Venezuelan approach known as “El Sistema,” offering free group-based music instruction 4–5 days a week to children from underserved communities of Los Angeles. The program emphasizes systematic, high intensity group music training, focusing on rhythm, melody, harmony, and ensemble practice with a goal of promoting social inclusion. Children (up to 20 per year) are selected by lottery from a list of interested families and, after selection, provided with a violin or viola. The curriculum includes group stringed instrument practice, group singing, Orff, and musicianship (ear training and theory skills), totaling 6–7 h of music instruction per week. Individual practice at home is left to discretion of the families.

The Sports Training Program

The soccer and swimming programs offer free or low-cost training in a community setting to all children whose parents choose to enroll. The soccer program consisted of a 2-h practice three times a week, and a 1-h game each weekend. Soccer practices included warmups, team cheers, skill training (dribbling, passing, etc.), and simulated games. The swimming program consisted of a 1-h practice, two times a week, with additional recreational sessions each weekend. Swimming practices included fitness, endurance, water safety, and stroke development. Both programs were taught by professional coaches.

The sports training group was selected as a comparison group to control for aspects of musical training that would likely be shared by those in a regular, extra-curricular activity such as social engagement, discipline, and sustained effort. Additionally, sports training was chosen due to its sensory motor learning, a component shared with music training. These aspects alone may have beneficial effects on development of both cognitive

and social skills, and thus including an active comparison group is essential.

Exclusion Criteria

Children were excluded from the study if they had a history of psychiatric or neurological disease. At all assessment times, participants' parents were interviewed to ensure that their children had not been diagnosed with a developmental or neurological disorder in the previous year.

During the first visit of each year, children were asked to report any activities they had begun to participate in outside of school. Participants in the control group were excluded from analyses if they had been involved in an extra-curricular music or athletic activity three times a week for at least 6 months. Participants in the sports group were excluded from analyses if they had been involved in a music training program three times a week for at least 6 months. Participants in the music group were excluded from analyses if they had discontinued participation in the music program.

Five time points were used to assess the development of inhibition (**Figure 1**). Children were assessed annually, at approximately the same time each year. Hereafter, "baseline" will be used to refer to the initial assessment, prior to the start of any training (in case of music and sports and no training in case of the control), and subsequent testing times will be referred to as "Year 1" through "Year 4."

Socio-Economic Status

All participants came from equally underprivileged communities in Los Angeles (Habibi et al., 2014). Socioeconomic status (SES) was assessed through parental interviews conducted by research assistants who were native speakers of the parents' preferred language (i.e., English, Spanish, or Korean). Parent interviews included questions ascertaining maternal and paternal education and occupation and annual family household income and size. An SES score was calculated as the mean of each parent's education and annual income. Education level was 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 (Ph.D., MD, JD). Annual income was also 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.

The ethnic distribution included children of Latino, Korean, and African-American backgrounds. 97.7% of participants were bilingual, raised in English–Spanish (93.2%) or English–Korean (4.5%) speaking households while attending English speaking schools that did not offer systematic music programs for their students.

Procedures

Recruitment and induction 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, at each year, was obtained

from each child. 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). To recognize continuous participation, participants were given a \$50 bonus after each year's testing and sent small holiday and birthday gifts. All children were tested individually at the Brain and Creativity Institute at the University of Southern California, or at Heart of Los Angeles in a designated private room.

Behavioral Assessment

Cognitive development was assessed using the Wechsler Abbreviated Scale of Intelligences (WASI-II) for children (Wechsler, 2011). Cognitive inhibition was measured with a child-friendly version of the Flanker task (Davidson et al., 2006; Diamond et al., 2007) (**Figure 2**). An Animal Stroop task was used at Year 1 (Wright et al., 2003), and a computerized Color-Word Stroop task was used starting Year 2. Additionally, we measured inhibitory control using a version of the delayed gratification task (Mischel et al., 1989), where children were presented with increasing reward sizes that they could elect to take immediately, or wait in favor of a larger reward at a future time.

See **Supplementary Material** for a detailed description of each task.

Neuroimaging Assessment

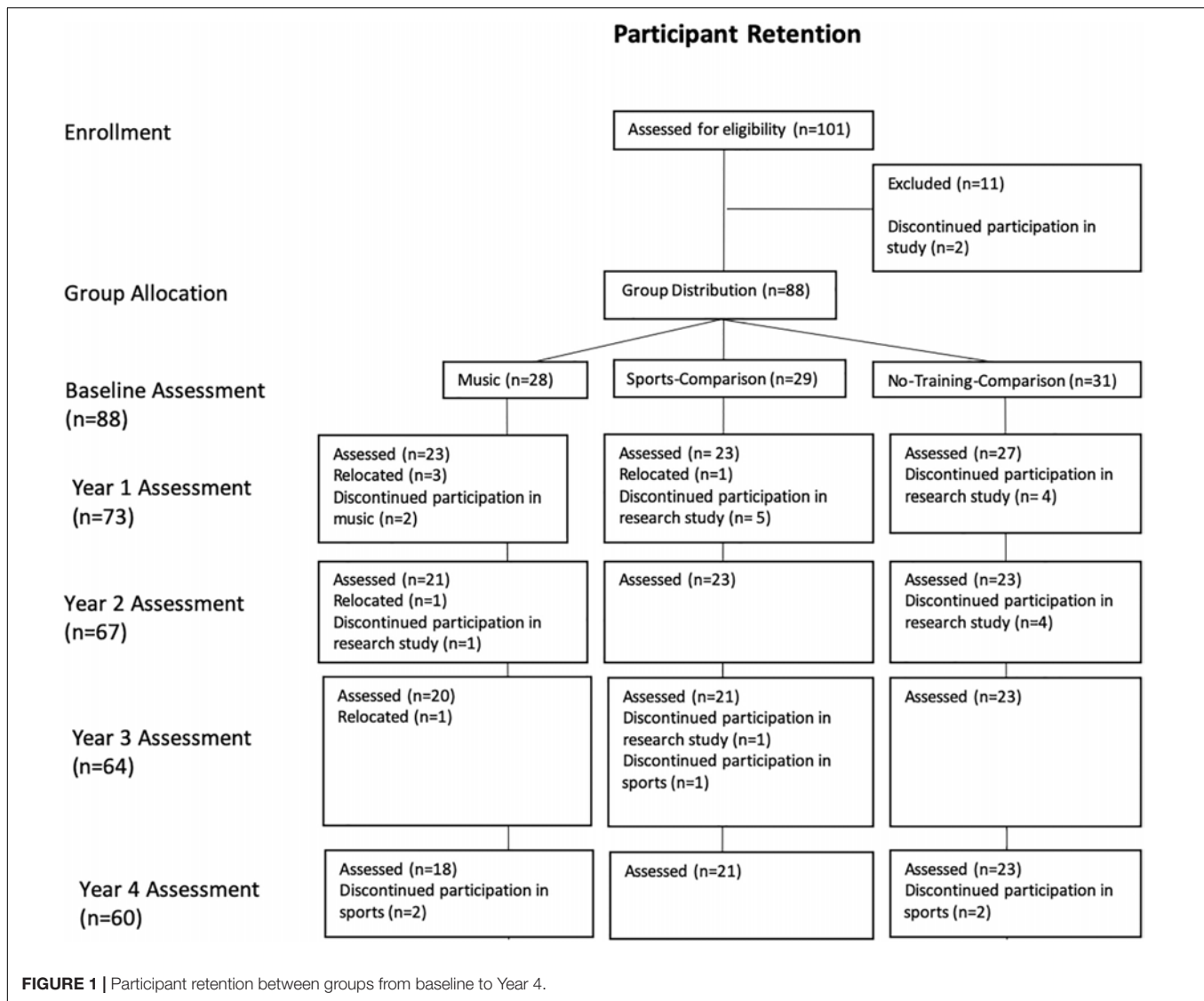
Children underwent anatomical, diffusion, and functional MR imaging of their brain. A child-friendly protocol was implemented that included a training session prior to scanning, where children learned about the scanner and familiarized themselves with the environment in a mock scanner. Results from structural and diffusion scans are discussed in a previous report (Habibi et al., 2017).

The fMRI consisted of a modified version of the Color-Word Stroop task, designed for performance inside the fMRI scanner (**Figure 3**), and performed at Years 2 and 4 (see **Supplementary Material** for detail). Children completed two functional runs, of six blocks each. Blocks were divided between word and color congruent, and incongruent conditions. Each block was followed by a 16 s rest period for a total scan time of 240 s (120 TRs).

A 3T MAGNETOM Prisma System was used to acquire high-resolution T1-weighted structural MRI images, using a 20-channel head coil (1 mm × 1 mm × 1 mm resolution over a 256 × 256 mm × 256 mm FOV; T1/TE/TR = 850/32.05/2300 ms; flip angle = 8°; GRAPPA acceleration factor $R = 2$). A gradient echo, echo-planar, T2*-weighted pulse sequence was used to acquire functional images (TR = 2000 ms, one shot per repetition, TE = 25 ms, flip angle = 90°, 64 × 64 in-plane resolution). Forty-one slices covering the entire brain were acquired (3 mm × 3 mm × 3 mm voxel resolution). Each run of the task consisted of 165 volumes.

Analysis

Statistical analyses were performed using R statistics (R Core Team, 2018). Only participants who had completed all time-point



assessments of a given task were included in the final analysis. Outliers, defined as scores two or more standard deviations above or below the mean, were removed. Not all tasks were performed every year (Table 1). When sphericity assumptions were violated, degrees of freedom were corrected using Greenhouse–Geisser epsilon adjustments. The alpha-level used in all analyses is 0.05.

Behavioral Analysis

Separate repeated measures ANOVAs were conducted for each task, with group as the between-groups factor and year as the within-groups factor. For tasks that included multiple conditions (e.g., reward size or congruency), condition was included as an additional within-groups factor. For tasks that included multiple variables (e.g., accuracy and reaction time), separate repeated measures ANOVAs were used for each variable.

The animal Stroop task, completed at Year 1, was analyzed with one-way ANOVAs for each reaction time and errors, obtained by subtracting congruent trials from incongruent trials.

For the color-word Stroop task, participants responded verbally Years 2 and 3, and through keypresses Year 4, thus analyses for Year 2 and 3 were conducted separately from Year 4. For Year 2 and 3, a repeated measures ANOVA was conducted for accuracy and reaction time, with group as the between-groups factor, and year and congruency as within-groups factors. For Year 4, one-way ANOVAs were conducted for accuracy and reaction time. Stroop reaction time performed during fMRI was analyzed with a repeated measures ANOVA, with between-subjects factors of congruency and year and group as the within-subjects factor. *Post hoc* tests for all tasks were computed using Tukey's HSD, and Cohen's *d* effect size was calculated and reported.

fMRI Analysis

Whole brain

All analyses of functional MR data were completed using the FMRIB Software Library (FSL). Pre-processing included removal

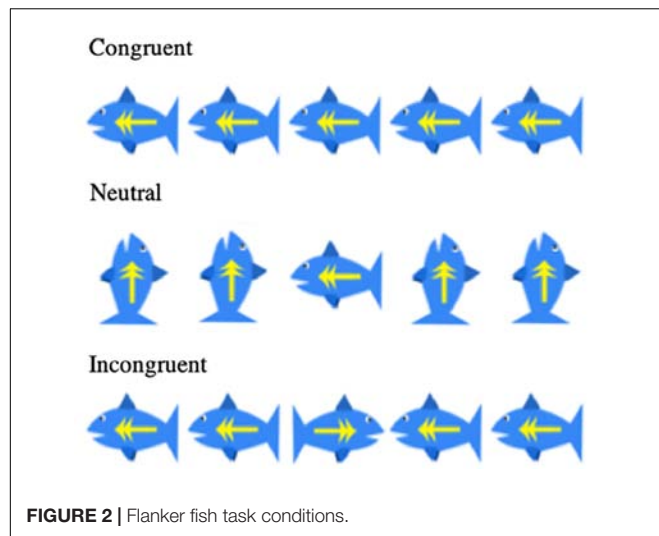


FIGURE 2 | Flanker fish task conditions.

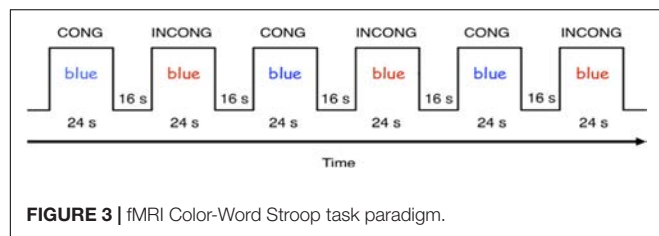


FIGURE 3 | fMRI Color-Word Stroop task paradigm.

TABLE 1 | Task completion per year.

Task	Baseline	Year 1	Year 2	Year 3	Year 4
Animal Stroop		<input type="checkbox"/>			
Color Stroop		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Delayed Gratification				<input type="checkbox"/>	<input type="checkbox"/>
Flanker Fish			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
WASI	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
fMRI Stroop			<input type="checkbox"/>		<input type="checkbox"/>

of non-brain tissue (BET), motion correction (MCFLIRT), slice-timing correction, spatial smoothing (5.0 mm FWHM Gaussian kernel), and high pass temporal filtering (140 s). Motion scrubbing was conducted for each functional run to correct for additional head motion, using root-mean-squares intensity differences (dvars) to determine which slices should be regressed out during GLM analysis (Power et al., 2012). Slices with dvars values greater than the 75th percentile $+1.5 \times$ inter-quartile range were included in a confound matrix added to the GLM. A fixed effects analysis was used to combine the two functional runs for each participant, and FLIRT was used to register images to a high-resolution structural and standard space with 12 DOF and a 2-mm MNI template. The task was modeled with a regressor for each congruency condition using a boxcar convolved with a double-gamma hemodynamic response function. BOLD signal between the task conditions (congruent, incongruent) were contrasted using a general linear model. Models were combined into a mixed-effects analysis, and independent two-sample *t*-tests were used to determine

brain activation differences during these contrasts between groups. A repeated measures ANOVA was used to compare contrasts between groups across years. Age at the time of scan was used as a covariate of non-interest. Statistical inference was completed using Z images and FSL's cluster thresholding, using a cutoff of $Z > 2.3$, and a cluster size probability of $p = 0.05$.

Region of interest

Region of interest (ROI) analysis was conducted to assess percent signal change in regions selected *a priori* due to their known involvement in cognitive control mechanisms (pre-SMA/SMA, ACC, and IFG). 8-voxel spheres were drawn with center coordinates located at the peak voxel base on significant clusters found in the group-level all subject, incongruent > congruent contrast from Year 2 results (Sachs et al., 2017). Percent signal change was calculated from beta values using Featquery in FSL. A repeated measures ANOVA was conducted for each ROI, with year as the within-subjects factor, and group as the between-subjects factor.

RESULTS

A chi-squared test revealed no significant group differences in sex [$\chi^2(2, N = 53) = 2.54, p > 0.05$]. Age at baseline was higher in the Control group ($M = 7.04$ years, $SD = 0.49$ years) than in the Music group ($M = 6.48$ years, $SD = 0.42$ years), [$F(2,52) = 5.34, p < 0.01$]. There were no differences between the sports group age ($M = 6.69$ years, $SD = 0.75$ years) and any other group ($p > 0.05$), and there were no differences between groups in age at any other year. The number of included subjects differed between tasks, thus age at baseline was used as a co-variate only in cases in which there was a significant group difference at onset in subjects that completed all years of a given task. There were no differences between the groups in gender, SES, or cognitive abilities at baseline assessment (Habibi et al., 2014) and thus these factors were not included in the analysis.

Behavioral Results

WASI-II

The WASI-II was completed at all five assessment times by 52 participants (Music $n = 17$, Sport $n = 17$, Control $N = 18$). For FSIQ-4, no main effect of group ($p > 0.05$) or year ($p > 0.05$) was observed. No year by group interaction effect was found ($p > 0.05$). For PRI, no main effect of group was observed ($p > 0.05$). A main effect of year was observed [$F(4,196) = 2.65, p < 0.05, \eta^2 = 0.01$], but *post hoc* comparisons did not reach significance ($p > 0.05$). No year by group interaction was observed ($p > 0.05$). No effects of group, year, or year by group interactions were observed in VCI (all $p > 0.05$), or FSIQ-II (all $p > 0.05$).

Stroop

The animal Stroop task was completed at Year 1 by 54 participants (Music $N = 19$, Sport $N = 14$, Control $N = 21$).

No significant group differences were observed ($p > 0.05$). The color-word Stroop task was assessed at Year 2, Year 3, and Year 4, and completed by 43 participants (Music $N = 17$, Sport $N = 12$, Control $N = 14$). For Year 2 and Year 3, a significant main effect of year on accuracy was found [$F(1,39) = 11.96$, $p < 0.01$, $\eta^2 = 0.24$], where accuracy was greater Year 3 ($M = 0.96$, $SD = 0.06$) than Year 2 ($M = 0.94$, $SD = 0.08$) ($p < 0.01$, $d = 0.40$). A significant main effect of condition was also observed [$F(1,39) = 36.74$, $p < 0.001$, $\eta^2 = 0.49$], where accuracy was greater on the congruent condition ($M = 0.99$, $SD = 0.03$) than the incongruent condition ($M = 0.92$, $SD = 0.08$) ($p < 0.001$, $d = 1.12$). No main effect of group was observed ($p > 0.05$). No year by group ($p > 0.05$) or condition by group ($p > 0.05$) interaction was observed. No significant year by condition by group interaction effect was found ($p > 0.05$). For reaction time (incongruent minus congruent trials), no main effect of year ($p > 0.05$) or group ($p > 0.05$) or year by group interaction ($p > 0.05$) was observed.

For Year 4, a main effect of condition on accuracy was observed in a one-way ANOVA [$F(1,40) = 22.27$, $p < 0.001$, $\eta^2 = 0.36$], where congruent trials ($M = 0.86$, $SD = 0.15$) were responded with more accuracy than incongruent trials ($M = 0.80$, $SD = 0.16$) ($p < 0.001$, $d = 0.37$). There was no main effect of group ($p > 0.05$), or an interaction effect between condition and group ($p > 0.05$). For reaction time (incongruent minus congruent trials), no main effect of group was found ($p > 0.05$).

Flanker Fish Task

The flanker task was completed by 48 participants at Year 2, Year 3, and Year 4 (Music $N = 17$, Sport $N = 15$, Control $N = 16$). There was a significant main effect of year on accuracy [$F(1.17,52.68) = 7.02$, $p < 0.01$, $\eta^2 = 0.14$], in which participants

improved their accuracy each year. There was a significant main effect of condition [$F(1.54,69.45) = 13.58$, $p < 0.001$, $\eta^2 = 0.23$], where participants responded more accurately on congruent trials ($M = 0.97$, $SD = 0.12$) than incongruent trials ($M = 0.91$, $SD = 0.15$) ($p < 0.001$, $d = 0.42$) and more accurately on neutral trials ($M = 0.96$, $SD = 0.13$) than incongruent trials ($p < 0.001$, $d = 0.30$). There was no main effect of group ($p > 0.05$), and no interaction effect of condition by group ($p > 0.05$), or year by group ($p > 0.05$). There was no significant year by condition interaction effect ($p > 0.05$). There was a significant year by condition by group interaction effect [$F(8,180) = 2.061$, $p < 0.05$, $\eta^2 = 0.08$]. *Post hoc* analysis revealed that the condition by year interaction effect was significant in the music group only, [$F(4,64) = 4.67$, $p < 0.01$, $\eta^2 = 0.06$], where during incongruent trials participants responded with higher accuracy at Year 4 ($M = 0.97$, $SD = 0.07$) ($p < 0.01$, $d = 0.98$) and Year 3 ($M = 0.95$, $SD = 0.08$) ($p < 0.05$, $d = 0.82$) compared to Year 2 ($M = 0.83$, $SD = 0.19$). No such year by condition effect was observed in the sport ($p > 0.05$) or control groups ($p > 0.05$) (see **Figure 4**).

For reaction time, there was a significant main effect of year [$F(2,90) = 41.43$, $p > 0.001$, $\eta^2 = 0.48$]. A main effect of group approached significance [$F(2,45) = 2.98$, $p = 0.06$, $\eta^2 = 0.12$], indicating that the Sport ($M = 680.64$ ms, $SD = 151.02$ ms) trended toward longer reaction times than Music ($M = 619.83$ ms, $SD = 138.76$ ms) ($p < 0.001$, $d = 0.42$) and Control ($M = 602.77$ ms, $SD = 137.37$ ms) ($p < 0.001$, $d = 0.54$). There was a significant main effect of condition on reaction time [$F(1.77,79.47) = 189.40$, $p < 0.001$, $\eta^2 = 0.81$]. No year by group interaction effect was observed ($p > 0.05$). A condition by group interaction effect approached significance [$F(4,90) = 2.41$, $p = 0.06$, $\eta^2 = 0.10$], indicating a trend that group differences

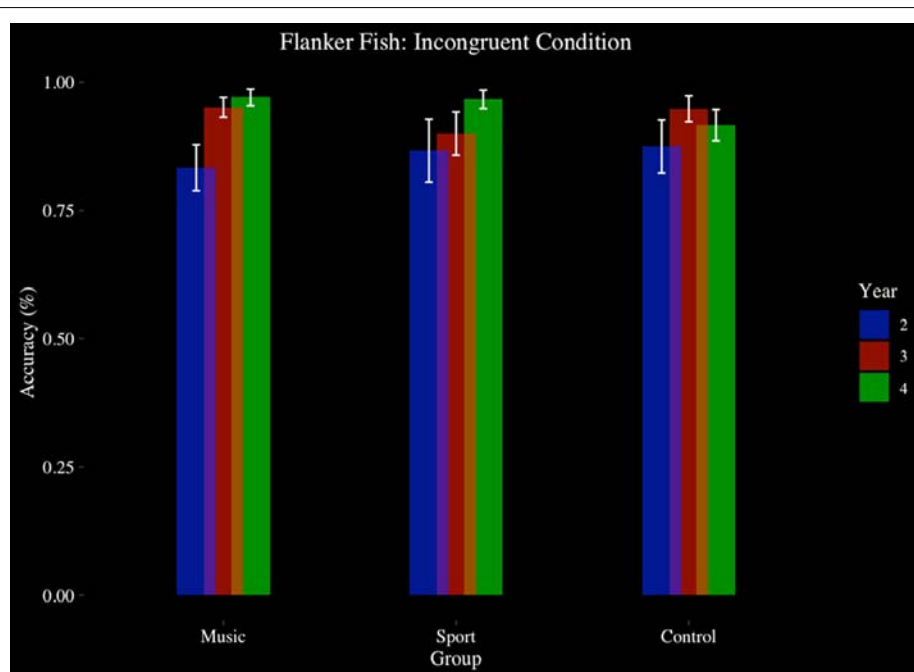


FIGURE 4 | Flanker fish task, accuracy, by group and year: incongruent condition.

were significant in the incongruent and congruent conditions, and not in the neutral condition. No year by condition by group interaction effect was observed ($p > 0.05$) (see **Figure 5**).

Delayed Gratification

The delayed gratification task was assessed at Year 3 and Year 4, completed by 59 participants (Music $n = 18$, Sport $n = 21$, Control $n = 20$). For quarters, no main effect of year ($p > 0.05$), or group was observed ($p > 0.05$). There was a significant main effect of reward size [$F(1.72, 96.2) = 4.88$, $p < 0.05$, $\eta^2 = 0.08$], where large rewards ($M = 0.89$, $SD = 0.32$) were saved more than small rewards ($M = 0.77$, $SD = 0.42$) ($p < 0.05$, $d = 0.32$). No year by group ($p > 0.05$) or reward size by group ($p > 0.05$) interaction was observed. No year by reward size by group interaction was observed ($p > 0.05$). Although group differences were non-significant, the music group consistently saved 100% of quarters in the large condition both Year 3 and Year 4, while the Sport and Control groups saw a decline in large reward delay, from Year 3 to Year 4.

For M&Ms, no main effects of year ($p > 0.05$), or group ($p > 0.05$) were observed. A main effect of reward size [$F(2, 112) = 15.42$, $p < 0.001$, $\eta^2 = 0.22$], where large ($M = 0.81$, $SD = 0.39$) ($p < 0.001$, $d = 0.62$) and medium rewards ($M = 0.76$, $SD = 0.43$) ($p < 0.01$, $d = 0.49$) were saved more than small rewards ($M = 0.54$, $SD = 0.50$). No year by group ($p > 0.05$) or reward size by group ($p > 0.05$) was found. No year by reward size interaction effect was observed ($p > 0.05$). A year by reward size by group interaction was found [$F(3.57, 99.81) = 2.80$, $p < 0.05$, $\eta^2 = 0.09$]. Follow up analyses indicated that the year by group interaction was significant only in the large reward condition

[$F(2, 56) = 3.49$, $p < 0.05$, $\eta^2 = 0.11$], where the Music group ($M = 1.00$, $SD = 0.00$) saved more M&Ms than the Control group ($M = 0.56$, $SD = 0.51$), ($p < 0.01$, $d = 1.19$) at Year 3 (see **Figure 6**). A non-significant trend indicated a stepwise increase in reward delay from small to large rewards that was present only in the Music group, but not the Sport or Control groups, where the music participants saved more large rewards than medium rewards and more medium rewards than small rewards. At Year 4, the group by reward size interaction was no longer significant.

fMRI Results

The fMRI protocol was completed by 40 participants at Year 2 and Year 4 (Music $N = 14$, Sport $N = 11$, Control $N = 15$).

Reaction Time

During the fMRI Stroop task, a main effect of condition on reaction time was observed [$F(1, 26) = 99.26$, $p < 0.001$, $\eta^2 = 0.22$], where participants performed faster at congruent trials ($M = 654.72$ ms, $SD = 129.21$ ms) than incongruent trials ($M = 810.51$ ms, $SD = 169.56$ ms) across years. There was no main effect of group or year ($p > 0.05$), and no interaction between group, year, or condition ($p > 0.05$).

Whole Brain

Whole brain analysis of the incongruent > congruent condition between Year 2 and Year 4 revealed no effect of year and no interaction between year and group on BOLD signal. Due to the small number of participants ($N = 25$) who were scanned in both Year 2 and Year 4, we additionally analyzed Year 4 ($N = 40$) separately. Whole-brain analysis for the incongruent > congruent condition contrast revealed significant signal differences across groups in the left IFG, bilateral ACC,

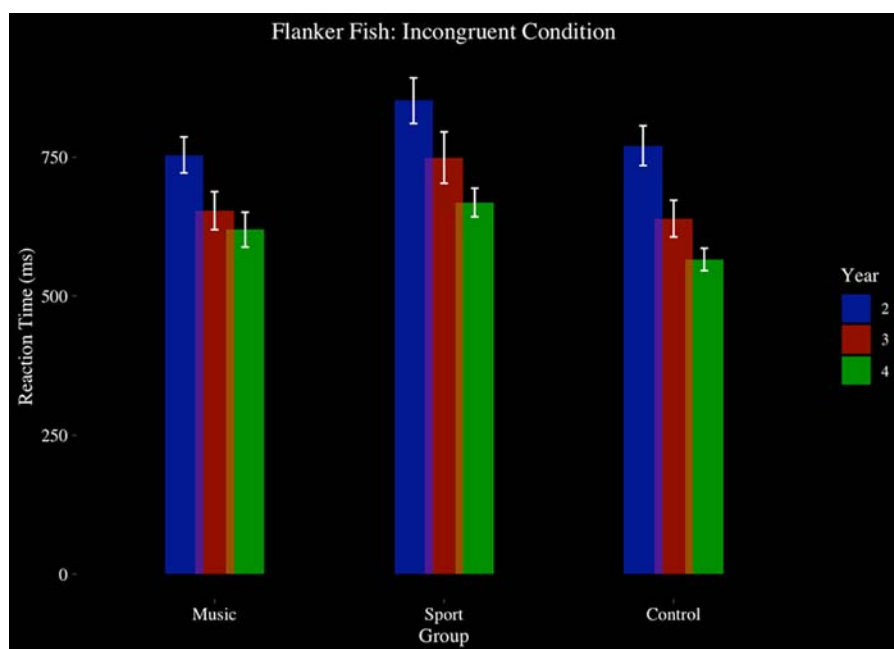


FIGURE 5 | Flanker fish task, reaction time, by group and year: incongruent condition.

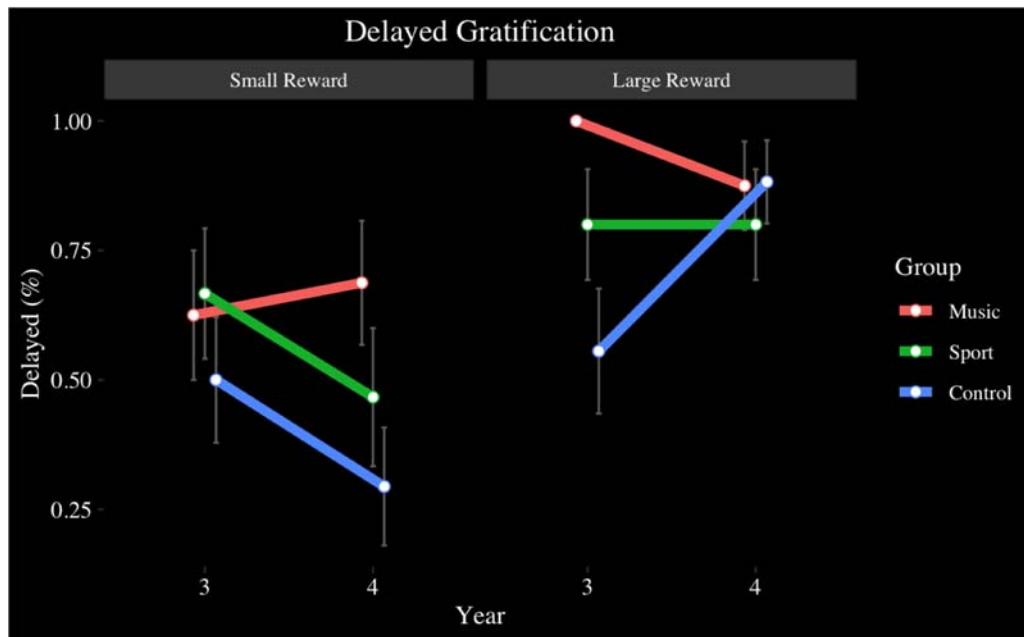


FIGURE 6 | Delayed Gratification task, candies, by group and year: small and large size conditions.

bilateral insula, and the left anterior intra-parietal sulcus at Year 4 (see **Figure 7**). There were no significant differences in BOLD signal found between groups. Analysis of the incongruent > rest condition revealed significant signal differences across groups in the bilateral premotor cortex, right frontal operculum cortex (Broca's area), bilateral occipital pole, left anterior intra-parietal

sulcus, and left angular gyrus (see **Figure 8**). Musicians compared to the control group showed significantly greater activation of the right inferior frontal gyrus in this contrast (see **Figure 9**). No other group differences in BOLD signal were observed at Year 4 (see **Table 2**).

Region of Interest

Region of interest analysis of the incongruent > congruent contrast (see **Figure 10**) in the SMA revealed an effect of group that approached significance ($p = 0.06$), where the music group trended toward greater signal change ($M = 0.19\%$) than the sport group ($M = 0.06\%$) across years. In the left IFG, a year by group interaction effect approached significance ($p = 0.09$), indicating that the music group ($M = 0.25\%$) trended toward greater signal change than the control group (0.19%) year 2, but not year 4 (music $M = 0.12\%$, sport $M = 0.19\%$). In the right ACC, a significant main effect of group was observed [$F(2,28) = 3.66$, $p < 0.05$, $\eta^2 = 0.07$], where the music group ($M = 0.18\%$) had greater percent signal change than the control group ($M = 0.06\%$), but this difference was not significant after correcting for multiple comparisons ($p = 0.07$, $d = 0.61$). No other significant main or interaction effects were observed for this contrast ($p > 0.05$).

The incongruent > rest contrast (see **Figure 11**) revealed a main effect of year in the SMA [$F(1,28) = 11.29$, $p < 0.001$, $\eta^2 = 0.19$], where participants had greater signal change at Year 4 ($M = 0.55\%$) than Year 2 ($M = 0.28\%$). In the left IFG, a significant main effect of group was observed [$F(2,28) = 3.87$, $p < 0.05$, $\eta^2 = 0.10$]. This finding indicated the music group ($M = 0.54\%$) had greater percent signal change than the control group (0.34%) but was not significant after correcting for multiple comparisons ($p = 0.08$, $d = 0.59$). A main effect of year was also observed

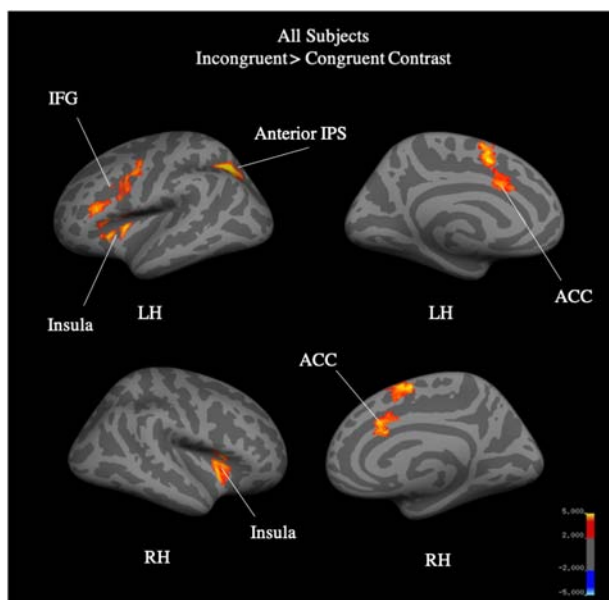


FIGURE 7 | All-subject, whole brain activation for incongruent > congruent contrast of fMRI Color-Word Stroop task at Year 4.

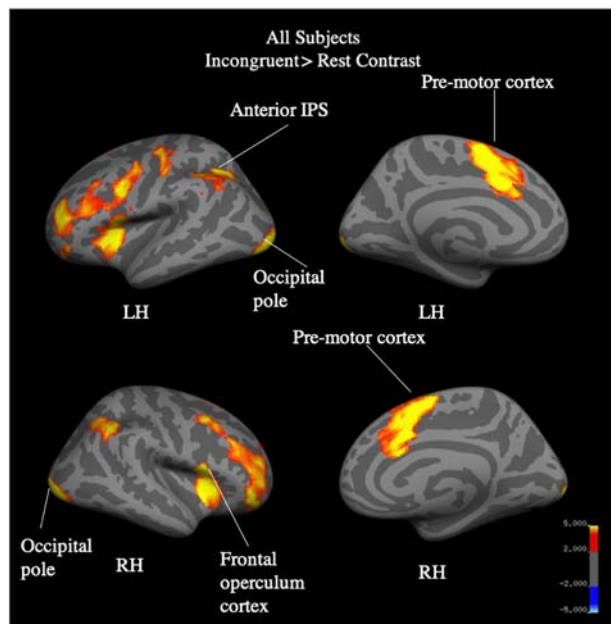


FIGURE 8 | All-subject, whole brain activation for incongruent > rest contrast of fMRI Color-Word Stroop task at Year 4.

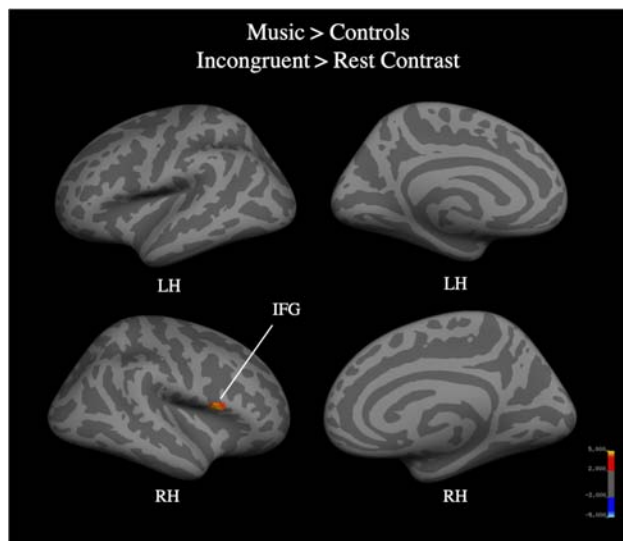


FIGURE 9 | Music group > control group, whole brain activation for incongruent > rest contrast of the fMRI Color-Word Stroop task at Year 4.

[$F(1,28) = 10.10$, $p < 0.01$, $\eta^2 = 0.28$], indicating that percent signal change was greater at year 2 ($M = 0.62\%$) than year 4 ($M = 0.28\%$). No other significant main or interaction effects were observed for this contrast ($p > 0.05$).

Neither response time nor accuracy of incongruent trials of the Color-Word Stroop task performed outside the scanner, on a separate day, was correlated with Incongruent - Congruent percent signal change in any ROI ($p > 0.05$).

DISCUSSION

The present study examined the effects of group-based music training on the development of inhibition skills in children from under-resourced communities. Using a longitudinal design, children involved in music training were compared with active and passive comparison groups (children involved in sports training, and no systematic after-school program, respectively). We assessed changes in children's inhibition skills and its neural correlates over the course of 4 years, using a delayed gratification task, a flanker task, and a Color-Word Stroop task that was performed both inside and outside of an MRI scanner.

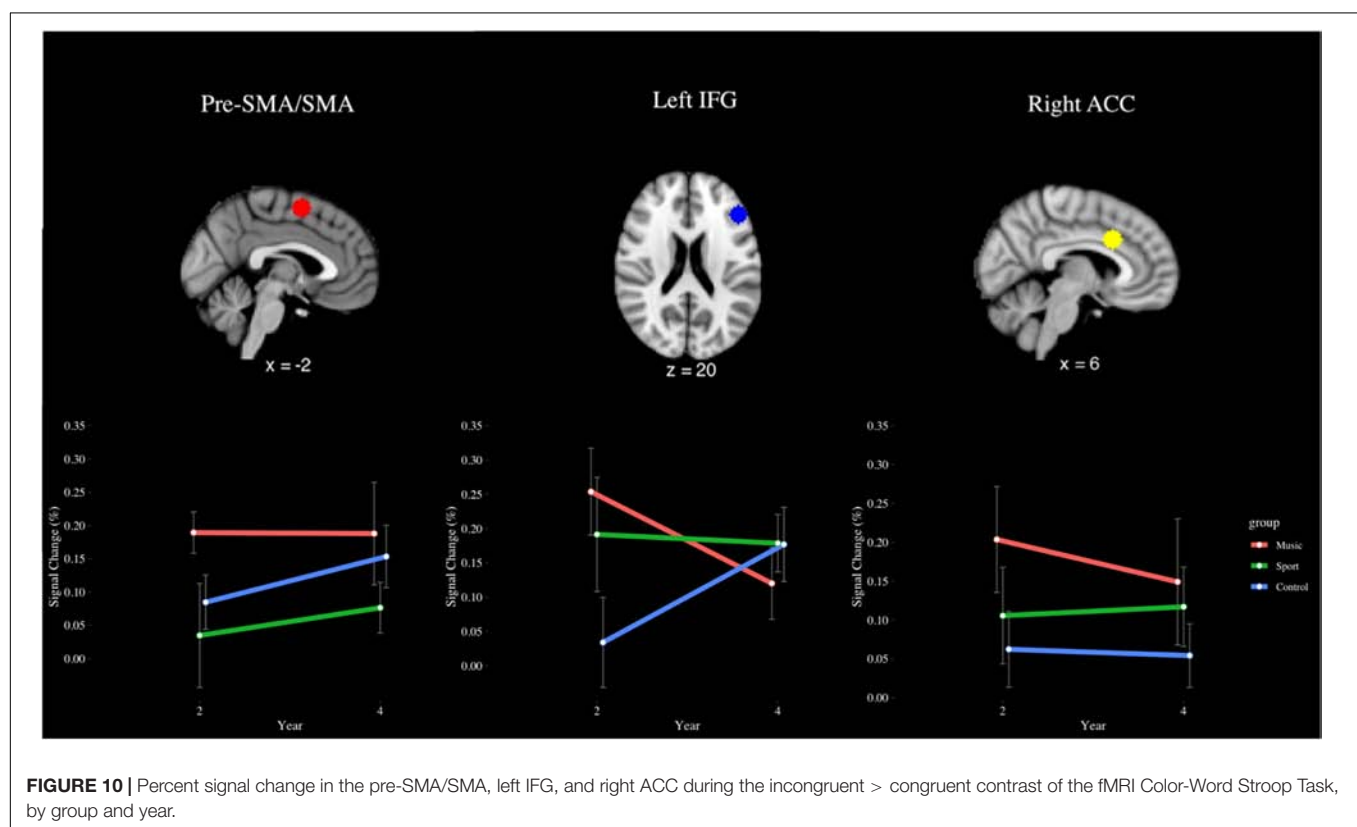
We observed gradual improved performance associated with music training in incongruent trials of the flanker task, in which participants were required to inhibit a dominant response. Music participants improved their accuracy significantly in incongruent trials after 3 and 4 years of training, whereas improvement was observed in the sports and control groups but did not reach a level of significant difference. Overall accuracy was not different among groups after 4 years, however. This may be explained by the music group's slightly, non-significant, lower level of accuracy at Year 2 than the sports and control groups, and that they therefore had more to gain from an intervention. While no group differences were observed at any time point, the development observed in the music group may suggest an effect of training unique to music instruction. This is similar to Holochwost et al.'s (2017) finding that children involved in a community-based orchestra showed improvement in go-no-go and flanker task accuracy after 2 and 3 years of training, and with Jaschke et al.'s (2018) finding of improved go-no-go scores in music participants after 2.5 years of training. However, here, the music group's improvement becomes evident after slightly later, after 3 years of training, as we did not obtain a baseline measure of this task, we do not know how participants would have performed at an earlier stage of training. Reaction time in the music group was similar to that of the control group, while sports participants showed a non-significant trend of slower response times. This was most apparent at the last year of participation, suggesting an effect related to the length of training on development of inhibition skills. These findings are consistent with previous work citing faster reaction times of child (Joret et al., 2017) and adult (Amer et al., 2013) musicians in comparison to controls in a Simon Arrows Task, where the effect is more pronounced in incongruent trials. These findings are additionally in line with research indicating adult musicians perform with faster reaction times overall than controls on inhibition measures such as Auditory and Color-Word Stroop tasks (Bialystok and Depape, 2009; Travis et al., 2011; Amer et al., 2013).

In the Delayed Gratification task, we observed an accelerated ability to reject a small, immediate reward in favor of a delayed, large reward in children who had been involved in music training in comparison to children who had not been involved in any training. This finding is not explained by other differences among the groups, as all groups were similar with regard to SES, age, and other cognitive tasks. Discounting of future rewards in

TABLE 2 | Significant clusters for the fMRI Color-Word Stroop task at Year 4.

Contrast	Group	z-score	x	y	z	Hemisphere	Region
Incongruent > Congruent	All	5.92	-50	30	20	L	IFG, pars triangularis
		5.27	10	18	30	R	ACC
		5.66	-28	-60	42	L	Anterior intra-parietal s
		5.96	40	12	0	R	Insula
Incongruent > Rest	All	7.64	-2	4	58	L/R	Premotor Cortex
		7.62	46	12	2	R	Frontal Operculum Cortex
		8.45	18	-98	-8	R	Occipital Pole
		8.59	-24	-98	-12	L	Occipital Pole
		5.79	-32	-54	38	L	Anterior intra-parietal s
		5.36	50	-50	56	R	Angular Gyrus
	Music > Control	5.08	50	14	6	R	IFG, pars opercularis

Coordinates represent the peak voxel of the cluster in MNI space.



favor of smaller, immediate rewards declines with age, where children increasingly make rational reward choices as they get older (Bettinger and Slonim, 2007; Steinberg et al., 2009). While the sport and control groups followed this pattern by delaying large rewards by the last year of assessment, the music group did so earlier and more often, demonstrating an ability to rationally delay rewards that was better than expected for their age group. Music-trained children were also less likely to delay discount small rewards as they got older, while sports-trained and control children were more likely to do so. While no other studies to

our knowledge have investigated the effect of music training on delayed gratification skills, it is a possibility that attention to detail and orientation to subtle musical cues had a role in extending inhibition to smaller rewards.

Group differences in the delayed gratification and flanker fish tasks, albeit not always statistically significant, indicate a trend toward enhanced inhibition skills in participants involved in music training and are consistent with other longitudinal findings citing overall greater inhibitory control improvements in musically trained children (Moreno et al., 2011; Holochwost

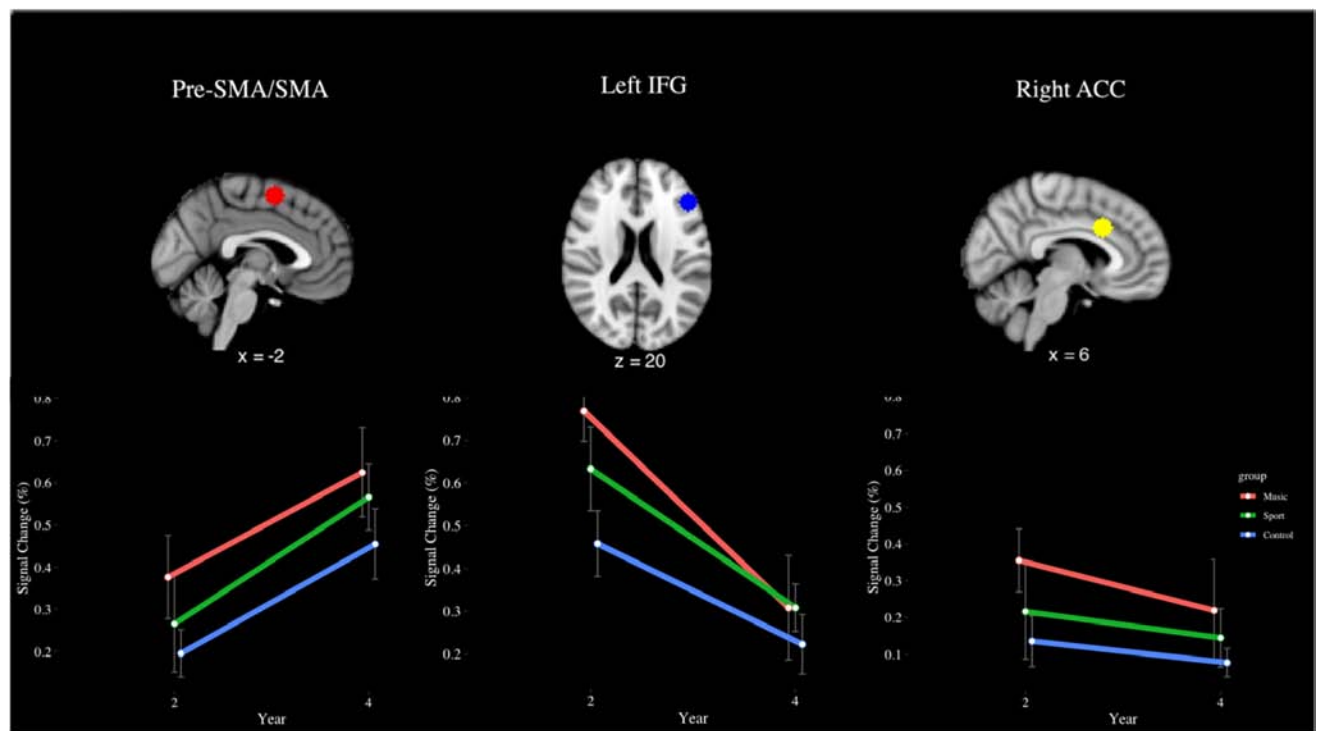


FIGURE 11 | Percent signal change in the pre-SMA/SMA, left IFG, and right ACC during the incongruent > rest contrast of the fMRI Color-Word Stroop Task, by group and year.

et al., 2017; Jaschke et al., 2018). Musicians' advantages in inhibition skills may relate to paralleled processes in disciplined instrumental practice; learning to play a musical instrument involves frequent stopping to correct mistakes and rehearsing small passages in isolation, which both delays the reward of playing a piece in its entirety and requires one to put aside immediate non-musical distractors for the larger reward of musical proficiency. These skills may additionally relate to musicians' regular practice of error monitoring, where children playing music must quickly adjust motor behavior in response to unanticipated musical demands and mistakes.

Results from the fMRI Stroop task support the trend observed in the Delayed Gratification task, where children involved in music training demonstrated an accelerated maturity that was not achieved until a later age by the Sports and Control groups. We previously reported that, at Year 2 (after 2 years of music training), music participants as compared to controls showed greater difference in BOLD signal in the cognitive control network in the incongruent versus congruent blocks, specifically in the IFG, pre-SMA/SMA, ACC, precentral gyrus, and insula (Sachs et al., 2017). In the current analysis, after 4 years, we found no differences between groups in this contrast, suggesting that the music group had matured on the task at an earlier age than the control group. In the incongruent > rest contrast, the music group showed greater BOLD signal in the right IFG, a finding not observed at Year 2. While color-word Stroop paradigm has been associated with only left IFG (Taylor et al., 1997; Fan et al., 2003;

Bernal and Altman, 2009), some studies have reported bilateral activation (Peterson et al., 1999; Banich et al., 2000; Adelman et al., 2002), often driven by differences in design (event-related vs. block; Leung et al., 2000). The right IFG has instead been linked to the stop-signal task and, specifically, cue detection and attentional monitoring, as opposed to the initiation of a motor response (Aron et al., 2003; Hampshire et al., 2010; Sharp et al., 2010). Given the conflicting literature on the right IFG implicated in the Stroop task, and the relatively small sample size here, we do not draw strong conclusions from this finding. Yet, we note that this may indicate improved inhibitory processing associated with length of music training, where music participants are more strongly marking incongruent stimuli as cues to stop a prepotent response.

Our ROI analysis similarly indicated that significant group differences in percent signal change, when present, occurred at Year 2 but not at Year 4. Such differences at Year 4 were present only as non-significant trends, or as significant effects that did not survive *post hoc* analysis. At Year 2, we found that the music group as compared to both control groups had significantly greater percent signal change in the left IFG, pre-SMA/SMA, and right ACC (Sachs et al., 2017). Here, with a notably smaller sample, we find the group by year interaction term approached significance only for the left IFG. This finding is notable due to the left IFG's role in inhibiting prepotent motor responses (Swick et al., 2008), which is a skill that is highly practiced by musicians when monitoring their performance. However, the

reduced number of participants from Year 2 to 4 may have contributed to the statistically non-significant trends.

It should be noted that several studies report increased engagement of the cognitive control network related to decreased efficiency in inhibition tasks (Tamm et al., 2002; Luna et al., 2010). Others have additionally reported activation decreases in these regions after interventions aimed at improving inhibition skills (Chaddock-Heyman et al., 2013; Nishiguchi et al., 2015). We acknowledge that, without a measure of accuracy inside the scanner, we cannot definitively interpret our BOLD signal findings as indicative of better inhibitory processing. Neural activation observed in the music group at Year 2 thus may have been due to greater cognitive effort required to complete the task and that, by Year 4, all groups were engaging the task with similar levels of effort. However, as accuracy in the Stroop task outside of the scanner trended toward a positive correlation with percent signal change in the pre-SMA/SMA and left IFG at Year 2 (Sachs et al., 2017), our interpretation of these findings aligns more with improved inhibitory control processing.

We note an important and possible confounding factor in this study: the high percentage of bilingualism among participants: 96% of participants retained through Year 4 were Spanish-English bilingual: 100% in the music group, 86.67% in the sports group, and 100% in the control group. Although formal bilingualism measures were not obtained, interview data at baseline indicated that all bilingual participants were fluent in both Spanish and English, primarily speaking Spanish at home and English in the classroom. All bilingual participants indicated that they watched television, read books, and listened to music in both English and Spanish. 66.23% of participants at baseline were enrolled in English as a Second Language classes at their school.

Bilingual individuals have demonstrated advanced inhibitory control skills, as evidenced on performance on the flanker task (Costa et al., 2008; Yang et al., 2011; Barac et al., 2016), stop-signal task (Bialystok et al., 2005; Colzato et al., 2008), Stroop task (Bialystok et al., 2008; Hernández et al., 2010), and Simon task (Martin-Rhee and Bialystok, 2008) when compared to monolingual counterparts, despite lower verbal abilities. These differences are evidenced to emerge as young as 7 months (Kovács and Mehler, 2009), indicating that bilingualism produces enhancements of inhibitory skills before the onset of speech production. Explanations of these findings cite evidence indicating that bilingual individuals activate both languages simultaneously, regardless of context (Hernandez et al., 1996; Costa et al., 1999; Kroll et al., 2008). It has thus been proposed that individuals must constantly suppress the context-dependent irrelevant language in order to communicate effectively, leading to an adaptation in neural attentional networks (Bialystok et al., 2009; Green and Abutalebi, 2013). Studies investigating functional brain organization indicate that bilinguals recruit the frontal regions typically associated with executive control, such as the DLPFC, when switching between languages (Hernandez et al., 2000; Hernandez, 2009; Luk et al., 2011). Despite these findings, a recent meta-analysis reported no evidence for a bilingual advantage in executive functioning in adults (Lehtonen et al., 2018), yet did not investigate studies involving children.

Music training and bilingualism may share similar mechanisms in relation to developing inhibitory control skills. While bilingualism and music playing both require the suppression of irrelevant stimuli, additional similarities exist at a neural level. Musicians, in comparison to musically un-trained individuals, activate the DLPFC when passively listening to music (Ohnishi et al., 2001), as bilinguals do when switching between languages (i.e., Hernandez et al., 2000). Koelsch et al. (2005) observed that, when listening to irregular versus regular musical chords, adults and children engage frontal regions, including the inferior frontal gyrus, and that this activity was positively correlated with music training in both age groups. Thus, both musicians and bilinguals appear to recruit regions associated with executive control when engaging in their respective domains, suggesting that intensive experience in either activity would contribute to enhanced inhibition skills.

Given the above evidence, and the fact that almost all of the participants in this study are bilingual, one possible explanation for discrepancies between our results and what we expected to be the effect of music training on inhibition is the contribution of bilingualism. While music training may have an effect, it is likely that multiple factors, including bilingualism, benefit inhibitory control skills. In studies comparing the effects of music training and bilingualism, bilinguals have been reported to perform similarly (Bialystok and Depape, 2009; Schroeder et al., 2016) or better than (Janus et al., 2016) musically trained individuals on measures of inhibition. Other research has indicated no interaction between music training and bilingualism on measures of inhibition, suggesting that both factors may independently contribute to the development of inhibitory control, but do not produce additional combined effects (Moradzadeh et al., 2015; Schroeder et al., 2016). A proposed explanation for the lack of additive benefits of music training and bilingualism in these studies was that each factor alone produced ceiling effects, restricting further benefit of the combination of experiences. Here, we observe differences, while not always significant, between groups in which nearly all participants are bilingual, suggesting that music training and bilingualism may have a small additive effect early in development. We do, however, acknowledge that, if Lehtonen et al. (2018)'s findings generalize to children, the possible explanation related to bilingualism may not be relevant in our findings.

A limitation of the present study is the inevitable decrease of sample size from baseline to Year 4. While we took preventive measures to reduce attrition, it is typical for longitudinal studies, in particular with a population from underrepresented communities, to experience a reduction in participant number over time. However, attrition was similar in all groups. Additionally, as this is not a randomized control trial, it is possible that our participants had pre-existing differences for which we could not account, such as home environment and parental motivation between groups. A strength of this study is that participants were assessed at baseline and demonstrated no differences between groups in measures of musical pitch

and rhythm discrimination (Ilari et al., 2016); it is nonetheless possible that children with inherent musical motivation persisted in the music training program. However, the absence of group differences at baseline in other key variables strongly suggests that any observed enhancements of inhibitory control are due to the intervention in contrast to pre-existing factors.

Despite these limitations, we provide evidence that 4 years of group-based music training leads to modest positive effects on inhibition skills, evidenced by a greater rate of accuracy improvement on a Flanker task and increased rational decisions on a Delayed Gratification task, and its neural correlates in children from underserved communities. The absence of consistent group differences across all inhibition measures point to the other possible explanations such as bilingualism. When present, advantages associated with music training manifest as early, as musicians demonstrate accelerated development of inhibition, but later diminish as non-musically trained children catch up in these skills.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by University of Southern California Institutional Review Board. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

REFERENCES

- Adelman, N. E., Menon, V., Blasey, C. M., White, C. D., Warsofsky, I. S., Glover, G. H., et al. (2002). A developmental fMRI study of the stroop color-word task. *Neuroimage* 16, 61–75. doi: 10.1006/nimg.2001.1046
- Allen, M., Dietz, M., Blair, K. S., van Beek, M., Rees, G., Vestergaard-Poulsen, P., et al. (2012). Cognitive-affective neural plasticity following active-controlled mindfulness intervention. *J. Neurosci.* 32, 15601–15610. doi: 10.1523/JNEUROSCI.2957-12.2012
- Alloway, T. P., Gathercole, S. E., Adams, A. M., Willis, C., Eaglen, R., and Lamont, E. (2005). Working memory and phonological awareness as predictors of progress towards early learning goals at school entry. *Br. J. Dev. Psychol.* 23, 417–426. doi: 10.1348/026151005x26804
- Amer, T., Kalender, B., Hasher, L., Trehub, S. E., and Wong, Y. (2013). Do older professional musicians have cognitive advantages? *PLoS One* 8:e71630. doi: 10.1371/journal.pone.0071630
- Aron, A. R., Fletcher, P. C., Bullmore, E. T., Sahakian, B. J., and Robbins, T. W. (2003). Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. *Nat. Neurosci.* 6, 115–116. doi: 10.1038/nn1003
- Bailey, C. E. (2007). Cognitive accuracy and intelligent executive function in the brain and in business. *Ann. N. Y. Acad. Sci.* 1118, 122–141. doi: 10.1196/annals.1412.011
- Banich, M. T., Milham, M. P., Atchley, R., Cohen, N. J., Webb, A., Wszalek, T., et al. (2000). fMRI studies of stroop tasks reveal unique roles of anterior and posterior brain systems in attentional selection. *J. Cogn. Neurosci.* 12, 988–1000. doi: 10.1162/08989290051137521

AUTHOR CONTRIBUTIONS

AH and BI: conceptualization and supervision. SH, BI, and AH: data curation. SH and AH: behavioral analysis and wrote the manuscript. SH and MS: fMRI analysis and visualization. AH: funding acquisition, investigation, resources, and software. SH: project administration.

FUNDING

This project was supported by the University of Southern California's Brain and Creativity Institute Research Fund.

ACKNOWLEDGMENTS

We thank all participating families and children. We additionally thank our team at the Brain and Creativity Institute who have helped make this project possible, including Priscilla Perez, Alison Wood, Katrina Heine. We also acknowledge the Los Angeles Philharmonic, Youth Orchestra of Los Angeles, Heart of Los Angeles, Brotherhood Crusade's Soccer for Success Program, Vermont Avenue Elementary School, Saint Vincent School, and MacArthur Park Recreation Center for their ongoing support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2019.01080/full#supplementary-material>

- Barac, R., Moreno, S., and Bialystok, E. (2016). Behavioral and electrophysiological differences in executive control between monolingual and bilingual children. *Child Dev.* 87, 1277–1290. doi: 10.1111/cdev.12538
- Berkman, E. T., Kahn, L. E., and Merchant, J. S. (2014). Training-induced changes in inhibitory control network activity. *J. Neurosci.* 34, 149–157. doi: 10.1523/jneurosci.3564-13.2014
- Bernal, B., and Altman, N. (2009). Neural networks of motor and cognitive inhibition are dissociated between brain hemispheres: an fMRI study. *Int. J. Neurosci.* 119, 1848–1880. doi: 10.1080/00207450802333029
- Bettinger, E., and Slonim, R. (2007). Patience among children. *J. Publ. Econ.* 91, 343–363. doi: 10.1016/j.jpubeco.2006.05.010
- Bialystok, E., Craik, F., and Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 859–873. doi: 10.1037/0278-7393.34.4.859
- Bialystok, E., Craik, F. I., Grady, C., Chau, W., Ishii, R., Gunji, A., et al. (2005). Effect of bilingualism on cognitive control in the simon task: evidence from MEG. *NeuroImage* 24, 40–49. doi: 10.1016/j.neuroimage.2004.09.044
- Bialystok, E., Craik, F. I., Green, D. W., and Gollan, T. H. (2009). Bilingual minds. *Psychol. Sci. Publ. Interest* 10, 89–129.
- Bialystok, E., and Depape, A. M. (2009). Musical expertise, bilingualism, and executive functioning. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 565–574. doi: 10.1037/a0012735
- Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436. doi: 10.1163/156856897x00357
- Chaddock-Heyman, L., Erickson, K. I., Voss, M., Knecht, A., Pontifex, M. B., Castelli, D., et al. (2013). The effects of physical activity on functional MRI activation associated with cognitive control in children: a randomized

- controlled intervention. *Front. Hum. Neurosci.* 7:72. doi: 10.3389/fnhum.2013.00072
- Chang, Y. K., Liu, S., Yu, H. H., and Lee, Y. H. (2012). Effect of acute exercise on executive function in children with attention deficit hyperactivity disorder. *Arch. Clin. Neuropsychol.* 27, 225–237. doi: 10.1093/arclin/acr094
- Chen, A. G., Yan, J., Yin, H. C., Pan, C. Y., and Chang, Y. K. (2014). Effects of acute aerobic exercise on multiple aspects of executive function in preadolescent children. *Psychol. Sport Exerc.* 15, 627–636. doi: 10.1016/j.psychsport.2014.06.004
- Colzato, L. S., Bajo, M. T., van den Wildenberg, W., Paolieri, D., Nieuwenhuis, S., La Heij, W., et al. (2008). How does bilingualism improve executive control? A comparison of active and reactive inhibition mechanisms. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 302–312. doi: 10.1037/0278-7393.34.2.302
- Costa, A., Hernández, M., and Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: evidence from the ANT task. *Cognition* 106, 59–86. doi: 10.1016/j.cognition.2006.12.013
- Costa, A., Miozzo, M., and Caramazza, A. (1999). Lexical selection in bilinguals: do words in the bilingual's two lexicons compete for selection? *J. Mem. Lang.* 41, 365–397. doi: 10.1006/jmla.1999.2651
- Davidson, M. C., Amso, D., Anderson, L. C., and Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia* 44, 2037–2078. doi: 10.1016/j.neuropsychologia.2006.02.006
- Davis, C. L., Tomporowski, P. D., McDowell, J. E., Austin, B. P., Miller, P. H., Yanasak, N. E., et al. (2011). Exercise improves executive function and achievement and alters brain activation in overweight children: a randomized, controlled trial. *Health Psychol.* 30, 91–98. doi: 10.1037/a0021766
- Degé, F., and Schwarzer, G. (2011). The effect of a music program on phonological awareness in preschoolers. *Front. Psychol.* 2:124. doi: 10.3389/fpsyg.2011.00124
- Diamond, A., Barnett, W. S., Thomas, J., and Munro, S. (2007). Preschool program improves cognitive control. *Science* 318, 1387–1388. doi: 10.1126/science.1151148
- Duckworth, A. L., and Seligman, M. E. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychol. Sci.* 16, 939–944. doi: 10.1111/j.1467-9280.2005.01641.x
- Eisenberg, N., Sadovsky, A., Spinrad, T. L., Fabes, R. A., Losoya, S. H., Valiente, C., et al. (2005). The relations of problem behavior status to children's negative emotionality, effortful control, and impulsivity: concurrent relations and prediction of change. *Dev. Psychol.* 41, 193–211. doi: 10.1037/0012-1649.41.1.193
- Fan, J., Flombaum, J. I., McCandliss, B. D., Thomas, K. M., and Posner, M. I. (2003). Cognitive and brain consequences of conflict. *Neuroimage* 18, 42–57. doi: 10.1006/nimg.2002.1319
- Flook, L., Smalley, S. L., Kitil, M. J., Galla, B. M., Kaiser-Greenland, S., Locke, J., et al. (2010). Effects of mindful awareness practices on executive functions in elementary school children. *J. Appl. Sch. Psychol.* 26, 70–95. doi: 10.1080/15377900903379125
- Green, D. W., and Abutalebi, J. (2013). Language control in bilinguals: the adaptive control hypothesis. *J. Cogn. Psychol.* 25, 515–530. doi: 10.1080/20445911.2013.796377
- Habibi, A., Damasio, A., Ilari, B., Veiga, R., Joshi, A. A., Leahy, R. M., et al. (2017). Childhood music training induces change in micro and macroscopic brain structure: results from a longitudinal study. *Cereb. Cortex* 28, 4336–4347. doi: 10.1093/cercor/bhx286
- Habibi, A., Ilari, B., Crimi, K., Metke, M., Kaplan, J. T., Joshi, A. A., et al. (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.* 8:690. doi: 10.3389/fnhum.2014.00690
- Hampshire, A., Chamberlain, S. R., Monti, M. M., Duncan, J., and Owen, A. M. (2010). The role of the right inferior frontal gyrus: inhibition and attentional control. *Neuroimage* 50, 1313–1319. doi: 10.1016/j.neuroimage.2009.12.109
- Henrich, J., Heine, S. J., and Norenzayan, A. (2010). The weirdest people in the world? *Behav. Brain Sci.* 33, 61–83. doi: 10.1017/s0140525x0999152x
- Hernandez, A. E. (2009). Language switching in the bilingual brain: what's next? *Brain Lang.* 109, 133–140. doi: 10.1016/j.bandl.2008.12.005
- Hernandez, A. E., Bates, E. A., and Avila, L. X. (1996). Processing across the language boundary: a cross-modal priming study of Spanish-English bilinguals. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 846–864. doi: 10.1037//0278-7393.22.4.846
- Hernandez, A. E., Martinez, A., and Kohnert, K. (2000). In search of the language switch: an fMRI study of picture naming in Spanish-English bilinguals. *Brain Lang.* 73, 421–431. doi: 10.1006/brln.1999.2278
- Hernández, M., Costa, A., Fuentes, L. J., Vivas, A. B., and Sebastián-Gallés, N. (2010). The impact of bilingualism on the executive control and orienting networks of attention. *Bilingualism* 13, 315–325. doi: 10.1017/s1366728909990010
- Holochwost, S. J., Propper, C. B., Wolf, D. P., Willoughby, M. T., Fisher, K. R., Kolacz, J., et al. (2017). Music education, academic achievement, and executive functions. *Psychol. Aesthet. Creat. Arts* 11, 147–166. doi: 10.1037/aca0000112
- Hughes, C., and Dunn, J. (1998). Understanding mind and emotion: longitudinal associations with mental-state talk between young friends. *Dev. Psychol.* 34, 1026–1037. doi: 10.1037//0012-1649.34.5.1026
- Ilari, B. S., Keller, P., Damasio, H., and Habibi, A. (2016). The development of musical skills of underprivileged children over the course of 1 year: a study in the context of an El Sistema-inspired program. *Front. Psychol.* 7:62. doi: 10.3389/fpsyg.2016.00062
- Janus, M., Lee, Y., Moreno, S., and Bialystok, E. (2016). Effects of short-term music and second-language training on executive control. *J. Exp. Child Psychol.* 144, 84–97. doi: 10.1016/j.jecp.2015.11.009
- Jaschke, A. C., Honing, H., and Scherder, E. J. (2018). Longitudinal analysis of music education on executive functions in primary school children. *Front. Neurosci.* 12:103. doi: 10.3389/fnins.2018.00103
- Jentsch, I., Mkrtchian, A., and Kansal, N. (2014). Improved effectiveness of performance monitoring in amateur instrumental musicians. *Neuropsychologia* 52, 117–124. doi: 10.1016/j.neuropsychologia.2013.09.025
- Joret, M. E., Germeyns, F., and Gidron, Y. (2017). Cognitive inhibitory control in children following early childhood music education. *Music. Sci.* 21, 303–315. doi: 10.1177/1029864916655477
- Keller, P. E. (2008). “Joint action in music performance,” in *Emerging Communication: Studies on New Technologies and Practices in Communication: Enacting Intersubjectivity: A Cognitive and Social Perspective on the Study of Interactions*, Vol. 10, eds F. Morganti, A. Carassa, and G. Riva (Amsterdam: IOS Press), 205–221.
- Kirby, K. N., Winston, G. C., and Santiesteban, M. (2005). Impatience and grades: delay-discount rates correlate negatively with college GPA. *Learn. Individ. Diff.* 15, 213–222. doi: 10.1016/j.lindif.2005.01.003
- Koelsch, S., Fritz, T., Schulze, K., Alsop, D., and Schlaug, G. (2005). Adults and children processing music: an fMRI study. *Neuroimage* 25, 1068–1076. doi: 10.1016/j.neuroimage.2004.12.050
- Kovács, Á.M., and Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *Proc. Natl. Acad. Sci. U.S.A.* 106, 6556–6560. doi: 10.1073/pnas.0811323106
- Kroll, J. F., Bobb, S. C., Misra, M., and Guo, T. (2008). Language selection in bilingual speech: evidence for inhibitory processes. *Acta Psychol.* 128, 416–430. doi: 10.1016/j.actpsy.2008.02.001
- Lakes, K. D., and Hoyt, W. T. (2004). Promoting self-regulation through school-based martial arts training. *J. Appl. Dev. Psychol.* 25, 283–302. doi: 10.1016/j.appdev.2004.04.002
- Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., de Bruin, A., and Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychol. Bull.* 144, 394–425. doi: 10.1037/bul0000142
- Lengua, L. J. (2003). Associations among emotionality, self-regulation, adjustment problems, and positive adjustment in middle childhood. *J. Appl. Dev. Psychol.* 24, 595–618. doi: 10.1016/j.appdev.2003.08.002
- Leung, H. C., Skudlarski, P., Gatenby, J. C., Peterson, B. S., and Gore, J. C. (2000). An event-related functional MRI study of the Stroop color word interference task. *Cereb. Cortex* 10, 552–560. doi: 10.1093/cercor/10.6.552
- Luk, G., Green, D. W., Abutalebi, J., and Grady, C. (2011). Cognitive control for language switching in bilinguals: a quantitative meta-analysis of functional neuroimaging studies. *Lang. Cogn. Process.* 27, 1479–1488. doi: 10.1080/01690965.2011.613209
- Luna, B., Padmanabhan, A., and O'Hearn, K. (2010). What has fMRI told us about the development of cognitive control through adolescence? *Brain Cogn.* 72, 101–113. doi: 10.1016/j.bandc.2009.08.005

- Martin-Rhee, M. M., and Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism* 11, 81–93. doi: 10.1017/s1366728907003227
- Miller, E. K., and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202. doi: 10.1146/annurev.neuro.24.1.167
- Miller, H. V., Barnes, J. C., and Beaver, K. M. (2011). Self-control and health outcomes in a nationally representative sample. *Am. J. Health Behav.* 35, 15–27.
- Mischel, W., Shoda, Y., and Rodriguez, M. I. (1989). Delay of gratification in children. *Science* 244, 933–938. doi: 10.1126/science.2658056
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734
- Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., et al. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2693–2698. doi: 10.1073/pnas.1010076108
- Moradzadeh, L., Blumenthal, G., and Wiseheart, M. (2015). Musical training, bilingualism, and executive function: a closer look at task switching and dual-task performance. *Cogn. Sci.* 39, 992–1020. doi: 10.1111/cogs.12183
- Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., and Chau, T. (2011). Short-term music training enhances verbal intelligence and executive function. *Psychol. Sci.* 22, 1425–1433. doi: 10.1177/0956797611416999
- Müntz, T. F., Altenmüller, E., and Jäncke, L. (2002). The musician’s brain as a model of neuroplasticity. *Nat. Rev. Neurosci.* 3, 473–478.
- Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., and Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cogn. Affect. Behav. Neurosci.* 12, 241–268. doi: 10.3758/s13415-011-0083-5
- Nishiguchi, S., Yamada, M., Tanigawa, T., Sekiyama, K., Kawagoe, T., Suzuki, M., et al. (2015). A 12-week physical and cognitive exercise program can improve cognitive function and neural efficiency in community-dwelling older adults: a randomized controlled trial. *J. Am. Geriatr. Soc.* 63, 1355–1363. doi: 10.1111/jgs.13481
- Ohnishi, T., Matsuda, H., Asada, T., Aruga, M., Hirakata, M., Nishikawa, M., et al. (2001). Functional anatomy of musical perception in musicians. *Cereb. Cortex* 11, 754–760. doi: 10.1093/cercor/11.8.754
- Peterson, B. S., Skudlarski, P., Gatenby, J. C., Zhang, H., Anderson, A. W., and Gore, J. C. (1999). An fMRI study of Stroop word-color interference: evidence for cingulate subregions subserving multiple distributed attentional systems. *Biol. Psychiatry* 45, 1237–1258. doi: 10.1016/s0006-3223(99)00056-6
- Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., and Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59, 2142–2154. doi: 10.1016/j.neuroimage.2011.10.018
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Saarikivi, K., Putkinen, V., Tervaniemi, M., and Huotilainen, M. (2016). Cognitive flexibility modulates maturation and music-training-related changes in neural sound discrimination. *Eur. J. Neurosci.* 44, 1815–1825. doi: 10.1111/ejn.13176
- Sachs, M., Kaplan, J., Der Sarkissian, A., and Habibi, A. (2017). Increased engagement of the cognitive control network associated with music training in children during an fMRI Stroop task. *PLoS One* 12:e0187254. doi: 10.1371/journal.pone.0187254
- Schellenberg, E. G. (2011). Examining the association between music lessons and intelligence. *Br. J. Psychol.* 102, 283–302. doi: 10.1111/j.2044-8295.2010.02000.x
- Schroeder, S. R., Marian, V., Shook, A., and Bartolotti, J. (2016). Bilingualism and musicianship enhance cognitive control. *Neural Plast.* 2016:4058620. doi: 10.1155/2016/4058620
- Seeyave, D. M., Coleman, S., Appugliese, D., Corwyn, R. F., Bradley, R. H., Davidson, N. S., et al. (2009). Ability to delay gratification at age 4 years and risk of overweight at age 11 years. *Arch. Pediatr. Adolesc. Med.* 163, 303–308. doi: 10.1001/archpediatrics.2009.12
- Sharp, D. J., Bonnelle, V., De Boissezon, X., Beckmann, C. F., James, S. G., Patel, M. C., et al. (2010). Distinct frontal systems for response inhibition, attentional capture, and error processing. *Proc. Natl. Acad. Sci. U.S.A.* 107, 6106–6111. doi: 10.1073/pnas.1000175107
- Slevc, L. R., Davey, N. S., Buschkuhl, M., and Jaeggi, S. M. (2016). Tuning the mind: exploring the connections between musical ability and executive functions. *Cognition* 152, 199–211. doi: 10.1016/j.cognition.2016.03.017
- Steinberg, L., Graham, S., O’Brien, L., Woolard, J., Cauffman, E., and Banich, M. (2009). Age differences in future orientation and delay discounting. *Child Dev.* 80, 28–44. doi: 10.1111/j.1467-8624.2008.01244.x
- Swick, D., Ashley, V., and Turken, U. (2008). Left inferior frontal gyrus is critical for response inhibition. *BMC Neurosci.* 9:102. doi: 10.1186/1471-2202-9-102
- Tamm, L., Menon, V., and Reiss, A. L. (2002). Maturation of brain function associated with response inhibition. *J. Am. Acad. Child Adolesc. Psychiatry* 41, 1231–1238. doi: 10.1097/00004583-200210000-00013
- Taylor, S. F., Kornblum, S., Lauber, E. J., Minoshima, S., and Koeppe, R. A. (1997). Isolation of specific interference processing in the Stroop task: PET activation studies. *Neuroimage* 6, 81–92. doi: 10.1006/nimg.1997.0285
- Travis, F., Harung, H. S., and Lagrosen, Y. (2011). Moral development, executive functioning, peak experiences and brain patterns in professional and amateur classical musicians: interpreted in light of a Unified Theory of Performance. *Conscious. Cogn.* 20, 1256–1264. doi: 10.1016/j.concog.2011.03.020
- Voss, M. W., Chaddock, L., Kim, J. S., VanPatter, M., Pontifex, M. B., Raine, L. B., et al. (2011). Aerobic fitness is associated with greater efficiency of the network underlying cognitive control in preadolescent children. *Neuroscience* 199, 166–176. doi: 10.1016/j.neuroscience.2011.10.009
- Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence*, 2nd Edn. Bloomington, MN: Pearson.
- Wright, I., Waterman, M., Prescott, H., and Murdoch-Eaton, D. (2003). A new Stroop-like measure of inhibitory function development: typical developmental trends. *J. Child Psychol. Psychiatry* 44, 561–575. doi: 10.1111/1469-7610.00145
- Yang, S., Yang, H., and Lust, B. (2011). Early childhood bilingualism leads to advances in executive attention: dissociating culture and language. *Bilingualism* 14, 412–422. doi: 10.1017/s1366728910000611
- Zatorre, R. J., Chen, J. L., and Penhune, V. B. (2007). When the brain plays music: auditory-motor interactions in music perception and production. *Nat. Rev. Neurosci.* 8, 547–558. doi: 10.1038/nrn2152
- Zuk, J., Benjamin, C., Kenyon, A., and Gaab, N. (2014). Behavioral and neural correlates of executive functioning in musicians and non-musicians. *PLoS One* 9:e99868. doi: 10.1371/journal.pone.0099868

Conflict of 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.

Copyright © 2019 Hennessy, Sachs, Ilari and Habibi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Selectively Enhanced Development of Working Memory in Musically Trained Children and Adolescents

Katri Annukka Saarikivi^{1,2*}, Minna Huotilainen^{2,3}, Mari Tervaniemi^{2,4} and Vesa Putkinen^{2,5}

¹Cognitive Brain Research Unit, University of Helsinki, Helsinki, Finland, ²Department of Psychology and Logopedics, Medicum, Faculty of Medicine, University of Helsinki, Helsinki, Finland, ³Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland, ⁴CICERO Learning, Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland, ⁵Turku PET Centre, Turku, Finland

In the current longitudinal study, we investigated the development of working memory in musically trained and nontrained children and adolescents, aged 9–20. We measured working memory with the Digit Span (DS) forwards and backwards tests ($N = 106$) and the Trail-Making A and B (TMT-A and B; $N = 104$) tests three times, in 2011, 2013, and 2016. We expected that musically trained participants would outperform peers with no musical training. Indeed, we found that the younger musically trained participants, in particular, outperformed their nontrained peers in the TMT-A, TMT-B and DS forwards tests. These tests all primarily require active maintenance of a rule in memory or immediate recall. In contrast, we found no group differences in the backwards test that requires manipulation and updating of information in working memory. These results suggest that musical training is more strongly associated with heightened working memory capacity and maintenance than enhanced working memory updating, especially in late childhood and early adolescence.

Keywords: musical training, longitudinal, working memory, updating, maintenance, development, trail-making test, Digit Span

OPEN ACCESS

Edited by:

Paul J. Colombo,
Tulane University, United States

Reviewed by:

Fabiana Silva Ribeiro,
Portuguese Catholic University,
Portugal
Lutz Jäncke,
University of Zurich, Switzerland

*Correspondence:

Katri Annukka Saarikivi
katri.saarikivi@helsinki.fi

Received: 17 June 2019

Accepted: 07 October 2019

Published: 06 November 2019

Citation:

Saarikivi KA, Huotilainen M, Tervaniemi M and Putkinen V (2019) Selectively Enhanced Development of Working Memory in Musically Trained Children and Adolescents. *Front. Integr. Neurosci.* 13:62. doi: 10.3389/fnint.2019.00062

INTRODUCTION

Musically trained individuals have been reported to outperform musically nontrained peers in various kinds of cognitive tests not directly related to music-making, including ones measuring long-term verbal and visual memory (Chan et al., 1998; Ho et al., 2003), executive functions (Bialystok and Depape, 2009; Degé et al., 2011; Moreno et al., 2011; Zuk et al., 2014; Saarikivi et al., 2016; however, see Schellenberg, 2011), and even intelligence (Schellenberg, 2004, 2006; Moreno et al., 2011). Executive functions (Stuss and Alexander, 2000; Jurado and Rosselli, 2007; Diamond, 2013) are cognitive processes typically divided into three related components: working memory, inhibition, and cognitive flexibility (Miyake et al., 2000; Lehto et al., 2003; Miyake and Friedman, 2012; Diamond, 2013). These three processes allow individuals to acquire, maintain, manipulate, and update representations of information of the environment, and monitor, direct and alter behavior according to these representations. Multicomponent models of working memory propose subprocesses for maintaining representations of information in memory and for manipulating this information. For instance, the influential

model of Baddeley and Hitch (1974) divided working memory into two components for storage and manipulation of verbal and visual material, and a cognitive control unit (for other models, see e.g., Cowan, 1988, 1999; Unsworth and Engle, 2007). Neuroimaging and lesion studies have found separate neural functions for memory representations and attention processes that govern manipulation of this information, supporting these modular views of working memory (Postle et al., 1999; Gerton et al., 2004; Owen et al., 2005; reviews: D'Esposito et al., 1995; Miller and Cohen, 2001; Linden, 2007; Nee et al., 2012; Rottschy et al., 2012; Eriksson et al., 2015; for a discussion on differences between short-term memory and working memory, see e.g., Unsworth and Engle, 2007; Cowan, 2008; Aben et al., 2012). Multicomponent models of working memory have been validated in child studies (Gathercole et al., 2004; Gray et al., 2017), and separate brain mechanisms for encoding, maintenance and retrieval of verbal information have also been found in neuroimaging studies of children, from the age 6 onwards, and in adolescents (Gathercole et al., 2004; Siffredi et al., 2017).

Working memory and other executive functions develop from early childhood until adolescence (Cepeda et al., 2001; De Luca et al., 2003; Vuontela et al., 2003; Luna et al., 2004; Zelazo et al., 2004; Huizinga et al., 2006), following the maturation of prefrontal areas (Casey et al., 2000; Fuster, 2002; Kwon et al., 2002; Bunge and Wright, 2007; Kharitonova et al., 2013). Different executive functions however mature at slightly different rates. Development of shifting ability, which is related to cognitive flexibility, has been found to continue until adolescence (Huizinga et al., 2006; Best and Miller, 2010; Huizinga and van der Molen, 2011), and development of working memory even further, until early adulthood (Kwon et al., 2002; Huizinga et al., 2006; Satterthwaite et al., 2013).

Several cross-sectional studies have reported varying musician advantage in working memory tasks. For example, in a study by George and Coch (2011), years of musical training correlated positively with scores in both verbal and visual span tests for memory in college-aged individuals. Similarly, in another study (Talamini et al., 2016), musically trained adults outperformed nontrained peers in auditory as well as visual span tests for working memory. Finally, Zuk et al. (2014) found better performance in the Digit Span backwards test in adult musicians compared to nonmusicians, but not in musically trained children compared to nontrained peers.

Longitudinal studies with children suggest that the putative musician advantage in memory tasks may be caused by training and does not solely reflect pre-existing differences (for a discussion on problems of inferring causation from these kinds of studies, see Schellenberg, 2015). In the study by Ho et al. (2003), verbal long-term memory improved in children who continued musical training during a year-long follow-up, but not in those who did not. Similarly, in a study following the development of musically trained and nontrained children (Bergman Nutley et al., 2014), musical training was associated with improvement of verbal working memory as measured by the backwards Digit Span test, but also visual working memory as measured by a visuo-spatial working memory task.

Another longitudinal study (Fujioka et al., 2006), comparing the development of children who undertook music lessons for 1 year to the development of children in a Control group, found significant improvement of working memory as measured by the Digit Span test only in the Music group. In another study (Roden et al., 2014), improvement of working memory was observed in preschool-aged children after 18 months of musical training, but not in an active Control group. In the study, effects were found specifically in tests measuring the phonological loop and the central executive subcomponents of working memory. The phonological loop was measured with the One Syllable Word Span Test, requiring participants to memorize and recite a sequence of words in the order they were presented, and the Nonword recall test, requiring participants to recite a nonword immediately after hearing it. The central executive was measured by complex span tasks requiring processing and storing information at the same time or requiring reversal of the order of a memorized sequence of information units. Last, in a recent study with quasi-random assignment of children into musical training and Control groups (Guo et al., 2018), it was found that 6 weeks of musical training improved working memory. Auditory working memory was assessed with the Digit Span forward and backward tests and with the Letter-Number Sequencing test. Both require working memory maintenance of aurally acquired information and updating and manipulation of that information in memory.

The notion that musical training might influence memory skills is further supported by findings of training-related changes in brain structures important for working memory. In their seminal study on structural differences, Gaser and Schlaug (2003) found that musicians had greater gray matter density in areas important for motor and auditory processes, and also a region of the cerebellum connected to working memory (Stoodley et al., 2012). Similarly, in a study by James et al. (2014), musical training correlated positively with gray matter density in a cerebellar areas and basal ganglia important for working memory. Another MRI study found increased thickness of frontal areas related to working memory in musicians, when compared to non-musicians (Bermudez et al., 2008).

Musical training has also been connected to changes in brain functions related to working memory. In the study by George and Coch (2011), musicians had shorter latencies of electrical brain responses (P3) to changes in visual as well as auditory stimuli, as well as larger P3 amplitudes to tonal changes. The P3 response is thought to reflect updating of working memory. In a recent study (Cheung et al., 2017), musically trained individuals outperformed nontrained peers in tasks for verbal memory and also differed in electrical brain activity measured during a verbal memory task. Specifically, musically trained individuals showed more intrahemispheric coherence in the theta band. In an fMRI study (Pallesen et al., 2010), musicians showed greater activation of brain networks for attention and working memory, including frontal, parietal, and subcortical areas than nonmusicians. In another study (Schulze et al., 2011), musicians showed different patterns of activation of brain areas during memory encoding and rehearsal of structured and unstructured tonal sequences. Nonmusicians did not show differences in activation patterns

during these tasks. Musicians also outperformed nonmusicians in learning the tonal sequences.

A recent meta-analysis on cross-sectional and longitudinal studies on music-related enhancement of working memory in children and adults (Talamini et al., 2017) concluded that musicians and musically trained individuals have a clear advantage across different memory tasks when compared to nontrained peers. However, according to the results, the effect size depended on the type of information that was processed and on the memory processes required by the task. In general, the musician advantage was stronger for working memory than long-term memory, and for auditory rather than visual stimuli.

Another recent meta-analysis (Sala and Gobet, 2017a) found only a weak musician advantage in memory tasks. In this meta-analysis, however, no distinction was made between working memory and long-term memory, which may have obscured the effects of musical training on working memory reported in several cross-sectional and longitudinal studies.

In sum, there is evidence of an association between musical training and specifically verbal working memory. However, longitudinal studies have focused on school- or preschool-aged children even though executive functions are known to develop long into adolescence. As a result, it is still unclear how musical training specifically augments the development of working memory, and for how long into adulthood the possible advantage persists.

In this longitudinal study, we compared the working memory skills of 114 musically trained and nontrained children and adolescents aged 9–20. During this age range, executive functions including working memory undergo significant development, owing to the protracted development of brain areas such as the frontal lobes that important for these skills, but also begin to reach maturity (Taylor et al., 2013, 2015). The sample allows for investigating the effects of musical training on working memory, and the persistence of these effects during a developmentally highly interesting window of time. The study aims at answering questions that remain unresolved in research examining the effects musical training may have on cognitive development: does musical training augment the development of working memory, does musical training produce an advantage in working memory tasks, does this advantage persist into adulthood?

To investigate working memory, we employed two broadly studied and well-established tests: the Digit Span backward and forwards tests and the Trail-Making Test A and B. Data on performance in Digit Span tests were collected during a 3-year follow-up and the TMT-A and B tests during a 2 year follow-up. Based on previous literature, we expected musically trained children and adolescents to perform better than nontrained peers in all tests.

MATERIALS AND METHODS

Participants

Altogether 106 children and adolescents aged 9–20 years participated in the study (Tables 1, 2). The musically trained participants ($N = 54$, 32 females) had started training on a

musical instrument approximately at age 7. They had attended or were currently attending a public elementary school that emphasizes music in the curriculum. In addition to weekly classes in classical instrumental training, school days contained music lessons such as choir and ensemble training and performances. Thus, at age 9, participants had a total of approximately 2 years of musical training and participation in the musical curriculum, at age 11, 4 years and so on. The nontrained participants ($N = 52$, 26 females) had no formal training on a musical instrument. They attended or had attended a standard elementary school with weekly group-based music lessons until the age of 13, but no instrumental tuition. No children reported hearing deficits or neurological impairments. The Music and Control groups were matched in SES and IQ (Putkinen et al., 2014).

Written informed consent for participation was obtained from guardians of underaged participants or from over 18-year old participants themselves before the experiment. All participants also gave verbal consent for their participation. Participants were rewarded three movie tickets for taking part in the measurement. The experiment protocol was approved by the Ethical Committees of the Department of Psychology and of the Faculty of Behavioural Sciences, both at the University of Helsinki, Finland.

Working Memory Tests

The Digit Span forwards and backwards (DS forwards, DS backwards) tests (WISC-IV, Wechsler, 2010) as well as the TMT-A and B (TMT-A, TMT-B; Poutiainen et al., 2010) were used to measure verbal working memory. In the Digit Span forwards test, subjects are aurally presented with a series of digits, and immediately recite them from memory. In the DS backwards test, participants are required to recite the presented digits in reverse order. There are multiple presentation rounds, with the experimenter always adding one to-be-memorized digit. The forwards test requires active maintenance of information in mind, and the backwards tests also manipulation of this information. Performance is evaluated by the total of digits that the participant is able to correctly recite.

The TMT-A requires the participant to connect digits printed randomly on paper by drawing a line from number to number in a sequential order. The Trail-Making Test B requires participants to alternate between connecting numbers and letters printed on the paper in order (1-A-2-B-3-C...). Both tests require maintenance of the rule of the task in mind and also maintaining awareness of where one is progressing on the sequence of digits (A) and both digits and letters (B). Performance is measured by the time taken to complete the test.

Procedure

This study is part of a longitudinal study that started in 2003 investigating the maturation of auditory processes and executive functions in children undergoing musical training and a control group. The study entailed also EEG measurements and other tests for various cognitive skills (these data are reported elsewhere). Measurements were conducted every 2 years, with a new group of 7-year-olds recruited every 2 years. The data, therefore, contains measurements from the same participants

but from different years. The data reported here include measurements conducted in the years 2011, 2013, and 2016 for the DS forwards and backwards tests and from the years 2011 and 2013 for the Trail-Making A and B Tests (The TMT was not conducted in 2016). Not all children took part in every measurement. For the tests, 25 children participated in only one measurement, 43 in two, and 38 in all three measurements. For the Trail-Making A and B Tests, 43 children participated in one measurement and 61 in both.

The cognitive tests were conducted before the EEG experiments and took a maximum of 1 h altogether. Upon arrival at the laboratory, written informed consent as well as oral consent was received from the participants. After this, the participants accompanied the experimenter to a room to complete the tests. Experimenters were graduate students, trained to work with children and adolescents and to administer the tests. The space was a comfortably lit sound-proofed room, previously used as an EEG lab, converted for testing use. The experimenter and the subject were orthogonally seated at a table. After the tests were completed, the subject was escorted to the EEG lab, where the EEG cap was attached, and the subject informed more closely about the EEG experiment. EEG measurement ensued. Participants were offered bathroom breaks when needed, and cookies and juice before the EEG measurement as well as half way through it.

Statistical Analysis

Completion times in the TMT-A and B, and span (number of correctly recited digits) in DS forwards and backwards were included separately in analyses of test performance. The effect of age and group membership on test performance was modeled with linear mixed modeling using the lmer function [Test Score \sim Age * Group + (1|Subject)] of the lme4 package in R (Bates, 2005; Bates et al., 2007). Age was mean centered so that the significant effect of Group indicates a group difference in the test score at average age of the participants (mean ages for the DS and TMTs were 14.39 and 13.44 years, respectively). Linear mixed modeling was selected as the analysis approach since it allows a different number of data points across subjects and takes into account the correlated nature of the data within a subject. Values below the $Q1 - 1.5 * IQR$ (inter-quantile-range) or above $Q3 + 1.5 * IQR$ were classified as outlier and replaced by the lower or upper cutoff values of this range, respectively. This procedure was applied twice for the DS backwards and Trail Making A data and five times for the Trail Making B data.

RESULTS

Performance of participants in all tests except for the DS backwards test improved with age (Figures 1, 2). Musically trained participants outperformed nontrained participants in the DS forwards test, but not in the backwards test. The musically trained individuals also outperformed nontrained peers in the TMT-A. However, the group difference depended on age. The difference between performance in the Music and Control groups decreased with age. A similar age-dependent effect was also

found for performance in Trail-Making Test B. The results are described in more detail below.

Digit Span Forwards and Backwards

Performance in the DS forwards test improved with age (estimated increase in span per year: 0.22, $p < 0.001$). The Music group outperformed the Control group (estimates for the Music and Control groups: Control difference in span: 0.56, $p < 0.05$). The model revealed no evidence that this group difference was dependent on age (Group*Age interaction, *ns*).

For the DS backwards test, there were no significant effects of Age or Group or and no significant interaction between these predictors.

Trail-Making A and B Tests

Subjects' performance in the TMT-A improved with age (estimate for the decrease in completion time per year: -2.87 , $p < 0.001$). There was a trend-level effect of group suggesting that the Music group outperformed the Control group in this test (estimate for the Music < Control difference: -2.50 , $p < 0.07$). However, there was also a significant interaction between Age and Group indicating that the group difference was more pronounced in the younger children and decreased with age (estimate for the Music < Control differences in the change in completion time per year: 1.11, $p < 0.05$).

The performance in the Trail-Making Test B also improved with age (Estimate for the decrease in completion time per year: -9.25 , $p < 0.001$). For this test, there was a significant effect of group indicating that the Music group outperformed the Control group (estimate for the Music < Control difference: -12.06 , $p < 0.05$) as well as a significant interaction between Age and Group indicating that this group difference decreased with age (estimate for the Music < Control differences in the change in completion time per year: 4.69, $p < 0.05$).

DISCUSSION

In this study, we investigated the development of working memory in musically trained and nontrained children and adolescents. Musically trained participants outperformed nontrained peers in the DS forwards test as well as the Trail-Making A and B tests. Furthermore, the group difference in the two latter tests decreased with age. We did not find a significant difference between the Music and the Control groups in the DS backwards test.

A Musician Advantage in the DS Forwards and Trail-Making A and B Tests

The better performance of the Music group in the DS forwards test concurs with previous research showing a musician advantage in tests for memory (Chan et al., 1998; Fujioka et al., 2006; George and Coch, 2011; Bergman Nutley et al., 2014; Roden et al., 2014; Zuk et al., 2014; Talamini et al., 2016; Guo et al., 2018; review: Talamini et al., 2017). It is noteworthy, however, that studies reporting memory enhancement in musicians have conceptual and methodological differences. In the study by Cheung et al. (2017), verbal memory enhancement was reported

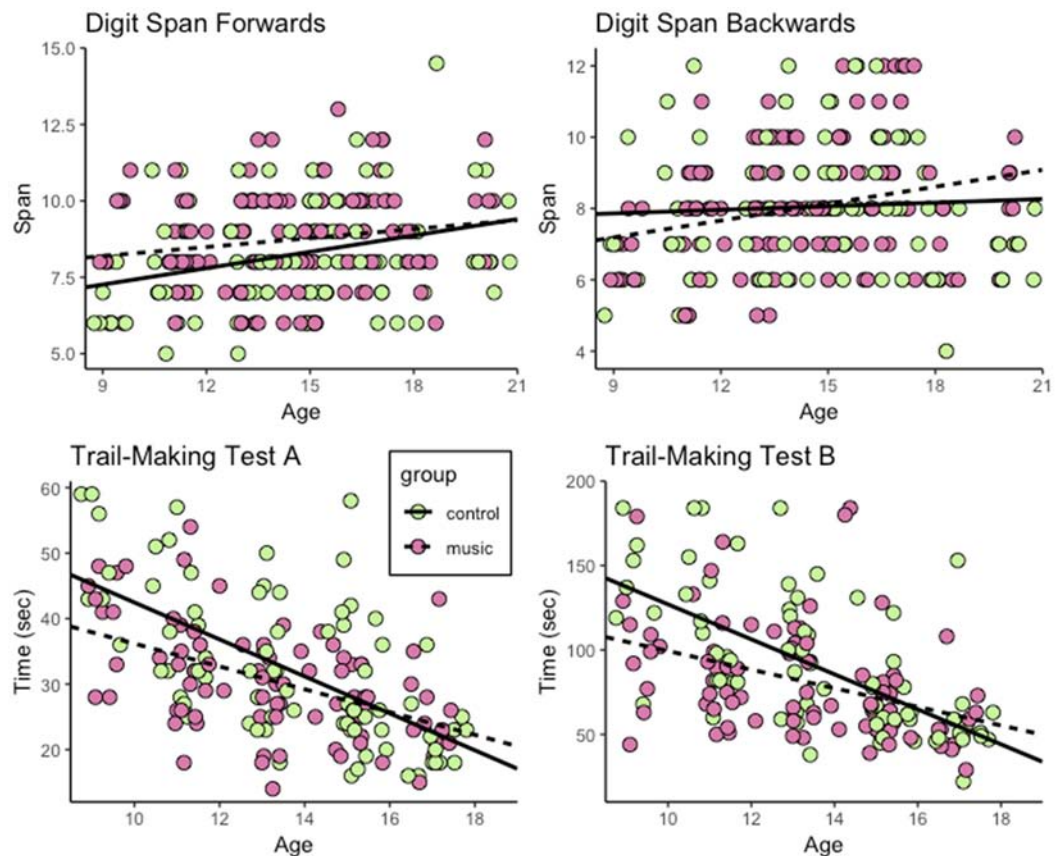


FIGURE 1 | Performance of participants in the Digit Span forwards and backwards, and the Trail-Making A and B tests across all age groups. Music and Control groups represented with different colors.

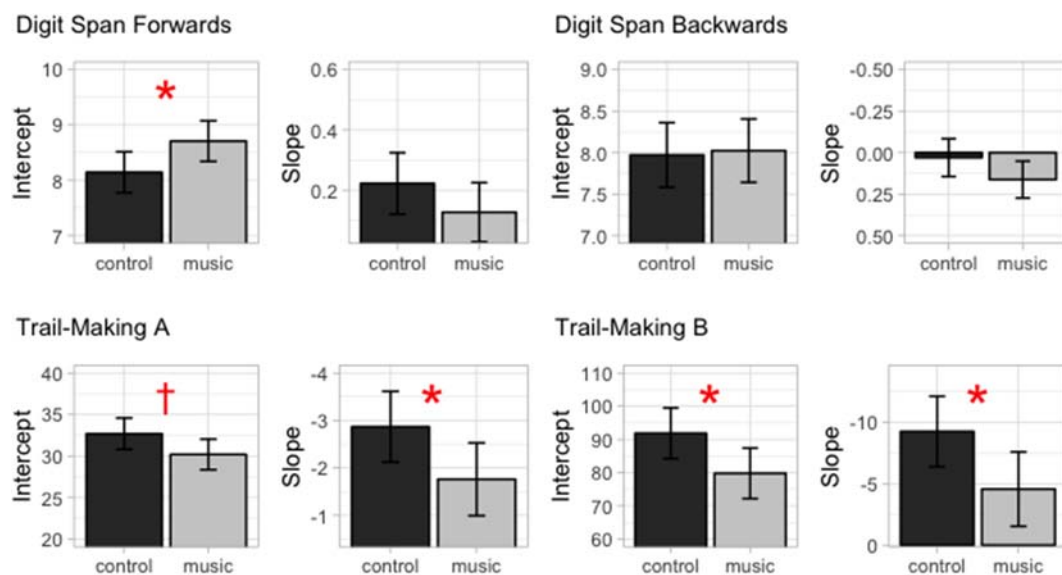


FIGURE 2 | The intercepts (i.e., estimated performance at mean age) and slopes (i.e., estimated change in performance with age) for each test separately for the Music and Control groups. The errorbars indicate 95% confidence intervals. * $p < 0.05$, † $p = 0.07$.

TABLE 1 | Ages of participants in the Music and Control groups per measurement year for the Digit Span Test.

Group	Year	Mean age	Minimum age	Maximum age	Standard deviation
Control	2011	12.77	8.75	15.92	2.26
Control	2013	14.61	10.43	17.81	2.45
Control	2016	17.25	13.56	20.78	2.51
Music	2011	12.46	8.92	15.83	2.31
Music	2013	14.09	10.60	17.44	2.20
Music	2016	16.76	13.50	20.29	2.30

TABLE 2 | Ages of participants in the Music and Control groups per measurement year for the Trail-Making Test.

Group	Year	Mean age	Minimum age	Maximum age	Standard deviation
Control	2011	12.77	8.75	15.92	2.26
Control	2013	14.61	10.43	17.81	2.45
Music	2011	12.46	8.92	15.83	2.31
Music	2013	14.09	10.60	17.44	2.20

based on performance in a task requiring immediate as well as delayed recall of a word list, i.e., working memory as well as long-term memory. In contrast, the current study focused on working memory and employed the classical DS measure. Therefore, this study adds to the evidence for enhanced working memory in musically trained children along with earlier longitudinal studies that have used the similar span tests (Fujioka et al., 2006; Bergman Nutley et al., 2014; Guo et al., 2018).

Interestingly, we found enhancement of performance in only the forwards and not the DS backwards test. Similar results have been obtained in the study by Hansen et al. (2013) who found that musical training was associated with better performance in the DS forwards, but not backwards test. Furthermore, in their study, DS forward performance was connected to performance in musical ability tests. Along the same lines, Lee et al. (2007) found that musically trained adults outperformed nontrained peers in forwards, but not DS backwards. In their study, musically trained children, aged 12 on average, however, outperformed nontrained peer both in the forwards and DS backwards tests. Guo et al. (2018), in turn, found enhancement of the backwards but not the DS forwards after a short-term instrumental training program. Likewise, Bergman Nutley et al. (2014) found only enhancement of DS backwards performance in musically trained adults and children, but unfortunately they did not include DS forwards to allow for comparison. Thus, the literature is mixed as to whether musical training is associated with enhancement of forwards or DS backwards or both.

In any case, the current study found longitudinal evidence in a large sample of subjects in favor of selective enhancement of DS forwards in musically trained children and adolescents. Although negative results cannot be taken as evidence for the null hypothesis that there is no difference between the groups in DS backwards, the substantial statistical power of the current study indicates that a putative undetected group difference in DS backwards would have to be very small and of little practical importance.

In this study, we also found that musically trained participants outperformed nontrained peers in both of the Trail-Making Tests. Previously, adult musicians have been found to

outperform nonmusicians in TMT-A and B (Bugos and Mostafa, 2011), or TMT B alone (Strong and Mast, 2019). However, for instance Bialystok and Depape (2009) and Virtala et al. (2014) found no differences between adult musicians and nonmusicians in span tests or the TMT-A or B. Our results concur with previous findings of enhanced performance in TMT-A and B in musically trained individuals but extend these findings to children and adolescents.

Working Memory Subcomponents Measured by the Digit Span Test and the Trail-Making Test

The Digit Span test has usually been categorized as a simple span test, requiring maintenance of information in memory. Complex span tests in turn require memory maintenance of information during another, unrelated cognitive operation (Wilhelm et al., 2013). However, in a meta-analysis conducted by Redick and Lindsey (2013), the correlation between DS backward performance and performance in n-back tasks as well as a verbal complex span tests was greater than the correlation between DS forward and these tests. Because the DS backward test requires subjects to reverse the order of the strings presented in mind, it also requires working memory updating. Furthermore, the DS forward and backward tests have both been found to recruit in part separate brain networks (Manan et al., 2014). Both activate areas connected to working memory, but with the backward test more strongly activating brain areas related to cognitive control and phonological processing (Gerton et al., 2004; Yang et al., 2015).

As the DS forwards and backwards tests have been found to recruit in part separate working memory processes, it is possible that performance in one but not the other could be enhanced through training. Indeed, selective enhancement of working memory updating (Linares et al., 2018, 2019) and maintenance (Carretti et al., 2007) skills has been found as a result of working memory training in adults. Another study achieved selective impairment of working memory maintenance, but not updating with tDCS (Wang et al., 2018). Our findings of musician advantage in DS forwards, but not backwards points

towards selective enhancement of working memory maintenance but not updating.

The Trail-Making Test is usually used to measure executive functions, and neuroimaging and lesion studies have identified that TMT recruits large-scale fronto-parietal brain networks related to these functions (Varjacic et al., 2018). However, there is evidence that performance in the TMT is related primarily to processing speed and working memory ability, as well as fluid intelligence (Sánchez-Cubillo et al., 2009; Satterthwaite et al., 2013). These findings are supported by evidence of genetic correlations between trail-making performance, reasoning ability and general cognitive ability, processing speed, and memory (Hagenaars et al., 2018). Research has also found differences between the cognitive processes underlying TMT-A and B performance. TMT-A is thought to rely mainly upon processing speed, and TMT-B to additionally require working memory and switching ability (Arbuthnott and Frank, 2000; Sánchez-Cubillo et al., 2009). According to a validation study of a computerized version of the TMT, TMT-B performance was explained to a large degree by inhibition and visual working memory skills (Fellows et al., 2017). Similar results were obtained in a factor analysis of TMT performance and several other neurocognitive measures in older individuals, where TMT-B performance was connected to measures of working memory and inhibition, and TMT-A to processing speed (Llinàs-Reglà et al., 2017).

There is also significant overlap between the cognitive processes that the TMT recruits. For instance, working memory skills and working memory capacity are tightly related to fluid intelligence (Kane et al., 2005; Kail, 2007; Demetriou et al., 2014; Salthouse, 2014; Heinzl et al., 2016). It has also been found that working memory predicts switching (Blackwell et al., 2009), presumably through supporting the maintenance of switching rules. Inhibitory control, in turn, may have a role in supporting working memory maintenance (Jonides et al., 1998; Zanto and Gazzaley, 2009; Getzmann et al., 2018).

In behavioral studies, DS backwards performance has been found to predict TMT-B performance, suggesting a partial overlap between the cognitive requirements of these tasks (Sánchez-Cubillo et al., 2009). Both DS backwards and TMT-B engage cognitive control more than DS forwards and TMT-A, but there are also obvious differences between test requirements. TMT-B requires switching attention from one rule and sequence of information in memory to another (letters or numbers). It also requires continuous updating of information about the respondent's position along the series of letters or numbers they are connecting. DS backwards requires recoding a string of digits in mind into reverse order, or updating the representation of the acquired information, but does not require switching between rules or response patterns during responding. The TMT-A and B also engage specifically working memory maintenance, by requiring the participant to keep the response rule and progression along the sequence of letters and numbers in mind. As in this study, we observed a musician advantage in DS forwards but not backwards, and both the TMT-A and B, our results point towards enhancement of skills that are required by these tests but not by the DS backwards test. These include working memory maintenance for DS forwards,

as well as TMT-A and B. TMT-B also requires switching ability, not required by the DS backwards test. In addition to working memory maintenance, the musician advantage in TMT-B can therefore also be explained by enhancement of switching ability.

In sum, while the Digit Span and the Trail-Making Test are routinely used to assess and connected to working memory ability, the task impurity problem complicates reaching conclusions about specifically which cognitive functions are measured and to what extent. Our results are best explained by enhancement of working memory maintenance, required by the TMT-A, B and the DS forward test. In addition, enhancement of switching ability may explain the musician advantage in TMT-B.

How Musical Training Could Exert Selective Effects on the Development of Working Memory

Learning to play a musical instrument or sing requires working memory in a multitude of ways. For example, memorizing and producing sequences of tones when learning music by heart, and responding to changes in music when playing together with others both require working memory. It is possible that musical training during childhood could enhance working memory to the extent that this could be seen as faster development of these skills.

Augmentation of memory skills has been obtained by working memory programs (Melby-Lervåg and Hulme, 2013; Sala and Gobet, 2017b). It has been suggested that programs focusing on core working memory skills are most effective (Morrison and Chein, 2011). These programs are characterized by tasks that contain stimuli in more than one modality, require working memory maintenance and interference control, quick memory encoding and retrieval, change according to the individual's skill level and require high engagement and focus (Morrison and Chein, 2011). Musical training matches these characteristics of core working memory training programs well. For instance, learning to play sheet music requires transformation of visual stimuli into motor actions, which produce sound stimuli. Playing from notes requires concurrent working memory maintenance and updating of visual information from notation and auditory information produced by the musician. Playing from memory adds to working memory updating and maintenance demands through requiring monitoring of the sounds and movements produced and matching them to the model of the musical piece in memory. In ensemble playing, interference control is needed in order to be able to segregate the stream of sound produced by the individuals from those produced by others. Ensemble playing also requires rapid working memory encoding and retrieval, as musicians need to follow not only their own stream of sound but also that of others, and respond to changes in others' output. In joint improvisation, these rapid working memory encoding and retrieval requirements are accentuated. Musical training increases in challenge according to the proficiency of the individual, and successful learning and playing of music requires great engagement and focus. It is

therefore feasible that musical training might influence working memory processes.

The results on selective enhancement of the participants' working memory maintenance, but not working memory updating skills, would mean that musical training selectively engages these mechanisms and perhaps selectively supports development of one more than the other. This explanation resonates with findings of different patterns of brain activation during memory encoding and rehearsal, reflecting differences in memory processes in musicians compared to nonmusicians (Schulze et al., 2011). It is feasible that musical training might exert powerful effects specifically on working memory maintenance. Learning to play by ear relies heavily on an individual's capability of acquiring and storing auditory information, melodies, and then reproducing this information immediately. Learning to play from notes, in turn, hones working memory maintenance in the visual domain. Conversely, classical musical training may not as much emphasize the ability to augment the presented information in mind, but rather reproduce it exactly as presented.

An alternative explanation for selective enhancement of working memory maintenance is that musical training improves selective attention. Indeed, there is evidence that selective attention underlies working memory maintenance (Sreenivasan and Jha, 2007; Berry et al., 2009; Gazzaley and Nobre, 2012). Selective attention seems to support encoding and maintenance of information in memory by shielding it from distracting information. This notion is supported by neuroimaging evidence of attenuated processing of distracting information during a working memory maintenance task (Sreenivasan and Jha, 2007). There is also tentative evidence of a musician advantage selective attention, indexed by decreased variability of frontal brain responses to attended stimuli (Strait and Kraus, 2011; Strait et al., 2015). It is possible that music training, for instance through playing in ensembles, and learning to focus on only the sound produced by one's instrument, could develop selective attention, which is of benefit in tasks requiring memorization of aurally presented information. Selective attention may also be required and therefore trained in learning to play sheet music. For instance, in learning to play piano, there are two notations to follow—one for the right and one for the left hand. Selectively attending to this visual information is required for successful production of sound. Trail-Making Test performance has been connected to selective attention skills, as measured by the ability to recognize speech in noise (Ellis et al., 2016), but to the knowledge of the authors, similar results on a connection between specifically selective attention and Digit Span performance have not been obtained. In future studies investigating working memory in musically trained individuals, including measures for selective attention would help further elucidate this possible connection.

Augmented Developmental Trajectories in Trained and Nontrained Participants

The difference in performance between the Music and Control group observed in this study diminished over time. It is possible

that musical training enhances the development of working memory maintenance or selective attention, which can be seen as faster maturation in the Music group, but with time the Control group children attain the same level of performance. This explanation is contrasted by studies that have found enhanced working memory still in musically trained adults (Chan et al., 1998; Bialystok and Depape, 2009; George and Coch, 2011; Zuk et al., 2014; Talamini et al., 2016; Ding et al., 2018). There are, however, also contrary findings. In one study, adult nonmusicians were found to outperform musicians in tests requiring immediate as well as delayed recall of newly acquired information, with no significant group differences in performance in the TMT-A or B or DS (Virtala et al., 2014). Thus, the existence of working memory benefits associated with musical expertise in adulthood should be considered with caution.

As stated before, the task impurity problem complicates understanding of which cognitive functions are putatively most affected by musical training. Furthermore, the maturation of other executive functions may influence the maturation of subprocesses of working memory. For instance, the protracted development of inhibitory control and shifting ability influence performance in complex working memory span tasks that require these skills in addition to maintaining information in working memory (Jonides et al., 1998; Schleepen and Jonkman, 2009). Future longitudinal studies investigating working memory development should include measures that allow for disentangling the unique contributions of development in these cognitive skills to the development of working memory.

Since our study lacks baseline measurement of working memory skills prior to musical training, our results may also be explained by pre-existing differences between the two groups, instead of developmental causal explanation (for a study pointing towards pre-existing differences in intelligence, which may explain better performance in executive functions, see Schellenberg, 2015). The lack of the baseline measurement is caused by our choice to minimize the length of the experimental session when the children were only 7 years old and about to start their instrumental training. We added more behavioral and ERP paradigms gradually when the children became older and could then better cope with longer sessions. By this arrangement, we were able to minimize the number of drop-out participants—a serious problem in all longitudinal studies (for discussion, see Tervaniemi et al., 2018; Barbaroux et al., 2019).

One might consider the lack of random group allocation also as a caveat of our study. However, in our view, it is not feasible to plan a longitudinal study for several years on children and adolescents, at least if a control group is included. If the participants are not motivated, they either quit the training, do not participate in the investigations, or both. Even in shorter longitudinal studies, it has been a challenge to maintain the motivation of the participants unless the study is conducted in special circumstances such as summer camp like the study environment in the innovative study by Moreno et al. (2011). Thus, the current choice of

having a longitudinal study on children who chose their music training based on their own and their family's initiative, gives solid evidence about the development of cognitive functions of music-oriented and control children obtained in an ecologically valid context.

SUMMARY AND CONCLUSIONS

In this study, we investigated the maturation of working memory in musically trained and nontrained children and adolescents. We found different patterns of development for different subcomponents of working memory in the trained and nontrained participants. Musically trained individuals had better performance in tests tapping working memory maintenance, but not updating, than musically nontrained individuals. However, the difference lessened over time, as nontrained participants attained a similar level of performance as trained participants. Our results extend previous findings of a musician advantage in tests for working memory by specifying which subcomponents of working memory may be most affected, and by clarifying the trajectory of enhancement from childhood into adolescence.

DATA AVAILABILITY STATEMENT

The datasets for this manuscript are not publicly available because the consent form signed by participants did not include

permission for distribution of data outside the research group. Requests to access the datasets should be directed to Katri Saarikivi, katri.sarikivi@helsinki.fi.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The University of Helsinki Ethical Review Board in the Humanities and Social and Behavioural Sciences. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

KS and VP participated in planning the experiment, conducting the measurements, analyzing the data and writing the manuscript. MT and MH were responsible for establishing the longitudinal study, planning the experiments and measurement paradigms included, and reviewing and writing the manuscript.

FUNDING

This research has been funded by The Jenny and Antti Wihuri Foundation, The Finnish Cultural Foundation and The Academy of Finland.

REFERENCES

- Aben, B., Stapert, S., and Blokland, A. (2012). About the distinction between working memory and short-term memory. *Front. Psychol.* 3:301. doi: 10.3389/fpsyg.2012.00301
- Arbuthnott, K., and Frank, J. (2000). Trail making test, part B as a measure of executive control: validation using a set-switching paradigm. *J. Clin. Exp. Neuropsychol.* 22, 518–528. doi: 10.1076/1380-3395(200008)22:4;1-0;FT518
- Baddeley, A. D., and Hitch, G. (1974). "Working memory," in *Psychology of Learning and Motivation* (Vol. 8), ed. G. H. Bower (New York, NY: Academic Press), 47–89.
- Barbaroux, M., Dittinger, E., and Besson, M. (2019). Music training with Démos program positively influences cognitive functions in children from low socio-economic backgrounds. *PLoS One* 14:e0216874. doi: 10.1371/journal.pone.0216874
- Bates, D. (2005). Fitting linear mixed models in R. *R News* 5, 27–30.
- Bates, D., Sarkar, D., Bates, M. D., and Matrix, L. (2007). The lme4 package. *R Package Version 2.74*.
- Bergman Nutley, S., Darki, F., and Klingberg, T. (2014). Music practice is associated with development of working memory during childhood and adolescence. *Front. Hum. Neurosci.* 7:926. doi: 10.3389/fnhum.2013.00926
- Bermudez, P., Lerch, J. P., Evans, A. C., and Zatorre, R. J. (2008). Neuroanatomical correlates of musicianship as revealed by cortical thickness and voxel-based morphometry. *Cereb. Cortex* 19, 1583–1596. doi: 10.1093/cercor/bhn196
- Berry, A. S., Zanto, T. P., Rutman, A. M., Clapp, W. C., and Gazzaley, A. (2009). Practice-related improvement in working memory is modulated by changes in processing external interference. *J. Neurophysiol.* 102, 1779–1789. doi: 10.1152/jn.00179.2009
- Best, J. R., and Miller, P. H. (2010). A developmental perspective on executive function. *Child Dev.* 81, 1641–1660. doi: 10.1111/j.1467-8624.2010.01499.x
- Bialystok, E., and Depape, A. M. (2009). Musical expertise, bilingualism, and executive functioning. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 565–574. doi: 10.1037/a0012735
- Blackwell, K. A., Cepeda, N. J., and Munakata, Y. (2009). When simple things are meaningful: working memory strength predicts children's cognitive flexibility. *J. Exp. Child Psychol.* 103, 241–249. doi: 10.1016/j.jecp.2009.01.002
- Bugos, J., and Mostafa, W. (2011). Musical training enhances information processing speed. *Bull. Council Res. Music Educ.* 187, 7–18.
- Bunge, S. A., and Wright, S. B. (2007). Neurodevelopmental changes in working memory and cognitive control. *Curr. Opin. Neurobiol.* 17, 243–250. doi: 10.1016/j.conb.2007.02.005
- Carretti, B., Borella, E., and De Beni, R. (2007). Does strategic memory training improve the working memory performance of younger and older adults? *Exp. Psychol.* 54, 311–320. doi: 10.1027/1618-3169.54.4.311
- Casey, B. J., Giedd, J. N., and Thomas, K. M. (2000). Structural and functional brain development and its relation to cognitive development. *Biol. Psychol.* 54, 241–257. doi: 10.1016/s0301-0511(00)00058-2
- Cepeda, N. J., Kramer, A. F., and Gonzalez de Sather, J. C. M. (2001). Changes in executive control across the life span: examination of task-switching performance. *Dev. Psychol.* 37, 715–730. doi: 10.1037/0012-1649.37.5.715
- Chan, A. S., Ho, Y. C., and Cheung, M. C. (1998). Music training improves verbal memory. *Nature* 396:128. doi: 10.1038/24075
- Cheung, M. C., Chan, A. S., Liu, Y., Law, D., and Wong, C. W. (2017). Music training is associated with cortical synchronization reflected in EEG coherence during verbal memory encoding. *PLoS One* 12:e0174906. doi: 10.1371/journal.pone.0174906
- Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychol. Bull.* 104, 163–191. doi: 10.1037/0033-2909.104.2.163
- Cowan, N. (1999). "An embedded-processes model of working memory," in *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (Vol. 20), eds A. Miyake and P. Shah (Cambridge, MA: Cambridge University Press), 506.
- Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Prog. Brain Res.* 169, 323–338. doi: 10.1016/s0079-6123(07)00020-9

- De Luca, C. R., Wood, S. J., Anderson, V., Buchanan, J.-A., Proffitt, T. M., Mahony, K., et al. (2003). Normative data from the CANTAB. I: development of executive function over the lifespan. *J. Clin. Exp. Neuropsychol.* 25, 242–254. doi: 10.1076/jcen.25.2.242.13639
- Degé, F., Kubicek, C., and Schwarzer, G. (2011). Music lessons and intelligence: a relation mediated by executive functions. *Music Percept.* 29, 195–201. doi: 10.1525/mp.2011.29.2.195
- Demetriou, A., Spanoudis, G., Shayer, M., Van der Ven, S., Brydges, C. R., Kroesbergen, E., et al. (2014). Relations between speed, working memory, and intelligence from preschool to adulthood: structural equation modeling of 14 studies. *Intelligence* 46, 107–121. doi: 10.1016/j.intell.2014.05.013
- D'Esposito, M., Detre, J. A., Alsop, D. C., Shin, R. K., Atlas, S., and Grossman, M. (1995). The neural basis of the central executive system of working memory. *Nature* 378, 279–281. doi: 10.1038/378279a0
- Diamond, A. (2013). Executive functions. *Annu. Rev. Psychol.* 64, 135–168. doi: 10.1146/annurev-psych-113011-143750
- Ding, Y., Gray, K., Forrence, A., Wang, X., and Huang, J. (2018). A behavioral study on tonal working memory in musicians and non-musicians. *PLoS One* 13:e0201765. doi: 10.1371/journal.pone.0201765
- Ellis, R. J., Molander, P., Rönnerberg, J., Lyxell, B., Andersson, G., and Lunner, T. (2016). Predicting speech-in-noise recognition from performance on the trail making test: results from a large-scale internet study. *Ear Hear.* 37, 73–79. doi: 10.1097/aud.0000000000000218
- Eriksson, J., Vogel, E. K., Lansner, A., Bergström, F., and Nyberg, L. (2015). Neurocognitive architecture of working memory. *Neuron* 88, 33–46. doi: 10.1016/j.neuron.2015.09.020
- Fellows, R. P., Dahmen, J., Cook, D., and Schmitter-Edgecombe, M. (2017). Multicomponent analysis of a digital Trail Making Test. *Clin. Neuropsychol.* 31, 154–167. doi: 10.1080/13854046.2016.1238510
- Fujioka, T., Ross, B., Kakigi, R., Pantev, C., and Trainor, L. J. (2006). One year of musical training affects development of auditory cortical-evoked fields in young children. *Brain* 129, 2593–2608. doi: 10.1093/brain/awl247
- Fuster, J. M. (2002). Frontal lobe and cognitive development. *J. Neurocytol.* 31, 373–385. doi: 10.1023/a:1024190429920
- Gaser, C., and Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *J. Neurosci.* 23, 9240–9245. doi: 10.1523/JNEUROSCI.23-09240.2003
- Gathercole, S. E., Pickering, S. J., Ambridge, B., and Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Dev. Psychol.* 40, 177–190. doi: 10.1037/0012-1649.40.2.177
- Gazzaley, A., and Nobre, A. C. (2012). Top-down modulation: bridging selective attention and working memory. *Trends Cogn. Sci.* 16, 129–135. doi: 10.1016/j.tics.2011.11.014
- George, E. M., and Coch, D. (2011). Music training and working memory: an ERP study. *Neuropsychologia* 49, 1083–1094. doi: 10.1016/j.neuropsychologia.2011.02.001
- Gerton, B. K., Brown, T. T., Meyer-Lindenberg, A., Kohn, P., Holt, J. L., Olsen, R. K., et al. (2004). Shared and distinct neurophysiological components of the digits forward and backward tasks as revealed by functional neuroimaging. *Neuropsychologia* 42, 1781–1787. doi: 10.1016/j.neuropsychologia.2004.04.023
- Getzmann, S., Wascher, E., and Schneider, D. (2018). The role of inhibition for working memory processes: ERP evidence from a short-term storage task. *Psychophysiology* 55:e13026. doi: 10.1111/psyp.13026
- Gray, S., Green, S., Alt, M., Hogan, T., Kuo, T., Brinkley, S., et al. (2017). The structure of working memory in young children and its relation to intelligence. *J. Mem. Lang.* 92, 183–201. doi: 10.1016/j.jml.2016.06.004
- Guo, X., Ohsawa, C., Suzuki, A., and Sekiyama, K. (2018). Improved digit span in children after a 6-week intervention of playing a musical instrument: an exploratory randomized controlled trial. *Front. Psychol.* 8:2303. doi: 10.3389/fpsyg.2017.02303
- Hagenaars, S. P., Cox, S. R., Hill, W. D., Davies, G., Liewald, D. C., Harris, S. E., et al. (2018). Genetic contributions to trail making test performance in UK Biobank. *Mol. Psychiatry* 23, 1575–1583. doi: 10.1038/mp.2017.189
- Hansen, M., Wallentin, M., and Vuust, P. (2013). Working memory and musical competence of musicians and non-musicians. *Psychol. Music* 41, 779–793. doi: 10.1177/0305735612452186
- Heinzel, S., Lorenz, R. C., Pelz, P., Heinz, A., Walter, H., Kathmann, N., et al. (2016). Neural correlates of training and transfer effects in working memory in older adults. *Neuroimage* 134, 236–249. doi: 10.1016/j.neuroimage.2016.03.068
- Ho, Y. C., Cheung, M. C., and Chan, A. S. (2003). Music training improves verbal but not visual memory: cross-sectional and longitudinal explorations in children. *Neuropsychology* 17, 439–450. doi: 10.1037/0894-4105.17.3.439
- Huizinga, M., Dolan, C. V., and van der Molen, M. W. (2006). Age-related change in executive function: developmental trends and a latent variable analysis. *Neuropsychologia* 44, 2017–2036. doi: 10.1016/j.neuropsychologia.2006.01.010
- Huizinga, M., and van der Molen, M. W. (2011). Task switching and shifting between stopping and going: developmental change in between-trial control adjustments. *J. Exp. Child Psychol.* 108, 484–503. doi: 10.1016/j.jecp.2010.10.004
- James, C. E., Oechslin, M. S., Van De Ville, D., Hauert, C. A., Descloux, C., and Lazeyras, F. (2014). Musical training intensity yields opposite effects on grey matter density in cognitive versus sensorimotor networks. *Brain Struct. Funct.* 219, 353–366. doi: 10.1007/s00429-013-0504-z
- Jonides, J., Smith, E. E., Marshuetz, C., Koeppe, R. A., and Reuter-Lorenz, P. A. (1998). Inhibition in verbal working memory revealed by brain activation. *Proc. Natl. Acad. Sci. U S A* 95, 8410–8413. doi: 10.1073/pnas.95.14.8410
- Jurado, M. B., and Rosselli, M. (2007). The elusive nature of executive functions: a review of our current understanding. *Neuropsychol. Rev.* 17, 213–233. doi: 10.1007/s11065-007-9040-z
- Kail, R. V. (2007). Longitudinal evidence that increases in processing speed and working memory enhance children's reasoning. *Psychol. Sci.* 18, 312–313. doi: 10.1111/j.1467-9280.2007.01895.x
- Kane, M. J., Hambrick, D. Z., and Conway, A. R. (2005). Working memory capacity and fluid intelligence are strongly related constructs: comment on Ackerman, Beier, and Boyle (2005). *Psychol. Bull.* 131, 66–71. doi: 10.1037/0033-2909.131.1.66
- Kharitonova, M., Martin, R. E., Gabrieli, J. D. E., and Sheridan, M. A. (2013). Cortical gray-matter thinning is associated with age-related improvements on executive function tasks. *Dev. Cogn. Neurosci.* 6, 61–71. doi: 10.1016/j.dcn.2013.07.002
- Kwon, H., Reiss, A. L., and Menon, V. (2002). Neural basis of protracted developmental changes in visuo-spatial working memory. *Proc. Natl. Acad. Sci. U S A* 99, 13336–13341. doi: 10.1073/pnas.162486399
- Lee, Y. S., Lu, M. J., and Ko, H. P. (2007). Effects of skill training on working memory capacity. *Learn. Instr.* 17, 336–344. doi: 10.1016/j.learninstruc.2007.02.010
- Lehto, J. E., Juujärvi, P., Kooistra, L., and Pulkkinen, L. (2003). Dimensions of executive functioning: evidence from children. *Br. J. Dev. Psychol.* 21, 59–80. doi: 10.1348/026151003321164627
- Linares, R., Borella, E., Lechuga, M. T., Carretti, B., and Pelegrina, S. (2018). Training working memory updating in young adults. *Psychol. Res.* 82, 535–548. doi: 10.1007/s00426-017-0843-0
- Linares, R., Borella, E., Lechuga, M. T., Carretti, B., and Pelegrina, S. (2019). Nearest transfer effects of working memory training: A comparison of two programs focused on working memory updating. *PLoS One* 14:e0211321. doi: 10.1371/journal.pone.0211321
- Linden, D. E. (2007). The working memory networks of the human brain. *Neuroscientist* 13, 257–267. doi: 10.1177/1073858406298480
- Llinàs-Reglà, J., Vilalta-Franch, J., López-Pousa, S., Calvó-Pexas, L., Torrents Rodas, D., and Garre-Olmo, J. (2017). The trail making test: association with other neuropsychological measures and normative values for adults aged 55 years and older from a Spanish-Speaking Population-based sample. *Assessment* 24, 183–196. doi: 10.1177/1073191115602552
- Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A., and Sweeney, J. A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Dev.* 75, 1357–1372. doi: 10.1111/j.1467-8624.2004.00745.x
- Manan, H. A., Franz, E. A., Yusoff, A. N., and Mukari, S. Z. M. S. (2014). Age-related brain activation during forward and backward verbal memory tasks. *Neurol. Psychiatry Brain Res.* 20, 76–86. doi: 10.1016/j.npbr.2014.08.001
- Melby-Lervåg, M., and Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Dev. Psychol.* 49, 270–291. doi: 10.1037/a0028228

- Miller, E. K., and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202. doi: 10.1146/annurev.neuro.24.1.167
- Miyake, A., and Friedman, N. P. (2012). The nature and organization of individual differences in executive functions four general conclusions. *Curr. Dir. Psychol. Sci.* 21, 8–14. doi: 10.1177/0963721411429458
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734
- Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., and Chau, T. (2011). Short-term music training enhances verbal intelligence and executive function. *Psychol. Sci.* 22, 1425–1433. doi: 10.1177/0956797611416999
- Morrison, A. B., and Chein, J. M. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychon. Bull. Rev.* 18, 46–60. doi: 10.3758/s13423-010-0034-0
- Nee, D. E., Brown, J. W., Askren, M. K., Berman, M. G., Demiralp, E., Krawitz, A., et al. (2012). A meta-analysis of executive components of working memory. *Cereb. Cortex* 23, 264–282. doi: 10.1093/cercor/bhs007
- Owen, A. M., McMillan, K. M., Laird, A. R., and Bullmore, E. (2005). N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. *Hum. Brain Mapp.* 25, 46–59. doi: 10.1002/hbm.20131
- Pallesen, K. J., Brattico, E., Bailey, C. J., Korvenoja, A., Koivisto, J., Gjedde, A., et al. (2010). Cognitive control in auditory working memory is enhanced in musicians. *PLoS One* 5:e11120. doi: 10.1371/journal.pone.0011120
- Postle, B. R., Berger, J. S., and D’Esposito, M. (1999). Functional neuroanatomical double dissociation of mnemonic and executive control processes contributing to working memory performance. *Proc. Natl. Acad. Sci. U S A* 96, 12959–12964. doi: 10.1073/pnas.96.22.12959
- Poutiainen, E., Kalska, H., Laasonen, M., Närhi, V., and Räsänen, P. (2010). *Trail Making Testi*. Helsinki: Psykologien Kustannus Oy.
- Putkinen, V., Tervaniemi, M., Saarikivi, K., Ojala, P., and Huotilainen, M. (2014). Enhanced development of auditory change detection in musically trained school-aged children: a longitudinal event-related potential study. *Dev. Sci.* 17, 282–297. doi: 10.1111/desc.12109
- Redick, T. S., and Lindsey, D. R. (2013). Complex span and n-back measures of working memory: a meta-analysis. *Psychon. Bull. Rev.* 20, 1102–1113. doi: 10.3758/s13423-013-0453-9
- Roden, L., Grube, D., Bongard, S., and Kreutz, G. (2014). Does music training enhance working memory performance? Findings from a quasi-experimental longitudinal study. *Psychol. Music* 42, 284–298. doi: 10.1177/0305735612471239
- Rottschy, C., Langner, R., Dogan, I., Reetz, K., Laird, A. R., Schulz, J. B., et al. (2012). Modelling neural correlates of working memory: a coordinate-based meta-analysis. *Neuroimage* 60, 830–846. doi: 10.1016/j.neuroimage.2011.11.050
- Saarikivi, K., Putkinen, V., Tervaniemi, M., and Huotilainen, M. (2016). Cognitive flexibility modulates maturation and music-training-related changes in neural sound discrimination. *Eur. J. Neurosci.* 44, 1815–1825. doi: 10.1111/ejn.13176
- Sala, G., and Gobet, F. (2017a). When the music’s over. Does music skill transfer to children’s and young adolescents’ cognitive and academic skills? A meta-analysis. *Educ. Res. Rev.* 20, 55–67. doi: 10.1016/j.edurev.2016.11.005
- Sala, G., and Gobet, F. (2017b). Working memory training in typically developing children: a meta-analysis of the available evidence. *Dev. Psychol.* 53, 671–685. doi: 10.1037/dev0000265
- Salthouse, T. A. (2014). Relations between running memory and fluid intelligence. *Intelligence* 43, 1–7. doi: 10.1016/j.intell.2013.12.002
- Sánchez-Cubillo, I., Periañez, J. A., Adrover-Roig, D., Rodríguez-Sánchez, J. M., Ríos-Lago, M., Tirapu, J. E. E. A., et al. (2009). Construct validity of the trail making test: role of task-switching, working memory, inhibition/interference control and visuomotor abilities. *J. Int. Neuropsychol. Soc.* 15, 438–450. doi: 10.1017/S155617709090626
- Satterthwaite, T. D., Wolf, D. H., Erus, G., Ruparel, K., Elliott, M. A., Gennatas, E. D., et al. (2013). Functional maturation of the executive system during adolescence. *J. Neurosci.* 33, 16249–16261. doi: 10.1523/JNEUROSCI.2345-13.2013
- Schellenberg, E. G. (2004). Music lessons enhance IQ. *Psychol. Sci.* 15, 511–514. doi: 10.1111/j.0956-7976.2004.00711.x
- Schellenberg, E. G. (2006). Long-term positive associations between music lessons and IQ. *J. Educ. Psychol.* 98, 457–468. doi: 10.1037/0022-0663.98.2.457
- Schellenberg, E. G. (2011). Examining the association between music lessons and intelligence. *Br. J. Psychol.* 102, 283–302. doi: 10.1111/j.2044-8295.2010.02000.x
- Schellenberg, E. G. (2015). Music training and speech perception: a gene-environment interaction. *Ann. N Y Acad. Sci.* 1337, 170–177. doi: 10.1111/nyas.12627
- Schleeper, T. M., and Jonkman, L. M. (2009). The development of non-spatial working memory capacity during childhood and adolescence and the role of interference control: an N-Back task study. *Dev. Neuropsychol.* 35, 37–56. doi: 10.1080/87565640903325733
- Schulze, K., Müller, K., and Koelsch, S. (2011). Neural correlates of strategy use during auditory working memory in musicians and non-musicians. *Eur. J. Neurosci.* 33, 189–196. doi: 10.1111/j.1460-9568.2010.07470.x
- Siffredi, V., Barrouillet, P., Spencer-Smith, M., Vaessen, M., Anderson, V., and Vuilleumier, P. (2017). Examining distinct working memory processes in children and adolescents using fMRI: results and validation of a modified Brown-Peterson paradigm. *PLoS One* 12:e0179959. doi: 10.1371/journal.pone.0179959
- Sreenivasan, K. K., and Jha, A. P. (2007). Selective attention supports working memory maintenance by modulating perceptual processing of distractors. *J. Cogn. Neurosci.* 19, 32–41. doi: 10.1162/jocn.2007.19.1.32
- Stoodley, C. J., Valera, E. M., and Schmahmann, J. D. (2012). Functional topography of the cerebellum for motor and cognitive tasks: an fMRI study. *Neuroimage* 59, 1560–1570. doi: 10.1016/j.neuroimage.2011.08.065
- Strait, D. L., and Kraus, N. (2011). Can you hear me now? Musical training shapes functional brain networks for selective auditory attention and hearing speech in noise. *Front. Psychol.* 2:113. doi: 10.3389/fpsyg.2011.00113
- Strait, D. L., Slater, J., O’Connell, S., and Kraus, N. (2015). Music training relates to the development of neural mechanisms of selective auditory attention. *Dev. Cogn. Neurosci.* 12, 94–104. doi: 10.1016/j.dcn.2015.01.001
- Strong, J. V., and Mast, B. T. (2019). The cognitive functioning of older adult instrumental musicians and non-musicians. *Neuropsychol. Dev. Cogn. B Aging Neuropsychol. Cogn.* 26, 367–386. doi: 10.1080/13825585.2018.1448356
- Stuss, D. T., and Alexander, M. P. (2000). Executive functions and the frontal lobes: a conceptual view. *Psychol. Res.* 63, 289–298. doi: 10.1007/s004269900007
- Talamini, F., Altoè, G., Carretti, B., and Grassi, M. (2017). Musicians have better memory than nonmusicians: a meta-analysis. *PLoS One* 12:e0186773. doi: 10.1371/journal.pone.0186773
- Talamini, F., Carretti, B., and Grassi, M. (2016). The working memory of musicians and nonmusicians. *Music Percept.* 34, 183–191. doi: 10.1525/mp.2016.34.2.183
- Taylor, S. J., Barker, L. A., Heavey, L., and McHale, S. (2013). The typical developmental trajectory of social and executive functions in late adolescence and early adulthood. *Dev. Psychol.* 49, 1253–1265. doi: 10.1037/a0029871
- Taylor, S. J., Barker, L. A., Heavey, L., and McHale, S. (2015). The longitudinal development of social and executive functions in late adolescence and early adulthood. *Front. Behav. Neurosci.* 9:252. doi: 10.3389/fnbeh.2015.00252
- Tervaniemi, M., Tao, S., and Huotilainen, M. (2018). Promises of music in education? *Front. Educ.* 3:74. doi: 10.3389/feduc.2018.00074
- Unsworth, N., and Engle, R. W. (2007). The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychol. Rev.* 114, 104–132. doi: 10.1037/0033-295x.114.1.104
- Varjadic, A., Mantini, D., Demeyere, N., and Gillebert, C. R. (2018). Neural signatures of Trail Making Test performance: evidence from lesion-mapping and neuroimaging studies. *Neuropsychologia* 115, 78–87. doi: 10.1016/j.neuropsychologia.2018.03.031
- Virtala, P., Huotilainen, M., Partanen, E., and Tervaniemi, M. (2014). Musicianship facilitates the processing of Western music chords—an ERP and behavioral study. *Neuropsychologia* 61, 247–258. doi: 10.1016/j.neuropsychologia.2014.06.028

- Vuontela, V., Steenari, M. R., Carlson, S., Koivisto, J., Fjällberg, M., and Aronen, E. T. (2003). Audiospatial and visuospatial working memory in 6–13 year old school children. *Learn. Mem.* 10, 74–81. doi: 10.1101/lm.53503
- Wang, J., Tian, J., Hao, R., Tian, L., and Liu, Q. (2018). Transcranial direct current stimulation over the right DLPFC selectively modulates subprocesses in working memory. *PeerJ* 6:e4906. doi: 10.7717/peerj.4906
- Wechsler, D. (2010). *Wechsler Intelligence Scale for Children IV*. Finland: Psykologien Kustannus.
- Wilhelm, O., Hildebrandt, A. H., and Oberauer, K. (2013). What is working memory capacity, and how can we measure it? *Front. Psychol.* 4:433. doi: 10.3389/fpsyg.2013.00433
- Yang, Z., Jutagir, D. R., Koyama, M. S., Craddock, R. C., Yan, C. G., Shehzad, Z., et al. (2015). Intrinsic brain indices of verbal working memory capacity in children and adolescents. *Dev. Cogn. Neurosci.* 15, 67–82. doi: 10.1016/j.dcn.2015.07.007
- Zanto, T. P., and Gazzaley, A. (2009). Neural suppression of irrelevant information underlies optimal working memory performance. *J. Neurosci.* 29, 3059–3066. doi: 10.1523/jneurosci.4621-08.2009
- Zelazo, P. D., Craik, F. I. M., and Booth, L. (2004). Executive function across the life span. *Acta Psychol.* 115, 167–183. doi: 10.1016/j.actpsy.2003.12.005
- Zuk, J., Benjamin, C., Kenyon, A., and Gaab, N. (2014). Behavioral and neural correlates of executive functioning in musicians and non-musicians. *PLoS One* 9:e99868. doi: 10.1371/journal.pone.0099868

Conflict of 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.

Copyright © 2019 Saarikivi, Huotilainen, Tervaniemi and Putkinen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



New Perspectives on Music in Rehabilitation of Executive and Attention Functions

Yuko Koshimori* and Michael H. Thaut

Music and Health Research Collaboratory, Faculty of Music, University of Toronto, Toronto, ON, Canada

OPEN ACCESS

Edited by:

Paul J. Colombo,
Tulane University, United States

Reviewed by:

Aline Moussard,
Centre de Recherche de l'Institut
Universitaire de Gériatrie de Montréal
(CRIUGM), Canada

Alfredo Raglio,
Scientific Clinical Institutes Maugeri
IRCCS (ICS Maugeri), Italy

*Correspondence:

Yuko Koshimori
yuko.koshimori@mail.utoronto.ca

Specialty section:

This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Neuroscience

Received: 03 June 2019

Accepted: 05 November 2019

Published: 19 November 2019

Citation:

Koshimori Y and Thaut MH (2019)
New Perspectives on Music
in Rehabilitation of Executive
and Attention Functions.
Front. Neurosci. 13:1245.
doi: 10.3389/fnins.2019.01245

Modern music therapy, starting around the middle of the twentieth century was primarily conceived to promote emotional well-being and to facilitate social group association and integration. Therefore, it was rooted mostly in social science concepts. More recently, music as therapy began to move decidedly toward perspectives of neuroscience. This has been facilitated by the advent of neuroimaging techniques that help uncover the therapeutic mechanisms for non-musical goals in the brain processes underlying music perception, cognition, and production. In this paper, we focus on executive function (EF) and attentional processes (AP) that are central for cognitive rehabilitation efforts. To this end, we summarize existing behavioral as well as neuroimaging and neurophysiological studies in musicians, non-musicians, and clinical populations. Musical improvisation and instrumental playing may have some potential for EF/AP stimulation and neurorehabilitation. However, more neuroimaging studies are needed to investigate the neural mechanisms for the active musical performance. Furthermore, more randomized clinical trials combined with neuroimaging techniques are warranted to demonstrate the specific efficacy and neuroplasticity induced by music-based interventions.

Keywords: executive function, attention processes, music neuroscience, music improvisation, Neurologic Music Therapy

INTRODUCTION

Brain and biomedical research involving music has shown that music is a highly structured auditory language engaging complex perception, cognition, and motor control in the brain (Peretz and Zatorre, 2005; Stewart et al., 2006; Alluri et al., 2011; Lee et al., 2011; Thaut et al., 2014) and has a distinct influence on the brain by stimulating complex cognitive (Leggieri et al., 2019), affective (Koelsch, 2014), and sensorimotor processes (Thaut et al., 2015; Crasta et al., 2018; Schaffert et al., 2019).

As musical sound patterns are dynamic, continuously unfold, and change in time, external auditory input may serve well for cognitive rehabilitation, in particular for the domains of executive function (EF) and attentional processes (AP) by helping to adapt to a changing environment. Music by its nature drives attention exogenously (Klein and Lawrence, 2011; Thaut and Gardiner, 2014). Furthermore, music is a multidimensional stimulus consisting of multiple acoustical elements (e.g., pitch, loudness, tempo, rhythms, timbre, melody, harmony) that exist in temporally patterned simultaneity and sequentiality to drive AP (Thaut and Gardiner, 2014). In addition, active music performance involves moment-to-moment feedback systems, which facilitates monitoring, adjusting, and updating (Gardiner and Thaut, 2014). Further, there is some evidence that musicians

perform better on EF/AP (e.g., Hanna-Pladdy and Mackay, 2011) as well as there are prefrontal structural and functional differences between musicians and non-musicians (e.g., Fauvel et al., 2014; Groussard et al., 2014). Furthermore, research shows that active musical performance extensively modulates the prefrontal activity in musicians (Erkkinen and Berkowitz, 2018).

Therefore, this sensory language may effectively be used as a therapy to induce neuroplasticity in the brain affected by disorders, diseases, and injuries (Särkämö et al., 2014). With this neuroscientific knowledge, music may be able to yield more specific therapeutic and stimulation outcomes for targeted functions. In this paper, we first briefly summarize existing studies investigating the effects of formal musical training on the EF/AP performance as well as on brain structure and function. Second, we discuss the neuroimaging/neurophysiological studies demonstrating the prefrontal neural correlates and neuroplasticity of active musical performance. Third, we present the body of literature in music-based interventions in healthy and clinical populations, focusing on the EF/AP findings. Finally, we present the future directions of research in this field to move forward the neuroscientific approach, which facilitates further advancement of music-based interventions for the EF/AP stimulation and rehabilitation.

MUSIC TRAINING ON EF/AP AND BRAIN STRUCTURE AND FUNCTION IN MUSICIANS

Some correlational and cross sectional studies have supported the beneficial effects of music training on EF/AP in older musicians (Hanna-Pladdy and Mackay, 2011; Hanna-pladdy and Gajewski, 2012; Liu et al., 2012; Amer et al., 2013; Moussard et al., 2016; Strong and Mast, 2019), younger musicians (Moradzadeh and Blumenthal, 2015; Okada and Slevc, 2018; Medina and Barraza, 2019), and musically trained children (Zuk et al., 2014). In addition, several studies demonstrated structural and functional changes associated with formal musical training in the prefrontal regions. For example, greater gray matter volume was associated with increasing musical practice in the multiple regions including supplementary motor area (SMA), superior/middle and medial frontal cortex, and insula (Groussard et al., 2014; James et al., 2014). Furthermore, resting-state functional connectivity revealed significantly greater connectivity density in the medial and lateral prefrontal regions as well as temporoparietal junction in musicians (Fauvel et al., 2014; Luo et al., 2014; Klein et al., 2016; Tanaka and Kirino, 2016).

A few studies investigated the brain activity during an EF task in musicians. One EEG study using a visual go/no-go task demonstrated that a group of 17 older musicians exhibited a significantly larger difference in the N2 amplitude (reflecting a conflict detect signal or inhibition of a prepotent response) in the central midline sites between go and no-go conditions compared to a group of 17 older non-musicians (Moussard et al., 2016). The music group also performed significantly better on the task, demonstrated with fewer no-go errors. Further, the music group showed a significant correlation between the N2

amplitude and the no-go task performance. However, there were no significant correlations between the N2 amplitude and any measures of musical background (age of first instruction, years of musical instruction, and hours of current practice). On the other hand, the measures of musical background, but not the task performance were significantly associated with the P3 amplitude that showed no significant group difference in amplitude but differential topography between groups (anterior shift in musicians). Two other studies demonstrated differential brain activities during EF tasks in younger musicians (Moreno et al., 2014) and musically trained children (Zuk et al., 2014) compared to non-musician groups without showing task performance differences between groups.

In summary, the cross-sectional or correlational studies in musicians have shed light on the potential benefit of formal musical training on EF/AP and brain changes in the prefrontal area. The task-based studies consistently demonstrated that individuals with formal musical training have differential brain activity during EF tasks in relative to those without it. However, these studies do not allow to determine the direct and causal effects of music training on EF/AP and brain changes. Furthermore, in these musician studies, the specific effects of different types of musical training on the EF/AP as well as on brain structure and function in the prefrontal area are still unclear.

NEURAL MECHANISMS FOR ACTIVE MUSICAL PERFORMANCE

Several neuroimaging/neurophysiological studies investigated the brain activity during active musical performance in musician. One study using near-infrared spectroscopy (NIRS) reported that piano playing engaged the frontal activity and playing more complex musical piece activated significantly wider frontal areas in musicians (Hashimoto and Okamoto, 2006). Furthermore, the brain activity during musical improvisation was investigated in amateur and professional musicians with varying experience of improvisation. These studies suggest that this creative musical activity extensively engages the brain activity in the prefrontal regions such as dorsal premotor area, pre-SMA, SMA, medial prefrontal cortex (mPFC), dorsolateral prefrontal cortex (DLPFC), anterior cingulate cortex (ACC), inferior frontal gyrus (IFG), and anterior insula (Brown et al., 2006; Pinho et al., 2014; Beaty, 2015; Erkkinen and Berkowitz, 2018; Loui, 2018). However, the directionality of improvisation-induced brain activity varied across the studies. The enhanced prefrontal brain activity and connectivity was interpreted as a reflection of goal-directed, top-down processing including motor planning, response selection, inhibition of competing stimuli, and conscious monitoring (Bengtsson, 2007; Villarreal et al., 2013). The dissociation of decreased lateral and increased medial prefrontal brain activity was interpreted as a reflection of a disconnection or disintegration of the lateral regions to suppress the top-down conscious control to generate spontaneous ideas, engaging the medial prefrontal regions, which was observed in more experienced improvisers (Limb and Braun, 2008;

Liu et al., 2012). Further, the decreased lateral prefrontal activity and functional connectivity was interpreted as indicating that experts could spare EF load for the highly automatized activity (de Manzano and Ullén, 2012; Pinho et al., 2014; Rosen et al., 2016). The observed activation or deactivation in the lateral prefrontal cortex may also depend on the degree of top-down control required by a given improvisation task (Berkowitz and Ansari, 2008; Pinho et al., 2014). For example, the deactivation in the fronto-parietal regions was greater as the freedom of improvisation task increased (Berkowitz and Ansari, 2008). On the other hand, the enhanced lateral prefrontal activity observed in experienced improvisers may be due to the constraints imposed on the improvisation task (Bengtsson, 2007). In addition, the existing studies demonstrated that musical improvisation induced greater brain activity in the prefrontal regions such as DLPFC compared to simple musical repetitive or patterned tasks (Berkowitz and Ansari, 2008; Villarreal et al., 2013) and reproducing music (de Manzano and Ullén, 2012).

In summary, piano playing and musical improvisation extensively engages the prefrontal activity. This is likely because these musical activities involve EF/AP (Hannon and Trainor, 2007; Beaty, 2015). In addition, musical improvisation induces greater activation of the prefrontal regions and functional connectivity in musicians with less experience of improvisation and compared to simple musical repetitive and patterned tasks. Therefore, these musical activities may have some potential for EF/AP stimulation and neurorehabilitation.

MUSIC-BASED INTERVENTIONS FOR EF AND AP STIMULATION IN NON-MUSICIANS

There are a few studies that have demonstrated benefits of learning to play a musical instrument on EF in older healthy non-musicians. In one randomized control trial (RCT), 16 older adults received individualized piano lessons for 6 months (Bugos et al., 2007). They exhibited improvement on the Trail Making Test-B (TMT-B) over time during the period. This was not observed in a control group consisting of 15 older adults. In another study, a 4-month weekly group piano lessons designed and implemented by a professional music teacher and pianist resulted in improved performance on the Color-Word Stroop test in 13 older adults. This cognitive improvement was not observed in 16 older adults of the control group (Seinfeld et al., 2013). Another study further supported the benefit of group piano lessons on EF such as verbal flexibility and inhibition control in 24 older adults (Bugos, 2010). The EF improvement was significantly greater compared to a music appreciation group consisting of 22 older adults who had learned musical elements while listening to music. However, both music groups significantly improved the EF performance.

In addition, one EEG study with a pseudo-randomized design investigated the effects of music making on inhibition control and interference (Alain et al., 2019). Sixty healthy older non-musicians received either 3-month musical, visual art, or no training. The music-based intervention included music making using body percussion, and voice and non-pitched

musical instruments, as well as learning basic music theory, and melody and harmony concepts by singing simple canons. It was provided by a professional music teacher. Transient differential neural activities were observed in the fronto-central sites in both intervention groups without showing improvement on the task performance.

In summary, learning to play the piano may have beneficial effects on EF in healthy older non-musicians. This may be because it is a complex process, requiring processes for the coordination of multiple sensory modalities, motor control, monitoring, working memory, inhibition, and attentional shifting (Hannon and Trainor, 2007; Seinfeld et al., 2013). However, the sample sizes of these studies are small with additional methodological limitations in some of the studies such as no randomization, no blind assessors, no active control group, and unknown intervention compliance.

MUSIC-BASED INTERVENTIONS FOR EF AND AP NEUROREHABILITATION IN CLINICAL POPULATIONS

Several clinical studies incorporated musical improvisation and instrument playing as their intervention techniques. A few pilot studies investigated the effects of Neurologic Music Therapy (NMT) training for EF/AP. It is guided by a NMT-certified music therapist and is based on group improvisation projects with a special emphasis on attention switching that requires participants to musically respond to specific musical cues (Gardiner and Thaut, 2014; Thaut and Gardiner, 2014). In a pseudo-experimental design (due to institutional constraints of U.S. Veterans Administration), 31 persons with acquired brain dysfunction received one 30-min NMT training session (Thaut et al., 2009). A paired *t*-test analysis showed that the music group significantly improved cognitive flexibility assessed on TMT-B with a large effect size ($d = 1.21$), but not memory functions. A control group ($N = 23$) without any intervention did not show any change on the test. In addition, the music group significantly increased their confidence in the EF skill. Other NMT studies included children/adolescents with neurodevelopmental disorders such as Attention Deficit Hyperactivity Disorder (Abrahams and van Dooren, 2018) and Autism Spectrum Disorder (Pasiali et al., 2014). In both studies, the interventions consisted of a 45-min weekly training session over a period of 6 weeks. Both studies reported some improvement on AP. However, both studies included small sample sizes (2 participants in the NMT group in the former study and 9 participants with varying severity in the latter study). Due to the methodological limitations, these NMT preliminary findings need to replicate in RCTs with large sample sizes.

In addition, one RCT included 35 older adults with Mini Mental State Examination scores ≥ 18 and employed a 12 bi-week group music cognitive training (Biasutti and Mangiacotti, 2018). The training was delivered by a specialist with music and neuropsychology background and consisted of voice and instrumental improvisation, which was created based on the

framework of different music-based interventions including NMT (Thaut, 2005). Repeated ANOVA showed that the music group ($N = 18$) showed significant improvement in mental flexibility assessed on verbal fluency and showed a trend toward significance in selective attention assessed on Attentional Matrices test compared to an active control group (gymnastic activities; $N = 17$).

In another RCT, 28 chronic stroke patients without extensive musical experience received either 30-h of music-supported therapy (MST) or conventional physical training over a 10-week period (Fujioka et al., 2018). The MST protocol included mapping functional movements on playing musical instruments, which was based on the NMT technique (Thaut, 2005), as well as music-making with a therapist. The music group showed significant improvement in cognitive flexibility measured by TMT-B in the mid-intervention whereas the active control group showed the improvement post-intervention.

In one RCT combined with functional near-infrared spectroscopy (fNIRS), 39 persons with mild cognitive impairment were assigned either to an instructor-guided 12-week movement music therapy (MMT) or a comparable active non-musical control therapy (Shimizu et al., 2018). MMT included light exercises synchronized to the background music and playing a percussion instrument accompanied with familiar songs. The MMT group showed enhanced mPFC activity and increased functional connectivity within the prefrontal area compared to the control group. ANOVA did not show a significant group effect on the Frontal Assessment Battery score. This is likely because both groups showed some improvement post-intervention in which the MMT group showed a significant improvement while the active control group showed a trend toward significance, shown using a paired *t*-test. A major limitation of this study was that the number of participants in each group was imbalanced (MMT: $N = 30$ vs. control: $N = 9$).

There are also RCTs that demonstrated the beneficial effects of other music-based interventions on EF/AP. For example, a 12-week singing (four different familiar songs) intervention guided by a professional choir instructor led to a significant improvement on inhibitory control in 31 persons with mild Alzheimer's disease (AD; Pongan et al., 2017). This positive effect was also observed in the active control group (painting) of 28 mild AD participants. Another RCT employed Sound Training for Attention and Memory in Dementia (STAM-Dem; Ceccato et al., 2012). STAM-Dem is a 12-week music-based manualized protocol delivered by a music therapist in which the participants perform specific movements to instructed sound stimuli that requires selective attention. It consists of multiple phases with increasing difficulty. Test score changes on selective attention were significantly greater (improved test score) in the music group of 27 participants compared to the control group of 23 participants who received standard care. In addition, 1-h daily listening to participant-selected favorite music guided by a music therapist for 2 months significantly improved AP recovery in 17 persons with acute post-stroke compared to both active and passive control groups of 19 and 17 participants, respectively (Särkämö et al., 2008). Furthermore, the AP improvement was also associated with a significant increase in

the prefrontal gray matter volume in those with left hemisphere stroke (Särkämö et al., 2014).

In summary, clinical studies showed some potential of active music-based interventions such as musical improvisation and instrument playing-based activities to enhance EF/AP. However, their specific effectiveness is still unclear as active control interventions are often effective as well. In addition, there is a paucity of research literature with various clinical conditions. With few neuroimaging studies, it is still to be determined whether the music-based interventions have engaged executive control areas or whether neuroplasticity had occurred in those areas. Further, some studies show critical limitations in the research design such as a small sample size, no randomization, no active control condition, no follow-up assessment, no analysis of a group \times time interaction effect, and no trained therapist.

FUTURE DIRECTIONS AND PERSPECTIVES FOR EF/AP STIMULATION AND NEUROREHABILITATION

The current paper summarizes the effects of formal musical training, active musical performance, and music-based interventions on EF/AP and associated brain activity in musicians, non-musicians, and clinical populations to present the future directions and perspectives of music-based interventions for the EF/AP stimulation and neurorehabilitation. Active musical performance such as piano playing and musical improvisation engages the prefrontal activity and some intervention studies showed the potential of these musical activities for a basis of the EF/AP stimulation and neurorehabilitation techniques. One study showed that enhanced inhibition control was associated with the extent of involvement of the musical activity in the motor system via the cognitive control system in musicians (Slater et al., 2017), suggesting that instrument playing may be more effective to enhance inhibition control compared to singing, for example. However, neuroimaging studies are needed to uncover the neural mechanisms for piano playing and musical improvisation in non-musicians and to demonstrate the direct link between these musical activities and the EF/AP performance. It is also useful for such studies to contrast images between musical activities and the resting state activity as well as between the two musical activities to show what areas are involved in each musical activity and how similar or different in the brain activity induced by these musical activities. Further, more RCTs with a rigorous design are warranted to replicate the findings. In addition, it is important to address whether these interventions are feasible and effective for different age groups and clinical populations with different clinical characteristics (e.g., severity, acute vs. chronic).

Furthermore, it is crucial to design music-based interventions that can tap into targeted processing components of EF/AP. Some of the clinical studies specifically targeted an enhancement of cognitive flexibility and selective attention, utilizing acoustic elements of music as cues or target stimuli (Thaut et al., 2009;

Ceccato et al., 2012). However, as music is tightly connected to the attention system (Särkämö et al., 2008), it can temporally enhance cognitive performance associated with some of the AP as well as learning and memory via increased arousal (Hommel et al., 1990; Thaut et al., 2005; Thompson et al., 2006; Särkämö et al., 2008). Therefore, future RCTs should include music listening as an active control condition to determine the specific efficacy of active music-based interventions above the music-induced arousal effect on EF/AP with follow-up assessments.

Lastly, neuroimaging techniques should be combined to determine whether music-based interventions engage the brain activity of the executive control regions and networks as well as whether neuroplasticity has occurred in the expected regions and networks following the interventions. These studies are also useful to determine whether the observed patterns of hyper- or hypo-activations are due to compensation, neural efficiency, or attempted compensation (Grady, 2012). Such research studies can help growth in this domain, shedding

an increasing light on the neural mechanisms of how music-based intervention techniques can tap into higher cognition to facilitate maintenance, enhancement, as well as recovery of cognitive functions.

AUTHOR CONTRIBUTIONS

YK conceived the concept of the manuscript. YK wrote the first draft of the manuscript and MT wrote the sections of the draft. Both authors contributed to manuscript revision, and read and approved the submitted version of the manuscript.

FUNDING

This research was supported by the Canada Research Chair Program.

REFERENCES

- Abrahams, T. P., and van Dooren, J. C. (2018). The arts in psychotherapy musical attention control training (MACT) in secure residential youth care: a randomised controlled pilot study. *Arts. Psychother.* 57, 80–87. doi: 10.1016/j.aip.2017.10.008
- Alain, C., Moussard, A., Singer, J., Lee, Y., Bidelman, G. M., and Francisco, G. (2019). Music and visual art training modulate brain activity in older adults. *Front. Neurosci.* 13:1–15. doi: 10.3389/fnins.2019.00182
- Alluri, V., Toivianen, P., Jääskeläinen, I. P., Gleran, E., Sams, M., and Brattico, E. (2011). Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. *NeuroImage* 59, 3677–3689. doi: 10.1016/j.neuroimage.2011.11.019
- Amer, T., Kalender, B., Hasher, L., Trehub, S. E., and Wong, Y. (2013). Do older professional musicians have cognitive advantages? *PLoS One* 8:e71630. doi: 10.1371/journal.pone.0071630
- Beaty, R. E. (2015). The neuroscience of musical improvisation. *Neurosci. Biobehav. Rev.* 51, 108–117. doi: 10.1016/j.neubiorev.2015.01.004
- Bengtsson, S. L. (2007). Cortical regions involved in the generation of musical structures during improvisation in pianists. *J. Cogn. Neurosci.* 19, 830–842. doi: 10.1162/jocn.2007.19.5.830
- Berkowitz, A. L., and Ansari, D. (2008). Generation of novel motor sequences: the neural correlates of musical improvisation. 41, 535–543. doi: 10.1016/j.neuroimage.2008.02.028
- Biasutti, M., and Mangiacotti, A. (2018). Assessing a cognitive music training for older participants: a randomised controlled trial. *Int. J. Geriatric Psychiatry* 33, 271–278. doi: 10.1002/gps.4721
- Brown, S., Martinez, M., and Parsons, L. (2006). Music and language side by side in the brain : a PET study of the generation of melodies and sentences. *Eur. J. Neurosci.* 23, 2791–2803. doi: 10.1111/j.1460-9568.2006.04785.x
- Bugos, J. A. (2010). The benefits of music instruction on processing speed, verbal fluency, and cognitive control in aging. *Music Educ. Res. Int.* 4, 1–9.
- Bugos, J. A., Perlstein, W. M., McCrae, C. S., Brophy, T. S., Bedenbaugh, P. H., Perlstein, W. M., et al. (2007). Individualized piano instruction enhances executive functioning and working memory in older adults. *Aging Ment. Heal.* 11, 464–471. doi: 10.1080/13607860601086504
- Ceccato, E., Vigato, G., Bonetto, C., Bevilacqua, A., Pizzolo, P., Crociani, S., et al. (2012). STAM Protocol in dementia: a and controlled trial. *Am. J. Alzheimers Dis. Other Dement.* 27, 301–310. doi: 10.1177/1533317512452038
- Crasta, J. E., Thaut, M. H., Anderson, C. W., Davies, P. L., and Gavin, W. J. (2018). Auditory priming improves neural synchronization in auditory-motor entrainment. *Neuropsychologia* 117, 102–112. doi: 10.1016/j.neuropsychologia.2018.05.017
- de Manzano, Ö., and Ullén, F. (2012). Goal-independent mechanisms for free response generation: creative and pseudo-random performance share neural substrates. *NeuroImage* 59, 772–780. doi: 10.1016/j.neuroimage.2011.07.016
- Erkkinen, M. G., and Berkowitz, A. L. (2018). “Brain research in music improvisation,” in *The Oxford Handbook of Music and the Brain*, eds D. A. Hodges, and M. H. Thaut (UK: Oxford University Press), 1–38. doi: 10.1093/oxfordhb/9780198804123.013.20
- Fauvel, B., Groussard, M., Chételat, G., Fouquet, M., Landeau, B., Eustache, F., et al. (2014). NeuroImage morphological brain plasticity induced by musical expertise is accompanied by modulation of functional connectivity at rest. *NeuroImage* 90, 179–188. doi: 10.1016/j.neuroimage.2013.12.065
- Fujioka, T., Dawson, D. R., Wright, R., Honjo, K., Chen, J. L., Chen, J. J., et al. (2018). The effects of music-supported therapy on motor, cognitive, and psychosocial functions in chronic stroke. *Ann. N. Y. Acad. Sci.* doi: 10.1111/nyas.13706 [Epub ahead of print].
- Gardiner, J. C., and Thaut, M. (2014). “Musical executive function training (MEET),” in *Handbook of Neurologic Music Therapy*, eds M. H. Thaut, and V. Hoemberg, (Oxford: Oxford University Press), 279–293.
- Grady, C. (2012). The cognitive neuroscience of ageing. *Nat. Rev. Neurosci.* 13, 491–505. doi: 10.1038/nrn3256
- Groussard, M., Viader, F., Landeau, B., and Desgranges, B. (2014). Brain and cognition the effects of musical practice on structural plasticity: the dynamics of grey matter changes. *Brain Cogn.* 90, 174–180. doi: 10.1016/j.bandc.2014.06.013
- Hanna-pladdy, B., and Gajewski, B. (2012). Recent and past musical activity predicts cognitive aging variability: direct comparison with general lifestyle activities. *Front. Hum. Neurosci.* 6:198. doi: 10.3389/fnhum.2012.00198
- Hanna-Pladdy, B., and Mackay, A. (2011). The relation between instrumental musical activity and cognitive aging. *Neuropsychology* 25, 378–386.
- Hannon, E. E., and Trainor, L. J. (2007). Music acquisition: effects of enculturation and formal training on development. *Trends Cogn. Sci.* 11, 466–472. doi: 10.1016/j.tics.2007.08.008
- Hashimoto, K., and Okamoto, T. (2006). Examination by near-infrared spectroscopy for evaluation of piano performance as a frontal. *Eur. Neurol.* 55, 16–21. doi: 10.1159/000091138
- Hommel, M., Peres, B., Pollak, P., Memin, B., Besson, G., Gaio, J. M., et al. (1990). Effects of passive tactile and auditory stimulation left visual neglect. *Arch. Neurol.* 47, 573–576. doi: 10.1001/archneur.1990.00530050097018
- James, C. E., Oechslin, M. S., Hauert, C. A., and Van De Ville, D. (2014). Musical training intensity yields opposite effects on grey matter density in cognitive versus sensorimotor networks. *Brain Struct. Funct.* 210, 353–366. doi: 10.1007/s00429-013-0504-z
- Klein, C., Liem, F., Elmer, S., and Lutz, J. (2016). The “Silent” imprint of musical training. *Hum. Brain Mapp.* 37, 536–546. doi: 10.1002/hbm.23045

- Klein, R., and Lawrence, M. (2011). "On the modes and domains of attention," in *Cognitive Neuroscience of Attention*, ed. M. I. Posner, (New York, NY: Guilford), 11–28.
- Koelsch, S. (2014). Brain correlates of music-evoked emotions. *Nat. Rev. Neurosci.* 15, 170–180. doi: 10.1038/nrn3666
- Lee, Y., Janata, P., Frost, C., Hanke, M., and Granger, R. (2011). Investigation of melodic contour processing in the brain using multivariate pattern-based fMRI. *NeuroImage* 57, 293–300. doi: 10.1016/j.neuroimage.2011.02.006
- Leggieri, M., Thaut, M. H., Fornazzari, L., Schweizer, T. A., Barfett, J., Munoz, D. G., et al. (2019). Music intervention approaches for alzheimer's disease: a review of the literature. *Front. Neurosci.* 13:132. doi: 10.3389/fnins.2019.00132
- Limb, C., and Braun, A. (2008). Neural substrates of spontaneous musical performance: a fMRI study of jazz improvisation. *PLoS One* 3:e1679. doi: 10.1371/journal.pone.0001679
- Liu, S., Chow, H. M., Xu, Y., Erkkinen, M. G., Swett, K. E., Eagle, M. W., et al. (2012). Neural correlates of lyrical improvisation: an fMRI study of freestyle rap. *Sci. Rep.* 2, 834. doi: 10.1038/srep00834
- Loui, P. (2018). Rapid and flexible creativity in musical improvisation: review and a model. *Ann. N. Y. Acad. Sci.* doi: 10.1111/nyas.13628 [Epub ahead of print].
- Luo, C., Tu, S., Peng, Y., Gao, S., Li, J., Dong, L., et al. (2014). Long-term effects of musical training and functional plasticity in salience system. *Neural Plast.* 2014, 13. doi: 10.1155/2014/180138
- Medina, D., and Barraza, P. (2019). Efficiency of attentional networks in musicians and non-musicians. *Heliyon* 5:e01315. doi: 10.1016/j.heliyon.2019.e01315
- Moradzadeh, L., and Blumenthal, G. (2015). Musical training, bilingualism, and executive function: a closer look at task switching and dual-task performance. *Cogn. Sci.* 39, 992–1020. doi: 10.1111/cogs.12183
- Moreno, S., Wodniecka, Z., Tays, W., Alain, C., and Bialystok, E. (2014). Inhibitory control in bilinguals and musicians: event related potential (ERP) evidence for experience-specific effects. *PLoS One* 9:e94169. doi: 10.1371/journal.pone.0094169
- Moussard, A., Bermudez, P., Alain, C., Tays, W., and Moreno, S. (2016). Life-long music practice and executive control in older adults: an event-related potential study. *Brain Res.* 1642, 146–153. doi: 10.1016/j.brainres.2016.03.028
- Okada, B. M., and Slevc, L. R. (2018). Individual differences in musical training and executive functions: a latent variable approach. *Mem. Cogn.* 46, 1076–1092. doi: 10.3758/s13421-018-0822-8
- Pasiali, V., Lagasse, A. B., and Penn, S. L. (2014). The effect of musical attention control training (MACT) on attention skills of adolescents with neurodevelopmental delays: a pilot study. *J. Music Ther.* 51, 333–354. doi: 10.1093/jmt/thu030
- Peretz, I., and Zatorre, R. (2005). Brain organization for music processing. *Annu. Rev. Psychol.* 56, 89–114. doi: 10.1146/annurev.psych.56.091103.070225
- Pinho, A. L., de Manzano, O., Fransson, P., Eriksson, H., and Ulle, F. (2014). Connecting to create: expertise in musical improvisation is associated with increased functional connectivity between premotor and prefrontal areas. *J. Neurosci.* 34, 6156–6163. doi: 10.1523/JNEUROSCI.4769-13.2014
- Pongan, E., Tillmann, B., Leveque, Y., and Etienne, S. (2017). Can musical or painting interventions improve chronic pain, mood, quality of life, and cognition in patients with mild alzheimer's disease? evidence from a randomized controlled trial. *J. Alzheimer's Dis.* 60, 663–677. doi: 10.3233/JAD-170410
- Rosen, D. S., Erickson, B., Kim, Y. E., Mirman, D., Hamilton, R. H., and Kounios, J. (2016). Anodal tDCS to right dorsolateral prefrontal cortex facilitates performance for novice jazz improvisers but hinders experts. *Front. Hum. Neurosci.* 10:579. doi: 10.3389/fnhum.2016.00579
- Särkämö, T., Ripollés, P., Vepsäläinen, H., Autti, T., Silvennoinen, H. M., Salli, E., et al. (2014). Structural changes induced by daily music listening in the recovering brain after middle cerebral artery stroke: a voxel-based morphometry study. *Front. Hum. Neurosci.* 8:245. doi: 10.3389/fnhum.2014.00245
- Särkämö, T., Tervaniemi, M., Laitinen, S., Forsblom, A., Soinila, S., Mikkonen, M., et al. (2008). Music listening enhances cognitive recovery and mood after middle cerebral artery stroke. *Brain* 131(Pt 3), 866–876. doi: 10.1093/brain/awn013
- Schaffert, N., Janzen, T. B., Mattes, K., and Thaut, M. H. (2019). A review on the relationship between sound and movement in sports and rehabilitation. *Front. Psychol.* 10:244. doi: 10.3389/fpsyg.2019.00244
- Seinfeld, S., Figueroa, H., Ortiz-gil, J., and Sanchez-vives, M. V. (2013). Effects of music learning and piano practice on cognitive function, mood and quality of life in older adults. *Front. Psychol.* 4:1–13. doi: 10.3389/fpsyg.2013.00810
- Shimizu, N., Umemura, T., and Matsunaga, M. (2018). Effects of movement music therapy with a percussion instrument on physical and frontal lobe function in older adults with mild cognitive impairment: a randomized controlled trial. *Aging Ment. Heal.* 22, 1614–1626. doi: 10.1080/13607863.2017.1379048
- Slater, J., Azem, A., Nicol, T., Swedenborg, B., and Kraus, N. (2017). Variations on the theme of musical expertise: cognitive and sensory processing in percussionists, vocalists and non-musicians. *Eur. J. Neurosci.* 45, 952–963. doi: 10.1111/ejn.13535
- Stewart, L., Kriegstein, K., Warren, J., and Griffiths, T. D. (2006). Music and the brain: disorders of musical listening. *Brain* 129(Pt 10), 2533–2553. doi: 10.1093/brain/awl171
- Strong, J. V., and Mast, B. T. (2019). The cognitive functioning of older adult instrumental musicians and non-musicians musicians and non-musicians. *Aging Neuropsychol. Cogn.* 26, 367–386. doi: 10.1080/13825585.2018.1448356
- Tanaka, S., and Kirino, E. (2016). Functional connectivity of the precuneus in female university students with long-term musical training. *Front. Hum. Neurosci.* 10:328. doi: 10.3389/fnhum.2016.00328
- Thaut, M., and Gardiner, J. (2014). "Musical attention control training," in *Handbook of Neurologic Music Therapy*, eds M. Thaut, and V. Hoemberg, (Oxford: Oxford University Press), 257–269.
- Thaut, M., Trimarchi, P., and Parsons, L. (2014). Human brain basis of musical rhythm perception: common and distinct neural substrates for meter, tempo, and pattern. *Brain Sci.* 4, 428–452. doi: 10.3390/brainsci4020428
- Thaut, M. H. (2005). *Rhythm, Music, and the Brain: Scientific Foundations and Clinical Applications*. New York, NY: Routledge.
- Thaut, M. H., Gardiner, J. C., Holmberg, D., Horwitz, J., Kent, L., Andrews, G., et al. (2009). Neurologic music therapy improves executive function and emotional adjustment in traumatic brain injury rehabilitation. *Ann. N. Y. Acad. Sci.* 1169, 406–416. doi: 10.1111/j.1749-6632.2009.04585.x
- Thaut, M. H., McIntosh, G. C., and Hoemberg, V. (2015). Neurobiological foundations of neurologic music therapy: rhythmic entrainment and the motor system. *Front. Psychol.* 5:1185. doi: 10.3389/fpsyg.2015.01185
- Thaut, M. H., Peterson, D. A., and McIntosh, G. C. (2005). Temporal entrainment of cognitive functions: musical mnemonics induce brain plasticity and oscillatory synchrony in neural networks underlying memory. *Ann. N. Y. Acad. Sci.* 1060, 243–254. doi: 10.1196/annals.1360.017
- Thompson, R. G., Moulin, C. J. A., Hayre, S., and Jones, R. W. (2006). Music enhances category fluency in healthy older adults and alzheimer's disease patients. *Exp. Aging Res.* 31, 91–99. doi: 10.1080/03610730590882819
- Villarreal, M. F., Cerquetti, D., Caruso, S., Schwarcz, V., and Aranguren, L. (2013). Neural correlates of musical creativity: differences between high and low creative subjects. *PLoS One* 8:1–11. doi: 10.1371/journal.pone.0075427
- Zuk, J., Benjamin, C., Kenyon, A., and Gaab, N. (2014). Behavioral and neural correlates of executive functioning in musicians and non-musicians. *PLoS One* 10:e0137930. doi: 10.1371/journal.pone.0099868

Conflict of 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.

Copyright © 2019 Koshimori and Thaut. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Short-Term Choir Singing Supports Speech-in-Noise Perception and Neural Pitch Strength in Older Adults With Age-Related Hearing Loss

Ella Dubinsky^{1*}, Emily A. Wood¹, Gabriel Nespoli¹ and Frank A. Russo^{1,2}

¹ Department of Psychology, Ryerson University, Toronto, ON, Canada, ² Toronto Rehabilitation Institute, Toronto, ON, Canada

OPEN ACCESS

Edited by:

Claude Alain,
Rotman Research Institute (RRI),
Canada

Reviewed by:

Jeanette Tamplin,
University of Melbourne, Australia
Emily B. J. Coffey,
Concordia University, Canada

*Correspondence:

Ella Dubinsky
ella.dubinsky@ryerson.ca

Specialty section:

This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Neuroscience

Received: 05 July 2019

Accepted: 11 October 2019

Published: 28 November 2019

Citation:

Dubinsky E, Wood EA, Nespoli G
and Russo FA (2019) Short-Term
Choir Singing Supports
Speech-in-Noise Perception
and Neural Pitch Strength in Older
Adults With Age-Related Hearing
Loss. *Front. Neurosci.* 13:1153.
doi: 10.3389/fnins.2019.01153

Prior studies have demonstrated musicianship enhancements of various aspects of auditory and cognitive processing in older adults, but musical training has rarely been examined as an intervention for mitigating age-related declines in these abilities. The current study investigates whether 10 weeks of choir participation can improve aspects of auditory processing in older adults, particularly speech-in-noise (SIN) perception. A choir-singing group and an age- and audiometrically-matched do-nothing control group underwent pre- and post-testing over a 10-week period. Linear mixed effects modeling in a regression analysis showed that choir participants demonstrated improvements in speech-in-noise perception, pitch discrimination ability, and the strength of the neural representation of speech fundamental frequency. Choir participants' gains in SIN perception were mediated by improvements in pitch discrimination, which was in turn predicted by the strength of the neural representation of speech stimuli (FFR), suggesting improvements in pitch processing as a possible mechanism for this SIN perceptual improvement. These findings support the hypothesis that short-term choir participation is an effective intervention for mitigating age-related hearing losses.

Keywords: aging, musical training, speech-in-noise, frequency, hearing

INTRODUCTION

As the population ages, and the expectation of longevity increases, a growing interest in healthcare is the promotion of *healthy aging* – the maintenance of mental, social, and physical wellbeing as one ages, in order to retain independence and lead a high-quality life. Aging is associated with declines in cognitive functioning (e.g., decreased working memory and attentional control; for review, see Fabiani, 2012), and deteriorating sensory-perceptual processes (e.g., Fozard, 1990). Declines in hearing can make it difficult for aging individuals to maintain personal relationships and engage socially, and have been linked to feelings of isolation and depression (Arlinger, 2003; Djernes, 2006). Although assistive technologies (e.g., hearing aids) can target aspects of peripheral hearing loss, persistent perceptual deficits are widely reported (e.g., Killian, 1997). One prevalent example is the loss of the ability to perceive speech in a noisy environment (Salomon, 1986; Chmiel and Jerger, 1996; Gomez and Madey, 2001; Ricketts and Hornsby, 2005; Betlejewski, 2006). Counseling programs may improve communication outcomes associated with age-related auditory declines, but they do not appear to influence speech-in-noise problems (Hickson et al., 2007, 2019). While some auditory rehabilitation programs have been shown to be moderately effective in mitigating speech-in-noise problems (Kricos and Holmest, 1996;

Sweetow and Sabes, 2006; Song et al., 2012), they require a high level of motivation and are not appropriate for all cases (Sabes and Sweetow, 2007; Saunders et al., 2016). As such, there is presently a great demand for complementary interventions that target age-related auditory declines, particularly ones that are engaging and scalable, and that show efficacy with regard to speech-in-noise perception. Developing and evaluating an intervention – and its proposed mechanism(s) for change – involves consideration of biological and experiential contributors to these abilities, beginning with age-related hearing loss and the role it plays in speech-in-noise perception.

Hearing loss can occur at different stages in the auditory system. *Peripheral hearing loss* refers to the reduction in efficient sound transmission through the bones of the middle ear (*conductive hearing loss*), and the deterioration of the outer and inner hair cells (*sensorineural hearing loss*; Arlinger, 2003; Yueh et al., 2003; Wingfield et al., 2005). *Central hearing loss* refers to the degradation of neural mechanisms that relay sound information from the cochlea to the brain, resulting from long-term attenuation of neural input from the cochlea, as well as age-related changes in neuronal responses to sound (Syka, 2002; Frisina and Walton, 2006; Yamasoba et al., 2013). Although peripheral losses can be remediated to some degree through the use of assistive technologies such as hearing aids (or, in extreme cases, cochlear implants), central processing deficits seem to persist in spite of such interventions (Chmiel and Jerger, 1996; Killion, 1997). These central processing deficits – including age-related declines in the synchrony of neural firing (Pichora-Fuller and Schneider, 1992; Frisina and Frisina, 1997; Pichora-Fuller et al., 2007), length of recovery time (Walton et al., 1998), and numbers of neurons in auditory nuclei (Frisina and Walton, 2006) – have been associated with age-related losses in key auditory perceptual abilities, such as sound localization (Abel et al., 2000), pitch discrimination (Raz et al., 1989), duration judgments (Fitzgibbons and Gordon-Salant, 1994; Schneider et al., 1994), mistuned harmonic detection (Alain et al., 2001), and speech-in-noise perception (Pichora-Fuller et al., 1995; Russo and Pichora-Fuller, 2008; Schneider et al., 2010). Of the perceptual deficits, the loss of speech-in-noise perception seems to have the most severe impact on the aging adult's quality of life (e.g., Pichora-Fuller et al., 1995, 2007; Anderson et al., 2011).

Speech-in-noise perception refers to the ability to track a voice in a complex acoustic environment, such as a crowded room with many people talking. Vital in social settings and everyday interactions, the loss of this skill can immensely impact an individual's ability to maintain independence, emotional wellbeing, and quality of life as they age (Salomon, 1986; Gomez and Madey, 2001; Betlejewski, 2006). This age-related decline also appears to persist in spite of peripheral remediation, and can even occur in adults with normal audiometric thresholds (Cruickshanks et al., 1998; Schneider B. A. et al., 2002; Tremblay et al., 2003; Gordon-Salant, 2005; Souza et al., 2007; Vermiglio et al., 2012; Alain et al., 2014); in research studies involving older individuals, pure-tone thresholds tend to be a poor predictor of speech-in-noise perception (Dubno et al., 1984; Hargus and Gordon-Salant, 1995; Kim et al., 2006; Souza et al., 2007).

One way to elucidate the neural underpinnings of speech-in-noise perception is through the use of electroencephalography (EEG recordings) to study cortical and subcortical responses to acoustic stimuli (Tremblay et al., 2003; Musacchia et al., 2008; Anderson et al., 2011). Of particular interest here, the *auditory brainstem* – a collection of nuclei involved in afferent and efferent auditory processing – has been shown to encode spectral and temporal acoustic information with a high degree of precision (Clinard et al., 2010; Skoe and Kraus, 2010).

One component of the auditory brainstem response (ABR; Skoe and Kraus, 2010) that has been implicated in perceptual deficits – in particular, speech-in-noise perception – is the *frequency following response* (Johnson et al., 2005; Skoe and Kraus, 2010). This response consists of phase-locked neural activation, wherein the inter-spike intervals correspond to the fundamental frequency (F0) of the sound input (Hoormann et al., 1992). On the basis of animal work involving ablations, the primary source of the FFR appears to be the *inferior colliculus* (Smith et al., 1975), however recent work also suggests cortical contributions (Lehmann and Schönwiesner, 2014; Coffey et al., 2016, 2017a).

The FFR provides a useful index of the auditory nervous system's representation of periodic sound – such as a vowel in speech – through sustained synchronous neural phase-locking. Spectral and temporal features of the FFR, obtained through signal analysis, are associated with different aspects of neural pitch encoding. A *fast Fourier transform (FFT)* of the signal yields a spectral analysis that can be used to assess the strength of the neural representation of periodic sound input (Skoe and Kraus, 2010). Another feature, the *inter-trial phase coherence (ITPC)*, can be used to assess the extent of consistency in the neural response to periodic sound input – i.e., the extent of phase alignment (synchronization) in oscillatory responses (e.g., Delorme and Makeig, 2004).

In the perception of speech cues, the ability to discern and track changes in pitch over time gets significantly more difficult when the signal-to-noise ratio (SNR) decreases (e.g., Killion et al., 2004). By the time an acoustic signal reaches the auditory cortex of an aging adult, it is likely to have undergone both peripheral and neural distortion (due to age-related declines in sensorineural hearing, and neural noise introduced as the signal is relayed through the ascending auditory pathway, respectively), leading to diminished preservation of key temporal and spectral characteristics (e.g., Yueh et al., 2003; Clinard et al., 2010). This suggests a possible mechanism for age-related declines in speech-in-noise perception (and other auditory perceptual abilities which rely on pitch discrimination), whereby age-related central processing deficits (such as reduced FFR fidelity) result in downstream perceptual impairments which tend to persist in spite of peripheral remediation. In terms of mitigating and preventing these declines, one activity that appears to confer some benefits against certain age-related auditory losses is musical experience (e.g., Alain et al., 2014).

Musicianship is purported to have some benefits outside the musical domain, but the most convincing positive effects have been observed in regards to the auditory system (for review, see Herholz and Zatorre, 2012). Over the course of

training, musicians are taught to attend to fine-grained acoustic features – including pitch, timing, and timbre – that contribute to human perception of sound (Kraus et al., 2009; Kraus and Chandrasekaran, 2010). This trained sensitivity to minute acoustic changes is thought to promote the enhancement of auditory perceptual abilities, including those which decline throughout the aging process (Musacchia et al., 2007; Parbery-Clark et al., 2012; Zendel and Alain, 2012; Alain et al., 2014). In studies comparing auditory perception in musicians and non-musicians, musical experience has been associated with a relative advantage in processing some of the same features that have been linked to age-related declines. These benefits include improved pitch discrimination (Kishon-Rabin et al., 2001; Micheyl et al., 2006; Schellenberg and Moreno, 2009; Bidelman et al., 2011b; Meha-Bettison et al., 2018), gap and duration judgments (Rammsayer and Altenmüller, 2006; Zendel and Alain, 2012; Habibi et al., 2014; Donai and Jennings, 2016), mistuned harmonic detection (Koelsch et al., 1999; Zendel and Alain, 2009), and perception of speech-in-noise (Parbery-Clark et al., 2009; for review see Coffey et al., 2017b).

Musicians also demonstrate structural and functional differences in the neural substrates of auditory, sensory-motor, and visuospatial processing (Musacchia et al., 2007; Kraus and Chandrasekaran, 2010; Schlaug, 2015). Among musicians, musical aptitude is correlated with an increase in gray matter volume in the primary auditory cortex, as well as somatosensory and motor areas, the inferior temporal gyrus, hippocampus, and corpus callosum regions (Schlaug et al., 1995; Schneider P. et al., 2002; Gaser and Schlaug, 2003; Herdener et al., 2010). In addition to structural changes in associated brain regions, musicians also demonstrate functional improvements in neural responses to sound, at cortical and subcortical levels in the auditory processing pathway. Compared with non-musicians, musicians demonstrate enhanced neural responses and activation in the auditory cortex (Koelsch et al., 1999; Schneider P. et al., 2002; Pantev et al., 2003; Shahin et al., 2003; Kuriki, 2006; Besson et al., 2007; Zendel et al., 2015; Habibi et al., 2016) and the auditory brainstem (Musacchia et al., 2007; Wong et al., 2007; Lee et al., 2009; Parbery-Clark et al., 2009; Strait et al., 2009; Bidelman and Krishnan, 2010). Notably, musicians demonstrate improvements in both FFR strength (Musacchia et al., 2007, 2008; Bidelman et al., 2011b; Slater et al., 2017) and consistency (Parbery-Clark et al., 2009; Strait et al., 2009; Bidelman et al., 2011a,b; Skoe and Kraus, 2013; Slater et al., 2017); these benefits appear largely resistant to normal age-related declines (Parbery-Clark et al., 2012; White-Schwoch et al., 2013). Because of the importance of pitch processing across auditory perceptual domains, FFR improvements have been suggested as one of the mechanisms through which musicianship enhances auditory perceptual abilities (Kraus and Chandrasekaran, 2010; Bidelman et al., 2011b; Coffey et al., 2017a).

While the aforementioned studies suggest that musical training can improve speech-in-noise perception, they are cross-sectional studies which should preclude causal inferences (Schellenberg, 2019); further, not all studies have found an effect (Ruggles et al., 2014; Boebinger et al., 2015; Madsen et al., 2017, 2019). One way to resolve these inconsistencies is through the evaluation of musical training outcomes in a controlled

experimental design (i.e., a longitudinal context), which a handful of studies have sought to do. These studies essentially provided musical training to non-musicians (with the extent and nature of musical training varying between studies), and administered pre- and post-training assessments to determine whether changes occurred in outcomes of interest. In terms of neural changes, musical training has been found to enhance both structure (Hyde et al., 2009) and function (Fujioka et al., 2006; Lappe et al., 2008; Lappe et al., 2011; Habibi et al., 2016) of the auditory cortex in young children and adults who received musical training, compared to those who did not (control participants). Consistency of the FFR has also been found to be enhanced in adolescents following musical training (Tierney et al., 2015). Children who took part in instrumental music training showed improvements in speech-in-noise perception following 2 years of training (Slater et al., 2015), and younger adults who participated in singing training demonstrated improvements in speech-in-noise perception after only 8 days (Jain et al., 2015). Older adults with 6 months of piano training demonstrated improved cortical responses and speech-in-noise perception following training, suggesting that neural and perceptual benefits can be conferred to aging adults in an intervention context (Zendel et al., 2019).

In addition to improvements in auditory processing, musical training has been linked to enhancements in different aspects of cognitive functioning in older adults – including improvements in working memory and executive control processes – in both cross-sectional (Parbery-Clark et al., 2011; Slevc et al., 2016; Grassi et al., 2017; Mansens et al., 2017) and longitudinal studies (Bugos et al., 2007; Särkämö et al., 2014; Biasutti and Mangiacotti, 2018; Fu et al., 2018). Musical experience has also been shown to alter neural structure and function in regions associated with cognition (Schulze et al., 2011; West et al., 2017); improvements in shared neural substrates for music and non-music domains have been suggested as one possible mechanism for transfer from musical training to speech-in-noise perception (e.g., Kraus et al., 2012).

Taken together, cross-sectional and longitudinal findings suggest that musical training may be able to alter brain structure and function; moreover, it appears to have the capacity to promote enhancements in the same auditory abilities that decline as we age (Solé Resano et al., 2010; Hanna-Pladdy and MacKay, 2011; Zendel and Alain, 2012; Alain et al., 2014), suggesting its use as an intervention to mitigate declines in older adults. Of the forms of music making available, singing may be particularly suited to this purpose.

Singing emerges spontaneously in the first months of life (Papoušek, 1996), and appears to be a universal form of expression (Mithen et al., 2006). Although considerable variability in accuracy exists, the vast majority of adults appear to be able to carry a tune (Dalla Bella et al., 2007; Pfordresher and Brown, 2007; Dalla Bella and Berkowska, 2009). Group singing has been shown to lead to improvements in cooperation (Good and Russo, 2016), social and emotional wellbeing (Hillman, 2002; Bailey, 2005; Hays and Minichiello, 2005; Clift and Morrison, 2011; Creech et al., 2013), and physical and creative outcomes (Beck et al., 2000; Clift and Hancox, 2001; Cohen et al., 2006). Some of these benefits may be mediated by changes in hormonal levels that occur during choral singing: after choir practice,

choristers demonstrate decreased cortisol (Beck et al., 2000) and enhanced immune system functioning more generally (Kreutz et al., 2004). Singing in a group can also be highly motivating for older adults (Hillman, 2002; Creech et al., 2013), which may promote intervention adherence. This is of particular import when singing is contrasted with existing auditory rehabilitation programs, which tend to be plagued by low compliance rates and high attrition (Sweetow and Sabes, 2010; Tye-Murray et al., 2012).

In addition to the social, cognitive, and emotional benefits, singing appears better positioned to confer *near-transfer* benefits to speech. All forms of vocal production involve the rapid integration of auditory and vocal-motor systems (Hickok, 2001; Zatorre et al., 2007; Pfordresher and Dalla Bella, 2011; Pruitt and Pfordresher, 2015); this integration requires feedback loops along the auditory dorsal stream that allow for real-time monitoring and adjustments (Houde and Jordan, 1998; Brainard and Doupe, 2000; Zheng et al., 2010). In a recent study that compared temporal lobe activations across perception of singing, instrumental music, and speech, it was found that compared with instrumental music, singing and speech both led to greater bilateral activations of the superior temporal sulcus (STS; Whitehead and Armony, 2018), a critical node in the auditory dorsal stream (Hickok et al., 2003).

In terms of perceptual processes, there is greater reliance on the vocal-motor system in more challenging listening environments, such as understanding speech-in-noise (Du et al., 2014); older adults rely on this to an even greater degree (Du et al., 2016). Training the vocal-motor system through singing could theoretically improve the resources upon which older adults draw to perceive degraded speech signals. An emphasis on pitch training, feedback, repetition, and the rewarding nature of improvements have been implicated as key components of successful auditory training paradigms, and in the transfer of musical experience to speech perceptual benefits (Besson et al., 2011; David et al., 2012; Herholz and Zatorre, 2012; Shepard et al., 2013; Patel, 2014; Pruitt and Pfordresher, 2015).

The current study investigated whether short-term choir participation and musical training could improve speech-in-noise perception in older adults, compared to an age- and audiometrically matched control group who were not taking part in musical training. Outcomes of interest included speech-in-noise perception (SIN) and pitch discrimination (FDL), strength and consistency of the neural response to sound (as indexed by features of the frequency following response [FFR] to a repeated speech stimulus), and exploratory cognitive measures of working memory (LSpan) and inhibitory control of attention (Flanker task). We hypothesized that older adults who took part in 10 weeks of group choral practice (2 hours weekly) and individual online musical training (up to 1 hour weekly) would demonstrate improved speech-in-noise perceptual abilities following training, which may be driven in part by enhancements in pitch processing and perception, as indexed by enhanced neural responses to sound (features of the FFR) and improved pitch discrimination thresholds. Exploratory cognitive measures of working memory and attention were assessed in relation to training outcomes, as potential dependent variables. We hypothesized that choir participants would experience greater

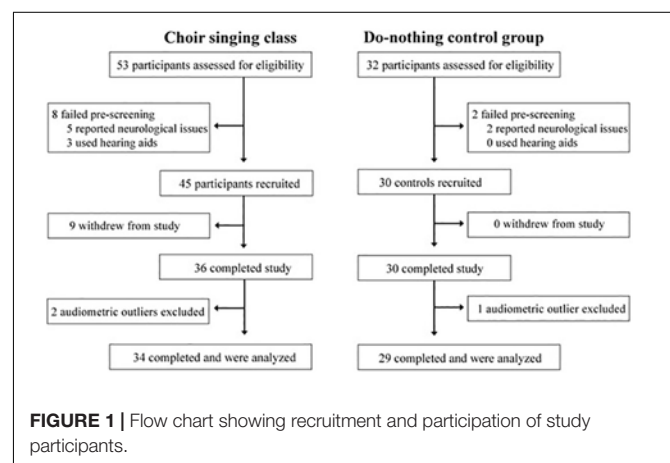
post-training gains than an age- and audiometrically matched do-nothing control group.

MATERIALS AND METHODS

Participants

The process of recruitment and participation is shown in **Figure 1**. Participants were recruited from the first class in a 10-week group singing course run through the 50+ Program at Ryerson University; interested participants came into the lab to undergo eligibility testing that week. Fifty three participants were screened (8 didn't meet eligibility criteria), 45 participants enrolled (9 withdrew from the study), and 36 participants completed choral training and all test sessions. Two participants were rejected as audiometric outliers, so the final analysis included 34 choir singers.

Thirty-four choir participants (31 female), aged 54–79 (mean age = 67.6, standard deviation [SD] = 6.1 years) underwent pre-testing data collection during the first week of the choir and post-testing data collection following the final choir class. Peripheral hearing loss was measured by an in-lab audiometric assessment at standard test frequencies between 0.25 and 8 kHz; average peripheral hearing loss ranged from 9.7 to 45.3 dB HL (mean = 23.1, SD = 9.9 dB HL). Participants were pre-screened to ensure that they did not have any neurological conditions and did not use assistive technology (e.g., hearing aids; see **Figure 1**). Twenty-nine age- and audiometrically matched *do-nothing control* participants (26 female) aged 60–76 (mean age = 67.7, SD = 4.9 years) were recruited through the Ryerson University Hearing Database. Control participants' average peripheral hearing loss ranged from 10.6 to 47.5 dB HL (mean = 24.1, SD = 10.3 dB HL); average pure-tone thresholds for both groups are shown in **Figure 2**. Groups were matched for the duration of time between test sessions, and there were no group differences in previous musical experience, as indexed by years of formal musical training. Informed consent was obtained from each volunteer prior to their participation in the study, in accordance with the Ryerson Research Ethics Board guidelines (REB 2013-128).



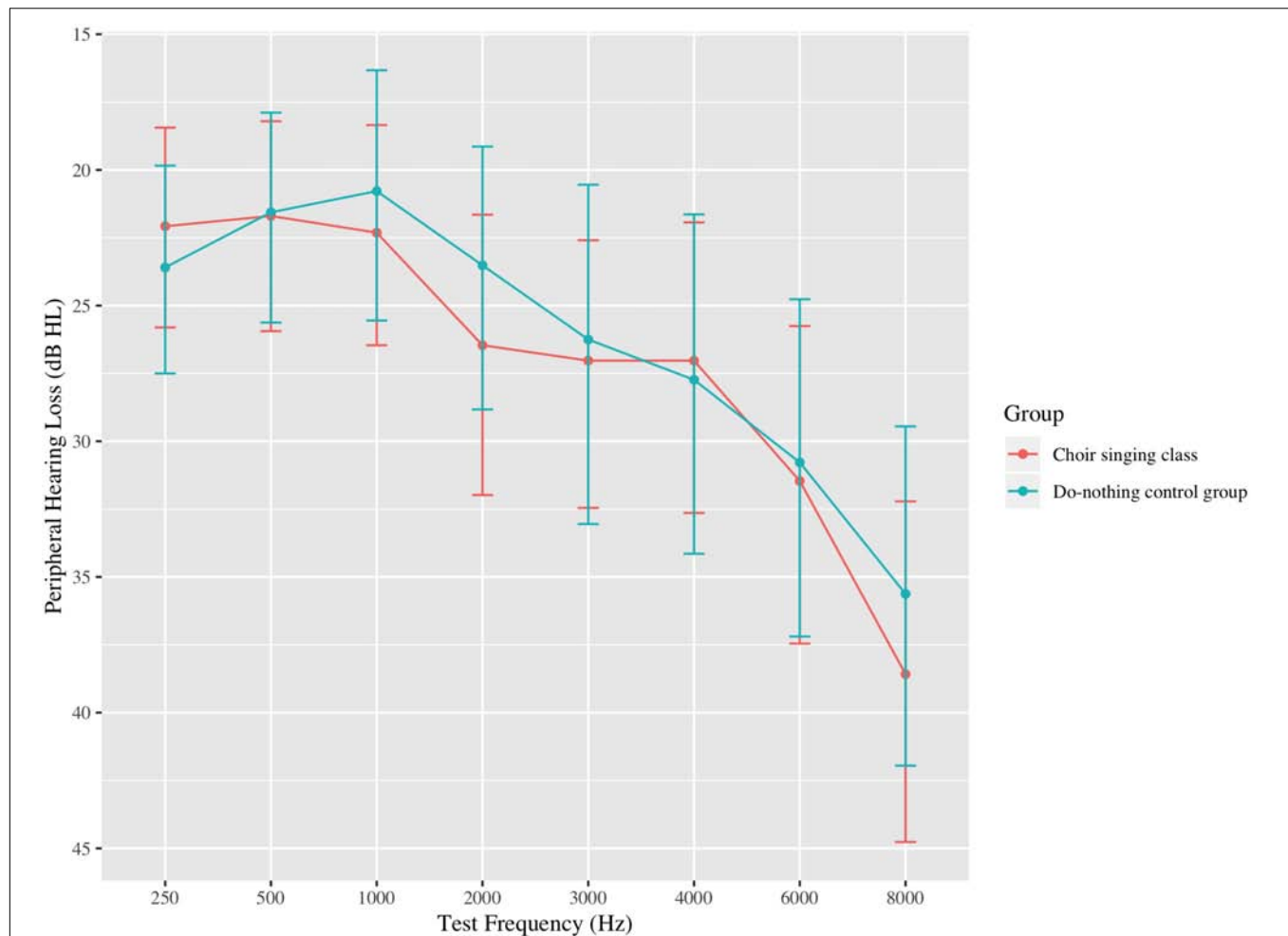


FIGURE 2 | Average bilateral pure-tone thresholds at standard test frequencies up to 8 kHz for participants in the choir singing class and the do-nothing control group. Error bars represent 95% confidence intervals around the means.

Study Design

Each choir participant ($n = 34$) visited the lab for a pre-training assessment that took approximately three hours, during which time they completed several questionnaires and auditory and cognitive assessments, and underwent an EEG during presentation of repeated auditory stimuli.

Choir-singing participants took part in weekly 2-hour group choral sessions over the course of 10 weeks, during which time they received pitch training and vocal direction in an open and encouraging environment. In addition to the weekly group choir sessions, participants were assigned weekly individual online musical and vocal training exercises (up to 1 hour weekly). This training consisted of pitch discrimination and vocal production exercises designed to target and improve the participants' abilities to perceive and produce small changes in pitch (Theta Music Trainer)¹.

After 10 weeks of choir participation and online musical training, each choir participant returned to the lab for a

post-training assessment that lasted approximately 2.5 hours. During this session, participants completed different versions of the original assessments, and underwent a post-training EEG of during auditory stimulus presentation. To account for possible differences in version difficulty within the matched behavioral tasks, participants were assigned one of four possible counterbalanced configurations of assessments.

The *do-nothing control group* ($n = 29$) underwent the same battery of pre- and post-testing, with 8–10 weeks between data collection sessions, but did not receive any active training during this time. The inclusion of this control group in the analysis intended to account for any practice effects within the repeated measures, enabling a controlled examination of the unique effects of the musical intervention on experimental outcomes.

Experimental Procedure

Apart from the questionnaires, all assessments were completed in an Industrial Acoustics Company (IAC) double-walled sound-attenuating booth. Computerized assessments were presented

¹<https://trainer.thetamusic.com>

using a Mac mini (Apple, 2010), with visual components of the experiment presented on a 24" Acer LCD display (Acer X243w, 1920 × 1200) placed at eye level approximately 0.5 m in front of the participant. Audiometric testing and FFR auditory stimuli were administered through binaural foam insert headphones (Electro-Medical Instruments, 3A) connected to a GSI 61 Clinical Audiometer (VIASYS Healthcare). All other auditory assessments were administered binaurally through Koss SB40 headphones at approximately 70 dB SPL.

Before the experiment began, participants were familiarized with task requirements and response methods for each assessment. Participants were monitored throughout the data collection session.

Questionnaires

After signing the consent form and going over experimental expectations and volunteer rights, participants were given background and music history questionnaires. These elicited demographics and medical history, and years of formal musical training.

Auditory Measures

Speech-in-noise perception: signal-to-noise ratio (SNR)

Ability to track speech in a noisy environment was assessed using the QuickSIN test (Speech-In-Noise; Etymotic Research; Killion et al., 2004). Participants were presented with four sets of six pre-recorded sentences, with five key words per sentence embedded in four-talker babble noise. In this assessment, the sentences were presented binaurally with a decreasing SNR: the first sentence was presented with an SNR of 25 dB (i.e., the target sentence was twenty-five dB above the background noise; *very easy*), each subsequent sentence was presented with a −5 dB SNR reduction, to an SNR of 0 dB for the final sentence. Participants were asked to repeat back the target sentences as closely to what they heard as possible, and were awarded one point for each correctly repeated target word, for a possible total of 30 points per set. The sentences in the QuickSIN do not contain many semantic or contextual cues, despite being syntactically correct (Wilson et al., 2007). Out of the four sets of sentences presented, the first two lists were treated as practice sets, to familiarize participants with the task requirements, and the second two lists were scored as experimental data. Mean SNR loss (dB) for each list was calculated by subtracting the total number of correct words from 25.5; Mean SNR loss (dB) represents the increased SNR required to correctly repeat 50% of key words on the QuickSIN test (Etymotic), above 2 dB SNR (the level required for normal hearing individuals to achieve 50% test accuracy; Killion, 1997; Killion et al., 2004). Final scores were calculated by averaging the scores of the two experimental lists; since this is a threshold assessment, a more negative SNR score indicates better performance. Participants' responses were scored online by a researcher, and were also recorded using Audacity software in case response ambiguity necessitated further review. The pre- and post-testing lists consisted of different sentence sets in order to avoid practice effects, and participants' exposure to the sets were counterbalanced across experimental conditions.

Pitch discrimination: frequency difference limens (FDL)

Participants' ability to distinguish different frequencies was measured using a computerized assessment of FDL. In this task, participants were presented with 3 pure tones, each lasting 200 ms, with amplitude envelopes of 20 ms rise and delay times. A three-alternative forced choice paradigm was used, in which each presented set contained two pure tones at the standard 500 Hz frequency, and one stimulus at a randomly selected higher frequency (Schneider, 1997; Parbery-Clark et al., 2009; Russo et al., 2012). The participant was instructed to identify which tone was higher than the other two by pressing the corresponding number on a computer keyboard (i.e., 1 = first tone is higher; 2 = second tone is higher; 3 = third tone is higher). An adaptive staircase procedure was used to determine the pitch discrimination threshold, whereby the difference between standard and comparison frequencies was halved after three correct responses, or doubled after one incorrect response. After five reversals, the step was changed, so that the frequency difference was divided by 1.414 after 3 correct responses or multiplied by 1.414 after one incorrect response. FDL was determined from the mean of the last 10 reversals.

EEG Measure: The Frequency Following Response (FFR)

Stimulus

Auditory presentation of a repeated/da/syllable ($F_0 = 100$ Hz) was used to elicit the FFR, following methodological conventions described by Skoe and Kraus (2010). This stimulus was selected because it is a speech sound that has been extensively used in this area of research, and robustly elicits clear FFRs (Russo et al., 2005; Parbery-Clark et al., 2009, 2012; Skoe and Kraus, 2010). Each participant heard 6000 repetitions of this 170 ms sound, presented at alternating polarities. Stimuli were presented binaurally through insert headphones; stimulus volume was set to 60 dB SPL for normal hearers. For individuals with hearing loss above 25 dB, presentation volume was set to 60 dB + (dB HL − 25 dB), controlling stimulus levels for sensory loss across all participants.

EEG administration and data collection

EEG data were collected using a vertical one-channel montage configuration, using three electrodes, in which active and reference electrodes were placed on the mastoids, and a ground electrode was placed on the forehead. A researcher applied 1" square cloth solid gel electrodes (EL504, BIOPAC Systems, Inc.) to the mastoids and forehead; electrodes were connected to a BIOPAC MP150 data acquisition system and ERS100C Evoked Response Amplifier (BIOPAC Systems, Inc.). Data were recorded at a sampling rate of 20 kHz, with an online low-pass filter of 10 kHz and a high-pass filter of 1 Hz; the signal was recorded using Acknowledge software (AcqKnowledge, version 4.1). Stimuli were presented for 25 minutes in total, during which time participants shown a silent film², to promote relaxation and stillness during the EEG.

²<http://www.openculture.com/free-silent-films>

TABLE 1 | Summary of linear mixed effects models for choir-singing and do-nothing control groups across key auditory measures.

Mean SNR Loss (dB)									
Predictors	Overall model			Choir-singing class			Do-nothing control group		
	Estimates	95% CI	p-value	Estimates	95% CI	p-value	Estimates	95% CI	p-value
(Intercept)	2.95	[2.15, 3.74]	<0.001	3.04	[2.32, 3.76]	<0.001	2.95	[2.14, 3.76]	<0.001
Session	0.45	[−0.24, 1.14]	0.205	−0.81	[−1.39, −0.23]	0.006	0.45	[−0.31, 1.21]	0.247
Group	0.10	[−0.98, 1.18]	0.862						
Session × Group	−1.26	[−2.20, −0.31]	0.009						
Random Effects									
σ^2	1.81			1.50			2.18		
τ_{00} Participant	2.95			3.09			2.79		
ICC Participant	0.62			0.67			0.56		
Observations	126			68			58		
Marginal R^2 /Conditional R^2	0.037/0.634			0.035/0.684			0.010/0.566		
Pitch Discrimination Thresholds (Frequency Difference Limens; FDL log Hz)									
(Intercept)	0.94	[0.84, 1.04]	<0.001	0.90	[0.81, 1.00]	<0.001	0.94	[0.84, 1.03]	0.001
Session	−0.03	[−0.10, 0.04]	0.455	−0.13	[−0.21, −0.03]	0.001	−0.03	[−0.08, 0.03]	0.322
Group	−0.03	[−0.17, 0.10]	0.623						
Session × Group	−0.11	[−0.20, −0.01]	0.036						
Random Effects									
σ^2	0.01			0.02			0.01		
τ_{00} Participant	0.04			0.04			0.04		
ICC Participant	0.77			0.70			0.85		
Observations	86			47			39		
Marginal R^2 /Conditional R^2	0.082/0.786			0.070/0.791			0.004/0.853		
FFR Strength at F0 (100 Hz; μV)									
(Intercept)	0.0118	[0.0091, 0.0144]	<0.001	0.0087	[0.0064, 0.0111]	<0.001	0.0118	[0.0091, 0.0145]	<0.001
Session	−0.0003	[−0.0034, 0.0027]	0.830	0.0033	[0.0008, 0.0058]	0.009	−0.0003	[−0.0035, 0.0028]	0.834
Group	−0.0030	[−0.0066, 0.0005]	0.096						
Session × Group	0.0037	[−0.0003, 0.0076]	0.072						
Random Effects									
σ^2	0.0025			0.0025			0.0029		
τ_{00} Participant	0.0023			0.0023			0.0021		
ICC Participant	0.45			0.47			0.43		
Observations	112			64			48		
Marginal R^2 /Conditional R^2	0.039/0.472			0.055/0.501			0.001/0.421		

The Session × Group interaction accounted for significant unique variance in speech-in-noise perception (mean SNR loss; dB) and pitch discrimination (Frequency Difference Limens; FDL, log Hz), and marginally significant variance in the strength of the FFR at F0 (μV). Group analyses showed that the choir-singing class demonstrated significant improvements in SNR, FDL, and FFR strength following training, while the control group showed no changes. Bold values indicate statistical significance ($p < 0.05$).

TABLE 2 | Mean scores compared with post hoc pairwise *t*-tests and related effect sizes (Cohen's *d*) for the choir-singing class ($n = 34$) and do-nothing control group ($n = 29$) at pre- and post-testing sessions, across auditory measures of interest.

		Pre-testing		Post-testing		<i>t</i> (df)	<i>p</i>	Cohen's <i>d</i>
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
Choir-singing class	Speech-in-noise perception (mean SNR loss; dB)	3.04	2.24	2.24	2.10	2.68(33)	0.0113	0.46
	Pitch discrimination (FDL; log Hz)	0.87	0.26	0.77	0.22	3.44(19)	0.0027	0.77
	FFR strength at F0 (μV)	0.0087	0.0065	0.0122	0.0075	−2.31(29)	0.0284	0.42
Do-nothing control group	Speech-in-noise perception (mean SNR loss; dB)	2.95	2.27	3.40	2.27	−1.14(28)	0.2651	0.21
	Pitch discrimination (FDL; log Hz)	0.92	0.25	0.90	0.21	0.98(17)	0.3402	0.23
	FFR strength at F0 (μV)	0.0120	0.0066	0.0117	0.0079	0.15(19)	0.8829	0.03

Bold values indicate statistical significance ($p < 0.05$).

EEG data processing

EEG data were processed in MATLAB, using the PHZLAB toolbox (Nespoli, 2016). A 75 Hz high-pass filter was applied, and data were segmented according to individual stimulus responses (i.e., 6000 segments), with epoch windows extending 40 ms pre- and post-stimulus, and the steady-state component extending from 60 to 170 ms post-stimulus onset (Skoe and Kraus, 2010). The 40 ms signal preceding stimulus onset was used as a baseline of ambient EEG activity, against which to compare the response activation. Peak amplitudes in the response waveform were compared to the baseline; response peaks with absolute amplitudes that did not exceed the baseline were not considered 'reliable' (Skoe and Kraus, 2010). Myogenic artifacts, which are many times larger than the neural response, were accounted for by rejecting all trials with amplitudes that exceeded a threshold of 50 μ V. Responses that remained after artifact rejection were averaged, using the addition method of inverse polarity processing in order to preserve the representation of the fundamental frequency while minimizing stimulus artifact in the signal. Peak amplitude of the fundamental frequency (a measure of the strength of the pitch representation in the signal, or *FFR strength*) was calculated by applying a FFT to the averaged signal, and inter-trial phase coherence (ITPC; a measure of response consistency or *FFR consistency*) was calculated by finding the latency variations across each participant's un-averaged signal.

Exploratory Cognitive Measures

Cognitive assessments were administered electronically on the stimulus computer (see section Experimental Procedure). Assessment scripts, coded in HTML-5, were retrieved from

the Millisecond online database³ (2016), adapted to have fewer blocks and runtime, and administered using Inquisit software (version 5.0.6). Working memory was assessed using a computerized version of the listening span task (LSpan), an auditory adaptation of the reading span task developed by Daneman and Carpenter (1980). Inhibitory control of attention was assessed using a computerized version of an adapted Flanker task (Ridderinkhof et al., 1997).

Statistical Analyses

Linear mixed effects analyses in a regression format were conducted on choir and control groups to examine the effects of choir participation on speech-in-noise perception (mean SNR loss; dB), pitch discrimination ability (FDL; Hz), and aspects of the FFR which represent the strength and consistency with which the speech fundamental frequency was represented. Exploratory cognitive measures of auditory working memory (LSpan) and inhibitory control of attention (Flanker effect) were also examined in the same format. In all models, measures were regressed on Session, Group, and the interaction between them (e.g., $\text{SNR} \sim \text{Session} \times \text{Group}$); contrasts were assigned such that the interaction effect represented the training effect of the choir group compared to the control group. Intercepts significantly varied across participants for all auditory measures; because the Session \times Group interaction was the main effect of interest for each outcome measure, individual variability in baseline scores across all dependent variables were included as random effects in the multilevel models. Session \times Group interactions are reported first in each section, and significant Session \times Group interactions were plotted and broken down in separate multilevel models of the choir and control groups. In these separate regressions conducted on each group, the models specified were the same as the main model for each variable, but excluded the main effect and interaction term involving Group. This was done in order to elucidate differential effects of choir participation vs. do-nothing control participation, on all outcomes of interest. Statistical analyses were conducted in R; the *nlme* and *lmer* packages were used to conduct linear mixed effects modeling in a regression format (Bates et al., 2015; Pinheiro et al., 2019).

RESULTS

Linear mixed effects models for key auditory measures are summarized in **Table 1**; pre- and post-testing group means, *post hoc* pairwise *t*-tests, and effect size calculations are reported in **Table 2**. Due to a computer error, FDL scores for 25 participants were spurious and removed from analyses.

Speech-in-Noise Perception

Pre-training and post-training SNR loss (dB) for choir singing and do-nothing control groups is plotted in **Figure 3**. The Session \times Group interaction accounted for significant variance in dB SNR loss [$b = -1.26$, $t(61) = -2.57$, $p = 0.009$]. Regressions conducted on each group showed that choir participants

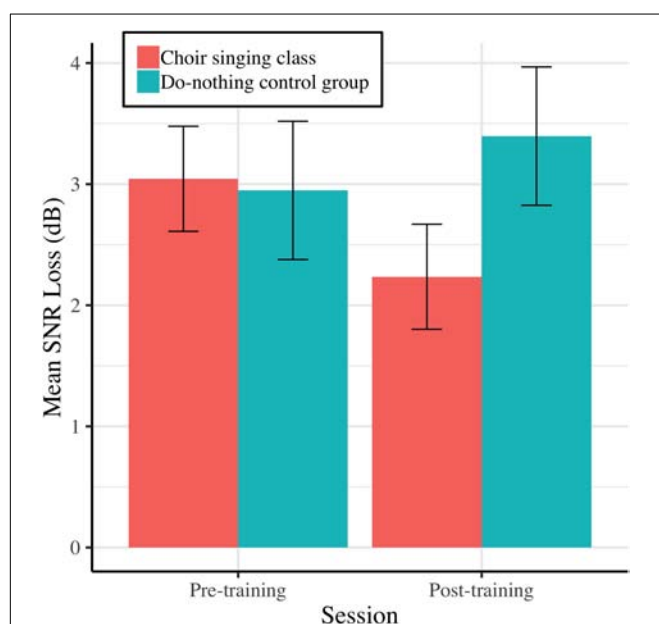


FIGURE 3 | Mean SNR loss (dB) before and after 10 weeks of choir singing or do-nothing control participation, plotted with 95% CIs for repeated measures.

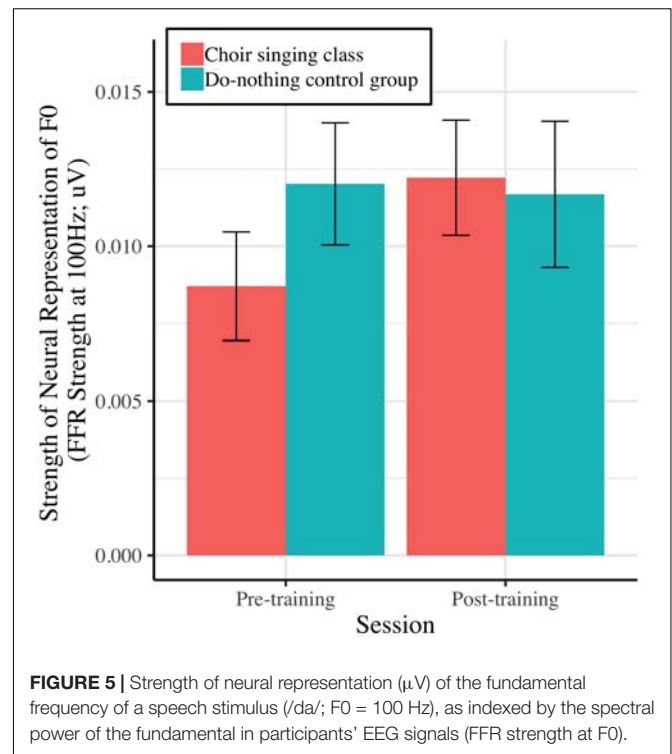
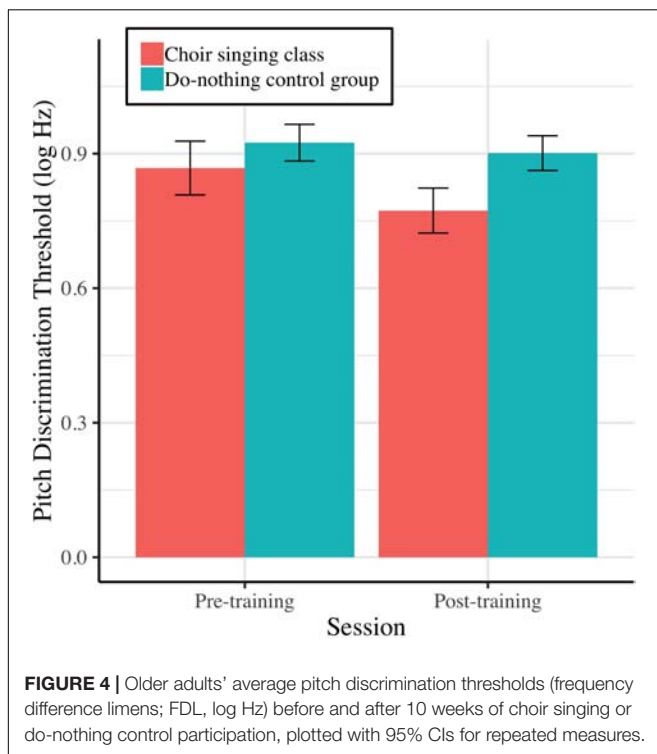
³<http://www.millisecond.com/download/library/>

demonstrated improvements of 0.81 dB SNR following training [$b = -0.81$, $t(33) = -2.68$, $p = 0.006$], while control participants showed no change [$b = 0.45$, $t(28) = 1.14$, $p = 0.247$; see **Table 1**].

Pitch Discrimination and Neural Representation of Frequency

Figure 4 shows mean pitch discrimination thresholds (FDL, log Hz) before and after 10 weeks of choir singing and do-nothing control participation. The Session \times Group interaction accounted for significant variance in frequency discrimination thresholds [log Hz; $b = -0.11$, $t(36) = -2.10$, $p = 0.036$]; regressions conducted on each group showed that choir participants demonstrated improved pitch discrimination thresholds following training [$b = -0.13$, $t(19) = -3.21$, $p = 0.0013$], while control participants showed no changes [$b = -0.03$, $t(17) = -0.99$, $p = 0.322$; see **Table 1**].

Figure 5 shows the strength of the neural representation of the fundamental frequency (F0) of the steady-state component of a complex sound (/da/; F0 = 100 Hz), before and after choir-singing or do-nothing control participation. The Session \times Group interaction accounted for marginally significant variance in the FFR strength at F0 [μV ; $b = 0.0037$, $t(48) = 1.77$, $p = 0.0721$]; regressions conducted on each group showed that following training, choir participants demonstrated improvements in the neural representation of pitch [$b = 0.0033$, $t(29) = 2.60$, $p = 0.009$], while control participants did not demonstrate significant changes [$b = -0.00034$, $t(19) = -0.21$, $p = 0.8346$; see **Table 1**]. The Session \times Group interaction did not account



for significant variance in the inter-trial phase coherence of the FFR [μV ; $b = 0.0025$, $t(50) = 0.38$, $p = 0.7066$].

Exploratory Cognitive Measures

There were no Session \times Group interaction effects on either listening span (auditory working memory) or the Flanker effect (inhibitory control of attention).

Possible Contributors to Choir-Driven Improvements in Speech-in-Noise Perception

Figure 6 shows the relationship between improvements in pitch discrimination ability, FFR strength at F0, and speech-in-noise perceptual gains. SNR scores of choir participants were analyzed in a multilevel model including Session (effect of training), FDL, and FFR strength as potential predictors of variance, in order to elucidate potential mechanisms for choir-related improvements in speech-in-noise perception. Predictors were added into the model hierarchically based on hypothesized contributions to speech-in-noise perception derived from previous research, and only predictors with significant Session \times Group interactions were included in the model (i.e., FDL and FFR strength at F0). The final choir model is reported in **Table 3**; with Session included in the model, the interaction between FDL and FFR strength accounted significant unique variance in SNR scores [$b = -239.74$, $t(21) = -3.18$, $p = 0.0045$]; inclusion of FDL accounted for training-related variance in SNR loss previously accounted for by Session, suggesting a potential mediating effect of pitch discrimination on training-related improvements in speech-in-noise perception.

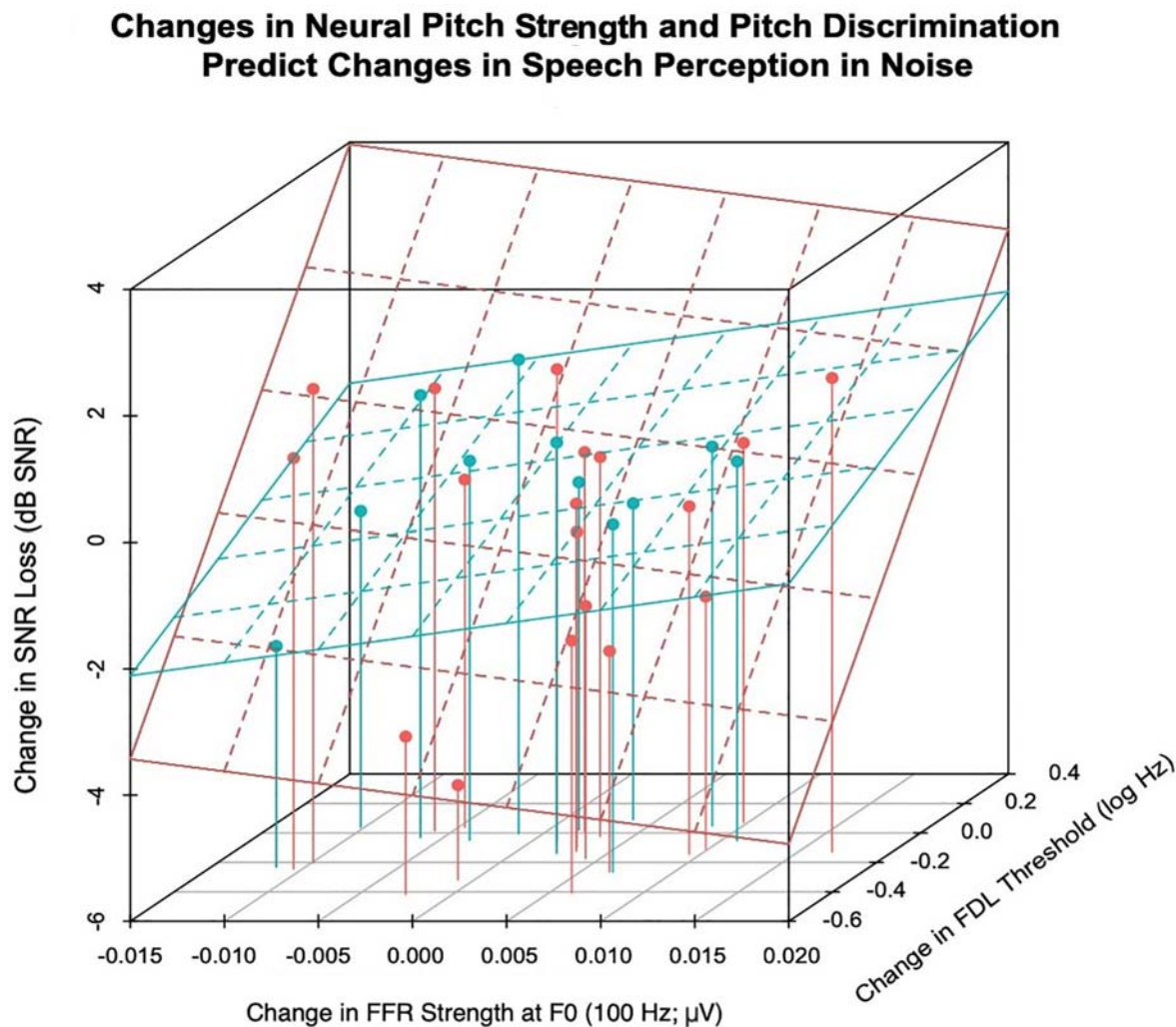


FIGURE 6 | The relationship between improvements in speech-in-noise perception (Δ mean SNR loss; dB), pitch discrimination thresholds (Δ FDL; log Hz), and the strength of the neural representation of frequency (Δ FFR strength at 100 Hz; μ V). Choir participants (red) who experienced greater improvements in the neural representation of F0 (more positive FFR strength) and pitch discrimination thresholds (more negative FDL) demonstrated greater improvements in speech-in-noise perception, as indexed by reduced SNR Loss (dB) following training. This relationship was non-significant for the control participants (blue).

Pitch Discrimination as a Potential Mechanism for Musicianship Improvements in Speech-in-Noise Perception: A Mediation Analysis

Figure 7 shows the distribution of the indirect effect of choir training on SNR, which shows that choir-related improvements in speech-in-noise perception were significantly mediated by improvements in pitch discrimination. The regression of mean SNR loss (dB) on Session, ignoring the mediator, was significant (see **Table 1**); when mean SNR loss (dB) was regressed on the mediator while controlling for Session, FDL significantly predicted SNR [$b = 3.57$, $t(18) = 3.22$, $p = 0.0047$], but training-specific effect was no longer a significant predictor of SNR [$b = -0.23$, $t(18) = -0.605$, $p = 0.553$]. Aroian's test revealed a significant mediating effect

of FDL on choir-related improvements in SNR (Aroian's test statistic = -2.20 , $p = 0.0280$), and a Monte Carlo resampling approach ($n = 20000$) confirmed that FDL fully mediated the relationship between choir participation and SNR improvements, 95% CI [-0.775 , -0.0894]; **Figure 7**.

The significant interaction effect of FDL \times FFR strength on SNR (**Table 3**) suggested a potential moderating effect of FFR on the relationship between pitch discrimination and SNR, so a moderated mediation analysis was conducted to assess the statistical significance of this effect. This analysis revealed a marginally significant moderation of the mediation by changes in FFR strength [$b = -277.14$, $t(14) = -1.885139$, $p = 0.0803$]; the model of this relationship is shown in **Figure 8**.

Exploratory analyses considered whether the relationship between choir-related improvements in pitch discrimination and speech-in-noise perception could be predicted by the

TABLE 3 | Hierarchical regression of possible contributors to speech-in-noise perceptual gains following choir (vs. control group) participation; marginal $R^2 = 0.325$, conditional $R^2 = 0.691$.

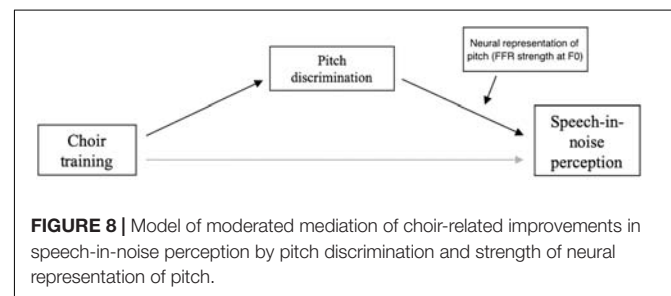
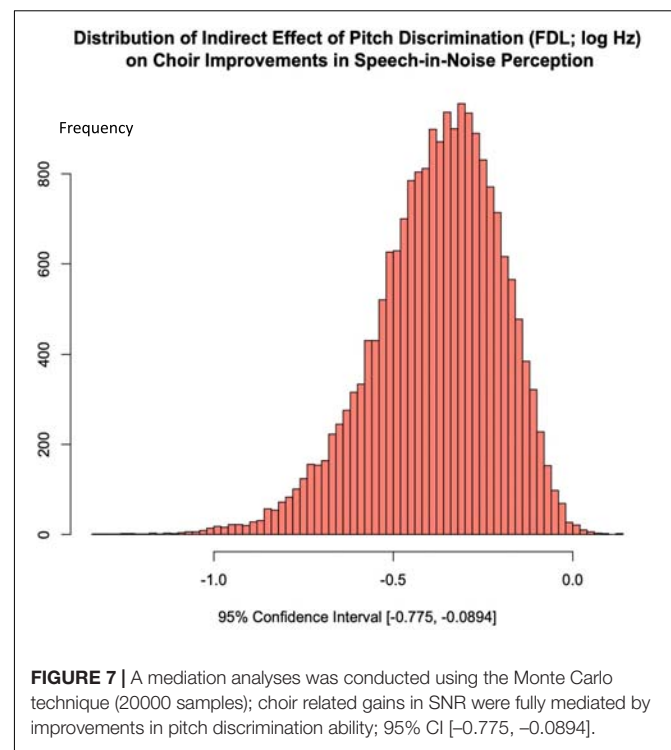
Choir SNR			
Predictors	Estimates	CI	p
(Intercept)	2.12	1.47 – 2.77	<0.001
Session	–0.22	–0.88 – 0.44	0.514
FDL	2.28	0.27 – 4.29	0.026
FFR strength	–102.93	–185.49 – –20.38	0.015
FDL × FFR strength	–277.14	–548.41 – –5.86	0.045
Random Effects			
σ^2	0.90		
τ_{00} Participant	1.07		
ICC Participant	0.54		
Observations	44		
Marginal R^2 /Conditional R^2	0.325/0.691		
Control SNR			
(Intercept)	2.87	1.82 – 3.91	<0.001
Session	0.63	–0.44 – 1.70	0.250
FDL	0.32	–3.38 – 4.02	0.864
FFR strength	–45.95	–190.40 – 98.51	0.533
FDL × FFR strength	164.51	–465.17 – 794.20	0.609
Random Effects			
σ^2	1.98		
τ_{00} Participant	2.62		
ICC Participant	0.57		
Observations	32		
Marginal R^2 /Conditional R^2	0.032/0.584		

Continuous predictor variables included in the moderation analysis (FDL and FFR strength at F0) were centered around the grand means for each variable before groups were subsetting. Bold values indicate statistical significance ($p < 0.05$).

strength of the representation of F0 in the FFR. Simple slopes analyses on low, average, and high FFR strength at F0 (centered variable \pm SD) showed that the relationship between pitch discrimination and SNR was strongest in *high FFR conditions* [i.e., when F0 is strongly represented in the FFR; $b = 4.28$, $t(14) = 3.29$, $p = 0.0054$], weaker in *average FFR conditions* [$b = 2.28$, $t(14) = 1.08$, $p = 0.0546$], and non-significant in *low FFR conditions* [$b = 0.28$, $t(14) = 0.16$, $p = 0.8714$], suggesting that when controlling for session, the strength of the FFR at F0 is predictive of the strength of the relationship between pitch discrimination and speech-in-noise perception. These analyses suggest that neural and perceptual pitch processes play a role in speech-in-noise perceptual ability in older adults, and could mechanistically contribute to a potential musicianship advantage in this domain.

Effects of Peripheral Hearing Loss on Perceptual and Neural Auditory Outcomes

Figures 9–11 demonstrate the differential effects of peripheral hearing loss (dB HL) on the efficacy of the choir-singing intervention on: speech-in-noise perception (Figure 9); pitch



discrimination (Figure 10); and FFR strength at F0 (Figure 11). There were significant main effects of Audiometry on SNR [$b = 0.07$, $t(60) = 2.66$, $p = 0.0101$] and FDL [$b = 0.008$, $t(45) = 2.52$, $p = 0.0153$]; across groups and sessions, worse peripheral impairments were predictive of worse performance on perceptual tasks. However, there were no significant effects of peripheral hearing loss on FFR strength at F0, indicating a potential differentiating effect of audiometry on neural vs. perceptual outcomes.

DISCUSSION

This study demonstrated experimentally that short-term choir participation can be used as an intervention to target and improve speech-in-noise perception in older adults, supporting the hypotheses that: (1) the musicianship advantage in speech-in-noise perception can be conferred to older adults through a relatively short training period, using choir singing and vocal training; and (2) enhancements in pitch processing contribute to improvements in this domain. This study lays the groundwork

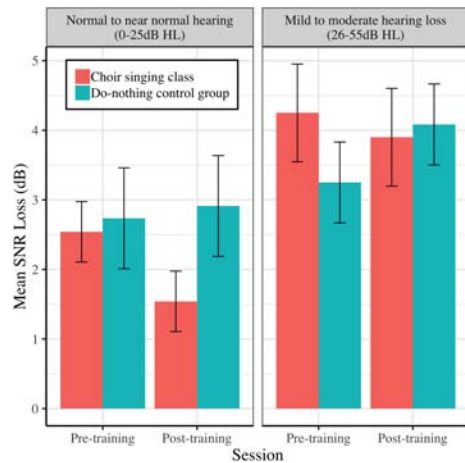


FIGURE 9 | The relationship between peripheral hearing loss and pre-post changes in mean SNR loss (dB), for choir and control groups. Degree of impairment is categorized on clinical audiometric criteria, where normal to near normal hearing = 0–25 dB HL, and mild to moderate = 26–55 dB HL. Error bars are within subjects CIs (95%) plotted around Session \times Group means.

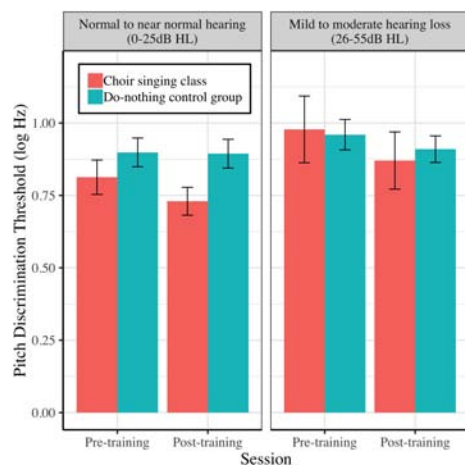


FIGURE 10 | The relationship between peripheral hearing loss and pre-post changes in pitch discrimination thresholds (FDL; log Hz) for choir and control groups. Degree of impairment is categorized on clinical audiometric criteria, where normal to near normal hearing = 0–25 dB HL, and mild to moderate = 26–55 dB HL. Error bars are within subjects CIs (95%) plotted around Session \times Group means.

for a highly scalable, cost-effective, and engaging intervention that can be used to mitigate declines in speech-in-noise perception in older adults, and importantly provides insight into potential neural and perceptual mechanisms underlying these changes. In particular, the relationship between auditory processing, pitch discrimination, and speech-in-noise perception suggested by this study elucidates one way in which musical experience – and specifically, singing and vocal training – can transfer to improvements in speech processing, through enhanced representation of pitch.

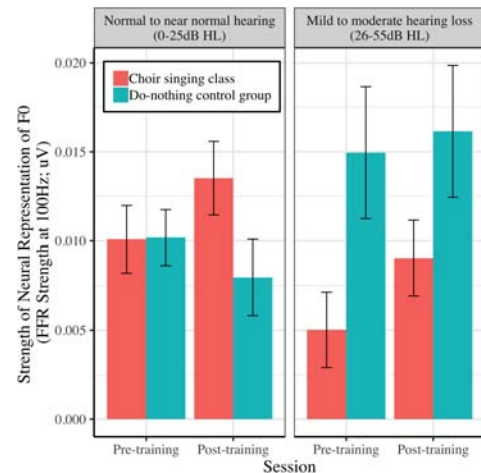


FIGURE 11 | The relationship between peripheral hearing loss and pre-post changes in the strength of neural representation of fundamental frequency (FFR strength at 100 Hz) for choir and control groups. Degree of impairment is categorized on clinical audiometric criteria, where normal to near normal hearing = 0–25 dB HL, and mild to moderate = 26–55 dB HL. Error bars are within subjects CIs (95%) plotted around Session \times Group means.

Compared with do-nothing control participants, choir singers demonstrated 1.26 dB improvements in mean SNR loss following training, a change that corresponds to a functional difference of 10–20% improvement in speech intelligibility (Middelweerd et al., 1990). Other forms of auditory rehabilitation for older adults yield similar improvements (1.5 dB with LACE training), require intensive practice in the target domain (30 min per day, 5 days per week for 4 weeks), which may account for the relatively poor compliance and high rates of attrition (Sweetow and Sabes, 2006; Song et al., 2012; Tyne-Murray et al., 2012). In contrast, group singing is reported to be highly engaging and motivating, provides many benefits outside of the focus of training, and promotes ongoing social involvement and activity (Hillman, 2002; Creech et al., 2013). It is important to note that in the current study, while nine participants withdrew from data collection, almost all of the original 53 participants surveyed remained in the choir class (two withdrew due to health issues), and many participants reported joining other choirs and singing groups after the study ended. As a proof of concept, this makes a strong case for the engagement and enjoyment of participants in a group singing class, and the sustainability of this type of intervention, along with its efficacy at improving speech-in-noise perception.

In terms of possible mechanisms accounting for changes in speech-in-noise perception, improvements appeared to be driven at least in part by enhancements in pitch processing. In addition to improved speech-in-noise perceptual abilities, choir singers demonstrated improved pitch discrimination thresholds (as indexed by lower FDL) and stronger neural representations of the speech fundamental frequency (F0) following training as (stronger FFR representation of F0 of the/da/stimulus; 100 Hz). Analyses showed that training-related improvements in

speech-in-noise perception were fully mediated by improvements in pitch discrimination, suggesting that the benefits afforded by choir-singing arose at least in part from enhancements in the perception of pitch. A moderated mediation analysis suggested that over the course of choir training, the strength of the neural representation of F0 was predictive of the strength of the relationship between pitch discrimination and speech-in-noise perception. This suggests that neural indices of pitch processing influence the extent to which older adults rely on pitch cues to support speech-in-noise perception. Taken together, these findings suggest that older adults who take part in 10 weeks of choir singing and vocal training demonstrate enhanced neural responses to and perception of subtle frequency cues, which lead to improved perception of speech-in-noise following training.

A number of previous studies have findings that converge with our mediation account of speech-in-noise via pitch perception. For example, musical experience has been correlated with improvements in speech-in-noise perception (Parbery-Clark et al., 2009; Zendel and Alain, 2012; Swaminathan et al., 2015; for review see Coffey et al., 2017a), pitch discrimination ability (Micheyl et al., 2006; Parbery-Clark et al., 2009; Schellenberg and Moreno, 2009; Bidelman et al., 2011a,b; Fuller et al., 2014; Boebinger et al., 2015; Yates et al., 2018), and subcortical encoding of F0 (Parbery-Clark et al., 2009, 2011; Bidelman et al., 2011a,b; Musacchia et al., 2017); relationships have also been demonstrated between pitch perception and subcortical encoding of F0 (Carcagno and Plack, 2011; Coffey et al., 2016; Bianchi et al., 2017), as well as between speech-in-noise perception and pitch processing (e.g., Coffey et al., 2017a). On the other hand, a number of correlational studies have not been able to replicate the musicianship advantage for speech-in-noise perception (e.g., Ruggles et al., 2014; Madsen et al., 2017). Some of this discrepancy may be based on methodological or sampling differences across studies. More generally, limited experimental work in this field has left the nature of these relationships somewhat uncertain (excepting some recent work by Zendel et al., 2019). Zendel et al. (2019) found that older adults (non-musicians) who took part in 6 months of piano training showed improvements in speech-in-noise perception compared with control and video game intervention groups showing no improvements in this domain. Importantly, these individuals were randomly assigned to the interventions in this study, lending credence to the use of musical training to support auditory abilities in older adults.

As a musical intervention, choir singing may be uniquely suited to hone pitch perceptual processes, through activation of existing vocal-motor systems, rapid integration of perceptual and productive processes, and shared neural architecture activated by speech and vocal song. Choir singers have the benefit of both intrinsic auditory and sensorimotor feedback, and can harness existing feedback loops between auditory perception and vocal production – which allow humans to monitor and dynamically alter speech – to rapidly alter and hone vocal output, including production of pitch. These integrative feedback loops and fine-tuned changes may allow choir singers to undergo rapid improvements in both productive

and perceptual processes, in a short period of time. Singing is also an intuitive and innate form of music-making, and may be learned (and improved upon) more quickly than learning to play an instrument. This innate quality, along with intrinsic auditory and vocal motor feedback loops, and extrinsic feedback (from the choir director and other singers) create the ideal circumstances to quickly and effectively improve pitch processing and downstream perceptual abilities such as speech-in-noise perception.

While the current study found that improvements in pitch processing fully mediated choir enhancements in speech-in-noise perception, this does not preclude the role of other neural, perceptual, and cognitive contributors to this ability. Previous work suggests that musical training also leads to enhancements in attentional processes involved with speech encoding (e.g., Zendel et al., 2019); auditory working memory has also been implicated in this ability (e.g., Kraus et al., 2012). Further experimental research is necessary to determine the unique contributions of various auditory, cognitive, and neural processes to music-related improvements in speech-in-noise perception in older adults. In addition, the contribution of productive musical training (i.e., singing practice) vs. perceptual training (i.e., learning to listen to differences in pitch) to auditory processing improvements is not clear from the current study. Notably, a recent study found that non-musicians who received targeted pitch perceptual training achieved pitch discrimination thresholds comparable to musicians in 4–8 hours (Micheyl et al., 2006); it is unknown whether these pitch improvements would be sustained over time, or transfer to speech-in-noise perceptual benefits; it is also unclear whether the mechanism by which pitch discrimination is improved – i.e., through targeted psychoacoustic training, vs. through a more naturalistic singing or music listening paradigm – would alter the degree to which pitch perception mediates speech perceptual processes. This further underscores the need for targeted experimental study of musical (and non-musical) perceptual and productive training on auditory abilities, to elucidate the roles and contributions of each. It is also unclear from the current study whether the auditory benefits of choir participation would persist after cessation of training, and whether these benefits would accumulate with additional/long-term choir or musical involvement. These are rich avenues for future research projects, especially in an experimental/longitudinal context.

In addition to the role of pitch, the degree of peripheral hearing loss appeared to influence the amount of gains choir participants experienced as a function of training, whereby participants with lower levels of peripheral hearing loss appeared to experience greater training-related improvements in speech-in-noise perception. This could be suggestive of a possible limit on the efficacy of this intervention at improving perceptual processes in individuals with levels of hearing loss approaching the need for peripheral assistance (i.e., 26–55 dB HL). One potential explanation is that these individuals may not have been able to hear well enough in the classes to pitch-match with other voices, and thus did not receive equivalent experiential benefits from this activity.

Another explanation is that greater peripheral impairments may have led to more substantial central deficits that were recalcitrant to a behavioral intervention in this capacity, at least in the current dose of 10 weeks of choir singing (2 hours/week of group singing, plus 1 hour/week of individual online exercises). However, there were no effects of peripheral hearing loss on the strength of FFR responses, as individuals within the upper range of peripheral hearing loss still showed improvements in the FFR representation of F0. This suggests that while individuals with greater peripheral hearing loss may not receive perceptual benefits from 10 weeks of choir participation, their neural responses may still be enhanced through this experience. An interesting line of inquiry for a future study would be to address whether individuals with greater hearing loss may be able to obtain similar benefits by participating in choir training in conjunction with the use of hearing aids.

Overall, group choral singing appears to be uniquely well suited for this training paradigm, as it encourages singers to produce (and hopefully, perceive) fine-tuned adjustments in pitch structure, which seem to play a major role in improving speech-in-noise perceptual outcomes in this population. The intrinsic relationship with speech, rapid sensory-motor integration, instantaneous feedback afforded by vocal production and auditory perception, and the innate nature of this ability suggest singing as an ideal candidate for improving a speech-related perceptual issue. Group singing is highly motivating, social, and emotionally fulfilling; this is of immense import in developing interventions that will encourage engagement and promote active involvement, especially with older adults. Overall, running a choir is an immensely scalable intervention, requiring minimal cost and equipment (and which could be implemented anywhere), and this study demonstrated that training-related improvements in auditory perception can appear after a very short intervention period (10 weeks of choir singing). The efficacy of this intervention can easily be assessed through experimental manipulations of dose or duration (e.g., using longer periods of choir singing, or assessing persistence of effects post-training), different study populations (e.g., hearing aid users vs. unaided individuals), and with different emphases during the class (e.g., focusing on pitch matching/singing in unison vs. attending to different melodic or harmonic lines). The ease of implementation and scalability of the choir singing paradigm, efficacy at improving auditory abilities in aging adults, and rich opportunity for further investigation suggest choir singing as an ideal framework for examining musical training as an auditory rehabilitation for aging adults.

CONCLUSION

Group singing is an intuitive, engaging, and motivating form of music making, that has in previous studies been shown to contribute to social, emotional, cognitive, and

physical well-being. The current findings suggest that choir singing can be used as an effective intervention to mitigate age-related losses in auditory perceptual abilities, in as short a time as 10 weeks. Importantly, these findings showed that this intervention improved older adults' abilities to perceive speech in noisy environments, a key concern in promoting healthy aging. This work provides an empirical basis for a highly scalable and effective intervention that could significantly improve quality of life in older adults.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ryerson University Ethics Board (REB). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

ED and FR conceived and designed the study. ED recruited the participants, collected the data, analyzed the results, and wrote up findings as Master's thesis with FR as supervisor. EW collected the data, revised the literature review, and contributed to the theoretical interpretation of findings. GN developed MATLAB toolbox for processing EEG data. All authors were involved in discussions about the interpretation of the results and the writing of the manuscript.

FUNDING

This work was supported by a Research Chair sponsored by Sonova Holding AG awarded to FR, as well as a grant from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2017-06969).

ACKNOWLEDGMENTS

We would like to thank Saul Moshe-Steinberg, James McGrath, and Paolo Ammirante for their work toward establishing a subset of the protocols implemented in the current investigation. We are also indebted to Sina Fallah for leading the singing classes and to the 50+ Program of the Chang School of Continuing Education (Ryerson University) for their partnership in all aspects of this research.

REFERENCES

- Abel, S. M., Giguere, C., Consoli, A., and Papsin, B. C. (2000). The effect of aging on horizontal plane sound localization. *J. Acoust. Soc. Am.* 108, 743–752. doi: 10.1121/1.429607
- Alain, C., McDonald, K. L., Ostroff, J. M., and Schneider, B. (2001). Age-related changes in detecting a mistuned harmonic. *J. Acoust. Soc. Am.* 109, 2211–2216. doi: 10.1121/1.1367243
- Alain, C., Zendel, B. R., Hutka, S., and Bidelman, G. M. (2014). Turning down the noise: the benefit of musical training on the aging auditory brain. *Hear. Res.* 308, 162–173. doi: 10.1016/j.heares.2013.06.008
- Anderson, S., Parbery-Clark, A., Yi, H.-G., and Kraus, N. (2011). A neural basis of speech-in-noise perception in older adults. *Ear Hear.* 32, 750–757. doi: 10.1097/AUD.0b013e31822229d3
- Arlinger, S. (2003). Negative consequences of uncorrected hearing loss - a review. *Int. J. Audiol.* 42, S17–S20. doi: 10.3109/14992020309074639
- Bailey, B. A. (2005). Effects of group singing and performance for marginalized and middle-class singers. *Psychol. Music* 33, 269–303. doi: 10.1177/0305735605053734
- Bates, D., Mächler, M., Bolker, B. M., and Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01
- Beck, R. J., Cesario, T. C., Yousefi, A., and Enamoto, H. (2000). Choral singing, performance perception, and immune system changes in salivary immunoglobulin a and cortisol. *Music Percep.* 18, 87–106. doi: 10.2307/40285902
- Besson, M., Chobert, J., and Marie, C. (2011). Transfer of training between music and speech: common processing, attention, and memory. *Front. Psychol.* 2:94. doi: 10.3389/fpsyg.2011.00094
- Besson, M., Schön, D., Moreno, S., Santos, A., and Magne, C. (2007). Influence of musical expertise and musical training on pitch processing in music and language. *Restorative Neurol. Neurosci.* 25, 399–410. doi: 10.1162/jocn.2010.21585
- Betlejewski, S. (2006). Age connected hearing disorders (presbycusis) as a social problem. *Otolaryngol. Polska Pol. Otolaryngol.* 60, 883–886.
- Bianchi, F., Hjortkjær, J., Santurette, S., Zatorre, R. J., Siebner, H. R., and Dau, T. (2017). Subcortical and cortical correlates of pitch discrimination: evidence for two levels of neuroplasticity in musicians. *NeuroImage* 163, 398–412. doi: 10.1016/j.neuroimage.2017.07.057
- Biasutti, M., and Mangiacotti, A. (2018). Assessing a cognitive music training for older participants: a randomised controlled trial. *Int. J. Geriatr. Psychiatr.* 33, 271–278. doi: 10.1002/gps.4721
- Bidelman, G. M., Gandour, J. T., and Krishnan, A. (2011a). Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. *Brain Cogn.* 77, 1–10. doi: 10.1016/j.bandc.2011.07.006
- Bidelman, G. M., Krishnan, A., and Gandour, J. T. (2011b). Enhanced brainstem encoding predicts musicians' perceptual advantages with pitch. *Eur. J. Neurosci.* 33, 530–538. doi: 10.1111/j.1460-9568.2010.07527.x
- Bidelman, G. M., and Krishnan, A. (2010). Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Res.* 1355, 112–125. doi: 10.1016/j.brainres.2010.07.100
- Boebinger, D., Evans, S., Rosen, S., Lima, C. F., Manly, T., and Scott, S. K. (2015). Musicians and non-musicians are equally adept at perceiving masked speech. *J. Acoust. Soc. Am.* 137, 378–387. doi: 10.1121/1.4904537
- Brainard, M. S., and Doupe, A. J. (2000). Auditory feedback in learning and maintenance of vocal behaviour. *Nat. Rev. Neurosci.* 1, 31–40. doi: 10.1038/35036205
- Bugos, J. A., Perlstein, W. M., McCrae, C. S., Brophy, T. S., and Bedenbaugh, P. H. (2007). Individualized piano instruction enhances executive functioning and working memory in older adults. *Aging Men. Health* 11, 464–471. doi: 10.1080/13607860601086504
- Carcagno, S., and Plack, C. J. (2011). Subcortical plasticity following perceptual learning in a pitch discrimination task. *J. Assoc. Res. Otolaryngol.* 12, 89–100. doi: 10.1007/s10162-010-0236-1
- Chmiel, R., and Jerger, J. (1996). Hearing aid use, central auditory disorder, and hearing handicap in elderly persons. *J. Am. Acad. Audiol.* 7, 190–202.
- Clift, S., and Hancox, G. (2001). The perceived benefits of singing: findings from preliminary surveys of a university college choral society. *J. R. Soc. Promot. Health* 121, 248–256. doi: 10.1177/146642400112100409
- Clift, S., and Morrison, I. (2011). Group singing fosters mental health and wellbeing: findings from the East Kent “singing for health” network project. *Men. Health Soc. Inclusion* 15, 88–97. doi: 10.1108/20428301111140930
- Clinard, C. G., Tremblay, K. L., and Krishnan, A. R. (2010). Aging alters the perception and physiological representation of frequency: evidence from human frequency-following response recordings. *Hear. Res.* 264, 48–55. doi: 10.1016/j.heares.2009.11.010
- Coffey, E. B. J., Chepesiuk, A. M. P., Herholz, S. C., Baillet, S., and Zatorre, R. J. (2017a). Neural correlates of early sound encoding and their relationship to speech-in-noise perception. *Front. Neurosci.* 11:1–14. doi: 10.3389/fnins.2017.00479
- Coffey, E. B. J., Herholz, S. C., Chepesiuk, A. M. P., Baillet, S., and Zatorre, R. J. (2016). Cortical contributions to the auditory frequency-following response revealed by MEG. *Nat. Commun.* 7:11070. doi: 10.1038/ncomms11070
- Coffey, E. B. J., Mogilever, N. B., and Zatorre, R. J. (2017b). Speech-in-noise perception in musicians: a review. *Hear. Res.* 352, 49–69. doi: 10.1016/j.heares.2017.02.006
- Cohen, G. D., Perlstein, S., Chapline, J., Kelly, J., Firth, K. M., and Simmens, S. (2006). The impact of professionally conducted cultural programs on the physical health, mental health, and social functioning of older adults. *Gerontologist* 46, 726–734. doi: 10.1093/geront/46.6.726
- Creech, A., Hallam, S., McQueen, H., and Varvarigou, M. (2013). The power of music in the lives of older adults. *Res. Stud. Music Educ.* 35, 87–102. doi: 10.1177/1321103X13478862
- Cruikshanks, K. J., Wiley, T. L., Tweed, T. S., Klein, B. E. K., Klein, R., Mares-Perlman, J. A., et al. (1998). Prevalence of hearing loss in older adults in beaver dam, wisconsin. *Am. J. Audiol.* 148, 879–886. doi: 10.1093/oxfordjournals.aje.a009713
- Dalla Bella, S., and Berkowska, M. (2009). Singing proficiency in the majority: normality and “phenotypes” of poor singing. *Ann. N. Y. Acad. Sci.* 1169, 99–107. doi: 10.1111/j.1749-6632.2009.04558.x
- Dalla Bella, S., Giguère, J.-F., and Peretz, I. (2007). Singing proficiency in the general population. *J. Acoust. Soc. Am.* 121, 1182–1189. doi: 10.1121/1.2427111
- Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory during reading. *J. Verb. Learn. Verb. Behav.* 19, 450–466. doi: 10.1016/S0022-5371(80)90312-6
- David, S. V., Fritz, J. B., and Shamma, S. A. (2012). Task reward structure shapes rapid receptive field plasticity in auditory cortex. *Proc. Natl. Acad. Sci. U.S.A.* 109, 2144–2149. doi: 10.1073/pnas.1117717109
- Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009
- Djernes, J. K. (2006). Prevalence and predictors of depression in populations of elderly: a review. *Acta Psychiatr. Scand.* 113, 372–387. doi: 10.1111/j.1600-0447.2006.00770.x
- Donai, J. J., and Jennings, M. B. (2016). Gaps-in-noise detection and gender identification from noise-vocoded vowel segments: comparing performance of active musicians to non-musicians. *J. Acoust. Soc. Am.* 139, EL128–EL134. doi: 10.1121/1.4947070
- Du, Y., Buchsbaum, B. R., Grady, C. L., and Alain, C. (2014). Noise differentially impacts phoneme representations in the auditory and speech motor systems. *Proc. Natl. Acad. Sci. U.S.A.* 111, 7126–7131. doi: 10.1073/pnas.1318738111
- Du, Y., Buchsbaum, B. R., Grady, C. L., and Alain, C. (2016). Increased activity in frontal motor cortex compensates impaired speech perception in older adults. *Nat. Commun.* 7, 1–12. doi: 10.1038/ncomms12241
- Dubno, J. R., Dirks, D. D., and Morgan, D. E. (1984). Effects of age and mild hearing loss on speech recognition in noise. *J. Acoust. Soc. Am.* 76, 87–96. doi: 10.1121/1.391011
- Fabiani, M. (2012). It was the best of times, it was the worst of times: a psychophysiology's view of cognitive aging. *Psychophysiology* 49, 283–304. doi: 10.1111/j.1469-8986.2011.01331.x
- Fitzgibbons, P. J., and Gordon-Salant, S. (1994). Age effects on measures of auditory duration discrimination. *J. Speech Hear. Res.* 37, 662–670. doi: 10.1044/jshr.3703.662
- Fozard, J. L. (1990). “Vision and hearing in aging,” in *The Handbooks of Aging. Handbook of the Psychology of Aging*, eds J. E. Birren, and K. W. Schaie, (San Diego, CA: Academic Press), 150–170. doi: 10.1016/b978-0-12-101280-9.50015-2

- Frisina, R. D., and Frisina, R. D. (1997). Speech recognition in noise and presbycusis: relations to possible neural mechanisms. *Hear. Res.* 106, 95–104. doi: 10.1016/S0378-5955(97)00006-3
- Frisina, R. D., and Walton, J. P. (2006). Age-related structural and functional changes in the cochlear nucleus. *Hear. Res.* 216–217, 216–223. doi: 10.1016/j.heares.2006.02.003
- Fu, M. C., Belza, B., Nguyen, H., Logsdon, R., and Demorest, S. (2018). Impact of group-singing on older adult health in senior living communities: a pilot study. *Arch. Gerontol. Geriatr.* 76, 138–146. doi: 10.1016/j.archger.2018.02.012
- Fujioka, T., Ross, B., Kakigi, R., Pantev, C., and Trainor, L. J. (2006). One year of musical training affects development of auditory cortical-evoked fields in young children. *Brain* 129, 2593–2608. doi: 10.1093/brain/awl247
- Fuller, C. D., Galvin, J. J., Maat, B., Free, R. H., and Başkent, D. (2014). The musician effect: does it persist under degraded pitch conditions of cochlear implant simulations? *Front. Neurosci.* 8:179. doi: 10.3389/fnins.2014.00179
- Gaser, C., and Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *J. Neurosci.* 23, 9240–9245. doi: 10.1523/jneurosci.23-27-09240.2003
- Gomez, R. G., and Madey, S. F. (2001). Coping-with-hearing-loss model for older adults. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 56, 223–225.
- Good, A., and Russo, F. A. (2016). Singing promotes cooperation in a diverse group of children. *Soc. Psychol.* 47, 1–5. doi: 10.1027/1864-9335/a000282
- Gordon-Salant, S. (2005). Hearing loss and aging: new research findings and clinical implications. *J. Rehabil. Res. Dev.* 42(4 Suppl. 2), 9–24. doi: 10.1682/JRRD.2005.01.0006
- Grassi, M., Meneghetti, C., Toffalini, E., and Borella, E. (2017). Auditory and cognitive performance in elderly musicians and nonmusicians. *PLoS One* 12:e0187881. doi: 10.1371/journal.pone.0187881
- Habibi, A., Cahn, B. R., Damasio, A., and Damasio, H. (2016). Neural correlates of accelerated auditory processing in children engaged in music training. *Dev. Cogn. Neurosci.* 21, 1–14. doi: 10.1016/j.dcn.2016.04.003
- Habibi, A., Wirantana, V., and Starr, A. (2014). Cortical activity during perception of musical rhythm: comparing musicians and nonmusicians. *Psychomusicology* 24, 125–135. doi: 10.1037/pmu0000046
- Hanna-Pladdy, B., and MacKay, A. (2011). The relation between instrumental musical activity and cognitive aging. *Neuropsychology* 25, 378–386. doi: 10.1037/a0021895
- Hargus, S. E., and Gordon-Salant, S. (1995). Accuracy of speech intelligibility index predictions for noise-masked young listeners with normal hearing and for elderly listeners with hearing impairment. *J. Speech Hear. Res.* 38, 234–243. doi: 10.1044/jshr.3801.234
- Hays, T., and Minichiello, V. (2005). The meaning of music in the lives of older people: a qualitative study. *Psychol. Music* 33, 437–451. doi: 10.1177/0305735605056160
- Herdener, M., Esposito, F., di Salle, F., Boller, C., Hilti, C. C., Habermeyer, B., et al. (2010). Musical training induces functional plasticity in human hippocampus. *J. Neurosci.* 30, 1377–1384. doi: 10.1523/JNEUROSCI.4513-09.2010
- Herholz, S. C., and Zatorre, R. J. (2012). Musical training as a framework for brain plasticity: behavior, function, and structure. *Neuron* 76, 486–502. doi: 10.1016/j.neuron.2012.10.011
- Hickok, G. (2001). Functional anatomy of speech perception and speech production: psycholinguistic implications. *J. Psychol. Res.* 30, 225–235. doi: 10.1023/A:1010486816667
- Hickok, G., Buchsbaum, B., Humphries, C., and Muftuler, T. (2003). Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area Spt. *J. Cogn. Neurosci.* 15, 673–682. doi: 10.1162/08989290322307393
- Hickson, L., Worrall, L., and Scarinci, N. (2007). A randomized controlled trial evaluating the active communication education program for older people with hearing impairment. *Ear Hear.* 28, 212–230. doi: 10.1097/AUD.0b013e31803126c8
- Hickson, L., Worrall, L., Scarinci, N., and Laplante-Lévesque, A. (2019). Individualised active communication education (I-ACE): another clinical option for adults with hearing impairment with a focus on problem solving and self-management. *Int. J. Audiol.* 0, 1–6. doi: 10.1080/14992027.2019.1587180
- Hillman, S. (2002). Participatory singing for older people: a perception of benefit. *Health Educ.* 102, 163–171. doi: 10.1108/09654280210434237
- Hoormann, J., Falkenstein, M., Hohnsbein, J., and Blanke, L. (1992). The human frequency-following response (FFR): normal variability and relation to the click-evoked brainstem response. *Hear. Res.* 59, 179–188. doi: 10.1016/0378-5955(92)90114-3
- Houde, J. F., and Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science* 279, 1213–1216. doi: 10.1126/science.279.5354.1213
- Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., et al. (2009). Musical training shapes structural brain development. *J. Neurosci.* 29, 3019–3025. doi: 10.1523/JNEUROSCI.5118-08.2009
- Jain, C., Mohamed, H., and Kumar, U. A. (2015). The effect of short-term musical training on speech perception in noise. *Audiol. Res.* 5, 5–8. doi: 10.4081/audiores.2015.111
- Johnson, K. L., Nicol, T. G., and Kraus, N. (2005). Brain stem response to speech: a biological marker of auditory processing. *Ear Hear.* 26, 424–434. doi: 10.1097/01.aud.0000179687.71662.6e
- Killion, M. C. (1997). Hearing aids: past, present, future: moving toward normal conversation in noise. *Br. J. Audiol.* 31, 141–148. doi: 10.3109/03005364000000016
- Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J., and Banerjee, S. (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* 116:2395. doi: 10.1121/1.1784440
- Kim, S. H., Frisina, R. D., Mapes, F. M., Hickman, E. D., and Frisina, D. R. (2006). Effect of age on binaural speech intelligibility in normal hearing adults. *Speech Commun.* 48, 591–597. doi: 10.1016/j.specom.2005.09.004
- Kishon-Rabin, L., Amir, O., Vexler, Y., and Zaltz, Y. (2001). Pitch discrimination: are professional musicians better than non-musicians? *J. Basic Clin. Physiol. Pharmacol.* 12, 125–143.
- Koelsch, S., Schröger, E., and Tervaniemi, M. (1999). Superior pre-attentive auditory processing in musicians. *Neuroreport* 10, 1309–1313. doi: 10.1097/00001756-199904260-00029
- Kraus, N., and Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nat. Rev. Neurosci.* 11, 599–605. doi: 10.1038/nrn2882
- Kraus, N., Skoe, E., Parbery-Clark, A., and Ashley, R. (2009). Experience-induced malleability in neural encoding of pitch, timbre, and timing: implications for language and music. *Ann. N. Y. Acad. Sci.* 1169, 543–557. doi: 10.1111/j.1749-6632.2009.04549.x
- Kraus, N., Strait, D. L., and Parbery-Clark, A. (2012). Cognitive factors shape brain networks for auditory skills: spotlight on auditory working memory. *Ann. N. Y. Acad. Sci.* 1252, 100–107. doi: 10.1111/j.1749-6632.2012.06463.x
- Kreutz, G., Bongard, S., Rohrmann, S., Hodapp, V., and Grebe, D. (2004). Effects of choir singing or listening on secretory immunoglobulin a, cortisol, and emotional state. *J. Behav. Med.* 27, 623–635. doi: 10.1007/s10865-004-0006-9
- Kricos, P. B., and Holmest, A. E. (1996). Efficacy of audiologic rehabilitation for older adults. *J. Am. Acad. Audiol.* 7, 219–229.
- Kuriki, S. (2006). Effects of musical experience on different components of MEG responses elicited by sequential piano-tones and chords. *J. Neurosci.* 26, 4046–4053. doi: 10.1523/JNEUROSCI.3907-05.2006
- Lappe, C., Herholz, S. C., Trainor, L. J., and Pantev, C. (2008). Cortical plasticity induced by short-term unimodal and multimodal musical training. *J. Neurosci.* 28, 9632–9639. doi: 10.1523/JNEUROSCI.2254-08.2008
- Lappe, C., Trainor, L. J., Herholz, S. C., and Pantev, C. (2011). Cortical plasticity induced by short-term multimodal musical rhythm training. *PLoS One* 6:e21493. doi: 10.1371/journal.pone.0021493
- Lee, K. M., Skoe, E., Kraus, N., and Ashley, R. (2009). Selective subcortical enhancement of musical intervals in musicians. *J. Neurosci.* 29, 5832–5840. doi: 10.1523/JNEUROSCI.6133-08.2009
- Lehmann, A., and Schönwiesner, M. (2014). Selective attention modulates human auditory brainstem responses: relative contributions of frequency and spatial cues. *PLoS One* 9:e85442. doi: 10.1371/journal.pone.0085442
- Madsen, S. M. K., Marschall, M., Dau, T., and Oxenham, A. J. (2019). Speech perception is similar for musicians and non-musicians across a wide range of conditions. *Sci. Rep.* 9, 1–10. doi: 10.1038/s41598-019-46728-1
- Madsen, S. M. K., Whiteford, K. L., and Oxenham, A. J. (2017). Musicians do not benefit from differences in fundamental frequency when listening to speech in competing speech backgrounds. *Sci. Rep.* 7, 1–10. doi: 10.1038/s41598-017-12937-9
- Mansens, D., Deeg, D. J. H., and Comijs, H. C. (2017). The association between singing and/or playing a musical instrument and cognitive functions

- in older adults. *Aging Men. Health* 22, 1–8. doi: 10.1080/13607863.2017.1328481
- Meha-Bettison, K., Sharma, M., Ibrahim, R. K., and Mandikal Vasuki, P. R. (2018). Enhanced speech perception in noise and cortical auditory evoked potentials in professional musicians. *Int. J. Audiol.* 57, 40–52. doi: 10.1080/14992027.2017.1380850
- Micheyl, C., Delhommeau, K., Perrot, X., and Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hear. Res.* 219, 36–47. doi: 10.1016/j.heares.2006.05.004
- Middelweerd, M. J., Festen, J. M., and Plomp, R. (1990). Difficulties with speech intelligibility in noise in spite of a normal pure-tone audiogram: original papers. *Int. J. Audiol.* 29, 1–7. doi: 10.3109/00206099009081640
- Mithen, S., Morley, I., Wray, A., Tallerman, M., and Gamble, C. (2006). The singing neanderthals: the origins of music, language, mind and body. *Cambridge Archaeol. J.* 16, 97–112. doi: 10.1017/S0959774306000060
- Musacchia, G., Ortiz-Mantilla, S., Choudhury, N., Realpe-Bonilla, T., Roesler, C., and Benasich, A. A. (2017). Active auditory experience in infancy promotes brain plasticity in theta and gamma oscillations. *Dev. Cogn. Neurosci.* 26, 9–19. doi: 10.1016/j.dcn.2017.04.004
- Musacchia, G., Sams, M., Skoe, E., and Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proc. Natl. Acad. Sci. U. S. A.* 104, 15894–15898. doi: 10.1073/pnas.0701498104
- Musacchia, G., Strait, D., and Kraus, N. (2008). Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. *Hear. Res.* 241, 34–42. doi: 10.1016/j.heares.2008.04.013
- Pantev, C., Ross, B., Fujioka, T., Trainor, L. J., Schulte, M., and Schulz, M. (2003). Music and learning-induced cortical plasticity. *Ann. N. Y. Acad. Sci.* 999, 438–450. doi: 10.1196/annals.1284.054
- Papoušek, H. (1996). “Musicality in infancy research: biological and cultural origins of early musicality,” in *Musical Beginnings*, eds I. Deliège, and J. Sloboda (Oxford: Oxford University Press), 37–55. doi: 10.1093/acprof:oso/9780198523321.003.0002
- Parbery-Clark, A., Anderson, S., Hittner, E., and Kraus, N. (2012). Musical experience offsets age-related delays in neural timing. *Neurobiol. Aging* 33, 1–4. doi: 10.1016/j.neurobiolaging.2011.12.015
- Parbery-Clark, A., Skoe, E., Lam, C., and Kraus, N. (2009). Musician enhancement for speech-in-noise. *Ear Hear.* 30, 653–661. doi: 10.1097/AUD.0b013e3181b412e9
- Parbery-Clark, A., Strait, D. L., Anderson, S., Hittner, E., and Kraus, N. (2011). Musical experience and the aging auditory system: implications for cognitive abilities and hearing speech in noise. *PLoS One* 6:e18082. doi: 10.1371/journal.pone.0018082
- Patel, A. D. (2014). Can nonlinguistic musical training change the way the brain processes speech? the expanded OPERA hypothesis. *Hear. Res.* 308, 98–108. doi: 10.1016/j.heares.2013.08.011
- Pfordresher, P. Q., and Brown, S. (2007). Poor-pitch singing in the absence of “tone deafness.” *Music Percep.* 25, 95–115. doi: 10.1525/mp.2007.25.2.95
- Pfordresher, P. Q., and Dalla Bella, S. (2011). Delayed auditory feedback and movement. *J. Chem. Inform. Model.* 53:160. doi: 10.1017/CBO9781107415324.004
- Pichora-Fuller, M. K., and Schneider, B. (1992). The effect of interaural delay of the masker on masking-level differences in young and old adults. *J. Acoust. Soc. Am.* 91(4 Pt 1), 2129–2135. doi: 10.1121/1.403673
- Pichora-Fuller, M. K., Schneider, B., MacDonald, E., Pass, H. E., and Brown, S. (2007). Temporal jitter disrupts speech intelligibility: a simulation of auditory aging. *Hear. Res.* 223, 114–121. doi: 10.1016/j.heares.2006.10.009
- Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *J. Acoust. Soc. Am.* 97, 593–608. doi: 10.1121/1.412282
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and R Core Team, (2019). *Nlme: Linear and Nonlinear Mixed Effects Models. R package Version 3.1-141*. Available at: <https://CRAN.R-project.org/package=nlme>
- Pruitt, T. A., and Pfordresher, P. Q. (2015). The role of auditory feedback in speech and song. *J. Exp. Psychol. Hum. Percept. Perform.* 41, 152–166. doi: 10.1037/a0038285
- Rammesayer, T., and Altenmüller, E. (2006). Temporal information processing in musicians and nonmusicians. *Music Percep.* 24, 37–48. doi: 10.1525/mp.2006.24.1.37
- Raz, N., Millman, D., and Moberg, P. J. (1989). Auditory memory and age-related differences in two-tone frequency discrimination: trace decay and interference. *Exp. Aging Res.* 15, 43–47. doi: 10.1080/03610738908259757
- Ricketts, T. A., and Hornsby, B. W. Y. (2005). Sound quality measures for speech in noise through a commercial hearing aid implementing digital noise reduction. *J. Am. Acad. Audiol.* 16, 270–277. doi: 10.3766/jaaa.16.5.2
- Ridderinkhof, K. R., van der Molen, M. W., Band, G. P., and Bashore, T. R. (1997). Sources of interference from irrelevant information: a developmental study. *J. Exp. Child Psychol.* 65, 315–341. doi: 10.1006/jecp.1997.2367
- Ruggles, D. R., Freyman, R. L., and Oxenham, A. J. (2014). Influence of musical training on understanding voiced and whispered speech in noise. *PLoS One* 9:e86980. doi: 10.1371/journal.pone.0086980
- Russo, F. A., Ives, D. T., Goy, H., Pichora-Fuller, M. K., and Patterson, R. D. (2012). Age-related difference in melodic pitch perception is probably mediated by temporal processing: empirical and computational evidence. *Ear Hear.* 33, 177–186. doi: 10.1097/AUD.0b013e318233acee
- Russo, F. A., and Pichora-Fuller, M. K. (2008). Tune in or tune out: age-related differences in listening to speech in music. *Ear Hear.* 29, 746–760. doi: 10.1097/AUD.0b013e31817bdd1f
- Russo, N. M., Nicol, T. G., Zecker, S. G., Hayes, E. A., and Kraus, N. (2005). Auditory training improves neural timing in the human brainstem. *Behav. Brain Res.* 156, 95–103. doi: 10.1016/j.bbr.2004.05.012
- Sabes, J. H., and Sweetow, R. W. (2007). Variables predicting outcomes on listening and communication enhancement (LACETM) training. *Int. J. Audiol.* 46, 374–383. doi: 10.1080/14992020701297565
- Salomon, G. (1986). Hearing problems and the elderly. *Danish Med. Bull.* 33(Suppl. 3), 1–22.
- Särkämö, T., Tervaniemi, M., Laitinen, S., Numminen, A., Kurki, M., Johnson, J. K., et al. (2014). Cognitive, emotional, and social benefits of regular musical activities in early dementia: randomized controlled study. *Gerontologist* 54, 634–650. doi: 10.1093/geront/gnt100
- Saunders, G. H., Smith, S. L., Chisolm, T. H., Frederick, M. T., McArdle, R. A., and Wilson, R. H. (2016). A randomized control trial: supplementing hearing aid use with listening and communication enhancement (LACE) auditory training. *Ear Hear.* 37, 381–396. doi: 10.1097/AUD.0000000000000283
- Schellenberg, E. G. (2019). Music training, music aptitude, and speech perception. *Proc. Natl. Acad. Sci.* 116, 2783–2784. doi: 10.1073/pnas.1821109116
- Schellenberg, E. G., and Moreno, S. (2009). Music lessons, pitch processing, and g. *Psychol. Music* 38, 209–221. doi: 10.1097/MAO.0000000000000213
- Schlaug, G. (2015). Musicians and music making as a model for the study of brain plasticity. *Prog. Brain Res.* 217, 37–55. doi: 10.1016/bs.pbr.2014.11.020
- Schlaug, G., Jäncke, L., Huang, Y. Y., and Steinmetz, H. (1995). *In vivo* evidence of structural brain asymmetry in musicians. *Science* 267, 699–701. doi: 10.1126/science.7839149
- Schneider, B. (1997). Psychoacoustics and aging: implications for everyday listening. *J. Speech Lang. Pathol. Audiol.* 21, 111–124.
- Schneider, B., Pichora-Fuller, K., and Daneman, M. (2010). Effects of senescent changes in audition and cognition on spoken language comprehension. *Springer Handb. Auditory Res.* 34, 167–210. doi: 10.1007/978-1-4419-0993-0
- Schneider, B., Pichora-Fuller, M. K., Kowalchuk, D., and Lamb, M. (1994). Gap detection and the precedence effect in young and old adults. *J. Acoust. Soc. Am.* 95, 980–991. doi: 10.1121/1.408403
- Schneider, B. A., Daneman, M., and Pichora-Fuller, M. K. (2002). Listening in aging adults: from discourse comprehension to psychoacoustics. *Can. J. Exp. Psychol.* 56, 139–152. doi: 10.1037/h0087392
- Schneider, P., Scherg, M., Dosch, H. G., Specht, H. J., Gutschalk, A., and Rupp, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nat. Neurosci.* 5, 688–694. doi: 10.1038/nn871
- Schulze, K., Zysset, S., Mueller, K., Friederici, A. D., and Koelsch, S. (2011). Neuroarchitecture of verbal and tonal working memory in nonmusicians and musicians. *Hum. Brain Mapp.* 32, 771–783. doi: 10.1002/hbm.21060
- Shahin, A., Bosnyak, D. J., Trainor, L. J., and Roberts, L. E. (2003). Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *J. Neurosci.* 23, 5545–5552. doi: 10.1523/jneurosci.23-13-05545.2003
- Shepard, K. N., Kilgard, M. P., and Liu, R. C. (2013). “Experience-dependent plasticity and auditory cortex,” in *Neural Correlates of Auditory Cognition*, eds

- Y. E. Cohen, A. N. Popper and R. R. Fay (Berlin: Springer), 293–327 doi: 10.1007/978-1-4614-2350-8_10
- Skoe, E., and Kraus, N. (2010). Auditory brainstem response to complex sounds: a tutorial. *Ear Hear.* 31, 302–324. doi: 10.1097/AUD.0b013e3181c8db272. Auditory
- Skoe, E., and Kraus, N. (2013). Musical training heightens auditory brainstem function during sensitive periods in development. *Front. Psychol.* 4:622. doi: 10.3389/fpsyg.2013.00622
- Slater, J., Azem, A., Nicol, T., Swedenborg, B., and Kraus, N. (2017). Variations on the theme of musical expertise: cognitive and sensory processing in percussionists, vocalists and non-musicians. *Eur. J. Neurosci.* 45, 952–963. doi: 10.1111/ejn.13535
- Slater, J., Skoe, E., Strait, D. L., O'Connell, S., Thompson, E., and Kraus, N. (2015). Music training improves speech-in-noise perception: longitudinal evidence from a community-based music program. *Behav. Brain Res.* 291, 244–252. doi: 10.1016/j.bbr.2015.05.026
- Slevc, L. R., Davey, N. S., Buschkuhl, M., and Jaeggi, S. M. (2016). Tuning the mind: exploring the connections between musical ability and executive functions. *Cognition* 152, 199–211. doi: 10.1016/j.cognition.2016.03.017
- Smith, J. C., Marsh, J. T., and Brown, W. S. (1975). Far-field recorded frequency-following responses: evidence for the locus of brainstem sources. *Electroencephalogr. Clin. Neurophysiol.* 39, 465–472. doi: 10.1016/0013-4694(75)90047-4
- Solé Resano, C., Mercadal-Brotons, M., Gallego Matas, S., and Riera, M. (2010). Contribution of music to aging adults' quality of life. *J. Music Ther.* 47:264. doi: 10.1093/jmt/47.3.264
- Song, J. H., Skoe, E., Banai, K., and Kraus, N. (2012). Training to improve hearing speech in noise: biological mechanisms. *Cereb. Cortex* 22, 1180–1190. doi: 10.1093/cercor/bhr196
- Souza, P. E., Boike, K. T., Witherell, K., and Tremblay, K. (2007). Prediction of speech recognition from audibility in older listeners with hearing loss: effects of age, amplification, and background noise. *J. Am. Acad. Audiol.* 18, 54–65. doi: 10.3766/jaaa.18.1.5
- Strait, D. L., Kraus, N., Skoe, E., and Ashley, R. (2009). Musical experience and neural efficiency - Effects of training on subcortical processing of vocal expressions of emotion. *Eur. J. Neurosci.* 29, 661–668. doi: 10.1111/j.1460-9568.2009.06617.x
- Swaminathan, J., Mason, C. R., Streeter, T. M., Best, V., Kidd, G., and Patel, A. D. (2015). Musical training, individual differences and the cocktail party problem. *Sci. Rep.* 5:11628. doi: 10.1038/srep11628
- Sweetow, R. W., and Sabes, J. H. (2006). The need for and development of an adaptive listening and communication enhancement (LACETM) program. *J. Am. Acad. Audiol.* 17, 538–558. doi: 10.3766/jaaa.17.8.2
- Sweetow, R. W., and Sabes, J. H. (2010). Auditory training and challenges associated with participation and compliance. *J. Am. Acad. Audiol.* 21, 586–593. doi: 10.3766/jaaa.21.9.4
- Syka, J. (2002). Plastic changes in the central auditory system after hearing loss, restoration of function, and during learning. *Physiol. Rev.* 82, 601–636. doi: 10.1152/physrev.00002.2002
- Tierney, A. T., Krizman, J., and Kraus, N. (2015). Music training alters the course of adolescent auditory development. *Proc. Natl. Acad. Sci. U.S.A.* 112, 10062–10067. doi: 10.1073/pnas.1505114112
- Tremblay, K. L., Piskosz, M., and Souza, P. (2003). Effects of age and age-related hearing loss on the neural representation of speech cues. *Clin. Neurophysiol.* 114, 1332–1343. doi: 10.1016/S1388-2457(03)00114-7
- Tye-Murray, N., Sommers, M. S., Mauzé, E., Schroy, C., Barcroft, J., and Spehar, B. (2012). Using patient perceptions of relative benefit and enjoyment to assess auditory training. *J. Am. Acad. Audiol.* 23, 623–634. doi: 10.3766/jaaa.23.8.7
- Vermiglio, A. J., Soli, S. D., Freed, D. J., and Fisher, L. M. (2012). The relationship between high-frequency pure-tone hearing loss, hearing in noise test (HINT) thresholds, and the articulation index. *J. Am. Acad. Audiol.* 23, 779–788. doi: 10.3766/jaaa.23.10.4
- Walton, J. P., Frisina, R. D., and O'Neill, W. E. (1998). Age-related alteration in processing of temporal sound features in the auditory midbrain of the CBA mouse. *J. Neurosci.* 18, 2764–2776. doi: 10.1523/jneurosci.18-07-02764.1998
- West, G. L., Zendel, B. R., Konishi, K., Benady-Chorney, J., Bohbot, V. D., Peretz, I., et al. (2017). Playing super mario 64 increases hippocampal grey matter in older adults. *PLoS One* 12:e0187779. doi: 10.1371/journal.pone.0187779
- Whitehead, J. C., and Armony, J. L. (2018). Singing in the brain: neural representation of music and voice as revealed by fMRI. *Hum. Brain Mapp.* 39, 4913–4924. doi: 10.1002/hbm.24333
- White-Schwoch, T., Carr, K. W., Anderson, S., Strait, D. L., and Kraus, N. (2013). Older adults benefit from music training early in life: biological evidence for long-term training-driven plasticity. *J. Neurosci.* 33, 17667–17674. doi: 10.1523/JNEUROSCI.2560-13.2013
- Wilson, R. H., McArdle, R. A., and Smith, S. L. (2007). An evaluation of the BKB-SIN, HINT, QuickSIN, and WIN materials on listeners with normal hearing and listeners with hearing loss. *J. Speech Lang. Hear. Res.* 50:844. doi: 10.1044/1092-4388(2007)059
- Wingfield, A., Tun, P. A., and McCoy, S. L. (2005). Hearing loss in older adulthood: what it is and how it interacts with cognitive performance. *Curr. Direct. Psychol. Sci.* 14, 144–148. doi: 10.1111/j.0963-7214.2005.00356.x
- Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., and Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat. Neurosci.* 10, 420–422. doi: 10.1038/nn1872
- Yamasoba, T., Lin, F. R., Someya, S., Kashio, A., Sakamoto, T., and Kondo, K. (2013). Current concepts in age-related hearing loss: epidemiology and mechanistic pathways. *Hear. Res.* 303, 30–38. doi: 10.1016/j.heares.2013.01.021
- Yates, K., Moore, D., Amitay, S., and Barry, J. (2018). Sensitivity to melody, rhythm, and beat in supporting speech-in-noise perception in young adults. *Ear Hear.* 40, 358–367. doi: 10.1097/AUD.0000000000000621
- Yueh, B., Shapiro, N., MacLean, C. H., and Shekelle, P. G. (2003). Screening and management of adult hearing loss in primary care: scientific review. *JAMA* 289, 1976–1985. doi: 10.1001/jama.289.15.1976
- Zatorre, R. J., Chen, J. L., and Penhune, V. B. (2007). When the brain plays music: auditory-motor interactions in music perception and production. *Nat. Rev. Neurosci.* 8, 547–558. doi: 10.1038/nrn2152
- Zendel, B. R., and Alain, C. (2009). Concurrent sound segregation is enhanced in musicians. *J. Cogn. Neurosci.* 21, 1488–1498. doi: 10.1162/jocn.2009.21140
- Zendel, B. R., and Alain, C. (2012). Musicians experience less age-related decline in central auditory processing. *Psychol. Aging* 27, 410–417. doi: 10.1037/a0024816
- Zendel, B. R., Tremblay, C.-D., Belleville, S., and Peretz, I. (2015). The impact of musicianship on the cortical mechanisms related to separating speech from background noise. *J. Cogn. Neurosci.* 27, 1044–1059. doi: 10.1162/jocn-a-00758
- Zendel, B. R., West, G., Belleville, S., and Peretz, I. (2019). Music training improves the ability to understand speech-in-noise in older adults. *Neurobiol. Aging* 81, 102–115. doi: 10.1016/j.neurobiolaging.2019.05.015
- Zheng, Z. Z., Munhall, K. G., and Johnsrude, I. S. (2010). Functional overlap between regions involved in speech perception and in monitoring one's own voice during speech production. *Otolaryngology* 22, 1770–1781. doi: 10.1162/jocn.2009.21324.Functional

Conflict of 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.

Copyright © 2019 Dubinsky, Wood, Nespola and Russo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The Effects of Bimanual Coordination in Music Interventions on Executive Functions in Aging Adults

Jennifer A. Bugos*

School of Music, Center for Music Education Research, University of South Florida, Tampa, FL, United States

OPEN ACCESS

Edited by:

Assal Habibi,
University of Southern California,
United States

Reviewed by:

Eckart Altenmüller,
Hanover University of Music Drama
and Media, Germany
Nils Henrik Pixa,
Chemnitz University of Technology,
Germany

*Correspondence:

Jennifer A. Bugos
bugosj@usf.edu

Received: 05 July 2019

Accepted: 05 November 2019

Published: 05 December 2019

Citation:

Bugos JA (2019) The Effects
of Bimanual Coordination in Music
Interventions on Executive Functions
in Aging Adults.
Front. Integr. Neurosci. 13:68.
doi: 10.3389/fnint.2019.00068

Music training programs have been shown to enhance executive functions in aging adults; however, little is known regarding the extent to which different types of bimanual coordination (i.e., fine and gross motor) in music instruction contribute to these outcomes. The aim of this study was to examine the effects of bimanual coordination in music interventions on cognitive performance in healthy older adults (60–80 years). Participants ($N = 135$) completed motor measures and battery of standardized cognitive measures, before and after a 16-week music training program with a 3 h practice requirement. All participants were matched by age, education, and estimate of intelligence to one of three training programs: piano training (fine motor); percussion instruction (gross motor), and music listening instruction (MLI) (no motor control condition). Results of a Repeated Measures ANOVA revealed significant enhancements in bimanual synchronization and visual scanning/working memory abilities for fine and gross motor training groups as compared to MLI. Pairwise comparisons revealed that piano training significantly improved motor synchronization skills as compared to percussion instruction or music listening. Results suggest that active music performance may benefit working memory, the extent of these benefits may depend upon coordination demands.

Keywords: music training, executive functions, piano training, cognitive intervention, older adult

INTRODUCTION

Bimanual coordination is a common element of instrumental music performance that requires collaboration between the hands. Often bimanual coordination is measured based upon in-phase, antiphase, or out- of phase movements calculated by phase angles of each limb at discrete time-points. However, in music performance, bimanual coordination can also be accounted for in hand synchronization, pattern accuracy, and rhythmic accuracy while performing hand movements in parallel, contrasting, and oblique motion. This research employs this musical definition of bimanual coordination with respect to pattern, synchronization, and rhythmic accuracy and examines its relationship to music training and cognitive performance.

Bimanually coordinated movements can be varied by level of task complexity, task difficulty, absence or presence of feedback, and level/amount of training or task practice (Maes et al., 2017). According to information processing theory, bimanual coordination is an example of a dual-task situation that includes interference between simultaneously performed tasks, placing demands upon cognitive and neural resources (Swinnen and Gooijers, 2015). Practicing a musical instrument can offer a level of complexity for novice instrumentalists using a dual-task situation of bimanual motor movements

coupled with auditory feedback which may exercise areas of cognition. Our hypothesis is that bimanual coordination tasks such as performance on a musical instrument, may contribute to improving and maintaining cognitive skills in aging adults.

The purpose of this study was to examine the effects of bimanual coordination in musical tasks with graded dual-task load on cognitive and motor outcomes in healthy older adults (60–80 years) as compared to a non-motor music training intervention, music listening instruction (MLI). Specifically, we evaluated the effects of music training with fine motor (i.e., group piano training), gross motor (i.e., group percussion ensemble), and no motor (i.e., music listening instruction) requirements on bimanual coordination skills and executive functions. We also examined the relationship between bimanual coordination performance as measured by pattern, synchronization, and rhythmic accuracy, and performance on measures of executive functions. Based upon the Supply and Demand framework in aging adults, fine motor training may contribute to corticocerebellar networks (Seidler et al., 2010). Therefore, we hypothesized that those allocated to group piano training who practice fine motor skills in piano training may exhibit greater enhancements in executive functions than the other two interventions.

Musicians devote many practice hours to the development of complex bimanual sensorimotor skills which vary based upon instrumentation (Krampe and Ericsson, 1996; Krampe, 2002; Kim et al., 2017). Gross motor skills refer to movements that require coordination of arms, legs, other large muscle groups. For example, percussionists with hand drums or mallets utilize gross motor skills. While fine motor skills such as those used by wind players and pianists, require the coordination of small muscles such as the digits of the hand. While all musical instruments require precise timed movements, differentiation in task requirements by instrument and types of motor coordination may impact transfer to motor skills and the degree of transfer to cognitive performance. For instance, Kim et al. (2017) showed correlations between synchronization errors committed on an electric drum task by older adults and errors committed on cognitive tasks of executive functions. This research explored relationships between motor performance and executive functions in the context of a training study.

Expert Musicians and Motor Control

Few studies examine the emergence of skilled bimanual coordination in novice musicians, however, studies of professional musicians can provide insight into the neural mechanisms implicated through intense musical training. In professional musicians with extensive practice, neuroplastic changes have been found that correspond to the trained instrument. For instance, in professional string players, larger cortical representations for fingers of the left hand were found as compared to non-musicians (Elbert et al., 1995). Expert musicians were shown to have a larger anterior corpus callosum as compared to non-musicians, suggesting more efficient interhemispheric communication (Schlaug et al., 1995).

Many examples of experience-dependent plasticity have been found in musicians who engage in fine motor control.

For example, wind players were found to have showed enlarged cortical thickness in lip-and tongue related brain areas when compared to non-musicians (Choi et al., 2015). Similar to these findings, research in professional pianists showed differential cortical activation patterns (Koenke et al., 2004), structural and functional activity in auditory and motor areas (Krings et al., 2000), and changes to white matter integrity and gray matter density (Han et al., 2009). If neurological evidence indicates neural changes resulting from instrumental training (Palomar-García et al., 2017), what are the functional differences in cognitive outcomes from engaging in different types of bimanual coordination in adults?

Bimanual Coordination and Cognition in Aging

Age-related cognitive decline is associated with deficits in executive functions and motor coordination (Seidler et al., 2010; Fujiyama et al., 2013; Bernard and Seidler, 2014; Rueda-Delgado et al., 2019). Motor deficits may be due to the gradual degeneration of the neuromuscular system, contributing to sensory and motor challenges. However, older adults can learn new motor skills despite the availability of fewer cognitive resources (Seidler, 2007).

Research suggests that older adults recruit additional neural resources when compared with younger adults (Heuninckx et al., 2005; Rueda-Delgado et al., 2019). For instance, neuroimaging research has shown that older adults completing coordination tasks, demonstrate increased activation patterns in the posterior cerebellum, an area associated with coordinated movements, when compared to younger adults (Heuninckx et al., 2005). Additional research by Rueda-Delgado et al. (2019) found functional reorganization after training in older adults as compared to younger adults who completed a novel bimanual gross motor task. Specifically, beta power was reduced after task training in older adults suggesting higher neural activity.

The “Supply and Demand” Framework (Seidler et al., 2010) provides a potential explanation for age-related changes in the neural control of movement. This framework accounts for structural and functional declines (demand), shown to increase with age, in the motor cortices, cerebellum, and basal ganglia pathways. Thus, older adults rely more heavily upon cognitive resources for motor control. In addition, aging adults demonstrate reductions in attentional capacity and cognitive resources (supply) due to the degradation of the prefrontal cortex and anterior corpus callosum. The framework suggests that the dopaminergic system acts on the corticocerebellar neural pathways projecting from the cerebellum to the frontal cortices, areas associated with higher cognitive processes. By engaging neural pathways associated with cerebellar activity in fine motor tasks such as piano training, the corticocerebellar pathway may be strengthened. Activations in the corticocerebellar pathway can influence neural architecture such as the dorsolateral prefrontal cortex, an area associated with executive functions (Watson et al., 2014). While Seidler et al. (2010) have suggested that training bimanual coordination skills may reduce the

potential for cognitive decline, few cognitive training programs consider bimanual motor coordination in the context of a music intervention.

Some interventions for aging adults include bimanual coordination tasks such as multisensory exercise programs (Moreira et al., 2018); and juggling interventions (Draganski et al., 2004; Scholz et al., 2009). Results of a randomized controlled trial showed benefits from multisensory exercise training with bimanual hand movements to cognition and motor outcomes in institutionalized older adults (Moreira et al., 2018). In addition, neuroimaging data in adults who participated in a 90-day juggling intervention showed increased gray matter volume in the mid-temporal area bilaterally and in the left posterior intraparietal sulcus (Draganski et al., 2004). Boyke et al. (2008) replicated the juggling intervention in a sample of older adults in which similar neuroplastic changes were found post-training. A significant increase in gray matter in the hippocampus and nucleus accumbens was found for older adult jugglers. Training on a musical instrument includes bimanual motor control in a temporal context, similar to complex motor activities such as juggling.

Piano Training and Cognition

Piano performance requires complex fine motor control and integrates auditory feedback in a temporal context. Research studies have found that piano training enhances several executive functions such as working memory (Bugos et al., 2007; Guo et al., 2018), spatial reasoning (Rauscher and Zupan, 2000), verbal fluency (Bugos and Kochar, 2017), and cognitive control (Seinfeld et al., 2013) in adults and children. Preschool children who received piano training programs were shown to increase spatial-temporal reasoning (Rauscher and Zupan, 2000). Results of a 6-week keyboard harmonica intervention improved working memory in young children as compared to no treatment controls (Guo et al., 2018). In addition, piano training was shown to significantly increase auditory word discrimination skills in young children as compared to a reading training program and no treatment controls (Nan et al., 2018). Collectively, studies in young children showed the impact of piano training on working memory, verbal skills, and reasoning abilities as compared to control conditions.

Additional research has been conducted in aging adults on the effects of piano training on cognitive performance in novice adult musicians. Bugos et al. (2007) found that older adults who were randomly assigned to a 6-month individualized piano training program outperformed those assigned to a no treatment control group on a series of standardized measures of executive functions and working memory. Data suggested that some areas of executive functions such as perceptual speed and visual-scanning, were maintained after a 3-month delay period.

In another research study, Seinfeld et al. (2013) examined the impact of a four-month piano intervention on cognitive control as measured by the Spanish version of the Stroop task, visual spatial performance by the Trail Making Card A, and well-being as measured by the *Beck Depression Inventory*, *Profile of Mood States*, and *World Health Questionnaire Quality of Life Brief Questionnaire*. While adults in Seinfeld et al. (2013) were

not randomly assigned, post-training data showed enhanced performance on measures of cognition, particularly in response inhibition for those engaged in group piano training compared to a group of older adults engaged in recreational activities. There is a need for additional experimental research on the impact of piano training on cognitive outcomes in older adults and to compare and contrast with outcomes in other percussion based interventions.

Gross Motor Coordination in Percussion Programs

Research in the music therapy literature showed that gross motor performance on non-pitched percussion instruments (e.g., drums, congas, or djembes) in drum circles can assist with health-related outcomes such as stress and anxiety. For example, patients with dementia who received a drumming program for 6 weeks (twice weekly) demonstrated significantly reduced anxiety compared to no treatment controls (Sung et al., 2012). Other variables affecting health such as blood pressure have been shown to change after drumming. Researchers in Africa found that older adults with high blood pressure who partook of three, 40 min djembe drumming sessions demonstrated reduced systolic blood pressure post-training (Smith et al., 2014).

Research has also found that drumming can impact psychosocial and cognitive outcomes. For example, patients with Parkinson's disease who received a 6-week West African drumming class (twice per week), self-reported higher quality of life than controls (Pantelyat et al., 2016). In another study, nine participants with Huntington's disease demonstrated increased performance on measures of executive functions after 2 months of a drumming intervention (Metzler-Baddeley et al., 2014). Neurological data further suggested that the training may have contributed to changes in the genu of the corpus callosum. While there are only a few studies that actively implement percussion performance in the music therapy literature, we seek to evaluate a different kind of percussion training program, mallet training in healthy older adults.

Few studies have researched the impact of percussion training on cognitive outcomes in healthy older adults. Dege and Kerkovius (2018) showed 15 weeks of drumming (60 min, weekly) enhanced working memory in aging females compared to a literature group and no treatment group. Another study found that beginning adult musicians who completed an 8 week mallet training program demonstrated no significant differences in cognitive outcomes, though a non-significant trend was noted for processing speed. Researchers found increased self-efficacy compared to an autobiographical writing group (Bugos and Cooper, 2019). While both studies acknowledged limitations in sample size, these studies included a progressively difficult curriculum and a social learning environment, factors positively associated with musical cognitive training programs (Bugos, 2014). Therefore, this research included training groups with a social learning environment (i.e., group-based format) and a progressively difficulty curriculum (see **Supplementary Material**).

The impact of fine motor and gross motor music training on cognitive and motor performance in aging adults remains unclear. The purpose of this study was to investigate the effects of bimanual coordination in fine motor training (Group Piano Training; GPI), gross motor training (Group Percussion Ensemble; GPeI), and a no motor control regimen (Music Listening Instruction; MLI) on cognitive performance and bimanual coordination in healthy adults. It was hypothesized that older adults assigned to instrumental interventions, group piano training, and group percussion instruction (GPeI), would outperform matched controls enrolled in MLI on measures of executive functions. Since fine motor activities such as piano training have been more closely linked to corticerebellar pathways associated with bimanual coordination (Herrojo Ruiz et al., 2017; Jirenghed et al., 2017), it was hypothesized that those in group piano instruction would demonstrate increased bimanual coordination as compared to GPeI. However, we predicted that those in GPeI would demonstrate increased rhythmic coordination.

MATERIALS AND METHODS

Participants

One hundred eighty non-musicians (60–80 years) were recruited from a mid-size city in the Southeastern United States and a series of surrounding rural communities. Recruitment included flyers and presentations to local Councils on Aging, church groups, and media coverage of the research. Criteria for participation consisted of those between the ages of 60–80, a native English speaker, no history of colorblindness, no prior history of neurological impairment (e.g., dementia or stroke), no difficulty with the movement of the hands or persistent tremor, not currently taking any psycho-reactive medications or those affecting memory performance (e.g., sleep medications, antidepressants), less than 3 years of prior musical training, and not currently engaged in music reading or musical performance. Participants were screened for cognitive impairment (scores ≥ 30) with the Telephone Interview for Cognitive Status (TICS, Brandt et al., 1988). The TICS is a short reliable screening for cognitive impairment. Those scores, ≥ 30 , suggest no cognitive impairment and correspond to those found on the Mini-Mental State Exam (Fong et al., 2009). Informed written consent was obtained from all participants in accordance with the procedures of the University Institutional Review Board.

One hundred thirty-five participants completed the research study (Table 1). While each music training intervention began with 60 adults, the final sample included: group piano instruction ($n = 49$), MLI ($n = 48$), or GPeI ($n = 38$). Attrition was due to lack of assignment to preferred group, personal illness, family illness, caregiver responsibilities, or a financial need to return to the workplace. The attrition rate was 24%, similar to other research employing interventions for older adults (Jancey et al., 2007).

Procedure

All participants completed baseline measures of music aptitude and estimate of intelligence. A series of dependent measures

TABLE 1 | Demographic data.

	GPI ($N = 49$)	MLI ($N = 48$)	GPeI ($N = 38$)
Age	67.90 (6.42)	68.83 (7.38)	69.13 (7.27)
Gender (Male/Female)	14/35	12/36	12/26
Education in years	14.86 (2.84)	14.94 (2.75)	15.87 (3.23)
VIQ	105.65 (10.36)	105.29 (13.35)	107.45 (19.79)
PIQ	109.9 (17.66)	108.5 (18.55)	110.47 (20.74)
FSIQ	107.8 (12.91)	107.25 (13.09)	111.5 (18.68)
AMMA	49.88 (7.35)	49.38 (6.64)	48.61 (6.95)

VIQ, Verbal Intelligence Quotient; PIQ, Performance Intelligence Quotient; FSIQ, Estimate of Full Scale Intelligence Quotient scores; AMMA, Advanced Measures of Music Audiation.

of executive functions (i.e., visual scanning/working memory, processing speed, verbal fluency, and cognitive control) and motor measures (i.e., finger dexterity and bimanual coordination) were repeated at post-testing. Alternate forms of measures (e.g., Form A and B for the Verbal Fluency subtest) were used whenever possible to avoid the potential for practice effects. All measures were administered by two highly trained research assistants with experience in psychological testing for aging adults. Both research assistants were blind to the participants' assigned condition. Participants were matched by age, education, and intelligence to three separate training interventions; however, placement of three individuals who matched these criteria were randomly assigned to one of three interventions. Sixty participants were assigned to each of three training interventions. Fifteen participants were allocated to each class session within each training intervention (i.e., music listening, group piano training, and group percussion ensemble). Trainers remained the same for the study duration and each conducted four separate classes for their prescribed intervention. All participants received 16 weeks of training sessions, which met for 45 min weekly.

Since success in music programs depends upon access to practice and consolidation of learning, all participants were required to practice training exercises for 30 min per day or 3 h per week. All practice homework was checked and practice times logged. The format of the training programs were structured similarly with 10 min of review (e.g., homework or practice log checks and review of concepts), 25 min of group performance of new skills/concepts/repertoire, and 10 min for self-practice and assessment. Training fidelity checks were conducted bi-weekly in each intervention by two separate research assistants. These checks insured that time was allocated equivalently between interventions and that the trainer adhered to the research protocols which included the training curriculum. A separate research assistant was solely responsible for collecting homework and checking practice logs for completion.

Description of Programs

Group Piano Instruction (GPI) consisted of a basic piano course which included technical exercises (e.g., major scales), finger dexterity exercises (e.g., Hanon exercises), basic piano repertoire selected from the *Alfred All-in-One* text (Palmer et al., 1995).

Classes were taught in a Yamaha Clavinova electronic piano lab by a certified music educator with 20 + years of piano teaching experience. Participants either borrowed keyboards for practice from the university or completed practice sessions at a local church. Weekly practice assignments were provided that consisted of finger exercises, written theory assignments, and two practice pieces.

Group Percussion Instruction (GPel) consisted of a basic mallet program which included technical exercises (e.g., major scales), music theory, ostinati exercises (e.g., *Music For Children: Volume 1*, Murray, 1976) and repertoire performed was chosen from *Get America Singing Again vol. 2* (Hal Leonard Publishing Corporation, 1997) and *Hot Marimba* (Hampton, 1995). Each participant borrowed a soprano or alto Peripole-Bergerault Orff rosewood xylophone that was brought to each class held at the local senior center. Classes were taught by the same music educator. Weekly homework assignments were provided that consisted of mallet exercises (e.g., scales, ostinati, and patterning), written music theory assignments, and two percussion ensemble pieces.

Music Listening Instruction (MLI) consisted of a basic music appreciation course with information about genre, composers, and forms in classical and world music based upon the *Music Listening Today* text (Hoffer, 2005). All music listening classes were held at a local senior center and were taught by the same music educator. Participants were loaned copies of the texts and accompanying enhanced CDs for the duration of the program. All enhanced CDs, not only provided audio practice, but CDs could also be inserted into computers (either at home or at the senior center) to view a listening map containing illustrated formal elements of the pieces unfolding over time. Weekly homework assignments included listening assignments and written questions that required written responses.

Description of Measures

Baseline Measures

The *Wechsler Abbreviated Scale of Intelligence* (WASI; Wechsler, 1999) was used to generate a Verbal Intelligence Quotient (VIQ), Performance Intelligence Quotient (IQ), and an estimate of Full-Scale Intelligence Quotient (FSIQ) (Table 1). The WASI consists of four subtests: *Vocabulary*, *Block Design*, *Similarities*, and *Matrix Reasoning*. The *Vocabulary* subtest is an untimed measure in which participants orally define 42 words based upon visual stimuli. The *Similarities* subtest involves 26 sets of paired words in which participants respond as to how the two words are similar. The *Block Design* subtest measures visuospatial skills with 13 timed block patterns constructed from nine cubes. Items of the *Matrix Reasoning* subtest measure non-verbal reasoning skills with the selection of a missing piece to complete a visual pattern from five potential items. The *Vocabulary* and *Similarities* subtests contribute to an estimate of Verbal Intelligence Quotient (VIQ), while the *Block Design* and *Matrix Reasoning* subtests may be used to estimate one's Performance Intelligence Quotient (PIQ). The WASI scores correspond to intelligence scores from the longer Wechsler Adult Intelligence Scale III (Axelrod, 2002); and provide norms for a broad age range (6–89 years).

To measure musical aptitude, participants completed the *Advanced Measures of Music Audiation* (AMMA; Gordon, 1989). The AMMA consists of 30 paired-melodic phrases and requires participants to differentiate changes in piano melodies. Melodies are either tonally altered, rhythmically altered, or the same. The AMMA was chosen for its reliability ($r = 0.81$) and content validity.

Dependent Measures of Executive Functions

Visual Scanning and Working Memory

The Trails Test Card A and B (TMT; Reitan and Wolfson, 1985) was administered as a measure of visual scanning and working memory. The test requires sequencing, visuomotor speed, and mental flexibility. Card A includes drawing a line connecting numeric stimuli in sequential order, while Card B required the switching between numeric and alphabetical stimuli in sequential order (1, A, 2, B, etc.). TMT scores reflected the time to complete the card (in seconds) and the number of errors made on the task. In order to examine the cognitive domain without the motor element, a TMT delta score was calculated; time to complete Card A was subtracted from time to complete Card B. Construct validity for the TMT was established by correlations (0.36–0.93) with an object-finding task and a hidden pattern task (Ehrenstein et al., 1982). Reported reliability coefficients for the TMT were ($r = 0.94$) for Part A and ($r = 0.90$) for Part B (Spren and Strauss, 1998).

Processing Speed

The *Paced Auditory Serial Addition Task* (PASAT; Gronwall, 1977) was administered to evaluate processing speed with complex attention. PASAT includes four trials, each consisting of 25 items in which a string of single digit numbers is presented aurally with progressively faster interstimulus intervals. Respondents summed the last number with the new number and verbal responses were recorded. The PASAT was chosen for its reliability ($r = 0.95$; Nikraves et al., 2017) and sensitivity in aging adults (Tombaugh, 2006).

Verbal Fluency

The *Delis Kaplan Executive Function Measure Verbal Fluency* subtest (Delis et al., 2001) includes two forms to minimize practice effects and measures three types of fluency conditions: letter fluency, category fluency, and category switching. Each of the three letter fluency trials consisted of naming as many words that begin with the letters F, A, S (pretesting) and B, H, R (post-testing) in 60 s. Words generated could not be names of people, places, or numbers. Category fluency consisted of naming items pertaining to a specific category such as animals or girl's names in 60 s. Category switching trials required switching between the generation of words belonging to two different categories such as fruit and furniture. Reliability coefficients for the conditions of the Verbal Fluency subtest are letter fluency ($r = 0.88$), category fluency ($r = 0.82$), and category switching ($r = 0.51$) (Delis et al., 2001).

Response Inhibition

The Cued Color Word Stroop (Perlstein et al., 2006) was used to measure response inhibition in the visual domain with Color and Word cues. Each trial began with a 750 ms visual cue (Word or Color) followed by an 1000 ms delay, and a visual stimulus presented for up to 2500 ms, and terminated by the participant's response. Words are presented in one of three colors (red, blue, and green). Participants chose between the ink color and physical word depending upon instructions of the cue. For the color condition, participants responded to the ink color of the word. For the word condition, participants responded to the name of the physical word on the screen. A total of 360 trials were presented in 4 blocks of 90 trials in the following order: mixed, color, mixed, and word with pseudorandomized stimuli. Visual stimuli were presented in the center of a black screen visual display, and delivered on an Apple Macintosh MacBook Pro computer using an E-Prime Professional 2.0 software (Schneider et al., 2002) with a Psych Scope button box response apparatus. Participants were instructed to respond as quickly and accurately as possible. Prior to administration, participants successfully completed a practice block of each condition (color, word, or mixed) with 70% accuracy.

Dependent Motor Measures

Manual Dexterity

The Finger Tapper Test (FTT; Reitan and Wolfson, 1993) was used to evaluate manual dexterity by depressing a key with the index finger for ten second increments until either a score within a five-tap range is reached or ten trials are completed per hand. Research in healthy older adults suggests that the FTT is a promising measure for upper body motor performance, and performance on the FTT was not correlated with performance on standardized measures of working memory, processing speed, or spatial organization (Ashendorf et al., 2009).

Bimanual Coordination

The Bimanual Coordination Task (BMCT; Bugos and Iordache, 2019) is a gross motor task that evaluates parallel motion,

contrary motion, and oblique motion with 15 separate color-coded visual patterns, containing 8 beats in two rows (numbered boxes represent beats in a measure), one for the right hand, and one for the left hand (Figure 1). The top row represents the right hand pattern and the bottom row represents the left hand pattern. Boxes on the visual represented the beats in two sets of four beat patterns. Respondents are allocated 1 min to study each visual pattern and then perform the pattern on a color coded four octave Yamaha synthesizer with a metronome (90 = bpm). All patterns utilize black keys on the piano keyboard in sets of twos or threes. All sets of two or three black keys must be held down and selected at the same time. The first four practice items encompass responses with one hand and are followed by the inclusion of both hands. Thus, no previous knowledge of music reading or the piano keyboard is necessary to complete the task. The task was chosen for its reliability ($r = 0.88$) in a group of healthy older adults with no previous musical training (Bugos and Iordache, 2019). Reliability for sub scores of rhythm ($r = 0.82$), synchronization ($r = 0.76$) and pattern ($r = 0.78$) demonstrate good psychometric properties.

Other Variables

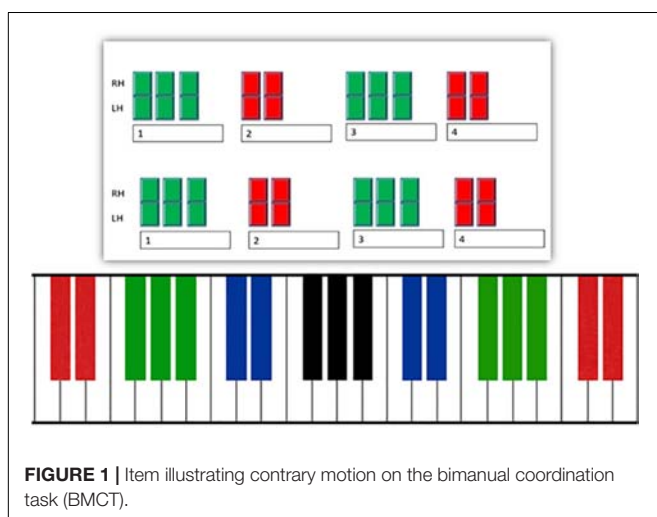
Mood was evaluated at baseline and at post-testing with the Beck Depression Inventory II (Beck et al., 1996) and the Geriatric Depression Scale (Yesavage, 1991), to ensure that mood had not changed significantly over the course of the intervention. While completion of the GDS has been shown to be less challenging for older adults, the sensitivity to change has been found to be more robust in the BDI. Thus, both were used as valuable instruments to evaluate changes in mood (Olin et al., 1992).

Statistical Analysis

Statistical analyses were carried out with IBM SPSS 24.0 (IBM Corp, 2016). Four factors were included into a Group \times Time ANOVA on executive functions: Trail Making Test Delta Scores (visual scanning/working memory), Paced Auditory Serial Addition Task mean number correct (processing speed), Category Switching mean correct (verbal fluency), and Stroop mean errors (response inhibition) with Bonferroni correction to control for Type I error. The analysis of motor outcomes included the mean number correct for the dominant and non-dominant hands for the Finger Tapper Test (manual dexterity), and each of three domains for the Bimanual Motor Coordination Task (motor coordination): Hand Synchronization, Pattern Accuracy, and Rhythmic Accuracy. Motor outcomes were entered into a separate Group \times Time ANOVA to evaluate differences across each domain of motor performance with Bonferroni correction.

RESULTS

Demographic data for participants ($N = 135$) can be found in Table 1. Results of an ANOVA across demographic variables showed no significant ($p < 0.05$) differences between age, $F(2, 132) = 0.381$, $p = 0.68$; education, $F(2, 132) = 1.52$, $p = 0.22$, estimate of full scale intelligence, $F(2, 132) = 0.99$, $p = 0.37$, or music aptitude, $F(2, 132) = 0.36$, $p = 0.70$.



Executive Functions

A Group (GPI, MLI, GPel) \times Time (Pretest, Posttest) Repeated Measures ANOVA was conducted across all measures of executive functions (Table 2): visual scanning/working memory (TMT Delta); processing speed (PASAT), verbal fluency (Category Switching-DKEFS), and response inhibition (Stroop). Results showed a main effect of time for processing speed, $F(1, 127) = 29.23$, $p < 0.001$, $\eta^2_p = 0.187$, $\omega^2_p = 0.179$, and for verbal fluency, $F(1, 127) = 12.04$, $p = 0.001$, $\eta^2_p = 0.086$, $\omega^2_p = 0.079$. Effect sizes for processing speed and verbal fluency were found to be moderate to large (Field, 2013). However, no main effect of time was found for response inhibition, $F(1, 127) = 1.66$, $p = 0.20$, $\eta^2_p = 0.013$, $\omega^2_p = 0.005$, or visual scanning/working memory, $F(1, 127) = 1.93$, $p = 0.17$, $\eta^2_p = 0.015$, $\omega^2_p = 0.007$.

Data showed a significant Group \times Time interaction for visual scanning and working memory as denoted by the TMT Delta scores, $F(2, 127) = 3.47$, $p = 0.03$, $\eta^2_p = 0.518$, $\omega^2_p = 0.037$ (Figure 2). No significant differences were found for verbal fluency in category switching, $F(2, 127) = 0.84$, $p = 0.21$, $\eta^2_p < 0.001$, $\omega^2_p < 0.001$, or for cognitive control as measured by errors committed on the Stroop task, $F(2, 127) = 0.84$, $p = 0.44$, $\eta^2_p < 0.001$, $\omega^2_p < 0.001$. However, a trend was found for processing speed on the PASAT, $F(2, 127) = 2.49$, $p = 0.08$, $\eta^2_p = 0.038$, $\omega^2_p = 0.022$ (Figure 3).

A secondary analysis that included independent trials was conducted with a Group \times Time ANOVA by cognitive domain was conducted to performance across specific areas of executive functions.

Visual Scanning and Working Memory

Results of an ANOVA on TMT Card A and B completion times showed a main effect of time on Card A, $F(132) = 3.86$, $p = 0.05$, $\eta^2_p = 0.028$, $\omega^2_p = 0.021$, and Card B, $F(1, 132) = 5.76$, $p = 0.02$, $\eta^2_p = 0.042$, $\omega^2_p = 0.034$. No group by time interaction was found for Card A, $F(2, 132) = 1.50$, $p = 0.23$, $\eta^2_p = 0.022$, $\omega^2_p = 0.007$, and a trend was found for Card B, $F(2, 132) = 2.67$, $p = 0.07$, $\eta^2_p = 0.039$, $\omega^2_p = 0.024$.

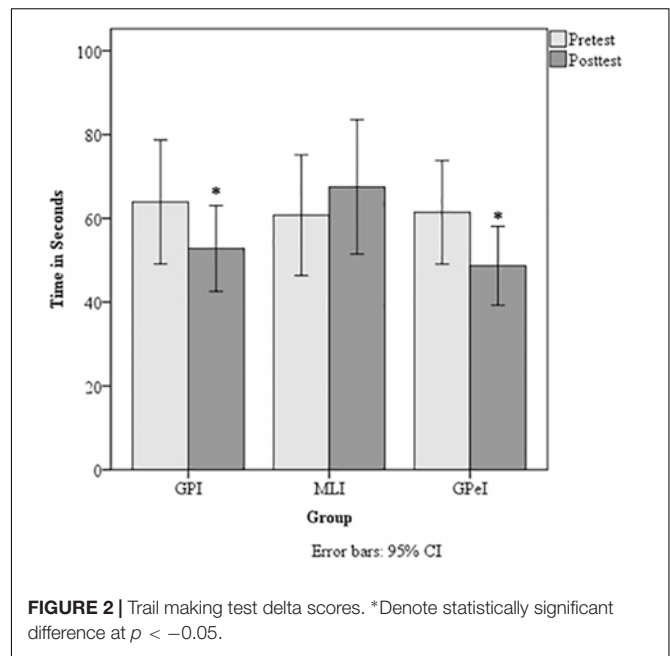


FIGURE 2 | Trail making test delta scores. *Denote statistically significant difference at $p < 0.05$.

Processing Speed

An independent trial analysis revealed a significant main effect of time for all four trials, $F(1, 132) = 11.09$, $p = 0.001$, $\eta^2_p = 0.078$, $\omega^2_p = 0.070$ (Trial 4, most challenging trial). However, only trial 3 showed a Group \times Time interaction, $F(2, 132) = 3.86$, $p = 0.02$, $\eta^2_p = 0.055$, $\omega^2_p = 0.041$. Pairwise comparisons reveal no significant interactions ($p = 0.538$) suggesting that the variance in scores may have contributed to this finding.

Dyad scores (i.e., the number of consecutive responses) were also analyzed (Table 2), since it is common for adults to skip items on this measure to compensate for task difficulty. Results of a Group \times Time ANOVA on dyad scores revealed a main effect of time, $F(1, 132) = 35.58$, $p = 0.001$, $\eta^2_p = 0.212$, $\omega^2_p = 0.205$, and no group by time interaction, $F(2, 132) = 1.76$, $p = 0.18$, $\eta^2_p = 0.026$, $\omega^2_p = 0.011$.

TABLE 2 | Measures of far transfer to executive functions.

	GPI (N = 49)		MLI (N = 48)		GPel (N = 38)	
	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
TMT (Card A) (in seconds)	36.27 (10.57)	34.82 (10.42)	40.69 (15.51)	36.46 (11.52)	36.58 (10.47)	36.45 (14.83)
TMT (Card B) (in seconds)	100.19 (56.98)	87.88 (39.11)	101.45 (58.76)	103.98 (59.50)	98.03 (46.29)	85.13 (33.71)
TMT delta scores*	63.93 (51.55)	53.06 (35.72)	60.76 (49.54)	67.53 (55.21)	61.44 (37.60)	48.68 (28.21)
Stroop mean errors	15.22 (12.35)	11.80 (10.82)	17.35 (14.95)	15.85 (15.15)	11.15 (11.54)	12.21 (10.13)
Stroop RT (in milliseconds)	1043.84 (140.83)	1010.41 (132.26)	1127.97 (174.55)	1116.36 (220.92)	1076.56 (196.25)	1059.84 (205.35)
Letter fluency (DKEFS)	39.51 (11.93)	39.35 (12.66)	36.90 (9.68)	37.52 (11.09)	35.03 (11.44)	40.71 (10.04)
Category fluency (DKEFS)	37.43 (6.33)	40.14 (8.55)	37.44 (7.49)	38.17 (6.78)	35.53 (6.89)	39.74 (8.19)
Category switching (DKEFS)	13.00 (7.85)	14.59 (3.03)	13.56 (3.20)	14.19 (2.99)	12.61 (2.84)	14.16 (2.55)
PASAT total correct	56.10 (17.93)	63.43 (17.29)	53.79 (17.71)	56.31 (19.79)	54.34 (18.69)	63.42 (17.47)
PASAT dyad scores	32.88 (21.73)	40.98 (22.99)	30.17 (20.64)	34.27 (22.64)	30.55 (22.99)	39.97 (24.27)

*Denotes significant ($p < 0.05$) group \times time differences.

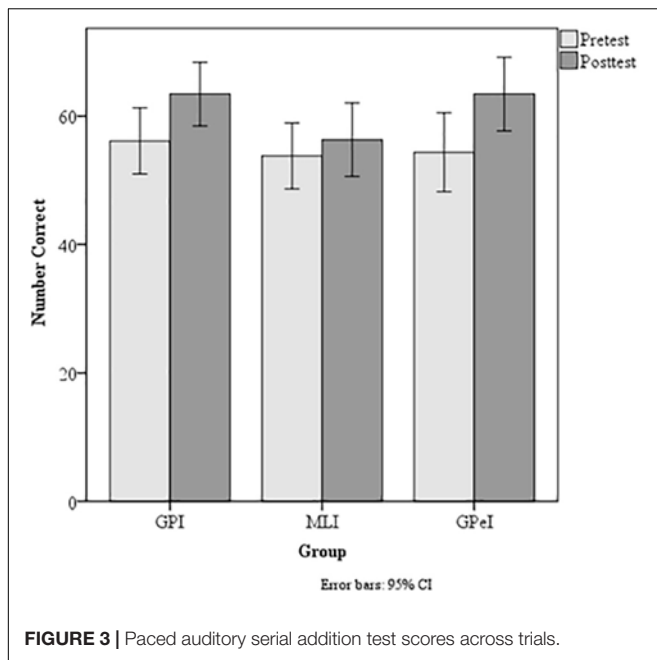


FIGURE 3 | Paced auditory serial addition test scores across trials.

Verbal Fluency

Results showed a significant main effect of time was found for category fluency, $F(1, 132) = 10.19$, $p = 0.002$, $\eta^2_p = 0.072$, $\omega^2_p = 0.064$, and category switching, $F(1, 132) = 12.87$, $p < 0.001$, $\eta^2_p = 0.089$, $\omega^2_p = 0.081$, however, no effect of time was found for letter fluency, $F(1, 132) = 2.51$, $p = 0.12$, $\eta^2_p = 0.019$, $\omega^2_p = 0.011$. No group \times time interactions were found for any of the trials of the verbal fluency subtest: letter fluency, $F(2, 132) = 1.86$, $p = 0.160$, $\eta^2_p = 0.027$, $\omega^2_p = 0.013$; category fluency, $F(2, 132) = 1.55$, $p = 0.217$, $\eta^2_p = 0.023$, $\omega^2_p = 0.008$; and category switching, $F(2, 132) = 0.86$, $p = 0.428$, $\eta^2_p < 0.001$, $\omega^2_p < 0.001$.

Response Inhibition

While the pattern of results showed decreases in mean errors committed by all groups at post-testing, analysis by blocks (color, mixed, word) showed no main effect for time and no group \times time interaction. Independent trial analysis of error rates by block: color, $F(2, 132) = 0.39$, $p = 0.676$, $\eta^2_p < 0.001$, $\omega^2_p < 0.001$;

word, $F(2, 132) = 1.02$, $p = 0.364$, $\eta^2_p = 0.015$, $\omega^2_p < 0.001$; or mixed, $F(2, 132) = 2.98$, $p = 0.055$, $\eta^2_p = 0.043$, $\omega^2_p = 0.028$; showed no significant main effect of time nor a Group \times Time interaction.

Motor Speed Analysis

Results of the two motor tests administered, the *Finger Tapper Test*, and the *Bimanual Motor Coordination Task* can be found on Table 3. A separate Repeated Measures ANOVA conducted on measures of motor speed showed a main effect for time for the dominant hand of the *Finger Tapper Test*, $F(1, 132) = 4.78$, $p = 0.03$, $\eta^2_p = 0.034$, $\omega^2_p = 0.027$, and all conditions of the *Bimanual Motor Coordination Task* including the Rhythm condition, $F(1, 132) = 22.18$, $p < 0.001$, $\eta^2_p = 0.144$, $\omega^2_p = 0.136$. Synchronization condition, $F(1, 132) = 21.60$, $p < 0.001$, $\eta^2_p = 0.141$, $\omega^2_p = 0.133$, and Pattern condition, $F(1, 132) = 19.63$, $p < 0.001$, $\eta^2_p = 0.129$, $\omega^2_p = 0.122$. No main effect of time was found for the *Finger Tapper Test*, non-dominant hand, $F(1, 132) = 0.08$, $p = 0.78$, $\eta^2_p < 0.001$, $\omega^2_p < 0.001$.

A significant Group \times Time interaction was found for synchronization on the *Bimanual Motor Coordination Task*, $F(2, 132) = 3.38$, $p = 0.03$, $\eta^2_p = 0.049$, $\omega^2_p = 0.034$ (Figure 4). No significant group \times time interactions were found for the *Finger Tapper Test* for dominant hand, $F(2, 132) = 0.78$, $p = 0.46$, $\eta^2_p < 0.001$, $\omega^2_p < 0.001$, or non-dominant hand, $F(2, 132) = 0.34$, $p = 0.71$, $\eta^2_p < 0.001$, $\omega^2_p < 0.001$. Nor were group \times time interactions found for the Rhythm condition, $F(2, 132) = 1.56$, $p = 0.21$, $\eta^2_p = 0.023$, $\omega^2_p = 0.008$ or Pattern conditions, $F(2, 132) = 1.03$, $p = 0.36$, $\eta^2_p = 0.015$, $\omega^2_p < 0.001$, of the *Bimanual Motor Coordination Task*.

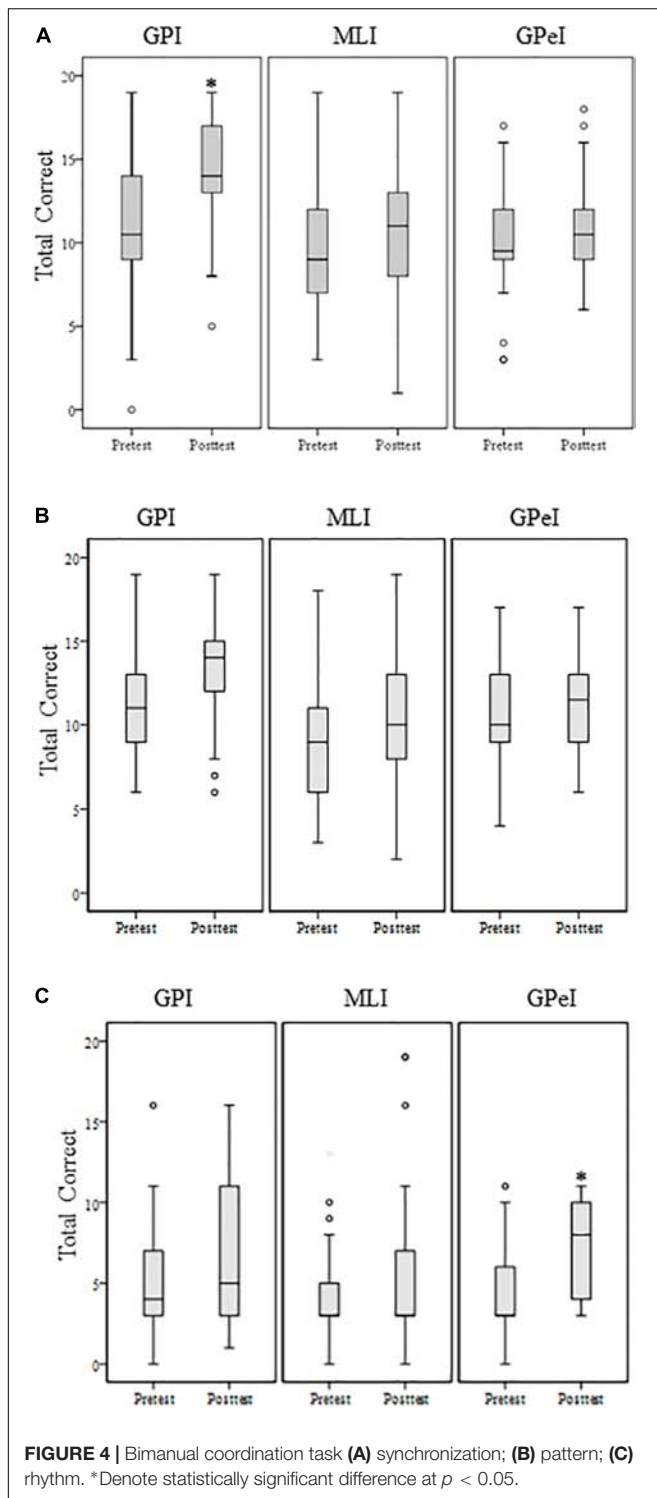
Correlational Analysis: Cognitive and Bimanual Coordination Outcomes

Correlations were conducted between executive function composite scores and bimanual coordination scores. Composites scores for measures of executive functions included: visual scanning/working memory (*Trail Making Card A* and *Card B*); processing speed (PASAT), verbal fluency (*Category Switching-DKEFS*), and response inhibition (*Stroop* errors and response times); and composite scores from the *Bimanual Motor Coordination Task*: Rhythm, Synchronization, and Pattern

TABLE 3 | Mood and Motor Measures.

	GPI (N = 49)		MLI (N = 48)		GPel (N = 38)	
	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
BDI	3.37 (3.02)	2.51 (3.49)	3.44 (3.55)	2.23 (2.27)	2.89 (2.73)	1.61 (1.93)
GDS	3.00 (3.26)	2.88 (2.71)	3.63 (4.40)	3.23 (3.11)	3.25 (2.90)	3.04 (3.10)
FT dominant hand	38.09 (9.22)	38.98 (9.12)	33.83 (10.31)	34.51 (9.69)	36.30 (8.06)	39.17 (6.39)
FT non-dominant hand	34.08 (6.61)	34.99 (7.23)	32.06 (7.40)	32.11 (6.47)	35.22 (5.93)	34.89 (6.43)
BMCT synchronization*	11.10 (3.80)	14.29 (3.14)	9.63 (4.36)	10.73 (4.55)	9.69 (3.67)	10.77 (3.22)
BMCT pattern	11.18 (3.12)	13.16 (3.04)	9.12 (3.83)	10.37 (3.41)	10.31 (3.47)	11.08 (3.10)
BMCT rhythmic*	5.20 (3.25)	6.59 (4.10)	4.22 (2.59)	5.24 (4.37)	4.62 (2.77)	7.04 (3.01)

BDI, Beck Depression Inventory; GDS, Geriatric Depression Scale; FT, Finger Tapper; BMCT, Bimanual Coordination Task. *Denotes significant ($p < 0.05$) group \times time differences.



Accuracy were included. Results revealed significant negative correlations between response inhibition (Stroop error rates) and Pattern accuracy at pretesting ($r = -0.208$, $p = 0.017$) and at post-testing ($r = -0.273$, $p = 0.002$). When the error rates increased in *Stroop* scores, pattern accuracy decreased in bimanual coordination. A similar significant negative correlation

was found for reaction times in Stroop performance and Rhythmic accuracy at post-testing ($r = -0.285$, $p = 0.001$), however, this relationship was not significant at pretesting ($r = -0.150$, $p = 0.096$).

Results of the correlational analysis showed a significant negative correlation in the time to complete the *Trail Making Test* (Card A) and all three composite scores from the *Bimanual Motor Coordination Task* at pretesting: Rhythm ($r = -0.173$, $p = 0.045$), Synchronization ($r = -0.184$, $p = 0.033$), and Pattern Accuracy ($r = -0.258$, $p = 0.002$). Increases in completion times for the Trail Making Test (Card A), resulted in reduced accuracy on the *Bimanual Motor Coordination Task*. However, no significant relationships were found at post-testing, Rhythm ($r = -0.107$, $p = 0.218$), Synchronization ($r = -0.027$, $p = 0.752$), and Pattern Accuracy ($r = -0.136$, $p = 0.115$).

Results of the correlational analysis showed a significantly negative correlation between the time to complete the *Trail Making Test* (Card B) and bimanual coordination scores on Pattern ($r = -0.253$, $p = 0.003$) and Synchronization ($r = -0.209$, $p = 0.015$) at pretesting. However, no significant correlation was found between the time to complete the Trail Making Test (Card B) and Rhythmic accuracy scores ($r = -0.129$, $p = 0.136$). Post-testing correlations showed a similar significantly negative correlation between completion times on the Trail Making Card B and Pattern accuracy ($r = -0.194$, $p = 0.024$); however, no significant relationships were found for Rhythmic accuracy ($r = -0.085$, $p = 0.324$) or Synchronization ($r = -0.091$, $p = 0.294$).

Similar findings were present for results of correlational analysis for the Total Correct on the PASAT, a measure of complex processing speed. Significant positive correlations between the PASAT number correct for pretesting and bimanual scores of Pattern ($r = 0.193$, $p = 0.025$) and Rhythmic accuracy ($r = 0.167$, $p = 0.045$), but not for Synchronization accuracy ($r = 0.103$, $p = 0.234$). However, no significant relationships were found between bimanual coordination scores at post-testing and PASAT performance, Pattern ($r = 0.150$, $p = 0.083$), Synchronization ($r = 0.001$, $p = 0.994$) and Rhythmic accuracy ($r = 0.089$, $p = 0.307$). Also, no significant correlations were found between composites for verbal fluency and bimanual coordination scores.

DISCUSSION

The purpose of the study was (1) to examine the effects of fine motor and gross motor bimanual coordination music interventions on cognitive and motor performance in healthy older adults and (2) to examine the relationships between performance on cognitive measures and areas of bimanual coordination. The experimental design of this research is unique as it is the first study to dissect complex behavioral interventions with three music training interventions in beginning older adult musicians. Specifically, this research differentiated gross and fine motor coordination in a stringent study design to discern potential benefits to cognitive and motor performance. Two music interventions involved instrumental performance in an

ensemble class that placed complex demands upon bimanual skill development, in-phase synchronization, antiphase coordination, precise timing, and sensorimotor integration (Repp and Su, 2013). Those assigned to GPI faced greater load on manual dexterity in movements that required distal musculature.

Cognitive Outcomes

Our results showed significantly enhanced performance by the GPI and GPel groups on visual scanning and working memory as compared to MLI. In addition, data showed a trend with a pattern of increased processing speed for the GPI group. These data are consistent with previous experimental studies examining the impact of active music participation in healthy older adults (Bugos et al., 2007; Bugos and Kochar, 2017; Dege and Kerkovius, 2018). Results may be due to the high level of sensorimotor integration involved in percussion and piano training programs. Sensorimotor synchronization in playing in an ensemble requires precise timing which involves anticipation of the beat and adaptation to stay with the ensemble (van der Steen and Keller, 2013; van der Steen et al., 2015). Synchronization to the beat was both challenging and rewarding for participants.

Results of this research that showed a significant main effect of time to visual scanning/working memory, processing speed, and verbal fluency, suggesting benefits related to participation in all active music interventions. These data are consistent with previous benefits stemming from drumming programs (Dege and Kerkovius, 2018), piano training programs (Bugos et al., 2007; Bugos and Kochar, 2017), and music listening (Mammarella et al., 2007). However, given research that examined factors contributing to executive functions on the *Trail Making Test* (Salthouse, 2011), we believe that differences found post-training may be related to processing speed.

Data analyzed by Salthouse (2011) examined areas of executive functions measured by the *Trail Making Test*, a common measure of visual scanning/working memory. Salthouse found abilities were related to spatial visualization and processing speed with no unique contribution of working memory. Recent research found after controlling for a general cognitive factor in a bifactor model in an older adult sample, only processing speed accounted for completion time on the *Trail Making Test* (MacPherson et al., 2019). Thus, changes found in this study may likely be due to enhanced processing speed, differentially measured by the *Trail Making Test* and PASAT. Furthermore, the correlational analysis includes supports the relationships between bimanual coordination and processing speed in the *Trail Making Test* and PASAT. Despite the low strength of the correlations found, replication of this research with alternate measures may help further differentiate the effects of music interventions to areas of executive functions.

No changes over time were found in response inhibition, a component of cognitive control. This result was in contrast to previous research by Seinfeld et al. (2013), who found enhanced cognitive control in older adults who received 16 weeks of group piano lessons. We believe these outcomes may be linked to practice quality and quantity. Practice requires cognitive control as the performer inhibits the incorrect note, in favor of the correct

note or rhythm. In the Seinfeld et al. (2013) study, practice requirements were 45 min per day or 5 h weekly, and in the current study 30 min per day or 3 h weekly. Further research is necessary to evaluate the role of practice, time duration of practice, and practice strategies that most efficiently contribute to cognitive control.

Motor Outcomes

Our results suggest that participants in all music interventions significantly improved bimanual motor outcomes of rhythmic accuracy, pattern accuracy and hand synchronization. In addition, the GPI group significantly improved in synchronization as compared to GPel and MLI group.

This result was not surprising in that piano training requires complex motor skills coupled by sensorimotor synchronization. Research has shown that even after short-term piano experiences, neurological and functional changes occur as a function of motor and auditory experiences (Lappe et al., 2008). In research comparing musicians and non-musicians, it was suggested that non-musicians showed greater auditory-to-visual transfer of information as compared to experts who demonstrate auditory-to-motor transfer of information (Brown and Penhune, 2018). While adults in the present study were not at the expert level, they certainly were able to demonstrate learning in the domain, as evidenced by a concluding recital (one for GPI, and GPel) which included memorized music. Future research is necessary to examine the trajectory of music instruction and to identify the point at which older beginners begin to demonstrate auditory-to-motor transfer of information.

Limitations and Potential Explanations

Measures in this study pertained to either cognitive or motor outcomes, this study did not measure learning in the trained domain with a measure of musical achievement or motivation. The lack of available music achievement measures with good psychometric properties and the lack of inclusion of measurement of musical achievement in training studies, renders it difficult to determine near transfer in the trained domain. In addition, the level of motivation may have influenced study outcomes. Anecdotal reports by many participants suggested an interest to enroll for the opportunity to learn musical notation; however, future research with standardized measures of motivation are necessary to explore the influence of motivation on music learning outcomes.

Practice logs were completed for all participants. While most individuals complied with practice requirements, some participants reported a longer practice period or less than 30 min per day. In addition, the quality of the practice session was not evaluated in this study. Future research is necessary to examine the quality and duration of practice sessions for novice adult musicians and the relationships to cognitive transfer. Recordings of each practice session and perhaps a practice coach, may contribute to the ability to evaluate the impact of music interventions delivered in a way that allows researchers to capture meaningful consolidation in music learning.

While the purpose of this study was to measure the effects of bimanual coordination in the context of music interventions, it is

unclear as to whether or not the benefits to executive functions are unique to music training or may be an outcome of tasks that require complex bimanual motor practice (e.g., juggling). Further research is necessary to examine the role of bimanual motor coordination interventions on executive functions.

This research included participants from a wide age-range (60–80 years). To adequately answer additional research questions regarding the impact of age-group differences on cognitive and motor outcomes, there is need for future research with larger samples of older adults. Nevertheless, this study showed that active engagement in music interventions contributes to cognitive and motor performance in healthy older adults. Interventions delivered in this research differed from many community music programs by including non-musicians and extending: a progressively difficult music curriculum, opportunities for social interaction, practice requirements, and bimanual coordination exercises. Instruments selected for training included pitched percussion instruments (i.e., pianos, xylophones, and metallophones) which were chosen to reduce any anxiety associated with tuning or embouchure to obtain a high quality sound. Group music programs can promote fine and gross motor coordination which may contribute to cognitive outcomes in aging.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

REFERENCES

- Ashendorf, L., Vanderslice-Barr, J. L., and McCaffrey, R. J. (2009). Motor tests and cognition in healthy older adults. *Appl. Neuropsychol.* 16, 171–176. doi: 10.1080/09084280903098562
- Axelrod, B. N. (2002). Validity of the Wechsler abbreviated scale of intelligence and other very short forms of estimating intellectual functioning. *Assessment* 9, 17–23. doi: 10.1177/1073191102009001003
- Beck, A. T., Steer, R. A., and Brown, G. K. (1996). *Beck Depression Inventory -II*. San Antonio, TX: Pearson.
- Bernard, J. A., and Seidler, R. D. (2014). Moving forward: age effects on the cerebellum underlie cognitive and motor declines. *Neurosci. Biobehav. Rev.* 42, 193–207. doi: 10.1016/j.neubiorev.2014.02.011
- Boyke, J., Driemeyer, J., Gaser, C., Buchel, C., and May, A. (2008). Training – induced brain changes in the elderly. *J. Neurosci.* 28, 7031–7035. doi: 10.1523/jneurosci.0742-08.2008
- Brandt, J., Spencer, M., and Folstein, M. (1988). The telephone interview for cognitive status. *Neuropsychiatry Neuropsychol. Behav. Neurol.* 1, 111–117.
- Brown, R. M., and Penhune, V. B. (2018). Efficacy of auditory versus motor learning for skilled and novice performers. *J. Cogn. Neurosci.* 30, 1657–1682. doi: 10.1162/jocn_a_01309
- Bugos, J., and Kochar, S. (2017). The effects of a short-term intense piano training program for cognitive aging: a pilot study. *Music. Sci.* 21, 137–150. doi: 10.177/1029864917690020
- Bugos, J. A. (2014). Community music as a cognitive training program for successful ageing. *Int. J. Commun. Music* 7, 319–331. doi: 10.1186/s12889-015-2395-9
- Bugos, J. A., and Cooper, P. (2019). The effects of mallet training on self-efficacy and processing speed in beginning adult musicians. *Res. Perspect. Music Educ.* 20, 21–32.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the East Carolina University Institutional Review Board. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JB conceived of the study, obtained funding, and carried out the activities associated with this study.

ACKNOWLEDGMENTS

I wish to thank the Retirement Research Foundation (RRF) for their generous support of this research. I wish to thank Rebecca Martin, Joy Fitzpatrick, and Wendy Mostafa for testing assistance, and Jamelle Simmons, Ayo Alexandra Gbadamosi and Miranda Rose Torres for data assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnint.2019.00068/full#supplementary-material>

MATERIAL S1 | Summary of the 16-week curriculum for GPI and GPel interventions.

- Bugos, J. A., and Iordache, A. (2019). “A rhythmic bimanual motor coordination task: reliability and validity in an adult population,” in *Paper Presented at the International Symposium on Assessment in Music Education*, (Gainesville, FL).
- Bugos, J. A., Perlstein, W. M., McCrae, C. S., Brophy, T. S., and Bedenbaugh, P. H. (2007). Individualized piano instruction enhances executive functioning and working memory in older adults. *Aging Mental Health* 11, 464–471. doi: 10.1080/13607860601086504
- Choi, U.-S., Sung, Y.-W., Hong, S., Chung, J.-Y., and Ogawa, S. (2015). Structural and functional plasticity specific to musical training with wind instruments. *Front. Hum. Neurosci.* 9:597. doi: 10.3389/fnhum.2015.00597
- Dege, F., and Kerkovius, K. (2018). The effects of drumming on working memory in older adults. *Ann. N. Y. Acad. Sci.* 1423, 242–250. doi: 10.1111/nyas.13685
- Delis, D., Kaplan, E., and Kramer, J. (2001). *Delis-Kaplan Executive Function Battery*. San Antonio, TX: Psychological Corporation.
- Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., and May, A. (2004). Neuroplasticity changes in grey matter induced by training. *Nature* 427, 311–312. doi: 10.1038/427311a
- Ehrenstein, W. H., Heister, G., and Cohen, R. (1982). Trail making test and visual search. *Arch. Psychiatry Nervenkr* 231, 333–338. doi: 10.1007/bf00345589
- Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., and Taub, E. (1995). Increased cortical representation of the fingers of the left hand in string players. *Science* 270, 305–307. doi: 10.1126/science.270.5234.305
- Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*, 4th Edn. London: Sage.
- Fong, T. G., Fearing, M. A., Jones, R. N., Shi, P., Marcantonio, E. R., Rudolph, J. L., et al. (2009). Telephone interview for cognitive status: creating a crosswalk with the mini-mental state examination. *Alzheimer's Dement.* 5, 492–497. doi: 10.1016/j.jalz.2009.02.007
- Fujiyama, H., Hinder, M. R., Garry, M. I., and Summers, J. J. (2013). Slow and steady is not as easy as it sounds: interlimb coordination at slow speed is

- associated with elevated attentional demand especially in older adults. *Exp. Brain Res.* 227, 289–300. doi: 10.1007/s00221-013-3511-7
- Gordon, E. E. (1989). *Advanced Measures of Music Audiation*. Chicago, IL: GIA Publications.
- Gronwall, D. M. (1977). Paced auditory serial-addition task: a measure of recovery from concussion. *Percept. Motor Skills* 44, 367–373. doi: 10.2466/pms.1977.44.2.367
- Guo, X., Ohsawa, C., Suzuki, A., and Sekiyama, K. (2018). Improved digit span in children after a 6-week intervention of playing a musical instrument: an exploratory randomized controlled trial. *Front. Psychol.* 8:2303. doi: 10.3389/fpsyg.2017.02303
- Hal Leonard Publishing Corporation (1997). *Get America Singing...Again Vol.1 Piano/Vocal/Guitar (A Project of the Music Educators National Conference)*. Milwaukee, WI: Hal Leonard Publishing.
- Hampton, W. (1995). *Hot Marimba! - Zimbabwean-Style Music for Orff Instruments*. New York, NY: Plank Road Publishing.
- Han, Y., Yang, H., Lv, Y.-T., Zhu, C.-Z., He, T., Tang, H.-H., et al. (2009). Gray matter density and white matter integrity in pianists' brain: a combined structural and diffusion tensor MRI study. *Neurosci. Lett.* 459, 3–6. doi: 10.1016/j.neulet.2008.07.056
- Herrojo Ruiz, M., Maess, B., Altenmüller, E., Curio, G., and Nikulin, V. V. (2017). Cingulate and cerebellar beta oscillations are engaged in the acquisition of auditory-motor sequences. *Hum. Brain Mapp.* 38, 5161–5179. doi: 10.1002/hbm.23722
- Heuninckx, S., Wenderoth, N., Debaere, P., Peeters, R., and Swinnen, S. P. (2005). Neural basis of aging: the penetration of cognition into action control. *J. Neurosci.* 25, 6787–6796. doi: 10.1523/jneurosci.1263-05.2005
- Hoffer, C. (2005). *Music Listening Today With Enhanced CDs*. Belmont, CA: Thomson Schirmer.
- IBM Corp (2016). *IBM SPSS Statistics for Windows, Version 24.0*. Armonk, NY: IBM Corp.
- Jancey, J., Lee, A., Howat, P., Clarke, A., Wang, K., and Shilton, T. (2007). Reducing attrition in physical activity programs for older adults. *J. Aging Phys. Activ.* 15, 152–165. doi: 10.1123/japa.15.2.152
- Jirenhed, D. A., Rasmussen, A., Johansson, F., and Hesslow, G. (2017). Learned response sequences in cerebellar Purkinje cells. *Proc. Natl. Acad. Sci. U.S.A.* 114, 6127–6132. doi: 10.1073/pnas.1621132114
- Kim, S. J., Cho, S. R., and Yoo, G. E. (2017). Age-related changes in bimanual instrument playing with rhythmic cueing. *Front. Psychol.* 8:1569. doi: 10.3389/fpsyg.2017.01569
- Koeneke, S., Lutz, K., Wüstenberg, T., and Jäncke, L. (2004). Long-term training affects cerebellar processing in skilled keyboard players. *NeuroReport* 15, 1279–1282. doi: 10.1097/01.wnr.0000127463.10147.e7
- Krampe, R. T. (2002). Aging, expertise and fine motor development. *Neurosci. Biobehav. Rev.* 26, 769–776. doi: 10.1016/S0149-7634(02)00064-7
- Krampe, R. T., and Ericsson, K. A. (1996). Maintaining excellence: deliberate practice and elite performance in young and older pianists. *J. Exp. Psychol.* 125, 331–359. doi: 10.1037/0096-3445.125.4.331
- Krings, T., Topper, R., Foltys, H., Erberich, S., Sparing, R., Willmes, K., et al. (2000). Cortical activation patterns during complex motor tasks in piano players and control subjects. A functional magnetic resonance imaging study. *Neurosci. Lett.* 278, 189–193. doi: 10.1016/S0304-3940(99)00930-1
- Lappe, C., Herholz, S. C., Trainor, L. J., and Pantev, C. (2008). Cortical plasticity induced by short-term unimodal and multimodal musical training. *J. Neurosci.* 28, 9632–9639. doi: 10.1523/JNEUROSCI.2254-08.2008
- MacPherson, S., Allerhand, M., Cox, S. R., and Deary, I. J. (2019). Individual differences underlying the trail making test-B performance in aging: the Lothian birth Cohort 1936. *Intelligence* 75, 29–32.
- Maes, C., Gooijers, J., de Xivry, J. O., Swinnen, S. P., and Boisgontier, M. P. (2017). Two hands, one brain, and aging. *Neurosci. Biobehav. Rev.* 75, 234–256. doi: 10.1016/j.neubiorev.2017.01.052
- Mammarella, N., Fairfield, B., and Cornoldi, C. (2007). Does music enhance cognitive performance in healthy older adults? the vivaldi effect. *Aging Clin. Exp. Res.* 19, 394–399. doi: 10.1007/BF03324720
- Metzler-Baddeley, C., Cantera, J., Coulthard, E., Rosser, A., Jones, D. K., and Baddeley, R. J. (2014). Improved executive function and callosal white matter microstructure after rhythm exercise in Huntington's Disease. *J. f Huntington's Dis.* 3, 273–283. doi: 10.3233/JHD-140113
- Moreira, N. B., Goncalves, G., da Silva, T., Zanardini, F. E. H., and Bento, P. C. B. (2018). Multisensory exercise programme improves cognition and functionality in institutionalized older adults: a randomized control trial. *Physiother. Res. Int.* 23, e1708. doi: 10.1002/pri.1708
- Murray, M. (1976). *Music for Children Volume 1: Pentatonic*. New York, NY: Schott Publishing.
- Nan, Y., Liu, L., Geiser, E., Shu, H., Gong, C. C., Dong, Q., et al. (2018). Piano training enhances the neural processing of pitch and improves speech perception in Mandarin-speaking children. *Proc. Natl. Acad. Sci. U.S.A.* 115, E6630–E6639. doi: 10.1073/pnas.1808412115
- Nikraves, M., Jafari, Z., Mehrpour, M., Kazemi, R., Amiri Shavaki, Y., Hossienifar, S., et al. (2017). The paced auditory serial addition test for working memory assessment: psychometric properties. *Med. J. Islamic Republic of Iran* 31:61. doi: 10.14196/mjiri.31.61
- Olin, J. T., Schneider, L. S., Eaton, E. M., Zemansky, M. F., and Pollock, V. E. (1992). The geriatric depression scale and the beck depression inventory as screening instruments in an older adult outpatient population. *Psychol. Assess.* 4, 190–192. doi: 10.1037/1040-3590.4.2.190
- Palmer, W. A., Manus, M., and Lethco, A. V. (1995). *Alfred-All-In-One Adult Piano Method*. New York, NY: Alfred Publishing.
- Palomar-García, M. -Á., Zatorre, R. J., Ventura-Campos, N., Bueichekú, E., and Ávila, C. (2017). Modulation of functional connectivity in auditory-motor networks in musicians compared with nonmusicians. *Cereb. Cortex* 27, 2768–2778. doi: 10.1523/JNEUROSCI.2018-08.2009
- Pantelyat, A., Syres, C., Reichwein, S., and Willis, A. (2016). DRUM-PD: the use of a drum circle to improve the symptoms and signs of Parkinson's disease (PD). *Mov. Disord. Clin. Pract.* 3, 243–249. doi: 10.1002/mdc3.12269
- Perlstein, W. M., Larson, M. J., Dotson, V. M., and Kelly, K. G. (2006). Temporal dissociation of components of cognitive control in severe TBI: ERPs and the cued-Stroop task. *Neuropsychologia* 44, 260–274. doi: 10.1016/j.neuropsychologia.2005.05.009
- Rauscher, F. H., and Zupan, M. A. (2000). Classroom keyboard instruction improves kindergarten children's spatial-temporal performance: a field experiment. *Early Child. Res. Q.* 15, 215–228. doi: 10.1016/S0885-2006(00)00050-8
- Reitan, R. M., and Wolfson, D. (1985). *The Halstead-Reitan Neuropsychological Test Battery*. Tucson, AZ: Neuropsychological Press.
- Reitan, R. M., and Wolfson, D. (1993). *The Halstead-Reitan Neuropsychological Test Battery: Theory and clinical interpretation*. Tucson, AZ: Neuropsychological Press.
- Repp, B. H., and Su, Y. H. (2013). Sensorimotor synchronization: a review of recent research (2006–2012). *Psychon. Bull. Rev.* 20:403. doi: 10.3758/s13423-012-0371-2
- Rueda-Delgado, L. M., Heise, K. F., Daffertshofer, A., Mantini, D., and Swinnen, S. P. (2019). Age-related differences in neural spectral power during motor learning. *Neurobiol. Aging* 77, 44–57. doi: 10.1016/j.neurobiolaging.2018.12.013
- Salthouse, T. A. (2011). What cognitive abilities are involved in trail-making performance? *Intelligence* 39, 222–232. doi: 10.1016/j.intell.2011.03.001
- Schlaug, G., Jäncke, L., Huang, Y., Staiger, J. F., and Steinmetz, H. (1995). Increased corpus callosum size in musicians. *Neuropsychologia* 33, 1047–1055. doi: 10.1016/0028-3932(95)00045-5
- Schneider, W., Eschman, A., and Zuccolotto, A. (2002). *E-Prime (Version 2.0). [Computer software and manual]*. Pittsburgh, PA: Psychology Software Tools Inc.
- Scholz, J., Klein, M. C., Behrens, T. E., and Johansen-Berg, H. (2009). Training induced changes in white-matter architecture. *Nat. Neurosci.* 12, 1370–1371. doi: 10.1038/nn.2412
- Seidler, R. D. (2007). Older adults can learn to learn new motor skills. *Behav. Brain Res.* 183, 118–122.
- Seidler, R. D., Bernard, J. A., Burutolu, T. B., Fling, B. W., Gordon, M. T., Gwin, J. T., et al. (2010). Motor control and aging: links to age-related brain structural,

- functional, and biochemical effects. *Neurosci. Biobehav. Rev.* 34, 721–733. doi: 10.1016/j.neubiorev.2009.10.005
- Seinfeld, S., Figueroa, H., Ortiz-Gil, J., and Sanchez-Vives, M. V. (2013). Effects of music learning and piano practice on cognitive function, mood and quality of life in older adults. *Front. Psychol.* 4:810. doi: 10.3389/fpsyg.2013.00810
- Smith, C., Vilijoen, J., and McGeachie, L. (2014). African drumming: a holistic approach to reducing stress and improving health? *J. Cardiovasc. Med.* 15, 441–446. doi: 10.2459/JCM.0000000000000046
- Spreeen, O., and Strauss, E. (1998). *A Compendium of Neuropsychological Tests: Administration, Norms and Commentary*, 2nd Edn. New York, NY: Oxford University Press.
- Sung, H. C., Lee, W. L., Li, T. L., and Watson, R. (2012). A group music intervention using percussion instruments with familiar music to reduce anxiety and agitation of institutionalized older adults with dementia. *Int. J. Geriatr. Psychiatry* 27, 621–627. doi: 10.1002/gps.2761
- Swinnen, S., and Gooijers, J. (2015). Bimanual coordination. *Brain Mapp.* 2, 475–482. doi: 10.1016/B978-0-12-397025-1.00030-0
- Tombaugh, T. N. (2006). A comprehensive review of the paced auditory serial addition test (PASAT). *Arch. Clin. Neuropsychol.* 21, 53–76. doi: 10.1016/j.acn.2005.07.006
- van der Steen, M. C., Jacoby, N., Fairhurst, M. T., and Keller, P. E. (2015). Sensorimotor synchronization with tempo-changing auditory sequences: modeling temporal adaptation and anticipation. *Brain Res.* 1626, 66–87. doi: 10.1016/j.brainres.2015.01.053
- van der Steen, M. C., and Keller, P. E. (2013). The adaptation and anticipation model (ADAM) of sensorimotor synchronization. *Front. Hum. Neurosci.* 7:253. doi: 10.3389/fnhum.2013.00253
- Watson, T. C., Becker, N., Apps, R., and Jones, M. W. (2014). Back to front: cerebellar connections and interactions with the prefrontal cortex. *Front. Syst. Neurosci.* 8:4. doi: 10.3389/fnsys.2014.00004
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence II*. San Antonio, TX: Pearson Corporation.
- Yesavage, J. A. (1991). Geriatric depression scale: consistency of depressive symptoms over time. *Percept. Motor Skills* 73:1032. doi: 10.2466/PMS.73.7.1032-1032

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Bugos. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Music Training, Working Memory, and Neural Oscillations: A Review

Kate A. Yurgil¹, Miguel A. Velasquez², Jenna L. Winston², Noah B. Reichman³ and Paul J. Colombo^{2,3*}

¹ Department of Psychological Sciences, Loyola University, New Orleans, LA, United States, ² Department of Psychology, Tulane University, New Orleans, LA, United States, ³ Brain Institute, Tulane University, New Orleans, LA, United States

OPEN ACCESS

Edited by:

Bruno L. Giordano,
UMR 7289 Institut de Neurosciences
de la Timone (INT), France

Reviewed by:

Peter Ciarani,
Harvard Medical School,
United States
Anne Keitel,
University of Dundee, United Kingdom

*Correspondence:

Paul J. Colombo
pcolomb@tulane.edu

Specialty section:

This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Psychology

Received: 13 June 2019

Accepted: 04 February 2020

Published: 21 February 2020

Citation:

Yurgil KA, Velasquez MA,
Winston JL, Reichman NB and
Colombo PJ (2020) Music Training,
Working Memory, and Neural
Oscillations: A Review.
Front. Psychol. 11:266.
doi: 10.3389/fpsyg.2020.00266

This review focuses on reports that link music training to working memory and neural oscillations. Music training is increasingly associated with improvement in working memory, which is strongly related to both localized and distributed patterns of neural oscillations. Importantly, there is a small but growing number of reports of relationships between music training, working memory, and neural oscillations in adults. Taken together, these studies make important contributions to our understanding of the neural mechanisms that support effects of music training on behavioral measures of executive functions. In addition, they reveal gaps in our knowledge that hold promise for further investigation. The current review is divided into the main sections that follow: (1) discussion of behavioral measures of working memory, and effects of music training on working memory in adults; (2) relationships between music training and neural oscillations during temporal stages of working memory; (3) relationships between music training and working memory in children; (4) relationships between music training and working memory in older adults; and (5) effects of entrainment of neural oscillations on cognitive processing. We conclude that the study of neural oscillations is proving useful in elucidating the neural mechanisms of relationships between music training and the temporal stages of working memory. Moreover, a lifespan approach to these studies will likely reveal strategies to improve and maintain executive function during development and aging.

Keywords: music training, working memory, neural oscillations, frequency bands, executive functions

INTRODUCTION

Music training engages some of our most complex cognitive abilities (Peretz and Zatorre, 2005; Zatorre and McGill, 2005) and induces brain plasticity in widely distributed cortical regions (Altenmüller and Schlaug, 2015). Musicianship requires selective and flexible attention, as well as inhibition of irrelevant auditory and visual stimuli. Musicians must readily manipulate stored information in accordance with complex hierarchies of rules and conventions. Reading musical notation requires strong spatial associations with percepts and symbols, and performing music requires effortful self-regulation and emotional expression.

Music training is emerging as an important model system for studying experience-dependent brain plasticity, and for the development of therapeutic interventions for healthy brain development and aging. Reports over the last few decades have firmly established that music training is associated with improvements on measures of executive functions, such as inhibitory

Glossary

Amplitude	Magnitude of an oscillatory signal, measured in microvolts
Coherence	A measure of oscillatory amplitude or phase consistency between a pair of electrodes
Cross-frequency coupling (CFC)	Interaction between two frequency bands such that one oscillation modulates the amplitude or phase of a second oscillation
Electroencephalography (EEG)	Non-invasive technique with millisecond temporal resolution used to measure changes in electrical cortical activity that accompanies sensory and cognitive processing
Event-related desynchronization (ERD)/synchronization (ERS)	Event-related difference in frequency band power relative to baseline; ERD reflects a decrease in post-stimulus EEG power, whereas ERS reflects an increase
Event-related potential (ERP)	Averaged EEG waveform time-locked to a stimulus or response event
Frequency band	Range of frequencies used in EEG analysis, denoted by different Greek letters: delta (1–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), beta (12–30 Hz), and gamma (> 30 Hz)
Neural oscillation	Rhythmic fluctuation in excitability of neural assemblies
Power	Squared amplitude of EEG signal within a frequency band
Phase	The position of an oscillation at a given point in time, measured in radians (0–2 π) or degrees (0–360)
Phase synchrony	A measure of the relationship between the phases of two oscillations, irrespective of amplitude
Transcranial magnetic stimulation (TMS)	Non-invasive neuromodulatory technique in which magnetic pulses are used to stimulate a targeted brain region
Transcranial alternating current stimulation (tACS)	Non-invasive neuromodulatory technique that delivers oscillating current at a desired frequency

control (Moreno and Farzan, 2015; D'Souza et al., 2018), working memory (Pallesen et al., 2010; Oechslin et al., 2013; Okhrei et al., 2016; Ding et al., 2018; D'Souza et al., 2018), and cognitive flexibility [see Okada and Slevc (2016) for a review], which coincide with structural and functional changes in brain regions implicated in these cognitive processes.

In studies using electroencephalography (EEG) and magnetoencephalography (MEG) to measure brain activity, inhibitory control is the most frequently reported executive function that is enhanced by music training, and it is frequently studied by measuring event-related potentials (ERPs), which coincide with discrete events. However, the relationship between music training and other components of executive function, such as working memory, has been studied less. In addition to ERPs, measurement of neural oscillations is a complementary method to study relationships between music training, working memory, and brain plasticity. In specific, the study of neural oscillations is well-suited for the study of events over periods of time from seconds to minutes, corresponding to the temporal stages of working memory. This review is focused on reports that link music training, working memory, and neural oscillations. The evidence that music training is associated with improvement in behavioral measures of working memory continues to increase. In addition, there is strong evidence linking behavioral measures of working memory to both localized and distributed patterns of neural oscillations (Tallon-Baudry et al., 1998; Jensen and Tesche, 2002; Kaiser et al., 2003, 2008, 2009; Herrmann et al., 2004; Grimault et al., 2009; Haegens et al., 2009; Moran et al., 2010; Van Dijk et al., 2010; Palva et al., 2011; Roux et al., 2012). And there is a small but growing number of reports of relationships between music training, working memory, and neural oscillations.

Taken together, we suggest that further study of neural oscillations will lead to significant progress in understanding the

neural mechanisms that support effects of music training on behavioral measures of executive functions. The current review is divided into five main sections as follows: (1) Discussion of behavioral measures of working memory, and effects of music training on working memory in adults; (2) effects of music training on neural oscillations during distinct temporal stages of working memory; (3) relationships between music training and working memory in children; (4) relationships between music training and working memory in older adults; and (5) effects of entrainment of neural oscillations on cognitive processing.

Behavioral Measures of Working Memory in Adults

There is increasing interest in the effects of music training on enhancement of general cognitive abilities and intelligence (Moreno et al., 2011; Schellenberg, 2011; Schellenberg and Weiss, 2013; Zuk et al., 2014; Costa-Giomi, 2015), and evidence suggests that working memory is an important component of the cognitive benefits derived from music training (Bergman Nutley et al., 2014; Suárez et al., 2015). For the purposes of this review, we define working memory as time- and capacity-limited storage of task-relevant information, which generally requires one or more of the following operations: mental manipulation, flexible use, or inhibition of distractors. Although working memory is sometimes equated with short-term memory, we consider it distinct for two reasons. First, working memory, as defined above, is distinct from short-term storage in the general requirement of mental manipulations of encoded information, or the inhibition of goal-irrelevant stimuli. Second, working memory requires integrity of medial temporal lobe regions, whereas short-term memory does not.

Working memory is often studied using behavioral tasks that implement a combination of stored information, cognitive manipulation, and interference. These tasks include the *N*-back (Pallesen et al., 2010; Oechslin et al., 2013; Ding et al., 2018), backward digit span (George and Coch, 2011; Zuk et al., 2014; Clayton et al., 2016), reading span (Franklin et al., 2008; D'Souza et al., 2018), and operation span tasks (Franklin et al., 2008; D'Souza et al., 2018). For the *N*-back task, a sequence of visual or auditory stimuli is presented, and the participant maintains the information while deciding whether each subsequent stimulus item matches the stimulus that came *N* letters previously (Owen et al., 2005). For the backward digit span, participants are presented with a series of digits, then asked to report the sequence of digits in reverse order (Hester et al., 2004). Both the *N*-back and the backward digit span require participants to maintain information in memory and manipulate the information in a certain order. The *N*-back may be a more demanding task as the participant must simultaneously keep track of current stimuli and determine whether or not they match the stimulus *N* turns back. For the reading span task, *N* number of sentences are presented one sentence at a time. As *N* increases, so does memory load required to perform the task. For *N* = 2, two sentences are presented sequentially, and after each sentence is presented, the participant writes the sentence verbatim, and the last word of each sentence in order (Daneman and Carpenter, 1980). Finally, the operation span task requires participants to memorize a sequence of unrelated words while simultaneously performing a series of math operations. After all of the operation-word strings are presented, the participant writes all of the words that were displayed in the order of presentation (Turner and Engle, 1989). Therefore, both the reading span and the operation span require participants to hold information while working on a secondary task, which may cause interference. As noted above, each of these tasks fulfill the criteria of maintenance and manipulation of information, which may occur with differing levels of interference (Aben et al., 2012). Working memory performance on these and related tasks will be discussed in the paragraphs that follow in relation to music training and neural oscillations.

RELATIONSHIP BETWEEN MUSIC TRAINING AND WORKING MEMORY IN ADULTS

Musicians reportedly outperform non-musicians on a variety of working memory tasks. For example, musicians showed better working memory capacity and duration than non-musicians using forward tonal discrimination tasks, whereas under atonal conditions, musicians outperformed non-musicians on working memory capacity, but not duration (Ding et al., 2018). Thus, musicians held more items in memory for both tonal and atonal auditory stimuli, but retained the items longer than non-musicians only for tonal stimuli. Furthermore, musicians have been reported to outperform non-musicians on *N*-back tasks using both auditory (Pallesen et al., 2010; Ding et al., 2018) and visual (Oechslin et al., 2013; Okhrei et al., 2016) stimuli. Therefore, the benefits of music training for working

memory are not necessarily limited to the auditory domain. Musicians have enhanced verbal working memory compared to non-musicians, as measured by both reading and operation span tasks (Franklin et al., 2008). However, this advantage may not extend to the visuospatial domain (Hansen et al., 2013), as no significant differences were reported between musicians and non-musicians on the Colorado Assessment for visual working memory (Strait et al., 2014), a computerized visual working memory task (Okhrei et al., 2016), and a spatial span task (Hansen et al., 2013). Discrepant reports of effects of musicianship are most likely due to the type of information to be retained, and less likely due to the age of participants, the sample size, or the types of tasks used. Slevc et al. (2016) reported that music training predicts individual differences in working memory, but that effect is not as strongly related to inhibitory control, and unrelated to cognitive flexibility. Using a similar approach to study the relationships between music training and executive functions, Okada and Slevc (2018) used a test battery consisting of tasks related to three subcomponents of executive functions: working memory, inhibitory control, and cognitive flexibility. They reported a positive correlation between individual variation in music training and working memory updating, but no relationships between music training and inhibitory control or shifting. In these contexts, inhibitory control refers to one's ability to regulate behavior, attention, and thoughts – especially in the face of conflicting information, or when doing so requires overriding a prepotent response. Cognitive flexibility refers to one's ability to successfully switch between task demands or mental sets (Slevc et al., 2016; Okada and Slevc, 2018). Relationships between music training and inhibitory control and cognitive flexibility have been reviewed in other sources and will not be taken up here (Moreno and Farzan, 2015; Okada and Slevc, 2016). It should be noted, however, that understanding the relationships between music training and behavioral measures of executive functions will progress more rapidly if a reliable battery of tests of executive functions is used consistently among investigators. For example, the NIH Toolbox cognitive measures include measures of executive function and have the added benefit of normative data.

Neural Oscillations and Working Memory in Adults

Neural oscillations are believed to represent synchronous changes in excitability in networks of cortical cells (Fries, 2005; Klimesch et al., 2007), and they are well-suited for the study of cognitive processes that occur over longer durations of time than a single event. This applies particularly to the study of the duration and capacity of the maintenance phase of working memory. Compared to intracranial recording techniques, EEG is a non-invasive and relatively inexpensive tool for measuring patterns of neuronal activity synchronized at the population level that may accompany sensory and cognitive processing. Electrical brain activity may be analyzed in the time-domain as ERPs, which are time-locked averages of EEG to a repeated stimulus or response event (see **Glossary**).

While there is abundant research related to ERPs and working memory (Drew et al., 2006; Yurgil and Golob, 2013; Pinal et al., 2015; Getzmann et al., 2018), the current review focuses on spectral (time-frequency) methods as a complementary approach to extracting ongoing oscillatory features that correspond to cognitive operations, such as the encoding or short-term storage of information. Various time-frequency analyses can be used to determine EEG magnitude or power within a given frequency band as well as the degree of coherence between different cortical regions (see **Glossary** for description of different frequency bands and spectral measures). Although time-frequency analyses are typically more computationally intensive than time-domain ERP methods, characterizing cortical oscillations and their synchronization is advantageous in examining the brain's distributed processes. In addition, there is strong evidence that different oscillations are supported by different functional networks that underlie sensory and cognitive processes, including working memory.

There is ample evidence that neural oscillations are important for memory formation. This was first demonstrated in studies of long-term potentiation (LTP) and long-term depression (LTD), which are neuronal models for memory formation involving biphasic changes in neuronal responses to synaptic inputs. Huerta and Lisman (1995) reported that a single burst of stimulation at the peak of theta phase induced LTP, whereas stimulation at the trough of theta induced LTD. The role of hippocampal theta has been investigated extensively in rodent models of memory formation (Buzsáki et al., 1994), and extended to interactions between the hippocampus and prefrontal cortex in humans (Anderson et al., 2010). Theta-band activity is also implicated in “chunking” of perceptual auditory information, which may contribute to hierarchical control representations necessary for complex, flexible behavior (Kikumoto and Mayr, 2018; Teng et al., 2018). In addition to theta, other neural oscillations, including alpha and gamma, are implicated in memory formation and maintenance. Indeed, Roux and Uhlhaas (2014) attributed distinct functional roles of alpha, theta, and gamma neural oscillations in working memory. In specific, they proposed that activity in the gamma-band subserves maintenance of working memory, whereas alpha-band oscillations may inhibit task-irrelevant information, and theta-band oscillations are important for temporal order of items in working memory. While significant advances have been made in understanding the roles of neural oscillations in working memory, less is known about the relationships between music training, neural oscillations, and their functional networks.

In this review, we offer a framework for examining oscillations related to different working memory components, and how music training may influence these underlying mechanisms. Specifically, we propose that oscillations may be used to: (1) dissociate temporal components of working memory, that is to distinguish between processing stages of encoding, maintenance, and retrieval; (2) investigate working memory processing demands, such as changes in memory load and inhibition; (3) measure short and long-range synchronous activity to

examine the relative contributions of distributed brain regions involved during working memory tasks; and (4) examine the degree to which music training influences oscillatory activity during working memory.

Music Training, Neural Oscillations, and Working Memory Maintenance

Stimulus-related oscillations may fluctuate over time and are therefore advantageous when examining neural activity over memory delays. There is considerable evidence that working memory maintenance, or the process of sustaining information in the absence of sensory input, is associated with enhanced activity or coherence in theta (4–8 Hz; Jensen and Tesche, 2002; Moran et al., 2010), alpha (8–12 Hz; Herrmann et al., 2004; Grimault et al., 2009; Haegens et al., 2009; Van Dijk et al., 2010), beta (12–30 Hz) in non-human primates (Tallon-Baudry et al., 2004) and humans (Palva et al., 2011), and gamma (>30 Hz; Tallon-Baudry et al., 1998; Kaiser et al., 2003, 2008, 2009; Palva et al., 2011; Roux et al., 2012) frequency bands [see Roux and Uhlhaas (2014) for extensive review of theta, alpha, and gamma]. See **Table 1** for a summary of results.

In addition, activity within each frequency band has been shown to vary with working memory load, or the number of items maintained over a brief delay (**Table 1**). High load conditions – that is, maintaining more information within working memory – are associated with increased power in alpha (Grimault et al., 2009; Palva et al., 2011; Samuel et al., 2018), beta (Palva et al., 2011), and gamma ranges (Howard et al., 2003; Palva et al., 2011). Increased load during memory delays is also associated with increased peak theta frequency (Moran et al., 2010), however findings on theta power and working memory load are mixed. Some studies show increased theta power with memory load over midline frontal (Jensen and Tesche, 2002; Meltzer et al., 2007) and other dispersed brain regions (Raghavachari et al., 2001, 2006), while others show decreased power with load over lateral frontal regions (Meltzer et al., 2007; van Vugt et al., 2010; Brzezicka et al., 2019).

While less studied, music training-related differences in working memory maintenance have been shown within the theta and beta ranges (**Table 2**). For example, music training may enhance theta-related activity during delay periods, affecting subsequent memory processing. Cheung et al. (2017) reported that musicians exhibited increased intra-hemispheric theta coherence during verbal memory encoding. Coherence is a measure of consistency within a given frequency band in amplitude or phase between different brain areas, thus higher coherence indicates greater synchrony of activity across regions. Higher theta coherence in musicians also correlated with subsequent memory performance (Cheung et al., 2017). Likewise, compared to non-musicians, musicians show increased left hemispheric theta coherence when judging the semantic relatedness of a new stimulus to that of previously learned information (Dittinger et al., 2017). Although not working memory maintenance *per se*, these tasks required sustained activation of memory representations for later processing; thus, music training may enhance theta coherence during

TABLE 1 | This table shows the results of studies on young adults, working memory, and oscillatory activity.

Study	Task	EEG measure	FB	Behavioral results	Oscillatory results			
					Encoding	Maintenance	Retrieval	Load
Van Dijk et al., 2010	aDMS	Power	α	–	$\uparrow\alpha$ (T)	$\uparrow\alpha$ (T)	–	–
Grimault et al., 2009	vSTM	Power	α	–	–	$\uparrow\alpha$	–	$\uparrow\text{Load} = \uparrow\alpha$ (F, P)
Haegens et al., 2009	DMS	Power	α β γ	73.7% correct responses	$\uparrow\gamma$ (contralateral, somatosensory) $\uparrow\alpha$ (contralateral, somatosensory) $\downarrow\beta$ (contralateral, somatosensory)	$\uparrow\gamma$ (bilateral SII) $\uparrow\alpha$ (O)	–	–
Herrmann et al., 2004	vDMS	Power	α	–	–	$\uparrow\alpha$ (P)	–	–
Jensen and Tesche, 2002	Sternberg	Power	θ	$\uparrow\text{Load} = \uparrow\text{RT}$	–	$\uparrow\text{Load} = \uparrow\theta$ (F) $\uparrow\text{Load} = \downarrow\theta$ (C, P)	–	$\uparrow\text{Load} = \uparrow\theta$ (F) $\uparrow\text{Load} = \downarrow\theta$ (C, P)
Kaiser et al., 2003	DMS	Power, coherence	γ	–	–	$\uparrow\gamma$ (LT) $\uparrow\gamma$ (Coh. LT - F)	–	–
Meltzer et al., 2007	Sternberg		θ α	$\uparrow\text{Load} = \uparrow\text{RT}$	–		–	$\uparrow\text{Load} = \uparrow\theta$ (FM)
Moran et al., 2010	vDMS	Frequency	θ	Items remembered: 1.5 - 5	–	$\uparrow\theta$ (F, P)	–	$\uparrow\text{Load} = \uparrow\theta$ (F, P)
Palva et al., 2011	vDMS	Power	β γ	$\uparrow\text{Load} = \downarrow\text{Accuracy}$	–	$\uparrow\text{Load} = \uparrow\gamma$ (O) $\uparrow\text{Load} = \uparrow\beta$ (O)	–	$\uparrow\text{Load} = \uparrow\gamma$ (O) $\uparrow\text{Load} = \uparrow\beta$ (O)
Roux et al., 2012	vDMS	Power	α γ	No difference in RT.	–	$\uparrow\gamma$ (P)	–	$\uparrow\text{Load} = \uparrow\gamma$ (P)
Samuel et al., 2018	vDMS	Frequency	α	–	$\downarrow\alpha$ (Freq.)	$\uparrow\alpha$ (Freq.)	$\uparrow\alpha$ (Freq.)	–
Tallon-Baudry et al., 2004	vDMS	Phase synchrony	β	–	–	$\uparrow\beta$ (P)	–	–

DMS, delay match-to-sample task; aDMS, auditory DMS; vDMS, visual DMS; F, frontal; C, central; T, temporal; P, parietal; O, occipital; M, medial; L, left; R, right. SII, supplementary somatosensory; RT, reaction time. Frequency bands: δ , delta (1–4 Hz); θ , theta (4–8 Hz); α , alpha (8–12 Hz); β , beta (12–30 Hz); γ , gamma (>30 Hz).

tasks that require maintaining stimulus representations over time. Increased theta coherence may support connections between brain areas important for memory processing, including medial temporal lobe and prefrontal cortex (Jones and Wilson, 2005; Anderson et al., 2010). In addition, Hsu et al. (2017) compared beta activity in musicians and non-musicians during an auditory *N*-back task. Enhanced beta activity has been shown to facilitate maintenance of information over delays in non-human primates (Tallon-Baudry et al., 2004) and humans (Lundqvist et al., 2016), and is predictive of individual visual working memory capacity (Palva et al., 2011). Group differences in beta activity indicated a processing advantage for musicians over non-musicians. As these differences were observed during the first 30 s of the 0-back condition, music training may promote more efficient encoding and maintenance of information within working memory, in the absence of interference or distraction.

There is also evidence that music training alters oscillatory coherence between different brain areas that are active during spatial working memory tasks (Table 2). Bhattacharya et al. (2001) examined the effects of music training on gamma coherence during mental rotation, a task that requires participants to discriminate between a 3D object and its rotated mirror image (Shepard and Metzler, 1971). Gamma coherence between frontal and right parietal cortex increased during mental rotation for all participants; however, musicians showed higher coherence overall and greater phase synchrony

in left hemisphere compared to non-musicians. Consistent with visuospatial behavioral findings discussed above (Hansen et al., 2013; Strait et al., 2014; Okhrei et al., 2016), there were no behavioral group differences in mental rotation. However, differences in gamma synchrony and hemispheric recruitment suggest music training alters activity within functional networks that support spatial working memory processes, including parietal and prefrontal cortices (Alivisatos, 1992; Tagaris et al., 1996; Alivisatos and Petrides, 1997). In a more recent study, musicians demonstrated enhanced gamma coherence between frontal and temporal regions during a spatial working memory task, but reduced coherence across all other frequency bands compared to non-musicians (Boutorabi and Sheikhan, 2018). Enhanced coherence was observed across the entire trial, and thus not strictly limited to working memory maintenance. However, it appears that enhanced coherence observed in musicians mediates cortical interactions important for working memory.

Music Training, Neural Oscillations, and Working Memory Encoding and Retrieval

Fluctuations in oscillatory activity may drive not only the active maintenance of items over brief delays, but the encoding and retrieval of such items. In a visual delayed match to sample paradigm, higher memory loads were associated with decreases in alpha frequency and power during encoding, and increases in alpha frequency during retrieval (Samuel et al., 2018). In another study by Myers et al. (2014),

TABLE 2 | This table summarizes the findings of the effects of music training on working memory and oscillatory activity.

Study	Task	FB	EEG measure	Behavioral results	Oscillatory results	
					Active recording	Resting state
Bhattacharya et al., 2001	Shepard–Metzler task of mental rotation	γ	Coherence	–	γ (Overall): M > NM γ (L-hemi): M > NM γ (R-hemi): M < NM	–
Boutorabi and Sheikhan, 2018	Spatial working memory	δ θ α β γ	Coherence	–	δ (T, C, P): M < NM θ (F, C, P): M < NM α (C, P): M < NM β (F, T, C, P): M < NM γ (C): M < NM γ (F, T): M > NM	–
Cheung et al., 2017	- HKLLT (verbal memory) - Visual reproduction (visual memory)	θ	Coherence	Verbal memory: M > NM Visual memory: M = NM	During encoding: θ (long-range, intrahemi): M > NM	θ : M = NM
Dittinger et al., 2017	- Matching task - Semantic task - Backward digit span	θ	Coherence	Semantic task: M > NM BDS: M = NM	θ (L-hemi): M > NM \uparrow Mus. Experience = $\uparrow \theta$ (L dorsal pathways)	–
Hsu et al., 2017	N-back (tonal)	β	Feature analysis	–	β : Response Speed: M > NM Response Intensity: M > NM Response Power: M > NM	–
Klein et al., 2016	Visual Sternberg	θ α β	Covariance mapping analysis	Correct responses: M = NM RT: M = NM	α (F, P): Mus prestimulus activity neg. correlates with RT.	All FB: M = NM

M, musicians; NM, non-musicians; FB, frequency band; F, frontal; C, central; T, temporal; P, parietal; O, occipital; M, medial; L, left; R, right; RT, reaction time; BDS, backward digit span.

pre-stimulus fluctuations in alpha power predicted accuracy during working memory retrieval; specifically, alpha event-related desynchronization (ERD; reduced power) was correlated with better memory performance.

Music training may further modulate pre-stimulus alpha activity, consequently affecting working memory retrieval (Table 2). In a study by Klein et al. (2016), pre-stimulus alpha activity over anterior–posterior regions was negatively associated with reaction time during a visual Sternberg task for musicians compared to non-musicians. According to findings from functional imaging studies, working memory recruits fronto-parietal brain areas (Wager and Smith, 2003; Owen et al., 2005), thus reduced anterior–posterior alpha activity may facilitate processing task-relevant information in musicians compared to non-musicians (Klein et al., 2016). Group differences in pre-stimulus alpha activity may reflect experience-dependent engagement of different neural networks involved in working memory.

Desynchronized oscillatory activity has been observed in other learning and memory paradigms. In a study by Silva et al. (2018), non-musicians learned to identify different melodic intervals, and were tested after an initial training session (baseline), and after 5 days of at-home training sessions (post-training). Improved

accuracy and reaction time to learned intervals was accompanied by reductions in alpha, beta, and gamma activity from baseline to post-training. Taken together, these findings suggest that changes in oscillatory activity and their functional networks may reflect individual variability during encoding, maintenance, and retrieval of items within working memory.

Music Training, Neural Oscillations, and Distractor Inhibition

It is important to note that the majority of findings discussed above were derived from working memory tasks that probe temporary storage of information in the absence of processing task-irrelevant information. However, given that working memory is a capacity-limited resource, active inhibition of task-irrelevant information is an important and necessary executive function. Neural oscillations are a useful tool in dissociating between storage and processing components of working memory and their underlying functional networks.

There is substantial evidence that alpha oscillations, first observed by Berger (1929), are involved in attentional suppression of task-irrelevant information. According to the alpha inhibitory hypothesis, alpha desynchronization (reduced alpha power) is associated with active information

processing while alpha synchronization (increased alpha power) is associated with active inhibition of task-irrelevant information [see Klimesch et al. (2007) and Foxe and Snyder (2011) for reviews].

Using a modified Sternberg task with weak and strong distracters, Bonnefond and Jensen (2012) found alpha activity increased with anticipation of strong distracters and was associated with faster reaction time to memory probes. Likewise, Sauseng et al. (2009) found increased alpha activity over posterior brain areas associated with distractor suppression, while maintaining relevant items was associated with increased theta-gamma interactions. Thus, high alpha activity facilitates inhibition of distracters during working memory. Furthermore, asymmetries in alpha power have been used to dissociate target selection vs. distractor inhibition functions within working memory (Schneider et al., 2019).

In tasks that require distractor suppression, individual differences in alpha activity may underlie behavioral performance during working memory tasks. Alpha activity has been shown to vary with working memory capacity, defined as the ability to actively maintain items in working memory in the face of distraction (Engle et al., 1999). Dong et al. (2015) found that independent of task difficulty, individuals with low working memory capacity exhibit greater alpha ERD (reduced alpha power), which may reflect the involvement of additional neural resources including those that are irrelevant to the task at hand.

To our knowledge, no literature exists on whether music training modulates oscillatory activity (alpha or otherwise) related to distractor inhibition. However, behavioral and ERP studies indicate better or more efficient inhibitory processing in musicians compared to non-musicians [see Moreno and Farzan (2015) for review]. While ERPs are useful in examining inhibition that is time-locked to a particular distracter event, measuring changes in oscillatory activity and coherence over longer periods of time may be useful to determine whether music training modulates distractor inhibition and associated functional networks over longer delays, as in working memory maintenance.

Music Training and Cross-Frequency Coupling During Working Memory

Neurons that oscillate at different frequencies may interact with each other, forming nested hierarchies in a time-dependent manner spanning multiple brain regions. Cross-frequency coupling, or the interaction of different frequency bands, is a useful tool in examining underlying generators and functional connectivity related to working memory.

It has been well established that working memory is associated with cross-frequency coupling between theta and gamma oscillations [Lisman and Jensen (2013) for review]. Studies show that the number of items maintained in working memory is related to the number of gamma rhythms nested within one theta cycle (Lisman and Idiart, 1995; Jensen and Lisman, 1998). This nesting predicts behavioral performance on working memory tasks, including response time (Axmacher et al., 2010) and individual differences in span (Kamiński et al., 2011). Theta-gamma coupling during working memory has been associated with activity in fronto-hippocampal networks in both

rodents (Belluscio et al., 2012) and humans (Axmacher et al., 2010). Furthermore, modulation of theta-gamma rhythms in prefrontal cortex improves working memory performance in adults (Aleksichuk et al., 2016).

Besides theta-gamma interactions, other cross-frequency couplings have been shown to support working memory processes. In monkeys, alpha-gamma coupling in the fronto-parietal network is sensitive to changes in working memory load (Pinotsis et al., 2019). In humans, interactions between visual and fronto-parietal theta and alpha/gamma, and between alpha and beta/gamma oscillations are enhanced during working memory maintenance, and reflect interactions between sensory and fronto-parietal networks (Siebenhüner et al., 2016). Additionally, cross-frequency coupling increased with working memory load and predicted individual differences in working memory capacity.

Previous sections of this review discussed differences in activity within frequency bands related to musical expertise. However, to our knowledge, there are no reports of whether cross-frequency coupling during working memory varies as a function of music training. While music training-dependent interactions between frequency bands remain to be investigated, we provided evidence that music training may alter neural connectivity in functional networks underlying working memory. During spatial working memory tasks, musicians show enhanced gamma coherence in fronto-parietal (Bhattacharya et al., 2001) and fronto-temporal (Boutorabi and Sheikhani, 2018) networks. During verbal memory tasks, musicians show enhanced intra-hemispheric theta coherence (Cheung et al., 2017; Dittinger et al., 2017). Group differences in spectral coherence complement fMRI findings that musicians engage different but overlapping brain regions during verbal vs. tonal working memory tasks, while non-musicians recruit the same regions regardless of task (Schulze et al., 2011). Thus, musicians may rely on stronger or alternative functional networks when engaged in different working memory tasks. These findings may be supported by future investigations of the effects of music training on cross-frequency coupling during working memory.

RELATIONSHIPS BETWEEN MUSIC TRAINING AND WORKING MEMORY IN CHILDREN

Several authors have reported that development of working memory may serve as a mechanism for emergence of cognitive abilities and other developmental outcomes (see Kraus et al., 2012). For example, there are reports of associations between working memory performance, music training or aptitude, and developmental outcomes such as neural encoding of speech (Strait et al., 2011, 2012; Christiner and Reiterer, 2018; Ireland et al., 2018), as well as cognitive skills related to reading abilities (Banai and Ahissar, 2013; Degé et al., 2015).

Behavioral investigations of children have shown that music training and musical aptitude are associated with enhanced auditory working memory. For example, in a longitudinal study of 6–8-year-old children, half of the sample was randomly

assigned to biweekly keyboard training for 6 weeks, while the other half received no training. Afterward, only the training group demonstrated a significant improvement in working memory capacity, measured with the backward digit span task (Guo et al., 2018). Advantages in working memory capacity have also been observed in school-aged children who received at least 6 months of music lessons (Degé and Schwarzer, 2017), as well as pre-schoolers who received 1 year of violin instruction (Fujioka et al., 2006). Bergman Nutley et al. (2014) measured working memory capacity every 2 years on two or three occasions among participants aged 6–25 years. They reported that musical practice was associated with better working memory capacity at each timepoint, and the increase was proportional to the hours of weekly musical practice, suggesting a dose–response relation between musical practice and working memory capacity.

The findings of cross-sectional studies of musical training and working memory in children are mixed. One report showed a musician advantage for auditory but not visual working memory (Strait et al., 2012), while other reports showed no effect of music training on working memory (Banai and Ahissar, 2013; Sachs et al., 2017). It should be noted, however, that while Sachs et al. (2017) reported no effect of music training on working memory, they did find differences between children with music training and those without training in measures of neural activity in brain regions associated with cognitive processes. In addition, cross-sectional studies have examined the relationship between musical aptitude and working memory capacity. Such investigations have shown that higher working memory capacity is associated with better scores on rhythmic (Strait et al., 2011; Degé et al., 2015; Ireland et al., 2018) and tonal (Christiner and Reiterer, 2018) subtests of musical aptitude in children as young as 5 years.

One issue in studies of relationships between musical training and cognitive outcomes is the direction of causality. For example,

those who are predisposed to succeed in music may also be predisposed to demonstrate enhanced cognitive abilities relative to age-matched peers (see Schellenberg, 2011). Alternatively, music training may cause enhancement of cognitive abilities. Longitudinal studies in which participants were randomly assigned to music training or control conditions clearly indicate a causal effect of training on working memory capacity (Fujioka et al., 2006; Degé and Schwarzer, 2017; Guo et al., 2018). Furthermore, evidence for a dose–response relationship between musical practice and working memory supports the view that training, rather than any predisposition, produces changes in working memory capacity (Bergman Nutley et al., 2014). However, associations between working memory capacity and measures of musical aptitude (Strait et al., 2011; Degé et al., 2015; Christiner and Reiterer, 2018; Ireland et al., 2018) suggest that training-related advantages are not independent of innate abilities. Thus the evidence presented here support the view that musical training and musical aptitude both contribute to working memory performance among children.

Working Memory and Neural Oscillations in Children

As indicated previously in this review, working memory is associated with neural oscillations in adults, and this relationship has also been reported in children (Table 3). The majority of research on neural oscillations in children has used the delayed-match-to-sample test to measure oscillations during memory maintenance. As with work in adults, alpha oscillations have been implicated in working memory maintenance in children. For example, one study of 10–13-year-olds, in which alpha activity was measured during working memory maintenance showed that the lateralization of alpha power changed with the number of items in working memory (Sander et al., 2012). Specifically,

TABLE 3 | This table shows the results of studies on development, working memory, and oscillatory activity.

Study	Task	FB	Behavioral results	Oscillatory results					
				Resting state	Pre-stimulus	Encoding	Maintenance	Retrieval	Load
Machinskaya and Kurgansky, 2012	vDMS	α θ β	–	–	θ (FM-T): A > Chil. β (F-C): A > Chil.	–	θ (T-P): A < Chil. α (F-P): A > Chil.	–	–
Sander et al., 2012	vDMS	α	YA > Chil. (low load)	–	–	–	–	α : YA < Chil.	High load: α : YA > Chil. Low load: α : YA < Chil.
Rodriguez-Martinez et al., 2013	WMTB-C	θ	\uparrow Age \rightarrow \uparrow WM	\uparrow Age, WM \rightarrow $\downarrow \theta$	–	–	–	–	–
Sato et al., 2018	vDMS	α	Accuracy = 80.15%	–	–	–	$\uparrow \alpha$ (whole brain connectivity during correct trials)	–	–
Walker et al., 2019	vDMS	γ	–	–	–	–	γ : Ado. > Chil.	–	–

YA, young adults; OA, adults; Ado, adolescents; Chil, children; F, frontal; C, central; T, temporal; P, parietal; O, occipital; M, medial; L, left; R, right.

when the number of items increased from low to medium, alpha power became more lateralized, whereas when the number of items increased from medium to high, alpha power became less lateralized. In contrast to the pattern seen in children, the adults showed an increase in alpha power lateralization as the number of items in working memory increased. In another study, Sato et al. (2018) reported higher alpha phase synchrony in 6-year-olds during the retention period of a delayed match to sample task for correct trials than for incorrect trials. Taken together, these reports indicate that the alpha network is engaged in memory maintenance during childhood.

In contrast to the view that alpha oscillations are important for memory maintenance in children, another study showed no difference in alpha coherence, but an increase in theta coherence, during maintenance among 7–8-year-old children (Machinskaya and Kurgansky, 2012). Notably, adults in the latter study showed an increase in alpha coherence, but no change in theta coherence during memory maintenance. Rodriguez-Martinez et al. (2013) also showed that theta oscillations play a role in working memory in children, and that this role diminishes throughout development. Specifically, the authors showed correlations between resting state theta oscillations and composite scores on a nine-subtest measure of working memory, conducted across an age range of 6–26 years. Power spectral density of the theta range negatively correlated with age across the sample, and this relationship, when included in a bivariate model with reaction time scores on the oddball task, accounted for 90% of variability in working memory due to age (Rodriguez-Martinez et al., 2013).

Taken together, these results indicate that alpha and theta oscillations are involved in working memory during childhood, and that oscillatory patterns during working memory processing differ between children and adults. Alpha peak frequency (APF) increases from infancy through adulthood (Klimesch, 1999), and may be a useful marker of brain maturation (Valdés et al., 1990), and cognitive development (Mierau et al., 2016). However, due to divergent findings on the role of alpha (Sander et al., 2012; Sato et al., 2018) and theta (Machinskaya and Kurgansky, 2012) during working memory maintenance, further studies are needed to understand the contributions of different oscillations during working memory maintenance.

It is also possible that gamma rhythms play an important role in working memory maintenance, specifically in regard to its maturation across development. In a recent study, 10–12-year-old children were compared with 15–17-year-old adolescents on the delayed match to sample task. The older children showed increased gamma power during the delay phase, and an increased gamma response to transcranial magnetic stimulation (TMS) applied to the prefrontal cortex. It is worth noting that the TMS-elicited gamma power was positively correlated with working memory capacity (Walker et al., 2019).

Overall, the present discussion shows the utility of measuring oscillations for investigating the neural mechanisms of working memory processing at different stages of development. Because working memory capacity and duration improve over the course of development, and oscillatory patterns for the same abilities

vary at different stages of development, studying how oscillations change throughout development could provide insights into the neural mechanisms of working memory. No study, to our knowledge, has been conducted in children to investigate relationships between music training, working memory, and neural oscillations. Thus, a significant gap in our knowledge is elucidation of how music training interventions and measures of musical aptitude might enhance or alter oscillatory activity throughout development.

RELATIONSHIP BETWEEN MUSIC TRAINING AND WORKING MEMORY IN OLDER ADULTS

In the sections above, we reported that adults with music training outperformed those without music training on several behavioral tasks of working memory. Consistent with those findings, older adults with music training also outperform older adults without music training. For example, Parbery-Clark et al. (2011) compared older musicians and non-musicians on auditory and visual working memory, and the ability to perceive speech in noise. They reported that musicians were significantly better at perceiving speech in noise and performed better in auditory, but not visuospatial, working memory capacity tasks. The study also revealed a linear relationship between auditory working memory and speech in noise performance, suggesting that these two functions are related. Grassi et al. (2018) also reported that older adult musicians outperformed older adult non-musicians on auditory and visuospatial working memory tasks, as well as auditory discrimination, but the groups did not differ on tests of short-term memory. In addition, Amer et al. (2013) reported that older adult musicians outperformed older adult non-musicians on several tests of executive functions, including visuospatial working memory. Taken together, these results reveal similarities and differences between adults and older adults with regard to relationships between music training and working memory. As with young adults, the benefits of music experience on working memory among older adults are not confined to the auditory domain. However, the finding that older musicians perform better than non-musicians on visuospatial working memory are in contrast to reports that young adult musicians and non-musicians do not differ on visuospatial working memory (Hansen et al., 2013). Also, life-long music training in older adults is not always associated with stronger working memory, as Hanna-Pladdy and MacKay (2011) reported strong correlations between music training and other executive functions, including cognitive flexibility, but not working memory. Some of the discrepancies in findings between young and older adults may be due to the effects of music training to compensate for age-associated decline in some executive functions.

Recently several studies have tested the potential of music intervention programs in reducing the deleterious effects of aging on cognition (Bugos et al., 2007; Hars et al., 2013; De Oliveira et al., 2014). In a study conducted by Bugos et al. (2007), older adults participated in individualized piano

instruction consisting of motor dexterity exercises and learning music theory. Participants had lessons each week for 6 months and were tested on cognitive and working memory measures across three time points: pre-training, post-training, and following a delay period of 3 months. The experimental group obtained significantly higher scores post-training on the Trail Making Test and Digit Symbols than the untrained controls, indicating an improvement in visoscanning, perceptual speed, and working memory.

Effects of Music Training and Interventions on Dementia

In addition to the benefits of music training on working memory in older adults, music training across the lifespan may serve as a protective factor against dementia (Verghese et al., 2003; Balbag et al., 2014). In a population based twin study, Balbag et al. (2014) showed older twins who played an instrument were 64% less likely to develop dementia than their co-twins. Contrary to these findings, a recent study (Kuusi et al., 2019) tracked the causes of death of classical musicians in Finland and found that musicians were just as likely to suffer from a neurodegenerative disorder as the general population. Taken together, these discrepant findings suggest that further studies are needed to determine whether music training is protective against neurodegenerative disorders.

Some studies have shown improvements in cognition and working memory in patients with dementia after active singing interventions (Särkämö et al., 2014; Maguire et al., 2015; Pongan et al., 2017). In a randomized controlled study, Särkämö et al. (2014) compared the effects of three different interventions on working memory on a group of patients with dementia. Each participant was assigned to one of three 10-week group-based interventions: singing, listening to music, or a usual care control group with physical or social activities. They reported that participants in the singing group showed a temporary improvement in working memory as measured by backward digit span. In contrast, Narme et al. (2014) found no cognitive improvements in patients with dementia after a 4-week music intervention. The discrepant results may be due to differences in intensity or duration of the interventions, suggesting that important parameters for interventions, such as the effective “dose” of training, are not yet well established.

Overall, these studies show that music training may protect against age-related decline in working memory, as well as improve performance among older adults who show decline in working memory. Importantly, music training may also be useful in the prevention and treatment of dementia.

Neural Oscillations in Older Adults

In comparison with young adults, the neural oscillations of older adults undergo several changes. As described in the paragraphs that follow, these changes may be related to age-related decline in working memory and cognition. However, there is a significant gap in our knowledge of the extent to which music training and music interventions can prevent or restore patterns of oscillatory activity and working memory among older adults. Some of the most often reported age-related changes in brain

wave activity are in: (1) activity during resting state recording (slowing of alpha frequency and theta changes); (2) changes in functional connectivity, or coherence; and (3) changes in power (or ERS/ERD) across different frequency bands and electrode sites during active recording. We discuss age-related changes in each of these measures below.

Resting State Alpha Frequency Slows Among Older Adults

Some of the most well studied age-related changes occur during resting-state recordings. Resting state alpha frequency increases throughout development (Lindsley, 1939; Stroganova et al., 1999; Marshall et al., 2002) reaching a peak in young adulthood (Chiang et al., 2011) and then slowing in older adults. The slowing of dominant alpha in older adults is a well-known phenomenon that may indicate various cognitive deficits (Obrist, 1954; Duffy et al., 1984; Köpruner et al., 1984; Aurlen et al., 2004; Chiang et al., 2011; Scally et al., 2018). Clark et al. (2004) found that age-related slowing of alpha, as measured by APF, was negatively correlated with working memory capacity. In the elderly, APF slowing is more dramatic in anterior than in posterior recording sites compared to young adults. In addition, variability in APF is negatively correlated with working memory across the lifespan, indicating the effectiveness of APF as a biomarker for working memory, a finding that has been replicated (Grandy et al., 2013). It is possible that normalizing the APF could restore some age-related decline in cognition. Training older individuals to increase APF through neurofeedback may improve processing speed, inhibitory control, and working memory (Angelakis et al., 2007).

Alpha peak frequency could also serve as a potential biomarker for mild cognitive impairment (MCI) (Babiloni et al., 2018) as it is lower for the MCI population than for normal older adults, as well as being correlated with MMSE scores. Furthermore, APF in posterior sites is positively correlated with hippocampal volume, which is reduced in the mild cognitively impaired population (Garcés et al., 2013). Likewise, a negative correlation has been found between frontal theta power and hippocampal volume (Grunwald et al., 2001). Overall, studies seeking to investigate the effects of music training on cognition and working memory in older adults should consider the relationship between aging, cognition, and changes in oscillatory activity during resting state, like APF.

Recording of neural oscillations during working memory may elucidate neural changes associated with aging (Table 4). Young and elderly adults show different patterns of ERS/ERD in alpha (Sander et al., 2012) and theta (Karrasch et al., 2004) during active recording in working memory tasks, even though the two groups may perform similarly. Older adults also tend to show lower theta power during working memory maintenance (Cummins and Finnigan, 2007; Kardos et al., 2014; Tóth et al., 2014) in comparison with young adults. Older adults also have lower theta (Kardos et al., 2014) and alpha power (McEvoy et al., 2001; Sander et al., 2012) in high working memory load tasks. Moreover, Sander et al. (2012) reported that older adults and children demonstrate similar activity during working memory tasks. On average, both older adults and children have higher

TABLE 4 | This table shows the results of studies on aging, working memory, and oscillatory activity.

Study	Task	FB	Behavioral results	Oscillatory results			
				Encoding	Maintenance	Retrieval	Load
Cummins and Finnigan, 2007	Verbal Sternberg	θ	YA = OA	–	$\theta(F, C)$: YA > OA	$\theta(F, C)$: YA > OA	–
Kardos et al., 2014	vDMS	θ	YA > OA	–	$\theta(P, O)$: YA > OA	–	High load: $\theta(P, O)$: YA > OA
Karrasch et al., 2004	Auditory Sternberg	$\alpha \beta \theta$	YA = OA	$\theta(P)$: YA > OA $\alpha(RT)$: YA < OA $\beta(C)$: YA < OA	–	$\alpha(C, T, O)$: YA < OA $\beta(C, T, O)$: YA < OA $\theta(O)$: YA > OA	–
McEvoy et al., 2001	Spatial working memory	$\alpha \theta$	YA = OA	–	–	–	High load: $\theta(FM)$: YA > OA α : YA > OA
Sander et al., 2012	vDMS	α	YA > OA (low load)	–	–	α : YA < OA	High load: α : YA > OA Low load: α : YA < OA
Tóth et al., 2014	vDMS	θ	YA > OA	–	θ : YA > OA	–	–

YA, young adults; OA, older adults; F, frontal; C, central; T, temporal; P, parietal; O, occipital; M, medial; L, left; R, right.

alpha power in comparison to young adults for low working memory loads. In children, higher alpha power is likely due to less well-developed neural mechanisms of working memory. In contrast, higher alpha power, and lower theta, in older adults is likely due to compensatory mechanisms.

Connectivity Decreases Among Older Adults

Age-related reductions of interhemispheric coherence have been observed in the delta, theta, alpha, and beta oscillations (Duffy et al., 1996; Kikuchi et al., 2000) during resting state recording. Global connectivity is also affected with age. Global alpha connectivity during resting state decreases with old age (Scally et al., 2018). Testing a large sample of 17,722, Vysata et al. (2014) reported an age-related decrease in global theta and alpha coherence. In a longitudinal study Fjell et al. (2016) found that reduction in connectivity is related to decline in inhibitory control. While there are not many studies that have explored the relationship between age-related declines in functional connectivity and working memory, it has been shown that working memory may improve by restoring theta synchronization in frontotemporal regions through transcranial alternating current stimulation (tACS; Reinhart and Nguyen, 2019). As mentioned previously, young adult musicians show higher theta synchrony during working memory tasks as well as better performance in those tasks (Cheung et al., 2017; Dittinger et al., 2017). Improved coherence as a result of music training may underlie the cognitive and working memory benefits seen in older musicians. Music training may help in strengthening some connections that may become weaker due to aging.

Consistent with the findings reported for young adults, older adults with music training outperform older adults without

music training on working memory tasks. These findings indicate that cognitive benefits that occur as a result of music training persist throughout the lifespan. Musical interventions in older non-musicians may also help in ameliorating age-related cognitive decline. Both non-pathological and pathological aging are accompanied by disruptions in brain wave activity, and these disruptions may also reflect working memory decline. While there are no studies showing the effects of music training on oscillatory activity in older adults, studies in young adults show the potential of music training to alter neural oscillations in ways consistent with enhanced working memory. It remains to be determined whether or not music training may improve some of the functional connectivity that is lost during aging. Studies with mice have shown that entrainment of oscillations in the gamma range through visual (Iaccarino et al., 2016) and auditory (Martorell et al., 2019) stimulation improves spatial and recognition memory and reduces the neuropathology associated with Alzheimer's disease. The entrainment of neural oscillations is taken up in the following section.

ENTRAINMENT OF NEURAL OSCILLATIONS

As discussed throughout this review, music training may alter oscillatory activity and functional networks associated with working memory. These findings suggest that oscillations may be targeted selectively to modulate working memory performance. Some studies have used stimulation techniques to disrupt oscillatory activity underlying WM processes; however, the effects on behavior may be at the individual level. For example, when repetitive TMS is applied over parietal brain regions during

memory delays, alpha-tuned stimulation impairs performance for individuals with high baseline WM capacity (Li et al., 2017), or for those who show increased alpha power during the delay period (Hamidi et al., 2009). Transcranial electrical stimulation applied over frontal and parietal regions during consecutive days of WM training has negative effects on WM performance and resting state connectivity (Möller et al., 2017). These findings seem to suggest a causal link between oscillatory rhythms and WM, such that perturbation of these rhythms, particularly during memory delays, negatively impacts maintenance of information over time.

If disruption of oscillatory activity impairs WM processes, can entrainment of specific rhythms enhance behavior? Indeed, a recent study using TMS examined whether tuning TMS to a specific frequency could enhance entrainment of neural networks associated with working memory (Albouy et al., 2017). TMS tuned to theta frequency enhanced performance during an auditory memory manipulation task; furthermore, theta power in correct trials positively correlated with musical experience. Similarly, theta-tuned tACS improved performance on an *N*-back task (Pahor and Jaušovec, 2018).

Stimulation techniques such as TMS and tACS, while non-invasive, require specialized equipment and expert administration. However studies have shown that simply listening to or tapping along with the rhythmic structure of music entrains the brain's low-frequency oscillations (Lakatos et al., 2008; Besle et al., 2011). For example, passive listening to musical sequences induces changes within alpha (Bridwell et al., 2017) and increases coupling between delta and high beta frequency ranges (Adamos et al., 2018). Word lists that are sung rather than spoken increase alpha coherence in bilateral frontal areas (Thaut et al., 2005) known to support learning-related processes (Sato et al., 2018). Furthermore, enhanced coherence to rhythmic musical stimuli correlates with improved memory for subsequent speech stimuli (Falk et al., 2017).

Tapping along with music rhythms induces beat-related entrainment (Nozaradan et al., 2015) as well as changes in functional somatosensory networks (Daly et al., 2014) and behavior (Nozaradan et al., 2016; Crasta et al., 2018). Importantly, rhythm entrainment is not limited to the auditory modality (Okawa et al., 2017) and can occur in the absence of an external stimulus. For example, entrained oscillatory activity persists during brief pauses in a rhythmic pattern (Stupacher et al., 2016), and varies not only with rhythm-based predictions of directly observed sequences, but also with memory-based predictions of imagined sequences (Breska and Deouell, 2017; Okawa et al., 2017). As discussed previously in this review, musicians show enhanced oscillatory activity and coherence during working memory delays, which may facilitate improved music-related error processing and pattern predictions shown in musicians compared to non-musicians (Doelling and Poeppel, 2015; Stupacher et al., 2017; Harding et al., 2019).

In addition to rhythm entrainment, oscillatory activity associated with working memory may be targeted by presenting two tones of different frequencies, one frequency to each ear; the resulting percept is of a *binaural beat* equal to the difference between the two frequencies (Licklider et al., 1950). Recent studies show that listening to binaural beats induces

changes in cortical networks associated with information processing. For example, increased gamma and beta power over frontal and central regions in response to binaural beats improved short-term memory of middle-list items on a serial recall task (Jirakittayakorn and Wongswat, 2017). In addition, accuracy on visuospatial (Beauchene et al., 2016) and verbal *N*-back (Beauchene et al., 2017) working memory tasks increased in response to binaural beats, which were thought to strengthen cortical networks involved in maintenance and retrieval of task-related information. Even passive listening to binaural beats induced changes in alpha power over frontal, temporal, and parietal lobes, with greater power increases in participants with musical experience (Ioannou et al., 2015). Thus, binaural beat stimulation may be a promising tool to entrain cortical networks involved in working memory.

SUMMARY AND FUTURE DIRECTIONS

Music training requires storage, manipulation, and integration of complex pitch and temporal sequences. In this way, it shares several features with commonly used measures of working memory. Therefore, it is not surprising that music training is related to enhancements in executive functions, including working memory. This review provides evidence that the study of neural oscillations is important for understanding the neural mechanisms underlying relationships between music training and working memory. In specific, we provide evidence that measurement of neural oscillations is particularly useful for studying temporal stages of memory and cognition that may occur over spans of seconds to minutes. In addition, we show that a lifespan approach to the study of relationships between music training, working memory, and neural oscillations reveals similarities and differences in working memory and underlying neural events at different stages of life. The following are some of the important specific points raised in this review, as well as suggestions for further investigation.

First, behavioral studies in adults show that benefits of music training are not restricted to auditory working memory, but may extend to the visual or other sensory modalities. In addition, music training may influence working memory capacity more selectively than working memory duration.

Second, music training is related to distinct patterns of modulation of oscillations that are related to encoding, maintenance, and retrieval phases of working memory in adults. **Table 1** lists studies in which measures of neural oscillations were used to: (1) dissociate temporal components of working memory such as encoding, maintenance, and retrieval; (2) investigate working memory processing demands, such as changes in memory load and inhibition, and (3) measure short and long-range synchronous activity to examine the relative contributions of distributed brain regions involved during working memory tasks. **Table 2** lists studies that examine the degree to which music training influences oscillatory activity during working memory. While we have discussed differences in activity within frequency bands related to musical expertise, to our knowledge

there are no reports of whether cross-frequency coupling during working memory varies as a function of music training. Thus while music training-dependent interactions between frequency bands remain to be investigated, there is evidence that music training may alter neural connectivity in functional networks underlying working memory.

Third, neural oscillations that support working memory and related cognitive processes change across the lifespan, and may serve as targets for music training or selective entrainment. Investigators have not yet studied relationships between music training and neural oscillations in children or older adults. Thus it remains to be determined whether or not relationships between music training and neural oscillations observed in adults will be the same as those observed in children or older adults. It is notable that alpha and theta oscillations are involved in working memory during childhood, and that oscillatory patterns during working memory processing differ between children and adults (Table 3). In addition, theta coherence declines with age but is enhanced during working memory maintenance in young musicians compared to non-musicians. Alpha activity, which is associated with inhibition in young adults, is also shown to decline with age, as does inhibitory processing. Accordingly, selective modulation of theta and alpha may show age-related changes in working memory maintenance and distractor suppression, respectively (Table 4). Future studies could test this hypothesis by comparing

the effects of music training or targeted entrainment on theta maintenance or alpha inhibition in young and older adults.

Finally, recent findings that entrainment of neural oscillations can enhance memory and cognition, and may reverse markers of neuropathology in models of Alzheimer's disease, suggest that further study of the role of neural oscillations in these processes will be necessary to guide development of therapeutic interventions for enhancement of cognition and for treatment of neural disorders across the lifespan. Of importance, music training, or targeted applications of musical stimuli, may serve as natural and non-invasive interventions for altering or entraining neural oscillations.

AUTHOR CONTRIBUTIONS

All authors contributed to the conception, organization, drafting, and revision of this manuscript.

FUNDING

This study was supported by the Phyllis M. Taylor Center for Social Innovation and Design Thinking, the Louise & Leonard Riggio Professorship in Social Innovation and Social Entrepreneurship, and Carnegie Professor I Endowments.

REFERENCES

- Aben, B., Stapert, S., and Blokland, A. (2012). About the distinction between working memory and short-term memory. *Front. Psychol.* 3:301. doi: 10.3389/fpsyg.2012.00301
- Adamos, D. A., Laskaris, N. A., and Micheloyannis, S. (2018). Harnessing functional segregation across brain rhythms as a means to detect EEG oscillatory multiplexing during music listening. *J. Neural Eng.* 15:036012. doi: 10.1088/1741-2552/aaac36
- Albouy, P., Weiss, A., Baillet, S., and Zatorre, R. J. (2017). Selective entrainment of theta oscillations in the dorsal stream causally enhances auditory working memory performance. *Neuron* 94, 193–206.e5. doi: 10.1016/j.neuron.2017.03.015
- Alekseichuk, I., Turi, Z., Amador de Lara, G., Antal, A., and Paulus, W. (2016). Spatial working memory in humans depends on theta and high gamma synchronization in the prefrontal cortex. *Curr. Biol.* 26, 1513–1521. doi: 10.1016/j.cub.2016.04.035
- Alivisatos, B. (1992). The role of the frontal cortex in the use of advance information in a mental rotation paradigm. *Neuropsychologia* 30, 145–159. doi: 10.1016/0028-3932(92)90024-G
- Alivisatos, B., and Petrides, M. (1997). Functional activation of the human brain during mental rotation. *Neuropsychologia* 35, 111–118. doi: 10.1016/S0028-3932(96)00083-8
- Altenmüller, E., and Schlaug, G. (2015). Apollo's gift: new aspects of neurologic music therapy. *Prog. Brain Res.* 217, 237–252. doi: 10.1016/bs.pbr.2014.11.029
- Amer, T., Kalender, B., Hasher, L., Trehub, S. E., and Wong, Y. (2013). Do older professional musicians have cognitive advantages? *PLoS One* 8:e71630. doi: 10.1371/journal.pone.0071630
- Anderson, K. L., Rajagovindan, R., Ghacibeh, G. A., Meador, K. J., and Ding, M. (2010). Theta oscillations mediate interaction between prefrontal cortex and medial temporal lobe in human memory. *Cereb. Cortex* 20, 1604–1612. doi: 10.1093/cercor/bhp223
- Angelakis, E., Stathopoulou, S., Frymiare, J. L., Green, D. L., Lubar, J. F., and Kounios, J. (2007). EEG neurofeedback: a brief overview and an example of peak alpha frequency training for cognitive enhancement in the elderly. *Clin. Neuropsychol.* 21, 110–129. doi: 10.1080/13854040600744839
- Aurlen, H., Gjerde, I. O., Aarseth, J. H., Eldoen, G., Karlsen, B., Skeidsvoll, H., et al. (2004). EEG background activity described by a large computerized database. *Clin. Neurophysiol.* 115, 665–673. doi: 10.1016/j.clinph.2003.10.019
- Axmacher, N., Henseler, M. M., Jensen, O., Weinreich, I., Elger, C. E., and Fell, J. (2010). Cross-frequency coupling supports multi-item working memory in the human hippocampus. *Proc. Natl. Acad. Sci. U.S.A.* 107, 3228–3233. doi: 10.1073/pnas.0911531107
- Babiloni, C., Del Percio, C., Lizio, R., Noce, G., Lopez, S., Soricelli, A., et al. (2018). Abnormalities of resting state cortical EEG rhythms in subjects with mild cognitive impairment due to Alzheimer's and Lewy body diseases. *J. Alzheimers Dis.* 62, 247–268. doi: 10.3233/JAD-170703
- Balbag, M. A., Pedersen, N. L., and Gatz, M. (2014). Playing a musical instrument as a protective factor against dementia and cognitive impairment: a population-based twin study. *Int. J. Alzheimers Dis.* 2014:836748. doi: 10.1155/2014/836748
- Banai, K., and Ahissar, M. (2013). Musical experience, auditory perception and reading-related skills in children. *PLoS One* 8:e75876. doi: 10.1371/journal.pone.0075876
- Beauchene, C., Abaid, N., Moran, R., Diana, R. A., and Leonessa, A. (2016). The effect of binaural beats on visuospatial working memory and cortical connectivity. *PLoS One* 11:e0166630. doi: 10.1371/journal.pone.0166630
- Beauchene, C., Abaid, N., Moran, R., Diana, R. A., and Leonessa, A. (2017). The effect of binaural beats on verbal working memory and cortical connectivity. *J. Neural Eng.* 14:026014. doi: 10.1088/1741-2552/aa5d67
- Belluscio, M. A., Mizuseki, K., Schmidt, R., Kempter, R., and Buzsaki, G. (2012). Cross-frequency phase-phase coupling between theta and gamma oscillations in the hippocampus. *J. Neurosci.* 32, 423–435. doi: 10.1523/JNEUROSCI.4122-11.2012
- Berger, H. (1929). Über das Elektroencephalogramm des Menschen. *Arch. Psychiatr. Nervenkr.* 87, 527–570. doi: 10.1007/BF01797193
- Bergman Nutley, S., Darki, F., and Klingberg, T. (2014). Music practice is associated with development of working memory during childhood and adolescence. *Front. Hum. Neurosci.* 7:926. doi: 10.3389/fnhum.2013.00926

- Besle, J., Schevon, C. A., Mehta, A. D., Lakatos, P., Goodman, R. R., McKhann, G. M., et al. (2011). Tuning of the human neocortex to the temporal dynamics of attended events. *J. Neurosci.* 31, 3176–3185. doi: 10.1523/JNEUROSCI.4518-10.2011
- Bhattacharya, J., Petsche, H., Feldmann, U., and Rescher, B. (2001). EEG gamma-band phase synchronization between posterior and frontal cortex during mental rotation in humans. *Neurosci. Lett.* 311, 29–32. doi: 10.1016/S0304-3940(01)02133-4
- Bonnefond, M., and Jensen, O. (2012). Alpha oscillations serve to protect working memory maintenance against anticipated distracters. *Curr. Biol.* 22, 1969–1974. doi: 10.1016/j.cub.2012.08.029
- Boutorabi, S. Y., and Sheikhan, A. (2018). Evaluation of electroencephalogram signals of the professional pianists during iconic memory and working memory tests using spectral coherence. *J. Med. Signals Sens.* 8, 87–94.
- Breska, A., and Deouell, L. Y. (2017). Neural mechanisms of rhythm-based temporal prediction: delta phase-locking reflects temporal predictability but not rhythmic entrainment. *PLoS Biol.* 15:e2001665. doi: 10.1371/journal.pbio.2001665
- Bridwell, D. A., Leslie, E., McCoy, D. Q., Plis, S. M., and Calhoun, V. D. (2017). Cortical sensitivity to guitar note patterns: EEG entrainment to repetition and key. *Front. Hum. Neurosci.* 11:90. doi: 10.3389/fnhum.2017.00090
- Brzezicka, A., Kamiński, J., Reed, C. M., Chung, J. M., Mamelak, A. N., and Rutishauser, U. (2019). Working memory load-related theta power decreases in dorsolateral prefrontal cortex predict individual differences in performance. *J. Cogn. Neurosci.* 31, 1290–1307. doi: 10.1162/jocn_a_01417
- Bugos, J. A., Perlstein, W. M., McCrae, C. S., Brophy, T. S., and Bedenbaugh, P. H. (2007). Individualized Piano Instruction enhances executive functioning and working memory in older adults. *Aging Ment. Health* 11, 464–471. doi: 10.1080/13607860601086504
- Buzsáki, G., Bragin, A., Chrobak, J. J., Nádasdy, Z., Sik, A., Hsu, M., et al. (1994). “Oscillatory and intermittent synchrony in the hippocampus: relevance to memory trace formation,” in *Temporal Coding in the Brain*, eds G. Buzsáki, R. Llinás, W. Singer, A. Berthoz, and Y. Christen, (Berlin: Springer), 145–172. doi: 10.1007/978-3-642-85148-3_9
- Cheung, M., Chan, A. S., Liu, Y., Law, D., and Wong, C. W. Y. (2017). Music training is associated with cortical synchronization reflected in EEG coherence during verbal memory encoding. *PLoS One* 12:e0174906. doi: 10.1371/journal.pone.0174906
- Chiang, A. K. I., Rennie, C. J., Robinson, P. A., Van Albada, S. J., and Kerr, C. C. (2011). Age trends and sex differences of alpha rhythms including split alpha peaks. *Clin. Neurophysiol.* 122, 1505–1517. doi: 10.1016/j.clinph.2011.01.040
- Christiner, M., and Reiterer, S. M. (2018). Early influence of musical abilities and working memory on speech imitation abilities: study with pre-school children. *Brain Sci.* 8:E169. doi: 10.3390/brainsci8090169
- Clark, C. R., Veltmeyer, M. D., Hamilton, R. J., Simms, E., Paul, R., Hermens, D., et al. (2004). Spontaneous alpha peak frequency predicts working memory performance across the age span. *Int. J. Psychophysiol.* 53, 1–9. doi: 10.1016/j.ijpsycho.2003.12.011
- Clayton, K. K., Swaminathan, J., Yazdanbakhsh, A., Zuk, J., Patel, A. D., Kidd, G., et al. (2016). Executive function, visual attention and the cocktail party problem in musicians and non-musicians. *PLoS One* 11:e0157638. doi: 10.1371/journal.pone.0157638
- Costa-Giomi, E. (2015). The long-term effects of childhood music instruction on intelligence and general cognitive abilities. *Update* 33, 20–26. doi: 10.1177/8755123314540661
- Crasta, J. E., Thaut, M. H., Anderson, C. W., Davies, P. L., and Gavin, W. J. (2018). Auditory priming improves neural synchronization in auditory-motor entrainment. *Neuropsychologia* 117, 102–112. doi: 10.1016/j.neuropsychologia.2018.05.017
- Cummins, T. D. R., and Finnigan, S. (2007). Theta power is reduced in healthy cognitive aging. *Int. J. Psychophysiol.* 66, 10–17. doi: 10.1016/j.ijpsycho.2007.05.008
- Daly, I., Hallowell, J., Hwang, F., Kirke, A., Malik, A., Roesch, E., et al. (2014). “Changes in music tempo entrain movement related brain activity,” in *Proceedings of the 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, (Chicago, IL: IEEE), 4595–4598. doi: 10.1109/EMBC.2014.6944647
- Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verbal Learn. Verbal Behav.* 19, 450–466. doi: 10.1016/S0022-5371(80)90312-6
- De Oliveira, T. C. G., Soares, F. C., De Macedo, L. D. E. D., Diniz, D. L. W. P., Bento-Torres, N. V. O., and Picanço-Diniz, C. W. (2014). Beneficial effects of multisensory and cognitive stimulation on age-related cognitive decline in long-term-care institutions. *Clin. Interv. Aging* 9, 309–320. doi: 10.2147/CIA.S54383
- Degé, F., Kubicek, C., and Schwarzer, G. (2015). Associations between musical abilities and precursors of reading in preschool aged children. *Front. Psychol.* 6:1220. doi: 10.3389/fpsyg.2015.01220
- Degé, F., and Schwarzer, G. (2017). Music lessons and verbal memory in 10- to 12-year-old children: investigating articulatory rehearsal as mechanism underlying this association. *Psychomusicology* 27, 256–266. doi: 10.1037/pmu0000201
- Ding, Y., Gray, K., Forrence, A., Wang, X., and Huang, J. (2018). A behavioral study on tonal working memory in musicians and non-musicians. *PLoS One* 13:e0201765. doi: 10.1371/journal.pone.0201765
- Dittinger, E., Valizadeh, S. A., Jäncke, L., Besson, M., and Elmer, S. (2017). Increased functional connectivity in the ventral and dorsal streams during retrieval of novel words in professional musicians. *Hum. Brain Mapp.* 39, 722–734. doi: 10.1002/hbm.23877
- Doelling, K. B., and Poeppel, D. (2015). Cortical entrainment to music and its modulation by expertise. *Proc. Natl. Acad. Sci. U.S.A.* 112, E6233–E6242. doi: 10.1073/pnas.1508431112
- Dong, S., Reeder, L. M., Yao, Y., Liu, Y., and Chen, F. (2015). Individual differences in working memory capacity are reflected in different ERP and EEG patterns to task difficulty. *Brain Res.* 1616, 146–156. doi: 10.1016/j.brainres.2015.05.003
- Drew, T. W., McCollough, A. W., and Voge, E. K. (2006). Event-related potential measures of visual working memory. *Clin. EEG Neurosci.* 37, 286–291. doi: 10.1177/155005940603700405
- D’Souza, A. A., Moradzadeh, L., and Wiseheart, M. (2018). Musical training, bilingualism, and executive function: working memory and inhibitory control. *Cogn. Res. Princ. Implic.* 3:11. doi: 10.1186/s41235-018-0095-6
- Duffy, F. H., Albert, M. S., McAnulty, G., and Garvey, A. J. (1984). Age-related differences in brain electrical activity of healthy subjects. *Ann. Neurol.* 16, 430–438. doi: 10.1002/ana.410160403
- Duffy, F. H., Mcanulty, G. B., and Albert, M. S. (1996). Effects of age upon interhemispheric EEG coherence in normal adults. *Neurobiol. Aging* 17, 587–599. doi: 10.1016/0197-4580(96)00007-3
- Engle, R. W., Kane, M. J., and Tuholski, S. W. (1999). “Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex,” in *Models of Working Memory*, eds A. Miyake, and P. Shah, (New York, NY: Cambridge University Press), 102–134. doi: 10.1017/CBO9781139174909.007
- Falk, S., Lanzilotti, C., and Schön, D. (2017). Tuning neural phase entrainment to speech. *J. Cogn. Neurosci.* 29, 1378–1389. doi: 10.1162/jocn_a_01136
- Fjell, A. M., Sneve, M. H., Grydeland, H., Storsve, A. B., and Walhovd, K. B. (2016). The disconnected brain and executive function decline in aging. *Cereb. Cortex* 27, 2302–2317. doi: 10.1093/cercor/bhw082
- Foxe, J. J., and Snyder, A. C. (2011). The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. *Front. Psychol.* 2:154. doi: 10.3389/fpsyg.2011.00154
- Franklin, M. S., Sledge Moore, K., Yip, C.-Y., Jonides, J., Rattray, K., and Moher, J. (2008). The effects of musical training on verbal memory. *Psychol. Music* 36, 353–365. doi: 10.1177/0305735607086044
- Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn. Sci.* 9, 474–480. doi: 10.1016/j.tics.2005.08.011
- Fujioka, T., Ross, B., Kakigi, R., Pantev, C., and Trainor, L. J. (2006). One year of musical training affects development of auditory cortical-evoked fields in young children. *Brain* 129, 2593–2608. doi: 10.1093/brain/awl247
- Garcés, P., Vicente, R., Wibrál, M., Pineda-Pardo, J. Á., López, M. E., Aurtenetxe, S., et al. (2013). Brain-wide slowing of spontaneous alpha rhythms in mild cognitive impairment. *Front. Aging Neurosci.* 5:100. doi: 10.3389/fnagi.2013.00100
- George, E. M., and Coch, D. (2011). Music training and working memory: an ERP study. *Neuropsychologia* 49, 1083–1094. doi: 10.1016/j.neuropsychologia.2011.02.001
- Getzmann, S., Wascher, E., and Schneider, D. (2018). The role of inhibition for working memory processes: ERP evidence from a short-term storage task. *Psychophysiology* 55:e13026. doi: 10.1111/psyp.13026
- Grandy, T. H., Werkle-Bergner, M., Chicherio, C., Lövdén, M., Schmiedek, F., and Lindenberger, U. (2013). Individual alpha peak frequency is related to latent

- factors of general cognitive abilities. *Neuroimage* 79, 10–18. doi: 10.1016/j.neuroimage.2013.04.059
- Grassi, M., Meneghetti, C., Toffalini, E., and Borella, E. (2018). Auditory and cognitive performance in elderly musicians and nonmusicians: correction. *PLoS One* 13:e0192918. doi: 10.1371/journal.pone.0192918
- Grimault, S., Robitaille, N., Grova, C., Lina, J.-M., Dubarry, A.-S., and Joliceur, P. (2009). Oscillatory activity in parietal and dorsolateral prefrontal cortex during retention in visual short-term memory: additive effects of spatial attention and memory load. *Hum. Brain Mapp.* 30, 3378–3392. doi: 10.1002/hbm.20759
- Grunwald, M., Busse, F., Hensel, A., Kruggel, F., Riedel-Heller, S., Wolf, M., et al. (2001). Correlation between cortical theta activity and hippocampal volumes in health, mild cognitive impairment, and mild dementia. *J. Clin. Neurophysiol.* 18, 178–184. doi: 10.1097/00004691-200103000-00010
- Guo, X., Ohsawa, C., Suzuki, A., and Sekiyama, K. (2018). Improved digit span in children after a 6-week intervention of playing a musical instrument: an exploratory randomized controlled trial. *Front. Psychol.* 8:2303. doi: 10.3389/fpsyg.2017.02303
- Haegens, S., Osipova, D., Oostenveld, R., and Jensen, O. (2009). Somatosensory working memory performance in humans depends on both engagement and disengagement of regions in a distributed network. *Hum. Brain Mapp.* 31, 26–35. doi: 10.1002/hbm.20842
- Hamidi, M., Slagter, H. A., Tononi, G., and Postle, B. R. (2009). Repetitive transcranial magnetic stimulation affects behavior by biasing endogenous cortical oscillations. *Front. Integr. Neurosci.* 3:14. doi: 10.3389/fnro.07.014.2009
- Hanna-Pladdy, B., and MacKay, A. (2011). The relation between instrumental musical activity and cognitive aging. *Neuropsychology* 25, 378–386. doi: 10.1037/a0021895
- Hansen, M., Wallentin, M., and Vuust, P. (2013). Working memory and musical competence of musicians and non-musicians. *Psychol. Music* 41, 779–793. doi: 10.1177/0305735612452186
- Harding, E. E., Sammler, D., Henry, M. J., Large, E. W., and Kotz, S. A. (2019). Cortical tracking of rhythm in music and speech. *Neuroimage* 185, 96–101. doi: 10.1016/j.neuroimage.2018.10.037
- Hars, M., Herrmann, F. R., Gold, G., Rizzoli, R., and Trombetti, A. (2013). Effect of music-based multitask training on cognition and mood in older adults. *Age Ageing* 43, 196–200. doi: 10.1093/ageing/af163
- Herrmann, C. S., Senkowski, D., and Rötger, S. (2004). Phase-locking and amplitude modulations of EEG alpha. *Exp. Psychol.* 51, 311–318. doi: 10.1027/1618-3169.51.4.311
- Hester, R. L., Kinsella, G. J., and Ong, B. (2004). Effect of age on forward and backward span tasks. *J. Int. Neuropsychol. Soc.* 10, 475–481. doi: 10.1017/s1355617704104037
- Howard, M. W., Rizzuto, D. S., Caplan, J. B., Madsen, J. R., Lisman, J., Aschenbrenner-Scheibe, R., et al. (2003). Gamma oscillations correlate with working memory load in humans. *Cereb. Cortex* 13, 1369–1374. doi: 10.1093/cercor/bhg084
- Hsu, C.-C., Cheng, C.-W., and Chiu, Y.-S. (2017). Analyze the beta waves of electroencephalogram signals from young musicians and non-musicians in major scale working memory task. *Neurosci. Lett.* 640, 42–46. doi: 10.1016/j.neulet.2017.01.022
- Huerta, P. T., and Lisman, J. E. (1995). Bidirectional synaptic plasticity induced by a single burst during cholinergic theta oscillation in CA1 *in vitro*. *Neuron* 15, 1053–1063. doi: 10.1016/0896-6273(95)90094-2
- Iaccarino, H. F., Singer, A. C., Martorell, A. J., Rudenko, A., Gao, F., Gillingham, T. Z., et al. (2016). Gamma frequency entrainment attenuates amyloid load and modifies microglia. *Nature* 540, 230–235. doi: 10.1038/nature20587
- Ioannou, C. I., Pereda, E., Lindsen, J. P., and Bhattacharya, J. (2015). Electrical brain responses to an auditory illusion and the impact of musical expertise. *PLoS One* 10:e0129486. doi: 10.1371/journal.pone.0129486
- Ireland, K., Parker, A., Foster, N., and Penhune, V. (2018). Rhythm and melody tasks for school-aged children with and without musical training: age-equivalent scores and reliability. *Front. Psychol.* 9:426. doi: 10.3389/fpsyg.2018.00426
- Jensen, O., and Lisman, J. E. (1998). An oscillatory short-term memory buffer model can account for data on the Sternberg task. *J. Neurosci.* 18, 10688–10699. doi: 10.1523/JNEUROSCI.18-24-10688.1998
- Jensen, O., and Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. *Eur. J. Neurosci.* 15, 1395–1399. doi: 10.1046/j.1460-9568.2002.01975.x
- Jirakittayakorn, N., and Wongsawat, Y. (2017). Brain responses to 40-Hz binaural beat and effects on emotion and memory. *Int. J. Psychophysiol.* 120, 96–107. doi: 10.1016/j.ijpsycho.2017.07.010
- Jones, M. W., and Wilson, M. A. (2005). Theta rhythms coordinate hippocampal–prefrontal interactions in a spatial memory task. *PLoS Biol.* 3:e402. doi: 10.1371/journal.pbio.0030402
- Kaiser, J., Heidegger, T., Wibral, M., Altmann, C. F., and Lutzenberger, W. (2008). Distinct gamma-band components reflect the short-term memory maintenance of different sound lateralization angles. *Cereb. Cortex* 18, 2286–2295. doi: 10.1093/cercor/bhm251
- Kaiser, J., Lutzenberger, W., Decker, C., Wibral, M., and Rahm, B. (2009). Task- and performance-related modulation of domain-specific auditory short-term memory representations in the gamma-band. *Neuroimage* 46, 1127–1136. doi: 10.1016/j.neuroimage.2009.03.011
- Kaiser, J., Ripper, B., Birbaumer, N., and Lutzenberger, W. (2003). Dynamics of gamma-band activity in human magnetoencephalogram during auditory pattern working memory. *Neuroimage* 20, 816–827. doi: 10.1016/S1053-8119(03)00350-1
- Kamiński, J., Brzezicka, A., and Wróbel, A. (2011). Short-term memory capacity (7 ± 2) predicted by theta to gamma cycle length ratio. *Neurobiol. Learn. Mem.* 95, 19–23. doi: 10.1016/j.nlm.2010.10.001
- Kardos, Z., Tóth, B., Boha, R., File, B., and Molnár, M. (2014). Age-related changes of frontal-midline theta is predictive of efficient memory maintenance. *Neuroscience* 273, 152–162. doi: 10.1016/j.neuroscience.2014.04.071
- Karrasch, M., Laine, M., Rapinaja, P., and Krause, C. M. (2004). Effects of normal aging on event-related desynchronization/synchronization during a memory task in humans. *Neurosci. Lett.* 366, 18–23. doi: 10.1016/j.neulet.2004.05.010
- Kikuchi, M., Wada, Y., Koshino, Y., Nanbu, Y., and Hashimoto, T. (2000). Effect of normal aging upon interhemispheric EEG coherence: analysis during rest and photic stimulation. *Clin. Electroencephalogr.* 31, 170–174. doi: 10.1177/155005940003100404
- Kikumoto, A., and Mayr, U. (2018). Decoding hierarchical control of sequential behavior in oscillatory EEG activity. *eLife* 7:e38550. doi: 10.7554/eLife.38550
- Klein, C., Diaz Hernandez, L., Koenig, T., Kottlow, M., Elmer, S., and Jäncke, L. (2016). The influence of pre-stimulus EEG activity on reaction time during a verbal Sternberg task is related to musical expertise. *Brain Topogr.* 29, 67–81. doi: 10.1007/s10548-015-0433-7
- Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res. Rev.* 29, 169–195. doi: 10.1016/S0165-0173(98)00056-3
- Klimesch, W., Sauseng, P., and Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition–timing hypothesis. *Brain Res. Rev.* 53, 63–88. doi: 10.1016/j.brainresrev.2006.06.003
- Köpruner, V., Pfürtscheller, G., and Auer, L. M. (1984). Quantitative EEG in normals and in patients with cerebral ischemia. *Prog. Brain Res.* 62, 29–50. doi: 10.1016/S0079-6123(08)62168-8
- Kraus, N., Strait, D. L., and Parbery-Clark, A. (2012). Cognitive factors shape brain networks for auditory skills: spotlight on auditory working memory. *Ann. N. Y. Acad. Sci.* 1252, 100–107. doi: 10.1111/j.1749-6632.2012.06463.x
- Kuusi, T., Haukka, J., Myllykangas, L., and Järvelä, I. (2019). Causes of death of professional musicians in the classical genre. *Med. Probl. Perform. Art.* 34, 92–97. doi: 10.21091/mppa.2019.2016
- Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., and Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320, 110–113. doi: 10.1126/science.1154735
- Li, S., Jin, J. N., Wang, X., Qi, H. Z., Liu, Z. P., and Yin, T. (2017). Theta and alpha oscillations during the retention period of working memory by rTMS stimulating the parietal lobe. *Front. Behav. Neurosci.* 11:170. doi: 10.3389/fnbeh.2017.00170
- Licklider, J. C. R., Webster, J. C., and Hedlund, J. M. (1950). On the frequency limits of binaural beats. *J. Acoust. Soc. Am.* 22, 468–473. doi: 10.1121/1.1906629
- Lindsley, D. B. (1939). A longitudinal study of the occipital alpha rhythm in normal children: frequency and amplitude standards. *Pedag. Semin. J. Genet. Psychol.* 55, 197–213. doi: 10.1080/08856559.1939.10533190

- Lisman, J. E., and Idiart, M. A. (1995). Storage of 7 +/- 2 short-term memories in oscillatory subcycles. *Science* 267, 1512–1515. doi: 10.1126/science.7878473
- Lisman, J. E., and Jensen, O. (2013). The θ - γ neural code. *Neuron* 77, 1002–1016. doi: 10.1016/j.neuron.2013.03.007
- Lundqvist, M., Rose, J., Herman, P., Brincat, S. L., Buschman, T. J., and Miller, E. K. (2016). Gamma and beta bursts underlie working memory. *Neuron* 90, 152–164. doi: 10.1016/j.neuron.2016.02.028
- Machinskaya, R. I., and Kurgansky, A. V. (2012). A comparative electrophysiological study of regulatory components of working memory in adults and seven to eight-year-old children: an analysis of coherence of eeg rhythms. *Hum. Physiol.* 38, 1–13. doi: 10.1134/S0362119712010136
- Maguire, L. E., Wanschura, P. B., Battaglia, M. M., Howell, S. N., and Flinn, J. M. (2015). Participation in active singing leads to cognitive improvements in individuals with dementia. *J. Am. Geriatr. Soc.* 63, 815–816. doi: 10.1111/jgs.13366
- Marshall, P. J., Bar-Haim, Y., and Fox, N. A. (2002). Development of the EEG from 5 months to 4 years of age. *Clin. Neurophysiol.* 113, 1199–1208. doi: 10.1016/s1388-2457(02)00163-3
- Martorell, A. J., Paulson, A. L., Suk, H. J., Abdurrob, F., Drummond, G. T., Guan, W., et al. (2019). Multi-sensory gamma stimulation ameliorates alzheimer's-associated pathology and improves cognition. *Cell* 177, 256–271. doi: 10.1016/j.cell.2019.02.014
- McEvoy, L. K., Pellouchoud, E., Smith, M. E., and Gevins, A. (2001). Neurophysiological signals of working memory in normal aging. *Cogn. Brain Res.* 11, 363–376. doi: 10.1016/S0926-6410(01)00009-X
- Meltzer, J. A., Negishi, M., Mayes, L. C., and Constable, R. T. (2007). Individual differences in EEG theta and alpha dynamics during working memory correlate with fMRI responses across subjects. *Clin. Neurophysiol.* 118, 2419–2436. doi: 10.1016/j.clinph.2007.07.023
- Mierau, A., Felsch, M., Hülsmüller, T., Mierau, J., Bullermann, P., Weiß, B., et al. (2016). The interrelation between sensorimotor abilities, cognitive performance and individual EEG alpha peak frequency in young children. *Clin. Neurophysiol.* 127, 270–276. doi: 10.1016/j.clinph.2015.03.008
- Möller, A., Nemmi, F., Karlsson, K., and Klingberg, T. (2017). Transcranial electric stimulation can impair gains during working memory training and affects the resting state connectivity. *Front. Hum. Neurosci.* 11:364. doi: 10.3389/fnhum.2017.00364
- Moran, R. J., Campo, P., Maestu, F., Reilly, R. B., Dolan, R. J., and Strange, B. A. (2010). Peak frequency in the theta and alpha bands correlates with human working memory capacity. *Front. Hum. Neurosci.* 4:200. doi: 10.3389/fnhum.2010.00200
- Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., and Chau, T. (2011). Short-term music training enhances verbal intelligence and executive function. *Psychol. Sci.* 22, 1425–1433. doi: 10.1177/0956797611416999
- Moreno, S., and Farzan, F. (2015). Music training and inhibitory control: a multidimensional model. *Ann. N. Y. Acad. Sci.* 1337, 147–152. doi: 10.1111/nyas.12674
- Myers, N. E., Stokes, M. G., Walther, L., and Nobre, A. C. (2014). Oscillatory brain state predicts variability in working memory. *J. Neurosci.* 34, 7735–7743. doi: 10.1523/JNEUROSCI.4741-13.2014
- Narme, P., Clément, S., Ehrle, N., Schiaratura, L., Vachez, S., Courtaigne, B., et al. (2014). Efficacy of musical interventions in dementia: evidence from a randomized controlled trial. *J. Alzheimers Dis.* 38, 359–369.
- Nozaradan, S., Peretz, I., and Keller, P. E. (2016). Individual differences in rhythmic cortical entrainment correlate with predictive behavior in sensorimotor synchronization. *Sci. Rep.* 6:20612. doi: 10.1038/srep20612
- Nozaradan, S., Zerouali, Y., Peretz, I., and Mouraux, A. (2015). Capturing with EEG the neural entrainment and coupling underlying sensorimotor synchronization to the beat. *Cereb. Cortex* 25, 736–747. doi: 10.1093/cercor/bht261
- Obrist, W. D. (1954). The electroencephalogram of normal aged adults. *Electroencephalogr. Clin. Neurophysiol.* 6, 235–244. doi: 10.1016/0013-4694(54)90025-5
- Oechslin, M. S., Van De Ville, D., Lazeyras, F., Hauert, C.-A., and James, C. E. (2013). Degree of musical expertise modulates higher order brain functioning. *Cereb. Cortex* 23, 2213–2224. doi: 10.1093/cercor/bhs206
- Okada, B. M., and Slevc, L. R. (2016). "Musical training: contributions to executive function," in *An Integrative Approach to Cognitive and Working Memory Training: Perspectives from Psychology, Neuroscience, and Human Development*, eds M. Bunting, J. Novick, M. Dougherty, and R. W. Engle, (New York, NY: Oxford University Press), doi: 10.13016/M2GM81P70
- Okada, B. M., and Slevc, L. R. (2018). Individual differences in musical training and executive functions: a latent variable approach. *Mem. Cogn.* 46, 1076–1092. doi: 10.3758/s13421-018-0822-8
- Okawa, H., Suefusa, K., and Tanaka, T. (2017). Neural entrainment to auditory imagery of rhythms. *Front. Hum. Neurosci.* 11:493. doi: 10.3389/fnhum.2017.00493
- Okhrei, A., Kutsenko, T., and Makarchuk, M. (2016). Performance of working memory of musicians and non-musicians in tests with letters, digits, and geometrical shapes. *Biologija* 62, 207–215. doi: 10.6001/biologija.v62i4.3408
- Owen, A. M., Mcmillan, K. M., Laird, A. R., and Bullmore, E. (2005). N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. *Hum. Brain Mapp.* 25, 46–59.
- Pahor, A., and Jaušovec, N. (2018). The effects of theta and gamma tACS on working memory and electrophysiology. *Front. Hum. Neurosci.* 11:651. doi: 10.3389/fnhum.2017.00651
- Pallesen, K. J., Brattico, E., Bailey, C. J., Korvenoja, A., Koivisto, J., Gjedde, A., et al. (2010). Cognitive control in auditory working memory is enhanced in musicians. *PLoS One* 5:e11120. doi: 10.1371/journal.pone.0011120
- Palva, S., Kulashekhar, S., Hämäläinen, M., and Palva, J. M. (2011). Localization of cortical phase and amplitude dynamics during visual working memory encoding and retention. *J. Neurosci.* 31, 5013–5025. doi: 10.1523/JNEUROSCI.5592-10.2011
- Parbery-Clark, A., Strait, D. L., Anderson, S., Hittner, E., and Kraus, N. (2011). Musical experience and the aging auditory system: implications for cognitive abilities and hearing speech in noise. *PLoS One* 6:e18082. doi: 10.1371/journal.pone.0018082
- Peretz, I., and Zatorre, R. (2005). Brain organization for music processing. *Annu. Rev. Psychol.* 56, 89–114. doi: 10.1146/annurev.psych.56.091103.070225
- Pinal, D., Zurrón, M., and Díaz, F. (2015). An event related potentials study of the effects of age, load and maintenance duration on working memory recognition. *PLoS One* 10:e0143117. doi: 10.1371/journal.pone.0143117
- Pinotsis, D. A., Buschman, T. J., and Miller, E. K. (2019). Working memory load modulates neuronal coupling. *Cereb. Cortex* 29, 1670–1681. doi: 10.1093/cercor/bhy065
- Pongan, E., Tillmann, B., Leveque, Y., Trombert, B., Getenet, J. C., Auguste, N., et al. (2017). Can musical or painting interventions improve chronic pain, mood, quality of life, and cognition in patients with mild Alzheimer's disease? Evidence from a randomized controlled trial. *J. Alzheimers Dis.* 60, 663–677.
- Raghavachari, S., Kahana, M. J., Rizzuto, D. S., Caplan, J. B., Kirschen, M. P., Bourgeois, B., et al. (2001). Gating of human theta oscillations by a working memory task. *J. Neurosci.* 21, 3175–3183. doi: 10.1523/JNEUROSCI.21-09-03175.2001
- Raghavachari, S., Lisman, J. E., Tully, M., Madsen, J. R., Bromfield, E. B., and Kahana, M. J. (2006). Theta oscillations in human cortex during a working-memory task: evidence for local generators. *J. Neurophysiol.* 95, 1630–1638. doi: 10.1152/jn.00409.2005
- Reinhart, R. M. G., and Nguyen, J. A. (2019). Working memory revived in older adults by synchronizing rhythmic brain circuits. *Nat. Neurosci.* 22, 820–827. doi: 10.1038/s41593-019-0371-x
- Rodriguez-Martinez, E. I., Barriga-Paulino, C. I., Rojas-Benjumea, M. A., and Gomez, C. M. (2013). Spontaneous theta rhythm and working memory co-variation during child development. *Neurosci. Lett.* 550, 134–138. doi: 10.1016/J.NEULET.2013.06.054
- Roux, F., and Uhlhaas, P. J. (2014). Working memory and neural oscillations: alpha-gamma versus theta-gamma codes for distinct WM information? *Trends Cogn. Sci.* 18, 16–25. doi: 10.1016/j.tics.2013.10.010
- Roux, F., Wibral, M., Mohr, H. M., Singer, W., and Uhlhaas, P. J. (2012). Gamma-band activity in human prefrontal cortex codes for the number of relevant items maintained in working memory. *J. Neurosci.* 32, 12411–12420. doi: 10.1523/JNEUROSCI.0421-12.2012
- Sachs, M., Kaplan, J., Der Sarkissian, A., and Habibi, A. (2017). Increased engagement of the cognitive control network associated with music training in children during an fMRI Stroop task. *PLoS One* 12:e0187254. doi: 10.1371/journal.pone.0187254

- Samuel, I. B. H., Wang, C., Hu, Z., and Ding, M. (2018). The frequency of alpha oscillations: task-dependent modulation and its functional significance. *Neuroimage* 183, 897–906. doi: 10.1016/j.neuroimage.2018.08.063
- Sander, M. C., Werkle-Bergner, M., and Lindenberger, U. (2012). Amplitude modulations and inter-trial phase stability of alpha-oscillations differentially reflect working memory constraints across the lifespan. *Neuroimage* 59, 646–654. doi: 10.1016/j.NEUROIMAGE.2011.06.092
- Särkämö, T., Tervaniemi, M., Laitinen, S., Numminen, A., Kurki, M., Johnson, J. K., et al. (2014). Cognitive, emotional, and social benefits of regular musical activities in early dementia: randomized controlled study. *Gerontologist* 54, 634–650. doi: 10.1093/geront/gnt100
- Sato, J., Mossad, S. I., Wong, S. M., Hunt, B. A. E., Dunkley, B. T., Smith, M. L., et al. (2018). Alpha keeps it together: alpha oscillatory synchrony underlies working memory maintenance in young children. *Dev. Cogn. Neurosci.* 34, 114–123. doi: 10.1016/j.DCN.2018.09.001
- Sauseng, P., Klimesch, W., Heise, K. F., Gruber, W. R., Holz, E., Karim, A. A., et al. (2009). Brain oscillatory substrates of visual short-term memory capacity. *Curr. Biol.* 19, 1846–1852. doi: 10.1016/j.cub.2009.08.062
- Scally, B., Burke, M. R., Bunce, D., and Delvenne, J.-F. (2018). Resting-state EEG power and connectivity are associated with alpha peak frequency slowing in healthy aging. *Neurobiol. Aging* 71, 149–155. doi: 10.1016/j.NEUROBIOLAGING.2018.07.004
- Schellenberg, E. G. (2011). Examining the association between music lessons and intelligence. *Br. J. Psychol.* 102, 283–302. doi: 10.1111/j.2044-8295.2010.02000.x
- Schellenberg, E. G., and Weiss, M. W. (2013). “Music and cognitive abilities,” in *The Psychology of Music*, ed. D. Deutsch, (Cambridge, MA: Academic Press), 499–550. doi: 10.1016/B978-0-12-381460-9.00012-2
- Schneider, D., Göddert, A., Haase, H., Hickey, C., and Wascher, E. (2019). Hemispheric asymmetries in EEG alpha oscillations indicate active inhibition during attentional orienting within working memory. *Behav. Brain Res.* 359, 38–46. doi: 10.1016/j.bbr.2018.10.020
- Schulz, K., Zysset, S., Mueller, K., Friederici, A. D., and Koelsch, S. (2011). Neuroarchitecture of verbal and tonal working memory in nonmusicians and musicians. *Hum. Brain Mapp.* 32, 771–783. doi: 10.1002/hbm.21060
- Shepard, R. N., and Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science* 171, 701–703. doi: 10.1126/SCIENCE.171.3972.701
- Siebenhühner, F., Wang, S. H., Matias Palva, J., and Palva, S. (2016). Cross-frequency synchronization connects networks of fast and slow oscillations during visual working memory maintenance. *eLife* 5:e13451. doi: 10.7554/eLife.134
- Silva, V., Pereira, T., and Conde, J. (2018). Brain processing and coding responses to musical expertise: an EEG study. *Neurophysiol. Res.* 1, 23–31.
- Slevc, L. R., Davey, N. S., Buschkuhl, M., and Jaeggi, S. M. (2016). Tuning the mind: exploring the connections between musical ability and executive functions. *Cognition* 152, 199–211. doi: 10.1016/j.cognition.2016.03.017
- Strait, D. L., Hornickel, J., and Kraus, N. (2011). Subcortical processing of speech regularities underlies reading and music aptitude in children. *Behav. Brain Funct.* 7:44. doi: 10.1186/1744-9081-7-44
- Strait, D. L., O’Connell, S., Parbery-Clark, A., and Kraus, N. (2014). Musicians’ enhanced neural differentiation of speech sounds arises early in life: developmental evidence from ages 3 to 30. *Cereb. Cortex* 24, 2512–2521. doi: 10.1093/cercor/bht103
- Strait, D. L., Parbery-Clark, A., Hittner, E., and Kraus, N. (2012). Musical training during early childhood enhances the neural encoding of speech in noise. *Brain Lang.* 123, 191–201. doi: 10.1016/j.BANDL.2012.09.001
- Stroganova, T. A., Orekhova, E. V., and Posikera, I. N. (1999). EEG alpha rhythm in infants. *Clin. Neurophysiol.* 110, 997–1012.
- Stupacher, J., Witte, M., Hove, M. J., and Wood, G. (2016). Neural entrainment in drum rhythms with silent breaks: evidence from steady-state evoked and event-related potentials. *J. Cogn. Neurosci.* 28, 1865–1877. doi: 10.1162/jocn_a_01013
- Stupacher, J., Wood, G., and Witte, M. (2017). Neural entrainment to polyrhythms: a comparison of musicians and non-musicians. *Front. Neurosci.* 11:208. doi: 10.3389/fnins.2017.00208
- Suárez, L., Elangovan, S., and Au, A. (2015). Cross-sectional study on the relationship between music training and working memory in adults. *Aust. J. Psychol.* 68, 38–46. doi: 10.1111/ajpy.12087
- Tagaris, G. A., Kim, S. G., Strupp, J. P., Andersen, P., Uğurbil, K., and Georgopoulos, A. P. (1996). Quantitative relations between parietal activation and performance in mental rotation. *Neuroreport* 7, 773–776.
- Tallon-Baudry, C., Bertrand, O., Peronnet, F., and Pernier, J. (1998). Induced gamma-band activity during the delay of a visual short-term memory task in humans. *J. Neurosci.* 18, 4244–4254. doi: 10.1523/JNEUROSCI.18-11-04244.1998
- Tallon-Baudry, C., Mandon, S., Freiwald, W. A., and Kreiter, A. K. (2004). Oscillatory synchrony in the monkey temporal lobe correlates with performance in a visual short-term memory task. *Cereb. Cortex* 14, 713–720. doi: 10.1093/cercor/bhh031
- Teng, X., Tian, X., Doelling, K., and Poeppel, D. (2018). Theta band oscillations reflect more than entrainment: behavioral and neural evidence demonstrates an active chunking process. *Eur. J. Neurosci.* 48, 2770–2782. doi: 10.1111/ejn.13742
- Thaut, M. H., Peterson, D. A., and McIntosh, G. C. (2005). Temporal entrainment of cognitive functions: musical mnemonics induce brain plasticity and oscillatory synchrony in neural networks underlying memory. *Ann. N. Y. Acad. Sci.* 1060, 243–254. doi: 10.1196/annals.1360.017
- Tóth, B., Kardos, Z., File, B., Boha, R., Stam, C. J., and Molnár, M. (2014). Frontal midline theta connectivity is related to efficiency of WM maintenance and is affected by aging. *Neurobiol. Learn. Mem.* 114, 58–69. doi: 10.1016/j.nlm.2014.04.009
- Turner, M. L., and Engle, R. W. (1989). Is working memory capacity task dependent? *J. Mem. Lang.* 28, 127–154. doi: 10.1016/0749-596X(89)90040-5
- Valdés, P., Virués, S., Szava, L., Galán, R., and Biscay, (1990). High resolution spectral EEG norms topography. *Brain Topogr.* 3, 281–282.
- Van Dijk, H., Nieuwenhuis, I. L. C., and Jensen, O. (2010). Left temporal alpha band activity increases during working memory retention of pitches. *Eur. J. Neurosci.* 31, 1701–1707. doi: 10.1111/j.1460-9568.2010.07227.x
- van Vugt, M. K., Schulze-Bonhage, A., Litt, B., Brandt, A., and Kahana, M. J. (2010). Hippocampal gamma oscillations increase with memory load. *J. Neurosci.* 30, 2694–2699. doi: 10.1523/JNEUROSCI.0567-09.2010
- Verghese, J., Lipton, R. B., Katz, M. J., Hall, C. B., Derby, C. A., Kuslansky, G., et al. (2003). Leisure activities and the risk of dementia in the elderly. *N. Engl. J. Med.* 348, 2508–2516.
- Vysata, O., Kukal, J., Prochazka, A., Pazdera, L., Simko, J., and Valis, M. (2014). Age-related changes in EEG coherence. *Neurol. Neurochir. Pol.* 48, 35–38. doi: 10.1016/j.pjnns.2013.09.001
- Wager, T. D., and Smith, E. E. (2003). Neuroimaging studies of working memory: a meta-analysis. *Cogn. Affect. Behav. Neurosci.* 3, 255–274.
- Walker, C., Murphy, N., Ramakrishnan, N., Fraher, D., and Cho, R. (2019). The role of gamma oscillations for working memory development in the adolescent brain. *Brain Stimul.* 12:546.
- Yurgil, K. A., and Golob, E. J. (2013). Cortical potentials in an auditory oddball task reflect individual differences in working memory capacity. *Psychophysiology* 50, 1263–1274. doi: 10.1111/psyp.12140
- Zatorre, R., and McGill, J. (2005). Music, the food of neuroscience? *Nature* 434, 312–315. doi: 10.1038/434312a
- Zuk, J., Benjamin, C., Kenyon, A., and Gaab, N. (2014). Behavioral and neural correlates of executive functioning in musicians and non-musicians. *PLoS One* 9:e99868. doi: 10.1371/journal.pone.0099868

Conflict of 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.

Copyright © 2020 Yurgil, Velasquez, Winston, Reichman and Colombo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Mechanisms of Timing, Timbre, Repertoire, and Entrainment in Neuroplasticity: Mutual Interplay in Neonatal Development

Joanne Loewy^{1*} and Artur C. Jaschke^{2,3}

¹The Louis Armstrong Center for Music and Medicine, Mount Sinai Beth Israel, Icahn School of Medicine, New York, NY, United States, ²Department of Music Therapy, Beatrix Children's Hospital—University Medical Centre, ArtEZ University of the Arts, Groningen, Netherlands, ³Department of Neonatology and Clinical Neuropsychology, Amsterdam, Netherlands

Neonatal brain development relies on a combination of critical factors inclusive of genetic predisposition, attachment, and the conditions of the pre and postneonatal environment. The status of the infant's developing brain in its most vulnerable state and the impact that physiological elements of music, silences and sounds may make in the earliest stages of brain development can enhance vitality. However, little attention has been focused on the integral aspects of the music itself. This article will support research that has hypothesized conditions of music therapeutic applications in an effort to further validate models of neurobehavioral care that have optimized conditions for growth, inclusive of recommendations leading toward the enhancement of self-regulatory behaviors.

Keywords: music plasticity, neural music mapping, neonatal music mapping, music brain, neonatal music therapy

OPEN ACCESS

Edited by:

Paul J. Colombo,
Tulane University, United States

Reviewed by:

Nadia Justel,
Interdisciplinary Laboratory of
Cognitive Neuroscience, Argentina
Marcela Pena,
Pontifical Catholic University of Chile,
Chile

*Correspondence:

Joanne Loewy
joanne.loewy@mountsinai.org

Received: 28 September 2019

Accepted: 04 February 2020

Published: 02 March 2020

Citation:

Loewy J and Jaschke AC
(2020) Mechanisms of Timing,
Timbre, Repertoire, and Entrainment
in Neuroplasticity: Mutual Interplay in
Neonatal Development.
Front. Integr. Neurosci. 14:8.
doi: 10.3389/fnint.2020.00008

INTRODUCTION

Recent research has evaluated the effects of music training with benefits that have attributed music learning as a significant contributor to the improvement of cognitive function (Barrett et al., 2013). In particular, music's implicit audio-structural capacity (Lordier et al., 2019) seems to reflect notable influence in both developmental outcomes and restorative capabilities that stretch across the lifespan. While numerous studies have focused on outcomes of young children (Trainor et al., 2012; Moreno et al., 2011) and aging adults (Johnson et al., 2011), few have considered the impact of the pre-natal 'sound' environment in terms of ritual and context, nor the culturally relevant conditions and applications that organized purposeful sounds and music may stimulate. Elements of sound in and out of music and their structured or disorganized contexts can be considered as disruptive or as viable enhancements to the advancement of the Neonatal Intensive Care Unit (NICU) experience (Stewart and Schneider, 2000; Rossetti and Canga, 2013), particularly as the conditions of the infant's brain in its most vulnerable state are at stake. The impact that the physiological elements of music, silences and sounds, can make in the earliest stages of brain development is worth consideration. This article will support research that has hypothesized conditions of music therapeutic applications validating models of neurobehavioral care that have provided proven optimal conditions for growth inclusive of recommendations leading toward the enhancement of self-regulatory behaviors. Developmental markers based on research and clinical experiences related to current quantitative data of musical elements have been sparse. Clinical experiences substantiated through the First Sounds: Rhythm, Breath and Lullaby research (Loewy et al., 2013)

that lead to a model of practice and training will be explicated within this article (First Sounds, 2016). Specifically, we will address the optimal contexts for the essential foundational relationships between infants and caregivers based on music's neural plasticity factors. These can be enhanced through attachment conditions which are fortified through music to address its relationship in developing executive functions (EFs) across the lifespan (Sala and Gobet, 2017). Elements contributing to this context include live vital sign entrainment, social interaction *via* attunement, rhythmic assessment with the institution of indicated repertoire, and sleep/wake patterns inclusive of aural transition capacity.

FACTORS OF INFLUENCE

Genetic Predisposition

Multiple influences such as auditory, visual or somatosensory factors, show a trend in influencing brain dynamics in young clinical populations, and especially in prematurely born infants in the NICU. These factors have the tendency to translate into later childhood and even adulthood expression (Galvan, 2018). Some of these trends can be seen as factors of influence in cognitive and overall neural development. Neural differences in social cognition, which can be triggered by hormonal changes, or by experiences of an “unnatural” sensory environment to the infant, have been described (Galvan, 2018). Their influence requires more research, as larger cohort studies are still scarce.

Neonatal growth relies on a combination of critical factors that contribute to brain development. In our experience of studying neonates' responses to music therapy stimuli, and in working within dyads and triads involving father, mother, staff and neonate, the factors of influence where music's potential might enhance developmental prowess are worth defining. This is because the neonate's sensitivity to sound and music is distinctly and definably related to neurologic outcomes. These factors are inclusive of genetic predisposition (Montirosso, 2015) and the conditions of the pre and postneonatal environment. There are distinct periods of infant development whereby a genetically determined microcircuitry of key limbic-hypothalamic-midbrain structures are particularly sensitive to early environmental contexts. These influences directly formulate an individual's responsiveness to psychosocial stressors as well as the resiliency or susceptibility to difficult circumstances in later life circumstances (Leckman et al., 2004).

Caregiver Capacity for Secure Attachment and Aural Sound Conditions

Developmental factors are typically attributed to a caregiver's knowledge and capacity for creating opportunities for secure attachment (Lomanowska et al., 2017). Emerging research undertaken first by Bowlby (Bretherton, 1992) and subsequently his student, Ainsworth emphasized the significance of secure attachment, and its relationship as a most critical interplay between the biological and environmental conditions of two systems, that of the mother (parent) and of the infant (Bretherton, 1992; Zimmer, 2019).

This interplay, coupled with an aural sensitivity (Lahav and Skoe, 2014) within a context conducive to optimal sound conditions can foster neural organization. Aural sensitivity is the intricate perception and processing of sounds. These cover a variety of different frequencies and can be traced back to, however not exclusively to, the tonotopic map of the primary auditory cortex. Moreover, sensitivity to sound reaches further than the primary auditory cortex and can have implications beyond the temporal lobe resulting in changes in behavior (avoidance, interest) or emotional responses (e.g., pain).

One study used a repeated-measures research design to examine variables of neonates recorded every 2 min through two 60-min observation periods for each research day (1 h in the morning and 1 h in the afternoon) over 2 days. Thirty-seven preterm infants provided a total of 4,164 observations that were made rendering a statistically significant ($P < 0.05$) relationship between environmental stressors and changes in physiological signals. There were also statistically significant ($P < 0.05$) relationships between environmental stress and specific stress behaviors. Physiologic stress responses to acoustic events such as changes in heart rate, intracranial pressure, and oxygen saturation may have a significant impact on the preterm infant's future neurologic development due to altered perfusion and oxygenation of the brain tissue (Peng et al., 2009). These changes have been shown to influence neurologic development. This study has shown a significant correlation between environmental stressors and physiological signals of bio-dynamical changes.

Trauma and Pain

Another factor that has moved in and out of the literature base encasing elements respective to the aforementioned conditions, is related to the concept of trauma. There is a suspected neurobiological mechanism by which maternal unresolved trauma can modify maternal caregiving so as to disrupt the normative development of an infant. This was highlighted in a recent review culminating a plethora of research indicating the impact of childhood adversity and its effects on parenting, which may be perpetuated inter-generationally (Kim et al., 2018).

Music therapist Kristen Stewart developed a model for treating trauma with premature infants and parents-entitled: Preventative Approach to Traumatic Experience by Resourcing the Nervous System (PATTERNS). Instituting music therapy in a preventative treatment model that is based on latent human resiliency and trauma renegotiation principles, Stewart's model (Stewart, 2009a) involves integrative mechanisms that influence perception, nurturance, and opportunities for building resilience at all levels within—infant-parent—and staff contexts utilizing music therapy.

An additional factor that affects neonatal neural activity involves the distinct formation of neuronal connections, including dendritic spine development, synaptogenesis, axon myelination, and crucial cortical folding of the neocortex (Dubois et al., 2008). Importantly, developmental changes are not wholly determined by intrinsic prenatal biological factors, but by postnatal experiences as well. While axonal guidance and myelination, and axonal elements are thought to be influenced

by genetics, the distinct pathways of connections that are retained rather than pruned are within the axon myelination process as well and are sensitive to the sound environment (Benasich and Ribery, 2018). This brings to the forefront of the realization that neonatal neural activity is affected by pain (Anderson et al., 2011). The younger the gestational age of a premature infant's receptive fields are the more vulnerable the effects of pain and in this way, the result of pain can involve hypersensitivity (Fitzgerald and Walker, 2009). Since it is known that sensory afferents are not "hard-wired," but plastic (Ren and Dubner, 2007) and given the wider receptive fields and hypersensitivity of preterm neonates, the pain they experience in this vulnerable time may have profound long-lasting effects. Primary hypersensitivity is more easily provoked in preterm neonates (Johnson et al., 2011). Attention to the neonates' early experience of pain has ramifications for plasticity and neural development. A recent study utilized live lullaby singing (song of kin-as selected from parental preference), for the first time on behavioral and psychological pain responses during venipuncture of 38 preterm and full-term infants with some interesting findings towards higher oxygen saturation and calmer heart rates in the intervention groups as compared to the control group. This could be an important development for future neonatal and young infant pain protocols (Ullsten et al., 2017).

Pre-mature Brain Function

Our knowledge of brain functions during the perception and processing of music has increased in the past decade (for review, see Jaschke, 2019). Modern technological advances such as functional Magnetic Resonance Imaging, Diffusion Tensor Imaging, and functional Near Infrared Spectroscopy, have made it possible to track and analyze music multisensory stimulus as it activates several brain regions in its path (Jaschke, 2019). Music is comprised of complex stimuli that activate regions beyond the auditory cortex. Among others it activates the thalamus, hippocampus, and temporal as well as frontal regions; all of which are known to have additional primary functions such as language processing, EFs, and associative memory.

Recent developments in music and brain research justify music as a multi-sensory stimulus and the influence of its multisensory integrative properties on dynamic brain networks (Barrett et al., 2013). There is still more research needed to investigate these properties in neurotypical pediatric as well as adult brain function. Multisensory properties are even less studied in the neonatal/preterm brain and will certainly assist in our potential to develop formative hypotheses that may well have bearing on all aged populations. Neonatal brain access allows for the study of neural plasticity in light of multisensory integration in real-time.

The multiple brain areas involved in the perception and processing of music make it a potentially interesting clinical tool in NIC (Sa de Almeida et al., 2020). Throughout the development of the brain, coupling and synchronization of different brain areas are thought to mediate between the assembly and development of neural networks, which in turn contribute to the coordination and plasticity of the brain as well as play supporting roles in sensorimotor functions

such as cognition including perception, language, learning, executive functioning and memory (Benasich and Ribery, 2018; Sa de Almeida et al., 2020).

Concerning sensorial influences in the NICU, a specific interrupted process weighs upon auditory brain development, which starts early in gestation. By 23–25 weeks, all structures necessary for hearing, among which the cochlea, have developed (McMahon et al., 2012). As such, most infants admitted to the NICU are able to hear, unless they have a congenital anomaly.

From approximately 26 weeks' gestation, fetuses or preterm infants will have the capacity to react to auditory stimuli. Sounds a fetus hears within the womb include a mother's heartbeat, respiration and the maternal voice, and show recognition to the mother's voice, and in some circumstances, father's voice as well (Lee and Kisilevsky, 2014). From 30 weeks' gestation onward, the infant is able to distinguish between varying speech tones and timbres and is also able to process complex auditory sounds. This point likely marks the start of speech and language development (McMahon et al., 2012).

In preterm birth, this process is interrupted. On the one hand, this means that infants are exposed to noise in the NICU that they may not yet be able to process, which could be harmful and has been shown to alter respiratory and cardiac functions (Sa de Almeida et al., 2020). On the other hand, the deprivation of sounds heard *in utero* could have consequences for auditory brain maturation affecting speech and language development. A recent study of 136 (<30 weeks) infants hospitalized in a quiet private room NICU environment vs. an open active ward showed lower language and motor scores and a trend towards lower cerebral maturation at follow up- 2 years of age (Pineda et al., 2014). In a second study, a literature review of 88 studies inclusive of measurements related to the auditory environments in the NICU sought to determine medical, environmental, and socio-demographic factors that predict and further define influential protagonists and antagonists related to infant auditory exposure in the NICU.

This review stratified succinct elements of the sounds in both conditions, the open and private room NICU-such as Noise, Mechanical equipment, meaningful words, electronic noise, adult words, distant words, et cetera (Pineda et al., 2017). It was found that there was an average of 3 h more silence in a 16-h period in the private room-that equated to roughly 30% more silence than in the open ward. This study, among other reports (Jobe, 2014; Webb et al., 2015) makes a case for the importance of low-frequency sound exposure for neonates. At the same time recommendations on the importance of all NICU sound environments whether in open or private spaces should seek to define the specific rationale for implementing any specific modality, inclusive of the timing and frequency considerations of sensory-based interventions. The goal should include strategies that optimize and inform any environmental modifications.

Examining infant sound exposure and seeking opportune experiences for premature infants inclusive of music will provide assurances that they will have exposure to meaningful words (McGrath, 2013), inclusive of language or singing that is infant-directed. Sound making and holding can also provide optimal conditions for development (Teckenberg-Jansson et al., 2011).

This delicate balance of activity, inhibition and multimodality integration prompts our capacity to make sense of the world around us in an attempt to perceive through multi-signaling. These dynamic processes develop over time through learning, exposure, and experience existent through until adulthood and beyond (Freeman, 1997). Such mechanisms are part of a complex and demanding process, which can place a fair amount of strain on any neurotypically developing infant, child or adult. Neonates and especially preterm born children face aside from the aforementioned difficulties, additional challenges as their brains and bodies encounter the complexity of these developing networks prematurely (Arnon et al., 2006).

After birth, the brain not only develops as a result of internally driven mechanisms alone. Neural circuitry is shaped by external factors and biogenetical mechanisms as well. The impact of external factors on a not-yet-ready-brain can, therefore, have a significant impact on the development of a child during both childhood and adulthood (Bieleninik et al., 2016; Provenzi et al., 2018). Considerations taken between over and under stimulation when it comes to understanding and supporting the development of a prematurely born infant brain have been scrutinized historically and currently. In order to be able to stimulate the brain to provide the infant with the best possible start; sounds and voices can be used as therapeutic stimuli (Loewy, 2015). The signature sound of the womb is one that an early pioneer, Van de Carr and Lehrer (1988) examined and notoriously named a “prenatal university.” These sounds are thought to be the most critical aspect of co-regulation for the neonate and mother, and subsequently family and community adaptation (Van de Carr and Lehrer, 1988). The mother’s heart rhythm, and her voice pitch, tone and rhythm are hallmarks of an infant’s first auditive “memory” prior to birth (Webb et al., 2015). Taking this a step further, in pro-active music activity with infants, exaggerated speech, such as singing, where the tones, pitches and intervals are purposefully arranged and orchestrated to meet the expression of the neonates, enhanced responsiveness develops. The inclusion of fathers is advisable. Breathing rhythms and tonal/vocal expressions become an active means of co-regulation and form the basis of expressive language and emotional attachment.

Stimulating the premature brain in this stage with entrained sounds, language, or music has to be implemented and coordinated with conscious effort and rationale, as the cortical folding is still in flux and a great number of neurons are still migrating to their final cortical destinations (Provenzi et al., 2018). Applied cautiously and consciously, parental and therapeutically trained voices and music can enhance neural development and stimulate neural plasticity, which in turn has the potential to influence cognitive function, brain oscillation and stimulate dynamic brain plasticity and network formation, in short as well as long-term development (Bieleninik et al., 2016; Provenzi et al., 2018; Sa de Almeida et al., 2020).

Music and Music Therapy as Neural Stimulation

Early in gestation—at around 16 weeks, the fetus’s auditory system is formed. Between the 26th and 30th weeks, the fetus

is able to detect and react to sound stimuli. This period in the womb is considered to be a critical period for neurodevelopment (Benasich and Ribery, 2018). Vandormael et al. (2019) discuss how the “fine-tuning process takes place in the uterus where both internal (e.g., respiration, heart rhythm, and digestion) and external sounds (e.g., voices and music) can be perceived” (Vandormael et al., 2019). They also cite significant recent research in numerous NICU conditions showing evidence that too much chaotic noise or not enough sounds, largely people’s voices, may have detrimental, or deprived conditions which can show up in delayed language capacity in toddlerhood.

In a recent study, of 272 infants born prematurely, and serving as their own controls within a randomized 2-week intervention period in 11 hospitals, when exposed to entrained music therapy conditions, either in parental or music therapist singing conditions, vital signs improved. Implementing song of kin; parent-selected familiar melodies, or improvised to create in-the-moment, meaningful lyrics, within a simple melodic construct, or using womb sounds simulated through the use of a quietly entrained Remo ocean disc, sleep patterns were enhanced, and heart rate patterns were more even (regulated) and soothed (promoting sleep; Loewy et al., 2013).

The implementation of entrainment conditions, confirm and substantiate that live music in a music therapy context is best indicated. Live music ensures feedback that makes for reliable patterning of a sound-based relationship and a safe environment as music is regulated, in real-time, at the moment. The implementation of entrainment conditions, confirm and substantiate that live music in a music therapy context is best indicated. Live music ensures feedback that makes for reliable patterning of a sound-based relationship and a safe environment as music is regulated, in real-time, in a shared moment. In a recent review, results of 512 preterm infants among 15 recent clinical trials using live and recorded maternal voice interventions showed fewer cardiorespiratory episodes and significant improvement of physiologic and behavioral conditions (Filippa et al., 2017).

Examining the mother’s voice, and recordings of her heartbeat sounds have shown auditory cortex improvement (Webb et al., 2015), at 1 month in 40 infants born extremely prematurely (between 25- and 32-week gestation). Newborns were randomized to receive auditory enrichment in the form of audio recordings of maternal sounds (including their mother’s voice and heartbeat) or routine exposure to hospital environmental noise. They were shown to elicit strengthened auditory plasticity at 1 month compared to their controls. This study again used recorded sounds that were static, and not meeting the meter or condition of the premature infant’s signals in the here and now. However, the organized sounds (music) and familiarity rendered better outcomes when compared with noxious hospital noise.

A recent study of five infants with congenital heart disease explored entrainment on the physiologic measurements of withdrawal through live singing and guitar accompaniment. This single case withdrawal pilot study examined the effects of music therapy “entrainment” on heart rate, respiratory rate, blood pressure and oxygen saturation rate, of five infants suffering

from congenital heart disease, in the cardiac intensive care unit. Receiving music therapy “entrainment sessions” before and after heart surgery, and consistently, 3–5 times a week for up to 3 weeks, their physiological measures were recorded every 15 s after the music therapy intervention began, until 20 min after the intervention was complete. Although the outcomes showed improvements when “entrainment” was used from baseline to follow up, the songs implemented, when not informed by parents did not necessarily consider range, melodic content, nor musical elements. The entrainment was matched with the guitar accompaniment (many strings) rather than the singing (single-toned phrase). A metronome was instituted, which produces a static rhythm, likely not providing sensitivity to the actual shifts of heart rhythms that one would advantageously follow to match the infants’ rhythmic heart rate changes moment by moment (Yurkovich et al., 2018). Even so, the study had an 80% success rate, whereby four of the five infants experienced decreases in average heart rate and respiratory rate and improvements in the derivative of the heart rate signal as well.

Music and Executive Functioning

Studies addressing very early exposure to music in (extremely) preterm born infants and a possible outcome-transfer to long-term effects on cognitive functions such as EF or learning are rare. Therefore, it is crucial to understand *music* itself and how it affects, touches and possibly influences brain and cognitive development, in a broader context. Being able to trace musical stimuli through the brain and linking this activity to a possible cognitive effect provides a means toward understanding neural development across the life span. Research on the development of the human brain has indicated that the brain reaches its full level of maturity at an age of around 30 years (Tomlinson, 2015).

Within these crucial years of development, the frontal lobe is one of the last to reach its full maturity (Tomlinson, 2015). With the prefrontal cortex being the primary seat of EFs, correlated with a possible influence of music (a combination of listening, playing and improvising), this area has the potential to influence the developing neural structures through the very complexity of the music itself and the conditions within which it is provided.

Processing music engages long fiber tracts in the brain and overlaps with regions responsible for EFs, and we expect a stimulation of EF when perceiving and performing music (Klein et al., 2016). Understanding the influence of music on EF can, in turn, provide insight into pathologies later in life, which manifest in a form of executive dysfunction. Music therapeutically informed interventions hold the potential, to stimulate the brain, supporting processes of neural plasticity, and can offer an enhanced start into life for (extremely) prematurely born infants, who have entered this world with so often physical and cognitive disadvantages stemming from birth trauma and/or deprivation.

Music Learning and Brain Function

Early exposure to music can have strong influences on cognitive reserve and development later in childhood (Jaschke, 2019). As

there is limited research on the development of EFs in preterm born children (regarding their age for neuropsychological testing) we have reverse-engineered out arguments from the literature which indicate improvements in executive functioning in relation to music training and exposure (Moreno et al., 2011; Degé et al., 2011; Jaschke et al., 2018a,b). This approach affords the tracing back of steps potentially contributing to neural development as informed by neural plasticity, utilizing large published data sets; a method that borrows from historical and anthropological research. We can, therefore, build the argument that by understanding EFs in “neurotypical” or so-called healthy populations we can trace this development back to influences shaped in an NICU environment (Sa de Almeida et al., 2020).

Frontal brain regions mainly process anticipation and expectation and the execution of musical thought during the event of listening and execution. Furthermore, they relate to music improvisation, which also relates to the amygdala and the hippocampus. Of note is the dorsolateral prefrontal cortex (dlPFC), which plays a crucial role in improvisation as well as in learning and memory, connecting deeper brain regions with frontal regions *via* the thalamocortical-thalamic loop. When it comes to music learning and its likely associated executive functioning, it is crucial to examine the difference between music listening and music playing. Both listening and especially playing music, activate a wide array of brain areas. These neural activity networks, however, should be linked in kind to cognitive processes. Understanding the differences between mere exposure to music and playing music, within a collaborative signal-reading environment, can have a significantly different outcome, especially as studies propose causality between music and far transfer domains (Jaschke et al., 2018a).

Linking back to prematurely born infants, to have a deeper understanding of the impact of musical learning and its capacity to affect learning in neurologic development, it may be useful to distinguish four modalities of music participation: (1) passive listening; (2) active listening; (3) music-making; and (4) improvisation. All four, even though consecutively building upon each other, translate uniquely and distinctly to a possible effect on cognition and should, therefore, be considered as influential when analyzing the effects of music on brain function, emotion and behavior.

In premature infancy, the listening conditions might translate well within a quiet or active alert state context, and the musical vocalizing might translate to a contingent singing (Malloch et al., 2012; Shoemark, 2017) or infant-directed singing (Mehr and Krasnow, 2016) exploration. In such contexts, the in-the-moment improvisatory experience might be indicative of a developing repartee and one whereby the music therapist is singing back the premature infant’s vocalization, strategically on the exact pitch of the infant’s tone. Eventually, the premature infant might absorb the vibration percept and create a new tone or an interval of two tones outside of the tone of the formerly set condition. This may unfold without prompt and occur seemingly suddenly, such as a crying sound with accented phrasing, or a comforting sound wherein the rhythmic meter of the plosive mouthing sounds indicates exploration (Loewy, 1995).

MUSICAL PARAMETERS

Timing

A recent study of 35, <32 weeks GA neonates revealed that those who listened to 8 min of pre-composed music five times per week, based on nursing assessment of developmental need (assessed on a neonatal behavioral assessment scale), timed to be offered at the moment and distinctly related to need of sleep, wakefulness-interactive, or at a time of alertness, such as prior to feeding, led to functional brain architectures (shown through fMRI) that were more similar to those of full-term newborns (Lordier et al., 2019). This provided their evidence for a “beneficial effect of music on the preterm brain.” This example of identifying the timing of a music intervention is quite indicative of sensitivity to best practice needs and might provide a unique platform for the inclusion of the infant’s state and readiness for music, suggestive of a dynamic aspect of signal reading and its importance when providing intervention for the music applied.

However, the music intervention, and its meter and instruments “composed of a soothing background, bells, harp, and punji (charming snake flute) five times per week from a GA (gestational age) of 33 weeks until the MRI” may not provide aspects related to the most efficacious provisions of the music itself. Unfortunately, readers are not provided with the music applied-and when one seeks to gather the evidence and arrives at the website¹, the original music provided for this referenced study is not easily found.

Even if it were easily accessed, as a link on-line, because it is an obstructive variable and one that is not reliant on the infants’ vital signs, it is likely not best-indicated. Utilizing a recording, the conditions of timing, so central to a preterm infant’s neural plasticity cannot be sensitively provided as recorded music does not provide a platform for the entrainment features most critical to the infant’s neurologic brain effect. The results for the most ultimately safe and nurturing environment might include the mechanism of entrainment, whereby the parent and/or music therapist (whose training in NICU music therapy) is inclusive of the assessment of the infants’ vital signs and whose timing of intervention is inclusive of the rhythm and melody relevant to the parents’ cultural background/preference, or adapted to be inclusive of musical elements related to these vital conditions. Such cultural aspects necessarily determine whether the music that may be best indicated is meter-less, such as, for example, an Indian raga, or a classically formatted work, such as a baroque piece, or soft rock-set, or folk song, perhaps phrased in 3/4 or 6/8 as a ballad, or in lullaby style. Knowing the music and composers that the infant heard *in utero* and presenting this music live, and eventually constructed within a suitable framework that parents can utilize as their “song of kin” (Loewy, 2015) may ensure continuity of care (by parents, in NICU, and at home post-discharge), with enhanced reliability that might influence the potential for bonding as an incentive for what some researchers have studied to be a “neurologic brain effect” in premature infants. The case for timing of when music is provided, and the timing of the music itself (time signature),

and how it is applied (entrained) within its natural meter of culture may be the best-indicated elements one considers for making music therapeutic interventions that seek to strengthen the neurological development of plasticity in premature infants.

Timbre

The timbre element of sound and music is something rarely discussed in health music applications (such as neurology, music medicine and music therapy). Two innovative NICU studies using recorded music did include the actual womb-sounds of the placenta, combined with female voices. With attention to timbre, these studies incorporated a music program inclusive of the sounds of an intrauterine maternal pulse along with synthesized female voices at 65 db. Created by Fred Schwartz, an anesthesiologist, the program was named “Transitions” (Placenta Music, Inc., Atlanta, Georgia). The first study of this music’s effects included four premature infants with Bronchopulmonary dysplasia (BPD) on continuous ventilator support. Each infant was exposed to three different 15-min conditions: (1) music played *via* a Somatron mattress; (2) music played *via* a recorder at the foot of the infant’s crib; and (3) silent isolation room condition. This study included providing music immediately following suctioning. The conditions were counterbalanced and repeated so that each infant received 18 interventions in the following order ABC-BCA-CAB-(etc). They looked at heart rate, oxygen saturation, arousal state, facial expressions, limb movements, and autonomic states all within the 13 s (with 7 s to recover before starting again), every 15 min for 18 trials. The time period of the data collection ranged from 8 to 21 days, every 13 s within a 15 min period for 18 trials. They collected in total 4,860 pieces of data per infant. Oxygen saturation and sleeping time improved significantly during the interventions (Burke et al., 1995).

The second study utilizing Schwartz’s placenta soundscape program was also incorporated in another infants-as-their-own measurement design (Chou et al., 2003). In 30 premature infants with respiratory distress requiring endotracheal intubation and ventilation and endotracheal suctioning, this second trial showed statistically significant higher oxygen saturation during recovery periods compared to the controls, who did not receive the “Transitions” soundscape. Statistical significance was reflected in the period oxygen saturation was achieved, with quicker recovery time when this music was applied. This procedure has potentially disturbing side effects such as dysrhythmias, oxygen desaturation, cerebral blood flow fluctuations, and laryngospasms. Familiar timbres can increase tolerance and build resilience. The potential of watery timbres, reminiscent of womb sounds and their effect on vital sign enhancement should not be overlooked.

Studies of older infants have shown that the mother’s voice, is recognizable to the infant who has heard it distinctly and repetitively in the womb pre-birth (Thompson and Trevathan, 2009). Greater attention taken has achieved the conditions of timbre that replicate a womb-like timbre experience inclusive of water or breathing sounds. These are effectively metered to the premature infant’s rate of breathing. This can occur during kangaroo or skin to skincare, whereby a parent holds their

¹<http://vollenweider.com/en>

premature infant over their heart (on the left side, Salk, 1973), and can readily pay attention to their infants' breathing patterns.

The perception of timbre, recruits multiple brain areas, including regions in the frontal lobe, the thalamus—as central relay and multisensory integration region, the hippocampus, amygdala and especially the planum temporal and the temporal plane (Jaschke, 2019). The neural function of perceiving timbre is a combination of pitch decoding, pitch identification, the spatial location of sound and the “color” (warm, cold, squeaky, round) of the perceived sound.

An audible “ah” sound will not only provide a timboric effect of vibration but additionally will offer a means of breath connection, particularly if the parent or music therapist entrains the application of their self-breathing to the meter of the infant's breathing, as in “tonal-vocal holding” (Loewy et al., 2013). One early study inclusive of entrainment (Ingersoll and Thoman, 1994) included a devised “breathing bear” contraption that could reflect the breathing rate of the infant in proximity. The study included 36 premature infants; half were randomly assigned the breathing bear device vs. the control, a breathing bear device that did not entrain breathing sounds. At 35 weeks, the intervention infants showed slower and more regulated respiration during quiet sleep. At 45 weeks, these same infants showed more quiet sleep and less active sleep. At both ages, only the intervention infants showed a correlation between respiratory regulation and the amount of quiet sleep (Ingersoll and Thoman, 1994). Sleep regulation is a critical mechanism of strong neurologic function that can be influenced by music (Loewy, 2020).

The findings show an effect of positive neurobehavioral development. The author's credit “entrainment” effects related to optional stimulation. Central to this effect was the reflecting of the infant's own biological rhythms activated and directed by the infants' rhythms. Of note, this timbre sound was a water sound set by a machine. It is likely that when a live human being is brought into the response, the essential interplay lays a foundation for not only neurological function but as one of attachment as well.

Repetition: Predictability as Enhancement in Neonatal Music Therapy

The human brain as it formulates structure in the womb is reliant on the rhythmic conductor of another; that of the mother's heartbeat. This “conductor” is variable, and the fetus necessarily regulates growth in accordance with a wide range of meters based on the mother's activity level.

The practice of ritual in patterning music and rhythms of daily life inclusive of the sleep-wake cycles of developing infants is critical (Loewy, 2020). It is known that infants recognize familiar voices and melodies, and language patterns. Repetition, in particular as part of a music gestalt, provides space for interaction and safety within the music's spaced cue, as in a breath before a downbeat, or within a final repetitive cadence.

It is best practice to engage premature infants with music that has predictability. One study showed that young infants engage more rhythmically to music when compared to speech and that the cued timing related to the effectual mood of songs vs. speech was linked to greater rhythmic coordination (Zentner

and Eerola, 2010). A former study (Loewy et al., 2005) comparing chloral hydrate with parents' songs of kin arranged in lullaby format, incorporating repetition, lead toward a quicker, more profound sedation effect than pharmacologic sedation for infants and toddlers requiring EEG.

Melody-Simplicity of Range and Recognizable Intervals

It is important when instituting music to seek the melodic organization and sequences of songs that are familiar to the neonate. If parent/s are unavailable, the range of the mother and father's voice is good to render when providing interventions (eg soprano, alto, tenor, baritone, bass). The “song of kin” (Loewy et al., 2013; Loewy, 2015) is inclusive of a model that assists parents in creating a natural, easily accessed sung melody, and most particularly one that has meaning for the parent. The song's significance and relevance can be part of a music psychotherapeutic process. It does not have to be a melody that the neonate heard *in utero*. It does not have to be related to spiritual or historical aspects one the infants' family. The song should aim to be a favored song-culturally relevant to parental preference, and the melody can be entrained and best applied with a simplified single-line matching of sung vocal phrasing. Accompaniment, if and when applied, can be minimal. A Capella is well suited for neonates.

The primary auditory cortex and secondary auditory cortex are the primary preceptors of the incoming auditory information. Wernicke's area and Heschl's gyrus, process mimicking and associations, together with pitch intervals and melody (Jaschke, 2019). Music therapy literature has advised on important considerations to keep in mind when implementing best practices of musical melodies for neonates. Notwithstanding cultural nuances, it is useful to consider aspects of caution and simplicity when developing melodies with infants and caregivers: “Slow tempo, simplicity, minimal number of instruments and harmonics, quiet and stable dynamics: decibel levels not greater than 60–65 dB (A-weighted scale), repetition and consistency; rocking meters, one-octave tonal range maximum, unidirectional melodic contours, with limited changed in pitch direction; emphasis on descending tones to engage relaxation response (Stewart, 2009b, p. 127).

Song of Kin: Affective Culture as Attachment Incentive

Perceiving, processing and executing music recruit numerous neural areas. In combination, the varying facets of music such as rhythm, melody, and timbre are projected to distinct areas of the brain, which decode the stimuli to create what we understand as music. These areas, however, are not exclusive to music. The temporal, as well as frontal lobes that are equally involved in language and arithmetic, are also overlapping within the music-related regions (Jaschke, 2019). Each of the regions includes its own interpretation of the task at hand and therefore stimulating these, increases more than the actual understanding of the task, but can be transferred to other mental exercises such as an increase in empathy or working memory, which share overlapping regions.

The “song of kin” importantly, it is a song that is set to a simple construct, whereby a holding condition or lull (such as 3/4 or 6/8) is implied through an easily accessible form for the parents. It is provided as a developmental tool inclusive of attunement (Bowlby, 1998; Ainsworth, 1989) strategies in the growing relationship between parent and infant. It can be used when an infant is seeking connection-so as to build attachment, in an improvisatory fashion, for instance, such as in the key the infant is cooing within (melodic entrainment). It can be used to calm when an infant is fussy (breath entrainment). It can be used as a ritual, providing the safe assuring conditions of sleep (looping and repeating the cadence to sustain a “cadential effect” (Loewy, 2009, 2020). In the instance of enhancing sleep, the lyrics can be removed, and vowels will elongate melodies and suggest a slowing down of the mind-body-attention.

With so many play songs and activating conditions on the current market related to musical circumstances available for infants and parents, it is perhaps hard to believe that the most nurturing source and use of music can emulate from the parent uniquely, and be formulated to fit the conditions of the infants’ ever-changing physiological state. This was highlighted in a recent study on vocalizations of parents and its influence on eliciting sound-making in preterm infants (Caskey et al., 2011).

As infants do not speak, touch and emotional closeness expressed through vocalizing can forge the seeds of trust and musical repartee (Goulet et al., 1998). The absence of such opportunities within the critical first days of life can increase prenatal anxiety in the NICU (Franck et al., 2005). As musical expression enhances participatory and opportune moments for inter-relating, enhancing a parent’s confidence is essential for bonding. Maternal sensitivity in mothers with preterm infants is less optimal when compared with full-term controls (Forcada-Guex et al., 2006) and early separation from an infant at birth in one study was related to an increase in parents’ NICU-related stress (Franck et al., 2005).

Attachment can be the most intimate and personalized condition of mutual interplay (Ainsworth, 1989) when it is enriched with musical conditions that are personalized. Using the infant’s name and creating circumstances whereby a parent’s singing becomes strengthened and more intense within a response (Loewy, 1995, 2015), such as a silent moment and singing a tone at the same pitch as the infant, or hiding the face and finding it when the infant creates a sound are part of a personalized connecting music soliloquy that prolongs interplay and sustains attention and expectancy factors which are the seeds of brain development.

This is not an instantaneous condition. The experienced clinician or therapist with advanced training deemed fit to work with parents understands that fragile infants have fragile parents (Zimmer, 2019). It is therefore imperative that the music is not simply implemented as a condition that a parent is told is good for the infant. Work in music therapy with mother (Arnon et al., 2014), fathers and families can lend integral inclusion factors that will take a critical community role within developing families (Ettenberger, 2017).

Als et al. (2012) were among the first to follow preterm infants who participated in an individualized developmental

approach to neonatal care-currently named the Newborn Individualized Developmental Care and Assessment Program (NIDCAP). The infants provided with NIDCAP showed improved electroencephalogram coherence and more mature frontal brain structural development as evidenced by magnetic resonance imaging at term-corrected age. There were improved neurobehavioral outcomes at 2 weeks and 9 months corrected age (Als et al., 2012).

Moreover, the RBL model includes the parent prong as one of the most essential treatment areas in working with families in the NICU. In playing for and playing music with parents who experience the circumstances of premature birth, opportunities for soothing and nurturance encourage easeful feelings of release and eventually of empowerment (Haslbeck and Costes, 2011). Acknowledgment of the difficult circumstances can circumvent rage, shame, doubt and self-mistrust that are common with premature birth. The provisions of music therapy for parents will likely lead to better outcomes for their capacity to use music in their new relationship with their infant (Loewy, 2015). The effects of trauma on the premature infant in terms of its implications within the developing relationship with the parent/s (Als, 2009) is an understudied aspect of arrested or delayed capacity of neuroplasticity (Stewart, 2009a). Stewart addressed the potential traumatic experiencing for infants and parents and recommended the utilization of a preventative music therapy treatment, inclusive of staff, and based on resiliency and renegotiation principles. The six phases of the model are: (a) stabilization; (b) self-regulation; (c) integration of experience/resolution of traumatic memories; deconditioning; (d) establishment of secure social connections: repair and/or development of effective attachment and reciprocity; (e) accumulation of restorative emotional experience; (f) future planning: development of self-care plans and goals (Stewart, 2009a).

Working with trauma in the NICU requires advanced training, such as the incorporation of Peter Levine’s Somatic Experiencing, or of Stephen Porges’ Polyvagal theory. Aspects of these orientations have been incorporated into the three prong (First Sounds, 2016) music therapy advanced training for NICU music therapists.

CONCLUSION

The first 1,000 days from conception to delivery are among the most important periods in the development of an individual. During this time, the infant brain is rapidly developing and is extremely sensitive to environmental influences. For preterm infants, born before a gestational age of 37 weeks, and specifically, those born very preterm (i.e., before 30 weeks’ gestation), the period that should have been spent in the womb, is interrupted, often suddenly. Preterm infants often undergo trauma, that may be experienced by their parents (Aagard and Hall, 2009) as they, at the same time, are subsequently admitted to the NICU, where they face many challenges. Among these challenges is the stress of physical and sensorial influences coupled with maternal separation.

Music as an often complex and cognitively demanding modality can be over-stimulating particularly when considering the fragility of the neonatal brain developmentally. Therefore, a conservative “less is better” minimalistic approach to the music instituted is recommended. Opportunities for music therapy assessment that implement one stimulus at a time, with sensitivity to rhythm, decibel level, proximity to the infant (at mid-line to encourage fetal positioning) is indicated. Entrainment conditions where the parent/therapist is trained to follow the infant’s cues and clues, with breath (air/water timbre qualities), and melody, with small, repetitive short phrases comprised of the cultural sequences and sung by familiar voices have been substantiated (Loewy, 2015; Mondanaro et al., 2016).

Opportunities to treat fragile parents directly with music therapy relaxation techniques may provide a wealth of physiological sensorial conditions that in turn, may be sensitively and naturally translated to their infants. Adherence to reading the signals of the neonates’ readiness or disengagement, inactive alert, quiet-alert, or sleep state preferences will provide key information related to the most opportune times for live music that can stimulate interactive activity, or alternatively induce restful transitions for sleep that will ultimately enhance

neonatal brain function. The clinical applications of music are best integrated when elements of decision-making related to musical choices are substantiated by cultural, psychological and neurological mechanisms. In this way, clinicians, caregivers and researchers alike can incorporate elements of music that will assist in regulating, enhancing, adapting and fostering the music to meet the cues and clues of the ultimate conductors—the premature infant and parent/s we serve.

AUTHOR CONTRIBUTIONS

JL conceived the manuscript concept and created the outline, authoring the introductions and body-support of the hypotheses, based on research on neonatal music therapy in Mount Sinai Beth Israel. AJ co-contributed to the neuroscience theory, enhancing aspects related to music mechanisms in neural plasticity.

FUNDING

The Louis Armstrong Center for Music and Medicine Mount Sinai Health System.

REFERENCES

- Aagard, H., and Hall, E. O. C. (2009). Mothers’ experiences of having a preterm infant in the neonatal care unit: a meta-synthesis. *J. Pediatr. Nurs.* 23, 26–36. doi: 10.1016/j.pedn.2007.02.003
- Ainsworth, M. (1989). Attachments beyond infancy. *Am. Psychol.* 44, 709–716. doi: 10.1037/0003-066x.44.4.709
- Als, H. (2009). NIDCAP: testing the effectiveness of a relationship-based comprehensive intervention. *Pediatrics* 124, 1208–1210. doi: 10.1542/peds.2009-1646
- Als, H., Duffy, F., McAnulty, G., Butler, S., Lightbody, L., Kosta, S., et al. (2012). NIDCAP improves brain function and structure in preterm infants with severe intrauterine growth restriction. *J. Perinatol.* 32, 797–803. doi: 10.1038/jp.2011.201
- Anderson, V., Spencer-Smith, M., and Wood, A. (2011). Do children really recover better? Neurobehavioural plasticity after early brain insult. *Brain* 134, 2197–2221. doi: 10.1093/brain/awr103
- Arnon, S., Diamant, C., Bauer, S., Regev, R., Sirota, G., and Litmanovitz, I. (2014). Maternal singing during kangaroo care led to autonomic stability in preterm infants and reduced maternal anxiety. *Acta Paediatr.* 103, 1039–1044. doi: 10.1111/apa.12744
- Arnon, S., Shapsa, A., Forman, L., Regev, R., and Bauer S and Litmanovitz, I. (2006). Live music is beneficial to preterm infants in the neonatal intensive care unit environment. *Birth* 33, 131–136. doi: 10.1111/j.0730-7659.2006.00090.x
- Barrett, K. C., Ashley, R., Strait, D. L., and Kraus, N. (2013). Art and science: how musical training shapes the brain. *Front. Psychol.* 4:713. doi: 10.3389/fpsyg.2013.00713
- Benasich, A. A., and Ribery, U. (2018). “Emergent brain dynamics: prebirth to adolescence,” in *Strüngmann Forum Reports*, eds A. A. Benasich and U. Ribery (Cambridge, MA: MIT Press), 23–26.
- Bieleninik, M., Ghetti, C., and Gold, C. (2016). Music therapy for preterm infants and their parents: a meta-analysis. *Pediatrics* 138:e20160971. doi: 10.1542/peds.2016-0971
- Bowlby, J. (1998). *A Secure Base: Parent-Child Attachment and Healthy Human Development*. New York, NY: Basic Books.
- Bretherton, I. (1992). The origins of attachment theory: John Bowlby and Mary Ainsworth. *Dev. Psychol.* 28, 759–775. doi: 10.1037/0012-1649.28.5.759
- Burke, M., Walsh, J., Oehler, J., and Gingras, J. (1995). Music therapy following suctioning: four case studies. *Neonatal. Netw.* 14, 41–49.
- Caskey, M., Stephens, B., Tucker, R., and Vohr, B. (2011). Importance of parent talk on the development of preterm infant vocalizations. *Pediatrics* 128, 910–916. doi: 10.1542/peds.2011-0609
- Chou, L. L., Wang, R. H., Chen, S. J., and Pai, L. (2003). Effects of music therapy on oxygen saturation in premature infants receiving endotracheal suctioning. *J. Nurs. Res.* 11, 209–216. doi: 10.1097/01.jnr.0000347637.02971.ec
- Degé, F., Kubicek, C., and Schwarzer, G. (2011). Music lessons and intelligence: a relation mediated by executive functions. *Music Percept. Interdiscip. J.* 29, 195–201. doi: 10.1525/mp.2011.29.2.195
- Dubois, J., Benders, M., Cachia, A., Lazeyras, F., Ha-Vinh Leuchter, R., Sizonenko, S. V., et al. (2008). Mapping the early cortical folding process in the preterm newborn brain. *Cereb. Cortex* 18, 1444–1454. doi: 10.1093/cercor/bhm180
- Ettenberger, M. (2017). Music therapy in the neonatal intensive care unit (NICU): putting the families at the centre of care. *Br. J. Music Ther.* 31, 12–17. doi: 10.1177/1359457516685881
- Filippa, M., Panza, C., Ferrari, F., Frassoldati, R., Kuhn, P., Balduzzi, S., et al. (2017). Systematic review of maternal voice interventions demonstrates increased stability in preterm infants. *Acta Paediatr.* 106, 1220–1229. doi: 10.1111/apa.13832
- First Sounds. (2016). *Rhythm, Breath and Lullaby International Training Compendium*. Available online at: https://nicumusictherapy.com/Nicumusictherapy/Grandparents_Training_Roster_1.html.
- Fitzgerald, M., and Walker, S. M. (2009). Infant pain management: a developmental neurobiological approach. *Nat. Clin. Pract. Neurol.* 5, 35–50. doi: 10.1038/ncpneu0984
- Franck, L. S., Cox, S., Allen, A., and Winter, I. (2005). Measuring neonatal intensive care unit-related parental stress. *J. Adv. Nurs.* 49, 608–615. doi: 10.1111/j.1365-2648.2004.03336.x
- Freeman, W. J. (1997). Three centuries of category errors in studies of the neural basis of consciousness and intentionality. *Neural Netw.* 10, 1175–1183. doi: 10.1016/s0893-6080(97)00054-3
- Galvan, A. (2018). “What factors determine changes in the adolescent brain?” in *Emergent Brain Dynamics: Pre-birth to Adolescence*, eds A. A. Benasich and U. Ribery (Cambridge, MA: MIT Press), 225–241.
- Goulet, C., Bell, L., St-Cyr, D., Paul, D., and Lang, A. (1998). A concept analysis of parent-infant attachment. *J. Adv. Nurs.* 28, 1071–1081. doi: 10.1046/j.1365-2648.1998.00815.x

- Haslbeck, F., and Costes, T. (2011). Advanced training in music therapy with premature infants: impressions from the united states and a starting point for europe. *Br. J. Music Ther.* 25, 19–31. doi: 10.1177/135945751102500203
- Ingersoll, E. W., and Thoman, E. B. (1994). The breathing bear: effects on respiration in premature infants. *Physiol. Behav.* 56, 855–859. doi: 10.1016/0031-9384(94)90315-8
- Jaschke, A. C., Honing, H., and Scherder, E. J. A. (2018a). Longitudinal analysis of music education on executive functions in primary school children. *Front. Neurosci.* 12:103. doi: 10.3389/fnins.2018.00103
- Jaschke, A. C., Honing, H., and Scherder, E. J. A. (2018b). Exposure to a musically enriched environment; its relationship with executive functions, short-term memory and verbal IQ in primary school children. *PLoS One* 13:e0207265. doi: 10.1371/journal.pone.0207265
- Jaschke, A. C. (2019). Music, maestro, please: thalamic multisensory integration in music perception, processing and production. *Music Med.* 11, 98–107.
- Jobe, A. H. (2014). A risk of sensory deprivation in the neonatal intensive care unit. *J. Pediatr.* 164, 1265–1267. doi: 10.1016/j.jpeds.2014.01.072
- Johnson, J. K., Chang, C. C., Brambati, S. M., Migliaccio, R., Gorno-Tempini, M. L., Miller, B. L., et al. (2011). Music recognition in frontotemporal lobar degeneration and Alzheimer disease. *Cogn. Behav. Neurol.* 24, 74–84. doi: 10.1097/WNN.0b013e31821de326
- Kim, S., Gabel, C., Aguilar-Raab, C., Hillecke, T., and Warth, M. (2018). Affective and autonomic response to dynamic rhythmic entrainment: mechanisms of a specific music therapy factor. *Arts Psychother.* 60, 48–54. doi: 10.1016/j.aip.2018.06.002
- Klein, C., Liem, F., Hänggi, J., Elmer, S., and Jäncke, L. (2016). The “silent” imprint of musical training. *Hum. Brain Mapp.* 37, 536–546. doi: 10.1002/hbm.23045
- Lahav, A., and Skoe, E. (2014). An acoustic gap between the NICU and womb: a potential risk for compromised neuroplasticity of the auditory system in preterm infants. *Front. Neurosci.* 8:381. doi: 10.3389/fnins.2014.00381
- Leckman, J. F., Feldman, R., Swain, J. E., Eicher, V., Thompson, N., and Mayes, L. C. (2004). Primary parental preoccupation: circuits, genes and the crucial role of the environment. *J. Neural Transm.* 111, 753–771. doi: 10.1007/s00702-003-0067-x
- Lee, G., and Kisilevsky, B. (2014). Fetuses respond to father’s voice but prefer mother’s voice after birth. *Dev. Psychobiol.* 56, 1–11. doi: 10.1002/dev.21084
- Loewy, J. V. (1995). The musical stages of speech: a developmental model of pre-verbal sound making. *Music Ther.* 13, 47–73. doi: 10.1093/mt/13.1.47
- Loewy, J. (2009). “Musical sedation: mechanisms of breathing entrainment,” in *Music, the Breath and Health: Advances in Integrative Music Therapy*, eds R. Azoulay and J. V. Loewy (New York, NY: Satchnote Press), 223–232.
- Loewy, J. (2020). Music therapy as a potential intervention for sleep improvement. *Nat. Sci. Sleep* 12, 1–9. doi: 10.2147/NS.S194938
- Loewy, J. (2015). NICU music therapy: song of kin as critical lullaby in research and practice: rhythm, breath and lullaby NICU music therapy. *Ann. N Y Acad. Sci.* 1337, 178–185. doi: 10.1111/nyas.12648
- Loewy, J., Hallan, C., Friedman, E., and Martinez, C. (2005). Sleep/sedation in children undergoing EEG testing: a comparison of chloral hydrate and music therapy. *J. Perianesth. Nurs.* 20, 323–331. doi: 10.1016/j.jopan.2005.08.001
- Loewy, J., Stewart, K., Dassler, A.-M., Telsey, A., and Homel, P. (2013). The effects of music therapy on vital signs, feeding and sleep in premature infants. *Pediatrics* 131, 902–918. doi: 10.1542/peds.2012-1367
- Lomanowska, A., Boivin, M., Hertzman, C., and Fleming, A. S. (2017). Parenting begets parenting: a neurobiological perspective on early adversity and the transmission of parenting styles across generations. *Neuroscience* 342, 120–139. doi: 10.1016/j.neuroscience.2015.09.029
- Lordier, L., Loukas, S., Grouiller, F., Vollenweider, A., Vasung, L., Meskaldj, D., et al. (2019). Music processing in preterm and full-term newborns: a psychophysiological interaction (PPI) approach in neonatal fMRI. *NeuroImage* 185, 857–864. doi: 10.1016/j.neuroimage.2018.03.078
- Malloch, S., Shoemark, H., Črnčec, R., Newnham, C., Paul, C., Prior, M., et al. (2012). Music therapy with hospitalized infants—the art and science of communicative musicality. *Infant Ment. Health J.* 33, 386–399. doi: 10.1002/imhj.21346
- McGrath, J. M. (2013). Human factors: the importance of communication to outcomes in the NICU. *J. Perinat. Neonatal. Nurs.* 27, 108–109. doi: 10.1097/JPN.0b013e3182907e89
- McMahon, E., Wintermark, P., and Lahav, A. (2012). Auditory brain development in premature infants: the importance of early experience. *Ann. N Y Acad. Sci.* 1252, 17–24. doi: 10.1111/j.1749-6632.2012.06445.x
- Mehr, S., and Krasnow, M. (2016). Parent-offspring conflict and the evolution of infant-directed song. *Evol. Human Behav.* 38, 674–684.
- Mondanaro, J., Ettenberger, M., and Park, L. (2016). Mars rising: music therapy and the increasing presence of fathers in the NICU. *Music Med.* 8, 96–107.
- Montirosso, R. (2015). Relationship between feeding and early stress in premature infant: the role of epigenetic factors. *J. Pediatr. Gastroenterol. Nutr.* 61, S15–S66. doi: 10.1097/01.mpg.0000471459.01168.7b
- Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., and Chau, T. (2011). Short-term memory training enhances verbal intelligence and executive function. *Psychol. Sci.* 22, 1425–1433. doi: 10.1177/0956797611416999
- Forcada-Guex, M., Pierrehumbert, B., Borghini, A., Moessinger, A., and Muller-Nix, C. (2006). Early dyadic patterns of mother-infant interactions and outcomes of prematurity at 18 months. *Pediatrics* 118, 107–114. doi: 10.1542/peds.2005-1145
- Peng, N. H., Bachman, J., Jenkins, R., Chen, C. H., Chang, Y. C., Chang, Y. S., et al. (2009). Relationships between environmental stressors and stress biobehavioral responses of preterm infants in NICU. *J. Perinat. Neonatal. Nurs.* 23, 363–371. doi: 10.1097/JPN.0b013e3181bdd3fd
- Pineda, R., Guth, P., Herring, A., Reynolds, L., Oberle, S., and Smith, J. (2017). Enhancing sensory experiences for very preterm infants in the NICU: an integrative review. *J. Perinatol.* 37, 323–332. doi: 10.1038/jp.2016.179
- Pineda, R., Neil, J., Dierker, D., Smyser, C., Wallendorf, M., Kidokoro, H., et al. (2014). Alterations in brain structure and neurodevelopmental outcome in preterm infants hospitalized in different neonatal intensive care unit environments. *J. Pediatr.* 164, 52–60. doi: 10.1016/j.jpeds.2013.08.047
- Provenzi, L., Guida, E., and Montirosso, R. (2018). Preterm behavioral epigenetics: a systematic review. *Neurosci. Biobehav. Rev.* 84, 262–271. doi: 10.1016/j.neubiorev.2017.08.020
- Ren, K., and Dubner, R. (2007). Pain facilitation and activity-dependent plasticity in pain modulatory circuitry: role of BDNF-TrkB signaling and NMDA receptors. *Mol. Neurobiol.* 35, 224–235. doi: 10.1007/s12035-007-0028-8
- Rossetti, A., and Canga, B. (2013). “Environmental music therapy: rationale for ‘multi-individual’ music psychotherapy in modulation of the pain experience,” in *Music and Medicine: Integrative Models in the Treatment of Pain*, eds J. F. Mondanaro, and A. G. Sara (New York, NY: Satchnote Press, The Louis Armstrong Center for Music and Medicine), 451–468.
- Sa de Almeida, J., Lordier, L., Zollinger, B., Kunz, N., Bastiani, M., Gui, M., et al. (2020). Music enhances structural maturation of emotional processing neural pathways in very preterm infants. *NeuroImage* 207:116391. doi: 10.1016/j.neuroimage.2019.116391
- Sala, G., and Gobet, F. (2017). When the music’s over. Does music skill transfer to children’s and young adolescents’ cognitive and academic skills? A meta-analysis. *Edu. Res. Rev.* 20, 55–67. doi: 10.1016/j.edurev.2016.11.005
- Salk, L. (1973). The role of the heartbeat in the relations between mother and infant. *Sci. Am.* 228, 24–29. doi: 10.1038/scientificamerican0573-24
- Shoemark, H. (2017). Time together: a feasible program to promote parent-infant interaction in the NICU. *Music Ther. Perspect.* 36, 6–16. doi: 10.1093/mtp/mix004
- Stewart, K. (2009a). PATTERNS-A model for evaluating trauma in NICU music therapy: part 1- theory and design. *Music Med.*, 29–40. doi: 10.1177/1943862109338370
- Stewart, K. (2009b). PATTERNS-A model for evaluating trauma in NICU music therapy: part 2- treatment parameters. *Music Med.* 1, 123–128. doi: 10.1177/1943862109344858
- Stewart, K., and Schneider, S. (2000). “The effect of music therapy on the sound environment in the neonatal intensive care unit: a pilot study,” in *Music Therapy in the Neonatal Intensive Care Unit*, ed. J. V. Loewy (New York, NY: Satchnote Press), 85–100.
- Teckenberg-Jansson, P., Huotilainen, M., Polkki, T., Lipsanen, J., and Jarvenpaa, A. L. (2011). Rapid effects of neonatal music therapy combined with kangaroo care on prematurely-born infants. *Nordic J. Music Ther.* 20, 22–42. doi: 10.1080/08098131003768123

- Thompson, L. A., and Trevathan, W. R. (2009). Cortisol reactivity, maternal sensitivity and infant preference for mother's familiar face and rhyme in 6-month-old infants. *J. Reprod. Infant Psychol.* 27, 143–167. doi: 10.1080/02646830801918463
- Tomlinson, G. (2015). *A Million Years of Music: The Emergence of Human Modernity*. New York: Zone books.
- Trainor, L. J., Marie, C., Gerry, D., Whiskin, E., and Unrau, A. (2012). Becoming musically enculturated: effects of music classes for infants on brain and behavior. *Ann. N Y Acad. Sci.* 1252, 129–138. doi: 10.1111/j.1749-6632.2012.06462.x
- Ullsten, A., Hugoson, P., Forsberg, M., Forzelius, L., Klässbo, M., Olsson, E., et al. (2017). Efficacy of live lullaby singing during procedural pain in preterm and term neonates. *Mus. Med.* 9, 73–85.
- Van de Carr, F. R., and Lehrer, M. (1988). Prenatal university: commitment to fetal family bonding and the strengthening of the family unit as an educational institution. *Pre Perinatal Psychol. J.* 3, 87–102.
- Vandormael, C., Schoenhals, L., Hüppi, P., Filippa, M., and Tolsa, C. B. (2019). Language in preterm born children: atypical development and effects of early interventions on neuroplasticity. *Neural Plast.* 2019:6873270. doi: 10.1155/2019/6873270
- Webb, A. R., Heller, H. T., Benson, C. B., and Lahav, A. (2015). Mother's voice and heartbeat sounds elicit auditory plasticity in the human brain before full gestation. *Proc. Natl. Acad. Sci. U S A* 112, 3152–3315. doi: 10.1073/pnas.1414924112
- Yurkovich, J., Burns, D., and Harrison, T. (2018). The effect of music therapy entrainment on physiologic measures of infants in the cardiac intensive care unit: single case withdrawal pilot study. *J. Music Ther.* 55, 62–82. doi: 10.1093/jmt/thx017
- Zentner, M., and Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proc. Natl. Acad. Sci. U S A* 107, 5768–5773. doi: 10.1073/pnas.1000121107
- Zimmer, M.-L. (2019). "Premature babies have premature mothers: practical experiences with premature infants and their mothers using auditive stimulation with the mother's voice," in *Music Therapy for Premature and Newborn Infants*, ed. M. Nöcker-Ribaupierre (Gilsum, NH: Barcelona Publishers), 113–128.

Conflict of 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.

Copyright © 2020 Loewy and Jaschke. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers



OPEN ACCESS

Articles are free to read
for greatest visibility
and readership



FAST PUBLICATION

Around 90 days
from submission
to decision



HIGH QUALITY PEER-REVIEW

Rigorous, collaborative,
and constructive
peer-review



TRANSPARENT PEER-REVIEW

Editors and reviewers
acknowledged by name
on published articles

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org

Contact us: info@frontiersin.org | +41 21 510 17 00



REPRODUCIBILITY OF RESEARCH

Support open data
and methods to enhance
research reproducibility



DIGITAL PUBLISHING

Articles designed
for optimal readership
across devices



FOLLOW US

@frontiersin



IMPACT METRICS

Advanced article metrics
track visibility across
digital media



EXTENSIVE PROMOTION

Marketing
and promotion
of impactful research



LOOP RESEARCH NETWORK

Our network
increases your
article's readership