# COGNITION AND MUSIC PERFORMANCE

EDITED BY: Gary Edward McPherson, Oscar Casanova, Francisco Javier Zarza-Alzugaray, Guadalupe López-Íñiguez and Laura Herrero

PUBLISHED IN: Frontiers in Psychology and Frontiers in Neuroscience

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88976-499-0
DOI 10.3389/978-2-88976-499-0

## 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: frontiersin.org/about/contact

# COGNITION AND MUSIC PERFORMANCE

Topic Editors:
**Gary Edward McPherson,** The University of Melbourne, Australia
**Oscar Casanova,** University of Zaragoza, Spain
**Francisco Javier Zarza-Alzugaray,** University of Zaragoza, Spain
**Guadalupe López-Íñiguez,** Sibelius Academy, Finland
**Laura Herrero,** Camilo José Cela University, Spain

**Citation:** McPherson, G. E., Casanova, O., Zarza-Alzugaray, F. J., López-Íñiguez, G., Herrero, L., eds. (2022). Cognition and Music Performance. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88976-499-0

# Table of Contents


**04** ***Editorial: Cognition and Music Performance***
Laura Herrero, Guadalupe López-Íñiguez, Oscar Casanova, Francisco Javier Zarza-Azulgaray and Gary Edward McPherson

**07** ***The Interplay Between Chamber Musicians During Two Public Performances of the Same Piece: A Novel Methodology Using the Concept of "Flow"***
Eva Bojner Horwitz, László Harmat, Walter Osika and Töres Theorell

**19** ***Typologies of Adolescent Musicians and Experiences of Performance Anxiety Among Instrumental Learners***
Ioulia Papageorgi

**37** ***The State of Music Therapy Studies in the Past 20 Years: A Bibliometric Analysis***
Kailimi Li, Linman Weng and Xueqiang Wang

**51** ***Sex Differences are Reflected in Microstructural White Matter Alterations of Musical Sophistication: A Diffusion MRI Study***
Mohammad-Mehdi Mehrabinejad, Parnian Rafei, Hossein Sanjari Moghaddam, Zeinab Sinaeifar and Mohammad Hadi Aarabi

**62** ***The Communication of Timbral Intentions Between Pianists and Listeners and Its Dependence on Auditory-Visual Conditions***
Shen Li, Renee Timmers and Weijun Wang

**79** ***EEG Correlates of Middle Eastern Music Improvisations on the*** **Ney** ***Instrument***
Mohammad Yaghmour, Padmakumari Sarada, Sarah Roach, Ibrahim Kadar, Zhivka Pesheva, Ali Chaari and Ghizlane Bendriss

**90** ***Social Support as a Facilitator of Musical Self-Efficacy***
Santos Orejudo, Francisco Javier Zarza-Alzugaray, Oscar Casanova and Gary Edward McPherson

**103** ***Effect of Bow Camber and Mass Distribution on Violinists' Preferences and Performance***
Aurélie Tomezzoli, Benjamin Michaud, Eric Gagné, Mickaël Begon and Sonia Duprey

**110** ***Constraint-Based Sound-Motion Objects in Music Performance***
Rolf Inge Godøy

**127** ***Oral Tradition as Context for Learning Music From 4E Cognition Compared With Literacy Cultures. Case Studies of Flamenco Guitar Apprenticeship***
Amalia Casas-Mas, Juan Ignacio Pozo and Ignacio Montero

frontiers | Frontiers in Psychology

EDITORIAL
published: 09 June 2022
doi: 10.3389/fpsyg.2022.923452

# Editorial: Cognition and Music Performance

**Laura Herrero[1]*, Guadalupe López-Íñiguez[2], Oscar Casanova[3], Francisco Javier Zarza-Azulgaray[3] and Gary Edward McPherson[4]**

[1] Education Department, Camilo José Cela University, Madrid, Spain, [2] Sibelius Academy, Helsinki, Finland, [3] Musical, Plastic and Corporal Expression Department, University of Zaragoza, Zaragoza, Spain, [4] Conservatorium of Music, The University of Melbourne, Parkville, VIC, Australia

**Keywords: music research, cognition, education, perception, socio-cultural**

**Editorial on the Research Topic**

**Cognition and Music Performance**

Because it involves a host of different perceptions, concepts and practices that are linked to particular social considerations, research in music is multifaceted and varied (Cross, 2001). The understanding of music cognition requires the contributions of psychological, neuropsychological, social, cultural, and educational perspectives to provide a full approximation to the phenomenon.

Psychological and neuropsychological studies have reported valuable insights regarding both musical perception and music performance. The cognitive and emotional mechanisms underlying the perception of music and how pitch, time, meter, and harmony are processed are of particular concern to researchers (Justus and Bharucha, 2002), as are how music expectations are produced (Pearce and Wiggins, 2012), how musical imagery is related to music perception (Brodsky et al., 2008; Martin et al., 2018), or how performers can regulate their emotional and motivational processes when preparing for concerts (López-Íñiguez and McPherson, 2020, 2021).

Music performance research has reported on evidence related to motor (Meyer and Palmer, 2003), auditory (Zatorre et al., 2007), cognitive (Jaschke et al., 2018; Herrero and Carriedo, 2019, 2020), and expressive (Canazza et al., 2013) processes during the interpretation of rehearsed or unrehearsed repertoire. In addition, the skills of playing from memory, playing by ear, sight reading and improvising, together with performing rehearsed repertoire are considered the five basic types of solo and group performance (McPherson, 1995a,b).

Whereas music education studies have mainly focused on teaching methodologies (e.g., Pozo et al., 2022), or perceptions about learning (López-Íñiguez and Pozo, 2014; García-Gil et al., 2021), social and cultural research has provided an extension of music considerations by exploring the benefits of music therapy (Gómez-Romero et al., 2017; Shi and Zhang, 2020), analyzing different cross-cultural approximations to music (Cross, 2001), and suppo rting the role of music as a social activity in musicians (Volpe et al., 2016), across the general population (D'Ausilio et al., 2015).

As none of the above-mentioned perspectives work in isolation, the current Research Topic presents a broad approach to music research by encouraging additional novel and diverse contributions.

Thus, musical experiences are addressed from the point of view of performance analysis through physiological and psychological measures focused on whether internal and external parameters are related to performance efficiency. Bojner et al. explore togetherness performance in string quartet by analyzing individual and group "flow" experienced by the musicians. By combining heart rate variability (HRV) and qualitative data, the study shows an interplay of synchronized social interactions among players associated with a feeling of group state flow. Papageorgi explores music performance anxiety (MPA) of adolescent musician learners in order to analyze whether it is interpreted and responded in different situations. The results suggest the existence of three

OPEN ACCESS

**Edited and reviewed by:**
Andrea Schiavio,
University of Graz, Austria

***Correspondence:**
Laura Herrero
lherrero@ucjc.edu

**Specialty section:**
This article was submitted to Performance Science, a section of the journal Frontiers in Psychology

**Received:** 19 April 2022
**Accepted:** 02 May 2022
**Published:** 09 June 2022

**Citation:**
Herrero L, López-Íñiguez G, Casanova O, Zarza-Azulgaray FJ and McPherson GE (2022) Editorial: Cognition and Music Performance. Front. Psychol. 13:923452. doi: 10.3389/fpsyg.2022.923452

different profiles (low, moderately, and highly anxious students) which are discussed in order to its clinical and educational implications. On their behalf, Orejudo et al. also analyze students of different ages and how the role of social support regarding self-efficacy in learning and performance may be considered as a potential predictor of long-term careers in music. This study shows how as age of performers increases the external social agents related with self-efficacy perceptions is varying. A different approximation performance based on communication is provided by Li et al. who study timbrical expressions and perceptions from the perspective of performers and listeners. Specifically, this study reports whether the timbral intentions of pianists provide visual cues that influence the perceived timbre of listeners. Music performance is also the focus of Tomezzoli et al. who investigate how bow mechanical characteristics influence violinists' preferences. This study shows that both camber and mass impact not only on the appreciations of violinists but also on the evaluations of the experts. Finally, the review of Godøy focuses on the practical applications in performance, improvisation, and composition of the our understanding of generative processes in music through the analysis of the crucial role of the sound-motion objects.

This Research Topic also contributes to increase the evidences in the analysis of educational music performance across cultural contexts. Yaghmour et al. report relevant neuropsychological cues to understand improvisation in a cross-cultural context. Through the analysis of EEG correlates, this study provides the first observations of the tonal-spatial system charactaristic of Middle Eastern music. A different cultural context is explored by Casas-Mas et al.. The authors provide a qualitative-descriptive approach to Flamenco music through the analysis of discourse and practice of two flamenco guitarists in their natural contexts by pointing out the embodied mind as a result of instrumental learning.

Beyond cross-cultural approximations, different social implications of music cognition and performance are also explored in this Research Topic. Neuropsychological evidences using diffusion MRI are reported in the study of Mehrabinejad et al. which reveals that microstructures with functional connection with motor and somatosensory areas as well as language processing area have significant correlation with music engagement and training and that this associations differ by sex. Li et al. contributes to social implications of music performance though the consideration of music to improve the quality of life of people by presenting an extensive bibliographical review about the contributions of music therapy.

Taken together, the papers included in this collection provide a rich insight into current research interested in the relationship between music cognition and performance. Through the use of a wide range of perspectives and methodologies, the articles presented here delve into key issues that allow us to improve our understanding of said relationship as well as its educational and social implications.

## AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

## ACKNOWLEDGMENTS

The authors would like to thank the reviewers who contributed to this Research Topic. Particular thanks go to the Board of the Frontiers Publishing team, as well as to the staff of the Frontiers Publishing team, and in particular to Ms. Monika Minaroy for her continuous support.

## REFERENCES

Brodsky, W., Kessler, Y., Rubinstein, B. S., Ginsborg, J., and Henik, A. (2008). The mental representation of music notation: notational audiation. *J. Exp. Psychol. Hum. Percept. Perform.* 34, 427. doi: 10.1037/0096-1523.34.2.427

Canazza, S., De Poli, G., Roda, A., and Vidolin, A. (2013). "Expressiveness in music performance: analysis, models, mapping, encoding", in *Structuring Music through Markup Language: Designs and Architectures*. Hershey, PA: IGI Global. p. 156–186

Cross, I. (2001). Music, cognition, culture, and evolution. *Ann. N. Y. Acad. Sci.* 930, 28–42. doi: 10.1111/j.1749-6632.2001.tb05723.x

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

García-Gil, D., Casanova, O., and Zarza-Alzugaray, F. J. (2021). Musical skills in the Spanish Grado university degree in early childhood education. How do Spanish university students view their preparation? *Music Educ. Res.* 23, 77–89. doi: 10.1080/14613808.2020.1853084

Gómez-Romero, M., Jiménez-Palomares, M., Rodríguez-Mansilla, J., Flores-Nieto, A., Garrido-Ardila, E. M., and González-López-Arza, M. V. (2017). Benefits of music therapy on behaviour disorders in subjects diagnosed with dementia: a systematic review. *Neurología (English Edition).* 32, 253–263. doi: 10.1016/j.nrleng.2014.11.003

Herrero, L., and Carriedo, N. (2019). The contributions of updating in working memory sub-processes for sight-reading music beyond age and practice effects. *Front. Psychol.* 10, 1080. doi: 10.3389/fpsyg.2019.01080

Herrero, L., and Carriedo, N. (2020). The role of cognitive flexibility and inhibition in complex dynamic tasks: the case of sight reading music. Current Psychology. doi: 10.1007/s12144-020-00983-y

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

Justus, T. C., and Bharucha, J. J. (2002). "Music perception and cognition", in *Steven's Handbook of Experimental Psychology: Sensation and Perception,* Pashler, H. and Yantis, S. (Eds.). John Wiley and Sons. p. 453–492. doi: 10.1002/0471214426.pas0111

López-Íñiguez, G., and McPherson, G. E. (2020). Applying self-regulated learning and self-determination theory to optimize the performance of a concert cellist. *Front. Psychol.* 11, 385. doi: 10.3389/fpsyg.2020.00385

López-Íñiguez, G., and McPherson, G. E. (2021). Regulation of emotions to optimise classical music performance: a single-case quasi-experimental study of a cellist-researcher. *Front. Psychol.* 12, 627601. doi: 10.3389/fpsyg.2021.627601

López-Íñiguez, G., and Pozo, J. I. (2014). Like teacher, like student? Conceptions of children from traditional and constructive teachers regarding the teaching and learning of string instruments. *Cogn. Instr.* 32, 219–252. doi: 10.1080/07370008.2014.918132

Martin, S., Mikutta, C., Leonard, M. K., Hungate, D., Koelsch, S., Shamma, S., et al. (2018). Neural encoding of auditory features during music perception and imagery. *Cerebral Cort.* 28, 4222–4233. doi: 10.1093/cercor/bhx277

McPherson, G. E. (1995a). The assessment of musical performance: Development and validation of five new measures. *Psychol. Music*. 23, 142–161. doi: 10.1177/0305735695232003

McPherson, G. E. (1995b). Redefining the teaching of musical performance. *Int. J. Music. Educ*. 6, 56–64.

Meyer, R. K., and Palmer, C. (2003). Temporal and motor transfer in music performance. *Music Percept.* 21, 81–104. doi: 10.1525/mp.2003.21.1.81

Pearce, M. T., and Wiggins, G. A. (2012). Auditory expectation: the information dynamics of music perception and cognition. *Top Cogn Sci.* 4, 625–652. doi: 10.1111/j.1756-8765.2012.01214.x

Pozo, J. I., Pérez-Echeverría, M. P., López-Íñiguez, G., and Torrado, J. A. (2022). "Learning and teaching in the music studio", in *A student-centred approach. Landscapes: Arts, Aesthetics*. Springer. doi: 10.1007/978-981-19-0634-3

Shi, E. R., and Zhang, Q. (2020). A domain-general perspective on the role of the basal ganglia in language and music: benefits of music therapy for the treatment of aphasia. *Brain Lang.* 206, 104811. doi: 10.1016/j.bandl.2020.104811

Volpe, G., D'Ausilio, A., Badino, L., Camurri, A., and Fadiga, L. (2016). Measuring social interaction in music ensembles. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 371, 20150377. doi: 10.1098/rstb.2015.0377

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:** 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.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Psychology

ORIGINAL RESEARCH
published: 06 January 2021
doi: 10.3389/fpsyg.2020.618227

# The Interplay Between Chamber Musicians During Two Public Performances of the Same Piece: A Novel Methodology Using the Concept of "Flow"

*Eva Bojner Horwitz[1,2,3*†], László Harmat[4*†], Walter Osika[1,3,5] and Töres Theorell[2,6]*

*[1]Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden, [2]Department of Music, Pedagogy and Society, Royal College of Music, Stockholm, Sweden, [3]Department of Neurobiology, Care Sciences and Society, Center for Social Sustainability, Karolinska Institutet, Stockholm, Sweden, [4]Department of Psychology, Faculty of Health and Life Sciences, Linnaeus University, Växjö, Sweden, [5]Northern Stockholm Psychiatry, Stockholm, Sweden, [6]Stress Research Institute, Stockholm, Sweden*

OPEN ACCESS

**Edited by:**
*Gary Edward McPherson,
The University of Melbourne, Australia*

**Reviewed by:**
*Adina Mornell,
University of Music and Performing Arts Munich, Germany*
*Laura Bishop,
University of Oslo, Norway*

***Correspondence:**
*Eva Bojner Horwitz
eva.bojner-horwitz@kmh.se
László Harmat
laszlo.harmat@lnu.se*

*[†]These authors share first authorship*

**Specialty section:**
*This article was submitted to Performance Science, a section of the journal Frontiers in Psychology*

**Received:** 16 October 2020
**Accepted:** 30 November 2020
**Published:** 06 January 2021

**Citation:**
*Horwitz EB, Harmat L, Osika W and Theorell T (2021) The Interplay Between Chamber Musicians During Two Public Performances of the Same Piece: A Novel Methodology Using the Concept of "Flow". Front. Psychol. 11:618227. doi: 10.3389/fpsyg.2020.618227*

The purpose of the study is to explore a new research methodology that will improve our understanding of "flow" through indicators of physiological and qualitative state. We examine indicators of "flow" experienced by musicians of a youth string quartet, two women (25, 29) and two men (23, 24). Electrocardiogram (ECG) equipment was used to record heart rate variability (HRV) data throughout the four movements in one and the same quartet performed during two concerts. Individual physiological indicators of flow were supplemented by assessments of group "state flow" (means from standardized questionnaires) and a group interview in which the musicians provided qualitative data. A matrix was constructed for the characterization of different kinds of demands in the written music in each one of the four movements for each one of the musicians. HRV derived from ECG data showed non-significant trends for group state flow across the eight musical episodes. Individual-level analysis showed that compared to the other players the first violin player had the highest mean heart rate and the lowest increase in high frequency (HF) power in HRV during this particular movement, particularly during the second concert. The qualitative data illustrated how an interplay of synchronized social interactions between this player and their colleagues during the musical performance was associated with a feeling of group state flow and served to support the first violinist. The case illustrates that the proposed mixed methodology drawing on physiological and qualitative data, has the potential to provide meaningful information about experiences of a flow state, both at individual and group levels. Applications in future research are possible.

**Keywords: synchronization, musical performance, heart rate variability, flow, interpersonal interaction**

## INTRODUCTION

We know from previous research that the interaction between individual musicians and their social and material environments becomes part of their performance. The interplay between music ensembles and their audiences becomes more dynamic when the musicians start to play (Bishop, 2018). The various components involved in a performance are interrelated and interdependent and are grounded in the specific social activity through which one perceives music. In our wider research, we want to improve our understanding of the experience of "flow," a relaxed and at the same time intensely focused state. Flow is characterized by a sense of total immersion in an activity leading to a sense of control, lack of self-awareness, and enhanced performance (Csikszentmihalyi, 1990). We are simultaneously interested in the interactional synchronization experienced by groups and its relationship to flow. Part of this is to describe and understand the interplay among musicians during flow. In this paper, we illustrate our methodological contribution by examining a specific performance on two different occasions as an opportunity for evaluation. A novel set of measures are taken from these performances that combine physiological and qualitative data. We combine measurements of heart rate variability (HRV) with data collected *via* two scales: a shortened form of Flow State Scale (FSS-2) to assess flow experiences within a particular event (Jackson and Ekelund, 2004) – and the Flow Synchronization Questionnaire (FSyQ; Magyaródi and Oláh, 2015). A matrix was also constructed for the characterization of different kinds of demands in the written music. Our aim is to present a case study of possible methodologies that show how HRV, in combination with qualitative data provided by the scales, can be used as an indicator for a state of flow.

The study is set within *Musethica*, an international program which offers young chamber musicians advanced training. It aims to "*translate*" and bring to life musical text and musical information from paper to air and audience (musethica.org). One of the distinctive features of the program is that the musicians are given the opportunity to perform to audiences that may not be accustomed to classical music. The concerts in the program are free for the audience and take place in public and community spaces such as schools (school children), psychiatric departments (patients), factories (workers), and culture unions (migrants). In schools, the concerts are integrated into the school day with compulsory attendance and the pieces are played according to the full format in which the composer wrote them, rather than in an abbreviated version. Time is added after the concert for students to ask questions to the musicians. The children in the audience are not typically familiar with concert music of this type. The goal is to provide musicians with an opportunity to feel what it is like when a lay audience is touched by a performance; that is, fully engaged by the music. A supplementary goal of the program is to offer new audiences the chance to encounter classical music and to interact with young musicians.

Due to our interest in flow, we decided to collect data with a mixed set of parameters and took this opportunity to develop a novel methodology that might inform future research. We performed electrocardiographic (ECG) recordings for HRV in all four musicians for the duration of their performances. After each performance, the musicians filled in a short-standardized questionnaire regarding their individual experience of flow as well as experiences of synchronized social interactions, an indicator of group state flow. Finally, a qualitative interview was conducted with the group of musicians to inform and enhance the interpretation of the dataset. These measurements are explained in further detail in the following section.

## FLOW AND HEART RATE VARIABILITY

Our research takes an interest in how certain psychological states are transferred and shared among performers, in particular in the achievement of "flow" (Csikszentmihalyi, 1990). Flow is a form of optimal experience characterized as a subjective feeling wherein individuals are totally immersed in an activity; their attention becomes absorbed when challenges are in balance with skills (Csikszentmihalyi, 1990). Characteristics of the flow experience include a high degree of focus that feels effortless, a sense of control, loss of self-awareness, an altered experience of time, and enjoyment (Csikszentmihalyi and Nakamura, 2010).

It is believed that there is a connection between the autonomic nervous system (ANS) and the state of flow. Most studies of flow have indicated that there is an increased level of arousal observable in the individual, but interpretations can vary as to their explanation of how arousal states differ from experiences such as stress that would be considered as straining rather than enhancing (see review in Tozman and Peifer, 2016). Peifer et al. (2014) suggested that the relation of flow with sympathetic arousal follows an inverted U-curve rather than a linear function: moderate physiological arousal should facilitate flow-experience, whereas excessive physiological arousal should hinder flow.

Findings from previous studies have suggested that there is a nonreciprocal co-activation of the sympathetic and parasympathetic nervous system during flow experience (de Manzano et al., 2010; Harmat et al., 2015); and that therefore the flow experience is associated with increase in both parasympathetic and sympathetic arousal: a combination of relaxation and readiness for action. Such results support the notion that flow is a state of "effortless attention that can be experienced during task performance as a result of an interaction between emotional and attentional systems" (Ullén et al., 2010).

This response aligns with the function of the ANS, which is to adjust bodily systems in order to maintain or re-establish homeostasis when adapting to surrounding conditions (Cannon, 1935; Ulrich-Lai and Herman, 2009). Flow seems to necessitate the adaptive responses that the ANS performs; using sympathetic and parasympathetic systems. When exposed to stress, the body adjusts to the challenging situation in order to recover homeostasis. The parasympathetic branch and the sympathetic branch of the ANS respectively ensure adaptive adjustment at the physiological level (Cannon, 1935; Ulrich-Lai and Herman, 2009). During periods of acute stress, the sympathetic branch mobilizes energy reserves to cope with internal and external challenges. The stress response is characterized by changes such as increase in muscle tone and alterations of respiratory rate and heart rate (Cohen et al., 2007). The parasympathetic branch is related to the "rest and digest" function that promotes relaxation. The parasympathetic branch innervates its target organs including the heart *via* the vagus nerve. Vagal action occurs immediately, decreasing oxygen consumption, respiratory rate, blood pressure, and heart rate, leading to an increase in well-being, also known as the "relaxation response" (Benson, 1983; Shaffer et al., 2014). The parasympathetic response, which brings the body into balance through its responses to external stressors, is an important component of flow.

In order to measure flow at the physiological level, we turn to HRV. HRV is generated by the variation in consecutive heartbeat intervals. Changes in HRV result from the interaction between the sympathetic and parasympathetic arms of the autonomic nervous system and are modulated with, for example, changing environmental demands (Task Force, 1996; Pieper et al., 2010). A high HRV reading, that is greater variation in inter-beat intervals, indicates that the body has self-regulated itself by initiating a parasympathetic response. Since the flow state is found at the point where the parasympathetic system is activated to an extent that brings the body into balance with the sympathetic system. HRV analysis can be of help in the analysis of states of flow and other psychophysiological conditions.

A number of different systems regulate heart rate. These systems differ in speed (i.e., cycle time), and commonly three speed intervals – or frequency bands – are studied: the vagally mediated signals, where breathing is the main contributor, have the shortest cycle time of about 2.5–6.7 s, corresponding to the high frequency (HF) band (0.15–0.4 Hz). The low frequency (LF) band (0.04–0.15 Hz), corresponds to a cycle time of 6.7–25.0 s, and is most often claimed to be a marker of sympathetic tone and closely associated with blood pressure regulation. The longest cycle (>25 s) is referred to as very low frequency (VLF, <0.04 Hz) and seems to reflect thermoregulatory and baroreceptor systems (Kitney, 1980; Eckberg and Kuusela, 2005) and renin-angiotensin systems (Bonaduce et al., 1994) and also depends on physical activity (Bernardi et al., 1996) and parasympathetic outflow (Taylor, 1998). In HRV analysis, the signal strength (power) reflecting activity in the regulatory system corresponding to each of these three frequency bands is measured and expressed in milliseconds squared ($ms^2$; Task Force, 1996). Total power (TP) indicates the synthesized overall activity for all three frequency bands. HF in heart rate is mainly an indicator of vagal (parasympathic) activity.

The state of flow has been investigated in musicians with different levels of expertise and in a variety of contexts, such as during live performance auditions and musical jam sessions (Wrigley and Emmerson, 2013; Hart and Di Blasi, 2015), as well as in experimental settings (de Manzano et al., 2010). Further studies also confirm that moderate sympathetic arousal and a co-activation of both main branches of the autonomic nervous system characterize task-related flow-experience (Peifer et al., 2014; Harmat et al., 2015; Tozman et al., 2015).

Additionally, Sawyer (2006) described the existence of group flow during musical collaboration and defined it as a "collective state of mind" when the performers were in "interactional synchrony" (Sawyer, 2006, 2007). However, HRV data represents a novel approach to measuring the experience of flow in musicians by means of physiological assessment. In the scientific literature, individual EEG recordings have been used in combination with video data to understand the emotional interplay involved in music therapy (Fachner et al., 2019), but there has been little research about flow and its mechanisms in interactive situations such as in performances by groups of musicians (Csikszentmihalyi, 1990; Sawyer, 2007; Hart and Di Blasi, 2015; Magyaródi and Oláh, 2015, 2017).

Two sets of qualitative data for characterization of the psychological states of the participants were collected.

Accordingly, the purpose of this exploratory study is to investigate how the quantitative data from HRV recordings can be combined with qualitative data in the study of flow. Our research question is therefore: can the combination of physiological and psychological responses in musicians help us understand individual and group state flow?

## A Matrix for Characterization of Demands in the Music

There was a need for a tool with which we could analyze the composition from the point of view of demands that the music makes on the musicians. We are not aware of any such measurement tools therefore a matrix was constructed for the characterization of three types of demands from the composer in the music for each one of the four musicians (violin 1 = V1, violin 2 = V2, viola and cello) and for each movement. It was a way of objectifying the music from a technical performance perspective. This offered a way to hypothesize how the individual musician could be expected to react to the playing of different parts of the quartet and relate that to possible flow-related experiences.

For each type of demand a four-graded scale was used (1 = lowest and 4 = highest demands). The three types of demands were:

### Pitch

For string instruments, the general rule is that the higher the pitch, the more demanding the notes. This gets particularly difficult from the seventh position on the violin. In this quartet, there are requirements for the first violin in all the movements to play in positions above the seventh position, in the third movement up to the 10th position. Scores were defined in the following way in the matrix: 4 = At least seventh position, 3 = At least fourth–sixth position, 2 = At least second to third positions, and 1 = First position only.

### Rhythm

Fast rhythm is always more demanding than slow. Scores: 4 = Several bars following one another with sixteenths in allegro-presto or 30-s in at least andante, 3 = Same as 4 but only occasional bars, 2 = Eighths in allegro-presto or sixteenths in at least andante, and 1 = Slower tempi.

### Emotional Engagement

Scores: emotional engagement means that an emotionally charged theme has to be played in such a way that the audience is emotionally moved. 4 = During an entire movement, 3 = During most of the movement, 2 = During less than half of the movement, and 1 = Not at all.

The scores 1–4 estimated for all parts (V1, V2, viola, and cello) and all four movements in this particular quartet are presented in **Table 1**. The table shows that the first violin part is more demanding than the other parts (as in many of Haydn's string quartets). The first violin has high demands for pitch, particularly in the first three movements. Fast rhythm adds to the technical demands in the first, second, and fourth movements. The contrast between the four instrumental parts becomes particularly evident in the second movement in which the first violin plays long sequences of 30-s in an emotionally moving melody. The viola also plays a similar passage of 30-s, but it only lasts for a few bars. In fact, the second movement has the form of a movement in a short violin concert in which the other instruments only accompany the first violin – with the exception of the short viola section. This also increases the emotional engagement demands on the first violin while the other instrumentalists should serve as support.

Based upon the demands in the music the following prediction can be made:

1. V1 should have higher heart rate and lower HF (= high frequency power) in the HRV) than the others throughout
2. The second movement should, compared to the other movements, be associated with lower heart rate and higher HF in all musicians except in the V1 player due to the technical demands in that part. The viola player also has a technically demanding part during a short section of this movement, so the lowest heart rate and highest HF is expected to occur in the V2 and cello players. These two parts are mainly supporting the first violin in this movement.
3. During the second concert, when the music was repeated and the musicians were accustomed to this secondary school class audience, a lowered heart rate, corresponding to a lowered sympathetic activity, and an elevated HF, corresponding to increased parasympathetic activity, is expected.

The matrix provided us with a possibility to relate expected performance demands to flow-related experiences.

**TABLE 1** | This matrix presents the characterization of three types of demands in the music (pitch, rhythm, and emotional engagement) for each one of the four musicians (violin 1 = V1, violin 2 = V2, viola and cello) and for each movement.

| | I | II | III | IV |
|---|---|---|---|---|
| **Pitch** | | | | |
| V1 | 4 | 4 | 4 | 3 |
| V2 | 2 | 2 | 2 | 2 |
| Viola | 2 | 2 | 2 | 2 |
| Cello | 2 | 2 | 2 | 2 |
| **Rhythm** | | | | |
| V1 | 4 | 4 | 3 | 4 |
| V2 | 2 | 2 | 2 | 3 |
| Viola | 2 | 3 | 2 | 3 |
| Cello | 3 | 2 | 2 | 3 |
| **Emotional engagement** | | | | |
| V1 | 2 | 4 | 2 | 2 |
| V2 | 2 | 3 | 2 | 2 |
| Viola | 2 | 3 | 2 | 2 |
| Cello | 2 | 3 | 2 | 2 |

*For each type of demand a four-graded scale was used (1 = Least demanding and 4 = Most demanding).*

## MATERIALS AND METHODS

### Participants

For this methodological experiment, the participants were all healthy chamber music students: two women, 25 and 29 years old, and two men 23 and 24 years old. They were enrolled at the time of the study at the Royal College of Music in Stockholm, Sweden. The musicians were participating in an advanced program, *Musethica*, which provides multiple opportunities for the students to play chamber music to lay audiences (see Theorell and Bojner Horwitz, 2019). The participants were recruited by direct inquiry.

None of the musicians were on medication of relevance. They were all young healthy adults, two men and two women.

### Context of the Data Collection

The musicians performed the same string quartet (Haydn's string quartet 76, No 2 D-minor) to two different audiences of secondary school classes. The quartet has four movements and was unfamiliar to the audience.

The second movement of the piece of music performed in this study differs from the other movements in its character and composition: while the tempo is a relatively calm andante, the first violin plays a demanding segment of 30-S, incorporating high notes. This is quite technically difficult to perform. The difficult bars are interspersed with less demanding parts. Apart from the viola player, who also has some challenging rapid bars to play toward the end of the movement, the other instrumentalists have comfortable parts to play.

### ECG Recordings

Heart rate variability (HRV) was measured using a combined heart rate and movement sensor (Actiheart, Cambridge Neurotechnology, Ltd., Papworth, UK) that was able to record inter-beat intervals with good reliability and validity (Brage et al., 2005).[1] This sensor was a small portable rechargeable

[1]The movement measurements that were recorded do not form part of the data for this paper.

unit (Actiheart) applied to the chest with temporary adhesive. The sensor recorded both heart rate and HRV. Since the LF power of the HRV is influenced by both the sympathetic and parasympathetic systems, we decided to focus our analysis on heart rate, mirroring the physiological arousal that is largely governed by the sympathetic nervous system, and in HF, correspondingly mirroring the parasympathetic system activity.

In the present study, measurement and analysis comprise HRV and heart rate.

The Actiheart recorder was fixed to the participants' upper chest of the musicians by clips that fit standard ECG electrodes. The Actiheart recorder was firmly fixed to an electrode placed just below the apex of sternum while the wire running from the monitor was fixed to an electrode placed on the same horizontal level and as lateral as possible. In the Actiheart recorder, the analog signal was band-pass filtered (10–35 Hz), sampled with a frequency of 128 Hz, and processed by a real-time QRS-detection algorithm (Actiheart User Manual, 2010). During the recording, interpolated RR intervals with a resolution of 1 ms were stored in the memory. In the recording mode employed in the present study, the raw ECG signal is not stored by the Actiheart recorder (Kristiansen et al., 2011). The HF (0.15–0.4 Hz) and LF (0.04–0.15 Hz) were calculated by Actiheart 4 Software based on the recorded RR intervals. Missed beats in the IBI data were corrected according to the Actiheart User Manual (2010).

## The State Flow Scale

In previous studies (de Manzano et al., 2010; Harmat et al., 2015), flow has been measured using a subset of nine items from the FSS-2 (Jackson and Eklund, 2004) to assess flow experience within a particular event. Good psychometric properties of the FSS-2, as well as of the shorter English nine-item version of the test, similar to the one employed here, have been used in several studies with different samples (Jackson and Eklund, 2002, 2004; Jackson et al., 2008; Kawabata et al., 2008). In the Jackson and Eklund (2002) study, item identification sample reliability estimates for the FSS-2 ranged from 0.80 to 0.90, with a mean alpha of 0.85. Items are formulated as statements about subjective experiences of a performance (e.g., “I had feeling of total control.”), in response to which the respondent should agree or disagree. Answers are given on a Likert scale with nine steps ranging from 1 (strongly disagree) to 9 (strongly agree).

In our study, we used four of the nine questions that related to the state flow experiences (Csikszentmihalyi, 1990; Jackson and Eklund, 2004). These were challenge-skills balance, unambiguous feedback, feeling of control, and autotelic experience. This selection was justified based upon psychometric analysis conducted in a previous study (De Manzano et al., 2010) and informed by a judgment that it would be most conducive to the study to expose the instrumentalists to a minimum of questions during the intermissions between movements. Measurements were taken between each movement and during both public music performances.

## Flow Synchronization Scale

Interpersonal synchrony may increase prosocial behavior in which we become likely to trust and cooperate with one another (Cirelli et al., 2014). According to the Flow Synchronization Theory (Magyaródi and Oláh, 2015, 2017), flow synchronization is a psychological mechanism stimulating the group members to interact with each other, and to work for common goals in cooperation during an optimally challenging situation. Thus, interaction includes experience of cooperating as well as iterative exchange of initiatives, ideas, and views. People who experience this know their exact purpose and share a common strategy with others to achieve this goal. Partners help each other, integrate with one another consistently, motivate themselves, and learn from each other. Looking back on the experience, they realize how much they have developed during the activity and how they positively affected each other's performance. Group members also display a sharing of their knowledge, giving feedback to one another. This may support emergent motivation to engage in ongoing collaboration (Csikszentmihalyi and Nakamura, 2010). In order to study how experience during an interactive situation becomes optimized through such synchronization we used the Flow Synchronization Questionnaire (FSyQ) which was developed by a Hungarian research group at Eötvös Lórand University, Budapest (Magyaródi and Oláh, 2015). It enabled us to measure the perceived quality of social interactions during public music performances.

The development of the FSyQ measure was based on both the rational and empirical test establishment traditions. The questionnaire contains 28 items and 5 latent factors that focus on the motivational and coordination (task- and relationship-focus) aspects of the experience: (1) Synchronization and effective cooperation with the partner (12 items, $\alpha = 0.93$); (2) Experience of engagement and concentration (five items, $\alpha = 0.83$); (3) Motivation and positive impact on the partner (three items, $\alpha = 0.82$); (4) Motivation and learning for the person (four items, $\alpha = 0.80$); and (5) Coordination with the partner during the activity (four items, $\alpha = 0.81$). The internal consistencies of the subscales are adequate and the Cronbach's alpha reliability of the original questionnaire (total score) is $\alpha = 0.94$ (Magyaródi and Oláh, 2015). In a recent article (Olsson and Harmat under review) presented the Swedish version of the questionnaire (Olsson and Harmat, Frontiers in Psychology, under review). The Swedish version was developed by first translating the original FSyQ from Hungarian into Swedish, using an independent translator and then back into Hungarian. After this, the original and the back translated FSyQ versions were compared, and a Swedish version was created. This procedure is according to the classic back-translation method developed by Brislins (1970, referred to in Cha et al., 2007). We tested the Swedish version of the questionnaire on a small sample ($n = 62$). The internal consistency of the questionnaire is adequate with a Cronbach's $\alpha = 0.93$. However, we assessed only seven items out of the 28 items of the questionnaire after each music performance based on the highest factor loads of each dimension, i.e., (1) Synchronization and effective cooperation with the partner (1); (2) Experience of engagement and concentration (two items); (3) Motivation of and positive impact on the partner (one item); (4) Motivation and learning for the person (one item); (5) Coordination with the partner during the activity (two items).

## Qualitative Interview

A single semi-structured group interview (duration of 1 h) was conducted at the Royal College of Music in Stockholm with the four musicians after the culmination of both of their performances. The interview was recorded, transcribed and analyzed using a phenomenological hermeneutic method. The data takes the form of the transcribed participant discussion, sitting around a table together and sharing answers to the researcher's questions. Focus group questions probed how musicians evaluated their experiences of interaction with the audience and their experiences of each musical movement. We asked open questions to achieve this, such as: describe how you remember the concert?, How did you perceive the interplay among you?, Are there special moments you remember from the concert?, and Did you feel any flow during the concert?

The data were processed through developing a Naïve Reading – a summary of the data derived from detailed descriptions of participants' experiences – in line with our research aim (Bojner Horwitz et al., 2003, 2016; Lindseth and Norberg, 2004; Grape Viding et al., 2017). To conduct a Naïve Reading, the researcher reads the focus-group transcripts several times in order to recognize and synthesize the essence of the informants' discussion. The analytical prism of the research question (How do musicians' physiological and psychological responses indicate individual and group state flow?) was employed to delineate and distinguish relevant "meaning units" contained in the accounts of musicians' experiences. From this, we gained insights into the interplay among musicians that served to strengthen the physiological findings.

## Ethical Considerations

Ethical approval was obtained for the study (Dnr 2017/1009-31/1 Central Ethical Review Board in Stockholm, Sweden, and both oral and written informed consent was obtained from the participants.

# RESULTS

The results are presented in the following order: the focus group interview, the state flow scale, the flow state and synchronization scales, ECG recordings.

## The Focus Group Interview

The musicians' experiences of interacting with the audience, reading their responses to the performance, and their perceived interactions with other musicians in the ensemble are summarized in a Naïve Reading. The Naïve Reading is a synthesis derived from the following quotations from the group interview:

*"The second movement, with intonation: thought it was difficult. We practiced it pretty much."*

*"Each time there were the same difficulties. We worked a lot with intonation in the second movement. It became so clear that we needed to support each other a lot."*

*"The second movement is quite slow. Playing slow is more difficult to do in front of a school audience. They have not chosen to be an audience. I have to make an extra effort, to make them appreciate it."*

*"It might just be so in my head … but the school class may not think it is as cool with slow and calm tempo as in other audiences."*

*"I have a solo in the second movement and you are helping me (to play), and then you listen more to me then, 'maybe you would become even more compassionate if I made big mistakes, then probably your compassion would increase because of that'… Probably, your compassion will then increase because of that."*

*"Empathy in the second movement. The music is beautiful and captivating in some way and for them (the other musicians) to be able to share a good interplay, so you have to listen to me playing the melody. There is empathy."*

*"I felt the most of interaction in the second (movement)…"*

*"The interaction was best for me in the first movement, and maybe also in the second one. Maybe not right from the start (first time performance) but when we played it over again (the second time). I think we had practiced the second (movement) the most."*

*"Felt like we were breathing together in the second (movement)."*

These quotes regarding flow-related experiences provide a triangulation point for the physiological data reported below.

Moreover, the musicians seem to share some elements of their experiences in the second slow movement. Deriving from the Naïve Reading, the findings can be summarized as follows:

The second movement was the most difficult part of the whole piece and the musicians had practiced the second movement the most. They had also worked a lot with intonation in the second movement. Some of the observations the musicians made about the second movement were potentially related to the experience of group flow: they expressed the need to support each other because of the problems they had with intonation during practice of this movement. They felt that they were, as musicians, breathing together in the second movement.

The participants were also asked to outline the different elements of their experience that contributed to the dynamic of their interplay as musicians:

- External circumstances (time of day, early morning was difficult).
- Supporting arrangements (in this case, good teachers in the classroom).
- Distracting factors (our ECG recordings and filling out forms).
- The ensemble's beliefs about audience attitude (perhaps being too negative in this case).

- The attitude of the four musicians toward one another.
- The attitudes of the audience/audience acquaintance with the kind of ensemble and the music.
- The content and difficulty of the music.
- Individual preparations (practiced a lot/a little on the piece).
- Group exercise (questioning whether they had practiced for intonation sufficiently).

This synthesis was derived from quotations in the group interview, such as:

> *"We need to support the soloist."*
> *"We have practiced a lot, especially the second movement."*
> *"But here I did not feel that they (the audience) loved the second movement (in class)."*
> *"It took some time to get used to the ECG equipment."*
> *"I think we have found a compassionate attitude toward another."*

## Individual State Flow and Group Flow Synchronization

One-way ANOVA was performed in order to measure the significance of the variation in flow scores over time during the different movements and performances. Individual flow assessments taken after each movement of the performance did not show any significant variation over time (**Figure 1**).

We collected data from the FSyQ after the first and the second performances (Magyaródi and Oláh, 2015), in order to examine the psychological mechanisms involved when the group members were interacting with one another, during an optimally challenging situation. We did not find significant differences between the two performances with regard to flow synchronization. A paired sample t-test indicated that there was no statistically significant difference between the two performances with regard to flow synchronization.

## ECG Recordings of Heart Rate and HRV

Electrocardiogram (ECG) recordings were subjected to analysis of heart rate and HRV. The four movements in the string quartet on the two occasions (2*4) were analyzed as distinct periods. It was observed that the mean heart rate in the four musicians was lowest when they played the second movement on the second occasion. Using one-way ANOVA, variation over time was not significant ($F = 2.15$, $df = 1/7$, $p = 0.08$, see **Figure 2**).

HRV: no differences between the occasions or movements were discovered with regard to variations in LF power of the HRV (see **Figure 2**). For the HF power (which is reportedly the parameter that mirrors parasympathetic activity) the highest mean was found in the second movement on the second occasion. We also analyzed the LF power of HRV but found no systematic changes. Variation over time did not reach statistical significance ($F = 2.00$, $df = 1/7$, $p = 0.10$, (see **Figure 3**). Each one of the four musicians' mean heart rates and high frequency means are illustrated separately in **Figures 4**, **5**.

Comparing with our expectations (see above), we found that:

Prediction 1: Yes, the first violin player had the highest heart rate. The V1 player's mean heart rate during the four movements in the two concerts ranged from 103 to 113. The corresponding means for V2 were 74–81, for the viola player 97–106 and for the cello player 90–106. With regard to HF, the V1 player had low HF. However, an equally low HF was observed in the viola player.

Prediction 2: Yes, in general, the musicians had a lower heart rate during the second movement than in the other

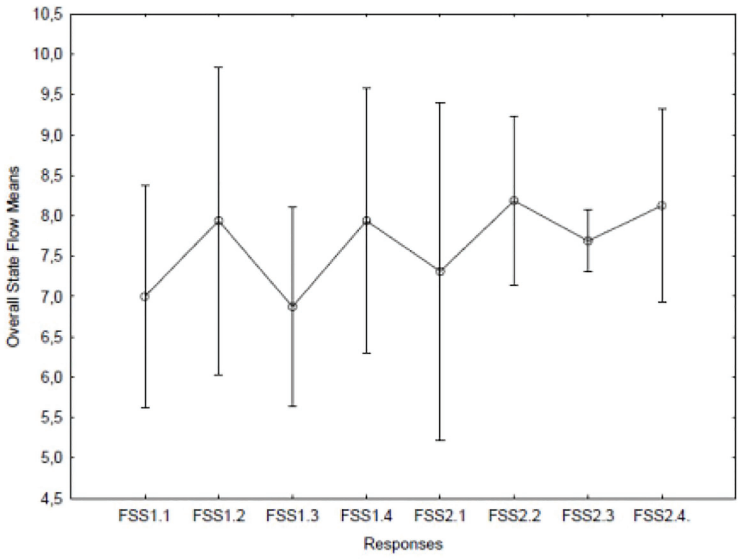



**FIGURE 1** | Flow means for the first concert with movements 1–4 (1.1–1.4) and correspondingly for the second concert with movements 1–4 (2.1–2.4). The average flow had its highest value after the second movement on the second occasion. The vertical lines are 95% CIs.

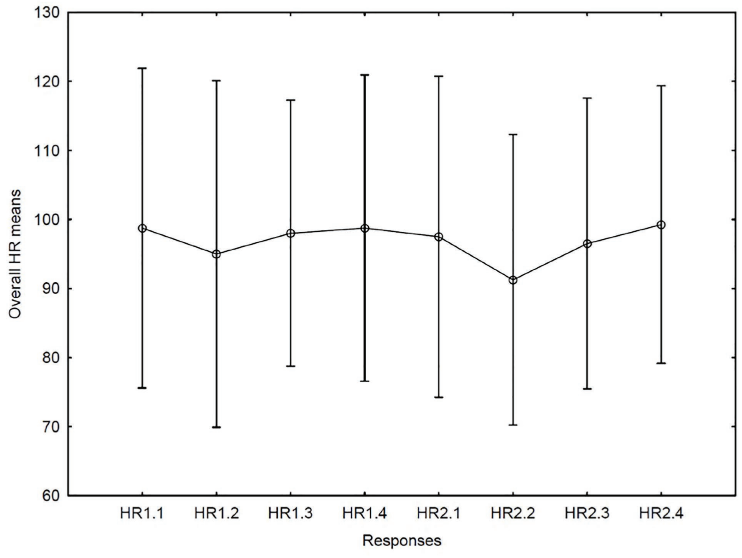



**FIGURE 2 |** Heart rate (HR) means for the first concert with movements 1–4 (1.1–1.4) and correspondingly for the second concert with movements 1–4 (2.1–2.4).

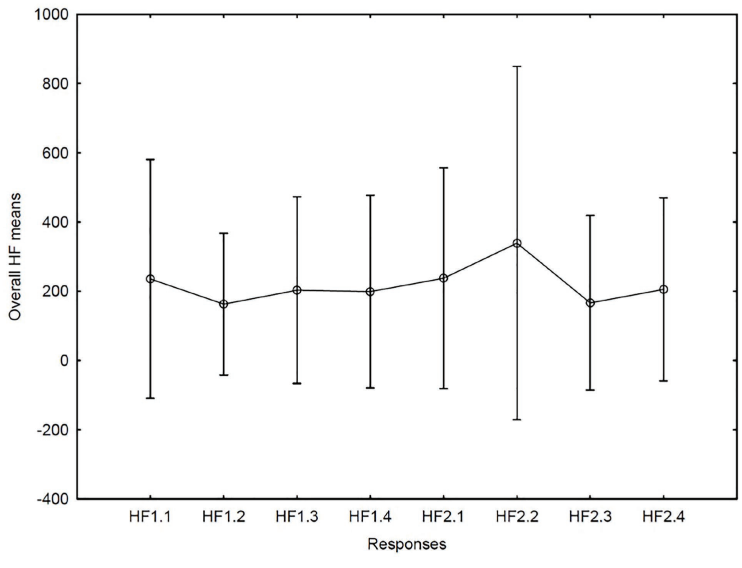



**FIGURE 3 |** High frequency (HF) means for the first concert with movements 1–4 (1.1–1.4) and correspondingly for the second concert with movements 1–4 (2.1–2.4).

movements both during the first and the second concert. The exception was the V1 player who did not show a lowered heart rate during this movement. For HF, the prediction was correct for the V2 and cello players but only during the second concert.

Prediction 3: The V1 and the viola player had lower heart rates during the second concert than during the first one, but for the other players there was no clear difference between the two concerts. During the first concert the mean HR during movements varied between 106 and 113 in the V1 player, between 76 and 80 in the V2 player, between 101 and 106 in the viola player, and between 90 and 104 in the cello player. The corresponding numbers during the second concert was 103–110, 74–81, 97–104, and 89–106.

HF: In the V1 and viola players, no clear changes were observed between the two concerts. These two players both had low HF. The two other players (V2 and cello) had much higher HF than the other players throughout, and a pronounced increase took place from the first to the second concert during this movement (327–598 in the V2 player and 204–635 in the cello player), which indicates a stimulated parasympathic system.

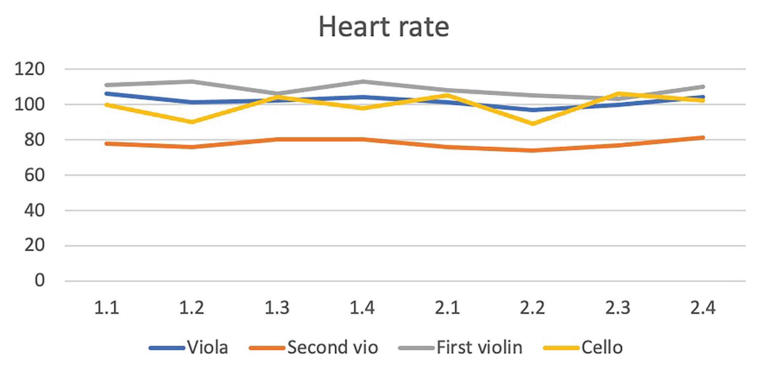



**FIGURE 4 |** The figure shows that the first violinist (gray line) had a relatively high heart rate (103–111 beats per minute) throughout both concerts (1.1–2.4) and when the second movement was performed in both concerts (1.2 and 2.2); with 113 beats per minute during the first concert's second movement (1.2) and 105 beats per minute during the second concert's second movement (2.2). The second violinist (orange line), and the cellist (yellow line) have their lowest heart rates during the second movement on both occasions.

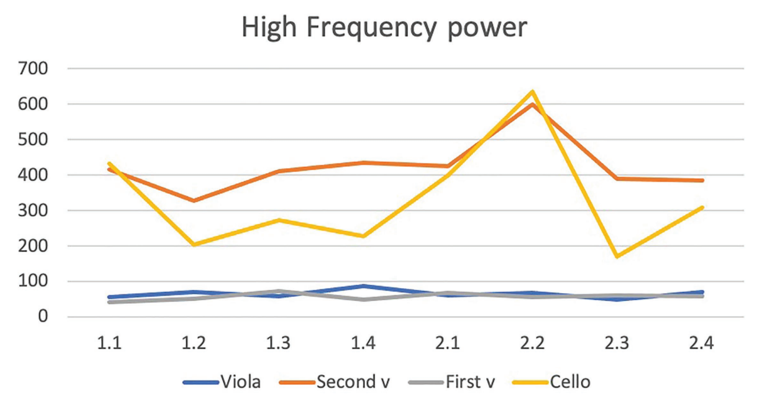



**FIGURE 5 |** The figure shows that the first violinist (gray line) had the lowest HF power during both concerts and during the second movement (2.2). From the first (1.1–1.4) to the second (2.1–2.4) concert, a dramatic increase in HF occurred in the second movement (2.2) in both the cellist (yellow line) and the second violinist (orange line). The viola player (blue line) shows data similar to the first violinist's data (gray line), and none of them have any increase in HF power during the second performance's second movement.

## DISCUSSION

In this case study, we used a combined methodology mirroring aspects of possible state flow in four musicians: perceived experience of synchronized social interactions and physiological states in the performance. By conducting recordings performed on a limited number of players, we have been able to explore a novel methodology that combines psychological and physiological data in order to characterize different aspects of flow.

The *Musethica* program with its repetition of concerts, under similar conditions, provided an opportunity for us to study behavioral and physiological parameters on two different occasions consecutively. It is important to remember in interpreting the value of this experiment and its potential usage on other studies, that there are of course physiological and psychological external factors not related to the playing itself that could have influenced our assessments. The findings on HR and HF only partly confirmed our predictions. Despite this, we were able to confirm predictions that the first violin player should have higher heart rate than the others and that the two players with a supportive role during second movement had low HR and high HF during the second concert when the group felt "safer." Despite our small sample size, the total between movement variation in heart rate and high frequency power approached significance. That the LF power did not show systematic changes may be due to the fact the LF power of HRV is influenced concomitantly by the sympathetic and the parasympathetic system (Porges, 2007).

Are the results of this study consistent with what one would expect from the musicians' experiences? Do they mirror the four elements of the performance outlined in the matrix constructed at the outset of the research? If we consider the patterns of the state flow response in the data, we do in fact see the same patterns during the two performances: the first and third movements obtained lower flow ratings compared to the second and the fourth movements in both performances. The mean ratings of the second performance were slightly higher than the fourth. In addition, the second time the slow movement was played the very lowest mean heart rate was recorded. Similarly, the highest mean high frequency power was found in the second movement but only during the second performance. This may indicate that the musicians experienced a physiologically more relaxed state during the performance of this movement on the second occasion. High HF and low HR also coincided with the peak mean flow rating.

Furthermore, the ECG analysis showed that the first violinist had the highest heart rate and the lowest HF during the second movement of the performance. However, this violinist also showed clearly decreased heart rate from the second movement in the first concert to the second movement in the second concert (113 vs. 105 beats per min). The most pronounced decrease in heart rate from the first to the second movement occurred in the two musicians with the least technically demanding parts.

These findings are consistent with the expectation, based upon the matrix scores (see above) that V1 should have higher heart rate and lower HF than the others throughout. They are also in line with our hypothesis that the second movement should, compared to the other movements, be associated with lower heart rate and higher HF except in the V1 player due to the technical demands in that part and also in the viola player during a short section of this movement. Accordingly, the lowest heart rate and highest HF were expected and also occurred in the V2 and cello players.

In addition, in keeping with our hypothesis, during the second concert when the music was repeated and the musicians were accustomed to this secondary school class audience, on average a lowered heart rate and an elevated HF, corresponding to increased parasympathetic activity, was found. The increase in parasympathetic activity was pronounced in the V2 and cello players in the second movement during that concert. Difference in the musical content cannot explain the differences between the patterns observed in the two concerts, and thus we must consider contextual factors, such as their playing the same music the second time during the same day; familiarization with the setting; a better general atmosphere in the room, and connection with the audience; or indeed a better overall sense of mastery during the second performance. Taken together these findings are consistent with our predictions and indeed de Manzano's and colleagues' previous work (2010) in which we used piano playing as a flow-inducing behavior in order to analyze the relationship between subjective flow reports and psychophysiological measures.

However, the results based upon the questionnaire in the present study contrasted with the study with professional pianists. This finding related to autotelic experience of each movement, the enjoyment of the task among musicians. In the previous study by de Manzano et al. (2010), in which professional pianists performed a difficult self-selected piece on five different occasions, the occasion with the highest flow score (measured according to Csikszentmihalyi) was compared with the other occasions. It was shown that the activity in the laughing muscle Zygomaticus major (assessed by means of continuous EMG recordings) was higher during the occasion with the highest flow rating than during the other occasions. The participants also displayed a deep slow breathing pattern when the mean flow score was high. What was surprising is that these musicians also showed a higher heart rate on that occasion than on the other four occasions. This was interpreted as demonstrative of flow because it was an example of combined sympathetic and parasympathetic activation; a state a high arousal manifested in a high heart rate, combined with a state of relaxation manifested in deep slow breathing stimulating the vagus nerve. According to the interpretations of these researchers, the findings of our study during the second movement would be more indicative of a relaxed happiness state.

Our qualitative data illustrating musicians' "experiences" of flow, suggested that they had a strong shared interpretation of the feeling of the music, its demands, and practical approaches to these features. This has been missing from previous studies and gives us greater insight into how the physiological measurements correlated with a potential flow state even when the demands on the players may be moderately unbalanced (Peifer et al., 2014).

## Implementation and Future Development of the Methodology

In our study, the recordings were made in real life conditions and there was only minimal time for baseline recording since the electrodes were applied immediately before the concert. In implementing this methodology in the future, it would be useful to establish robust baseline data. This would mean that more rigorous analysis in relation to baseline data could be performed. Sampling of ECG data was, however, performed continuously during the concerts and each movement was treated as a separate period. The Actiheart technology was able to identify movements in the same way as during walking and "normal" activities of daily life. We would recommend that measurements are taken continuously rather than at sampled intervals for accuracy. Regarding emotional "standard" ratings, i.e., our matrix for characterization of demands in the music, this assessment instrument will need systematic psychometric evaluation in the future.

We experienced some technical difficulties: the type of instrument played by each musician did not seem to influence the data overall; however, one of the instrumentalists had a problem with a lost electrode during the last part of one movement. A small segment of the ECG was lost as a result and the electrode was re-applied during the subsequent intermission. However, in general, the equipment functioned well, and the musicians felt comfortable with it during their performances. Furthermore, the elements of the Actiheart that measured muscular movements (although not reported on here) did not appear to interfere with the electrocardiographic analyses. Any ectopic beats in the cardiac rhythm, which are common even among perfectly healthy young people, were identified by the automated program and did not affect analysis.

In terms of future subjects for research, firstly, one could use this methodology as a basis for a study of physiological synchronization during flow. This would require specialist ECG analytical skills. We did not have access to this expertise. Secondly, the audience constitutes an important part of the interplay that occurs during a musical event and offers avenues for future investigation. In relation to this, one might consider for example the type of music performance and its impact on the dynamic interaction between musicians and audience. In our study, the musicians performed a classical piece by Joseph Haydn that did not require any musical improvisation (Hart and Di Blasi, 2015). According to Keeler et al. (2015), there were no differences in the intensity of the flow experience when a small group of singers performed pre-composed pieces or improvised together. However, there is a question how the different traditions and different types of musical events/genre shape the interactions between musicians and their audiences. For example, maybe jazz musicians perceive more concrete feedback from their audiences during the performance (e.g., hand clapping after an improvised moment) than musicians during other types of musical events such as classical music performances (Brand et al., 2012). However, silence can also be a positive feedback from the audience during a classical music performance in accordance with genre convention. Future research is needed to clarify the different types of musical events/genres. It is known from extant research that positive feedback and public acknowledgment may not only enhance performers' self-confidence and motivation (LeBlanc et al., 1997; O'Grady 2009; Baker and MacDonald, 2013), but also facilitate strong interplay between audiences and performers. Vice-versa too, when musicians are strongly engaged and have a feeling of mastery (ease-less effort or "flow"; de Manzano et al., 2010; Harmat et al., 2011, in press) their emotional state is also likely to have a strong emotional influence on the audience. This suggests a mutual relationship between musicians and their audiences that may be shaped by the specific nature of the performed piece.

## CONCLUSION

The combination of physiological data *via* ECG recordings, and psychological data *via* expressive narratives provided new insights and has contributed to a proposal of a novel methodology. Although the number of observations was limited in this pilot study, we can conclude that the proposed methods are indeed feasible to investigate the interplay between chamber musicians and may provide meaningful results in further studies. The qualitative data illustrating the musicians' "experiences" of flow, suggested that they had a strong shared interpretation of the feeling of the music, its demands, and practical approaches to these features. This has been missing from previous studies and gives us greater insight into how the physiological measurements correlated with a potential flow state even when the demands on the players may be moderately unbalanced. Our study indicates that it is possible to use these kinds of mixed-methodology techniques to examine perceived flow experiences at individual and group levels during musical performances.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Dnr 2017/1009-31/1 Central Ethical Review Board in Stockholm, Sweden. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

EH: ethical approval, project administration, conceptualization, writing, formal quant analyses, and editing. LH: data curation, methodological consideration, formal quant analyses, and writing. WO: conceptualization, supervision, writing, editing, and project administration. TT: ethical approval, supervision, writing, editing, methodological consideration, and formal analyses. All authors contributed to the article and approved the submitted version.

## REFERENCES


Actiheart User Manual (Jan 2010).

Baker, F. A., and MacDonald, R. A. (2013). Flow, identity, achievement, satisfaction and ownership during therapeutic songwriting experiences with university students and retirees. *Music. Sci.* 17, 131–146. doi: 10.1177/1029864913476287

Benson, H. (1983). The relaxation response: its subjective and objective historical precedents and physiology. *Trends Neurosci.* 6, 281–284. doi: 10.1016/0166-2236(83)90120-0

Bernardi, L., Valle, F., Coco, M., Calciati, A., and Sleight, P. (1996). Physical activity influences heart rate variability and very-low-frequency components in Holter electrocardiograms. *Cardiovasc. Res.* 32, 234–237. doi: 10.1016/0008-6363(96)00081-8

Bishop, L. (2018). Collaborative musical creativity: how ensembles coordinate spontaneity. *Front. Psychol.* 24:1285. doi: 10.3389/fpsyg.2018.01285

Bojner Horwitz, E., Rydwik, E., and Huss, E. (2016). Arts as an ecological method to enhance quality of work experience of health care workers: a phenomenological hermeneutic study. *Int. J. Qual. Stud. Health Well-being* 12:1333898. doi: 10.1080/17482631.2017.1333898

Bojner Horwitz, E., Theorell, T., and Anderberg, U. M. (2003). New technique for assessment of self-perception in fibromyalgia patients: a pilot study with video-interpretation. *Arts Psychother.* 31, 257–264. doi: 10.1016/S0197-4556(04)00031-0

Bonaduce, D., Marciano, F., Petretta, M., Migaux, M. L., Morgano, G., Bianchi, V., et al. (1994). Effects of converting enzyme inhibition on heart period variability in patients with acute myocardial infarction. *Circulation* 90, 108–113. doi: 10.1161/01.cir.90.1.108

Brage, S., Brage, N., Franks, P. W., Ekelund, U., and Wareham, N. J. (2005). Reliability and validity of the combined heart rate and movement sensor Actiheart. *Eur. J. Clin. Nutr.* 59, 561–570. doi: 10.1038/sj.ejcn.1602118

Brand, G., Sloboda, J., Saul, B., and Hathaway, M. (2012). The reciprocal relationship between jazz musicians and audiences in live performances: a pilot qualitative study. *Psychol. Music* 40, 634–651. doi: 10.1177/0305735612448509

Cannon, W. B. (1935). Stresses and strains of homeostasis. *Am. J. Med. Sci.* 189, 13–14. doi: 10.1097/00000441-193501000-00001

Cha, E. S., Kim, K. H., and Erlen, J. A. (2007). Translation of scales in cross-cultural research: issues and techniques. *J. Adv. Nurs.* 58, 386–395. doi: 10.1111/j.1365-2648.2007.04242.x

Cirelli, L. K., Einarson, K. M., and Trainor, L. J. (2014). Interpersonal synchrony increases prosocial behavior in infants. *Dev. Sci.* 17, 1003–1011. doi: 10.1111/desc.12193

Cohen, S., Janicki-Deverts, D., and Miller, G. E. (2007). Psychological stress and disease. *JAMA* 298, 1685–1687. doi: 10.1001/jama.298.14.1685

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper & Row Publishers.

Csikszentmihalyi, M., and Nakamura, J. (2010). "Effortless attention in everyday life: a systematic phenomenology" in *Effortless attention: A new perspective in the cognitive science of attention and action*. ed. B. Bruya (Cambridge, MA: The MIT Press).

de Manzano, Ö., Theorell, T., Harmat, L., and Ullén, F. (2010). The psychophysiology of flow during piano playing. *Emotion* 10, 301–311. doi: 10.1037/a0018432

Eckberg, D. L., and Kuusela, T. A. (2005). Human vagal baroreflex sensitivity fluctuates widely and rhythmically at very low frequencies. *J. Physiol.* 567, 1011–1019. doi: 10.1113/jphysiol.2005.091090

Fachner, J. C., Maidhof, C., Grocke, D., Nygaard Pedersen, I., Trondalen, G., Tucek, G., et al. (2019). "Telling me not to worry…" hyperscanning and neural dynamics of emotion processing during guided imagery and music. *Front. Psychol.* 10:1561. doi: 10.3389/fpsyg.2019.01561

Grape Viding, C., Osika, W., and Bojner Horwtiz, E. (2017). You can't feel healthier than your caregiver – the ripple effect of trust and empathy for patients and health care staff, cultivated through cultural activities. *J. Nurs. Care* 6:5. doi: 10.4172/2167-1168.1000422

Harmat, L., de Manzano, Ö., Theorell, T., Högman, L., Fischer, H., et al. (2015). Physiological correlates of the flow experience during computer game playing. *Int. J. Psychophysiol.* 97, 1–7. doi: 10.1016/j.ijpsycho.2015.05.001

Harmat, L., De Manzano, O., and Ullén, F. (in press). "Flow in music and arts" in *Advances in flow research. 2nd Edn.* eds. C. Peifer and S. Engeser (Springer).

Harmat, L., Ullén, F., de Manzano, Ö., Olsson, E., Elofsson, U., von Schéele, B., et al. (2011). Heart rate variability during piano playing: a case study of three professional solo pianists playing a self-selected and a difficult prima vista piece. *Music Med.* 3, 102–107. doi: 10.1177/1943862110387158

Hart, E., and Di Blasi, Z. (2015). Combined flow in musical jam sessions: a pilot qualitative study. *Psychol. Music* 43, 275–290. doi: 10.1177/0305735613502374

Jackson, S. A., and Eklund, R. C. (2002). Assessing flow in physical activity: the flow state scale-2 and dispositional flow scale-2. *J. Sport Exerc. Psychol.* 24, 133–150. doi: 10.1123/jsep.24.2.133

Jackson, S. A., and Eklund, R. C. (2004). *The flow scales manual*. Morgantown: Publishers Graphics.

Jackson, S. A., Martin, A. J., and Eklund, R. C. (2008). Long and short measures of flow: the construct validity of the FSS-2, DFS-2, and new brief counterparts. *J. Sport Exerc. Psychol.* 30, 561–587. doi: 10.1123/jsep.30.5.561

Kawabata, M., Mallett, C. J., and Jackson, S. A. (2008). The flow state scale-2 and dispositional flow scale-2: examination of factorial validity and reliability for Japanese adults. *Psychol. Sport Exerc.* 9, 465–485. doi: 10.1016/j.psychsport.2007.05.005

Keeler, J. R., Roth, E. A., Neuser, B. L., Spitsbergen, J. M., Waters, D. J. M., and Vianney, J. M. (2015). The neurochemistry and social flow of singing: bonding and oxytocin. *Front. Hum. Neurosci.* 9:518. doi: 10.3389/fnhum.2015.00518

Kitney, R. (1980). "An analysis of the thermoregulatory influences on heart-rate variability" in *The study of heart-rate variability*. eds. R. I. Kitney and O. Rompelman (Oxford, England: Clarendon Press).

Kristiansen, J., Korshøj, M., Skotte, J. H., Jespersen, T., Søgaard, K., Mortensen, O. S., et al. (2011). Comparison of two systems for long-term heart rate variability monitoring in free-living conditions-a pilot study. *Biomed. Eng. Online* 10:27. doi: 10.1186/1475-925X-10-27

LeBlanc, A., Jin, Y. C., Obert, M., and Siivola, C. (1997). Effect of audience on music performance anxiety. *J. Res. Music. Educ.* 45, 480–496. doi: 10.2307/3345541

Lindseth, A., and Norberg, A. (2004). A phenomenological hermeneutical method for researching lived experience. *Scand. J. Caring Sci.* 18, 145–153. doi: 10.1111/j.1471-6712.2004.00258.x

Magyaródi, T., and Oláh, A. (2015). A cross-sectional survey study about the most common solitary and social flow activities to extend the concept of optimal experience. *Eur. J. Psychol.* 11, 632–650. doi: 10.5964/ejop.v11i4.866

Magyaródi, T., and Oláh, A. (2017). The effect of social interaction on flow experience. *Int. J. Psychol. Behav. Anal.* 3:126. doi: 10.15344/2455-3867/2017/126

O'Grady, L. (2009). The therapeutic potentials of creating and performing music with women in prison: a qualitative case study [Doctoral dissertation]. Barcelona Publisher.

Peifer, C., Schulz, A., Schächinger, H., Baumann, N., and Antoni, C. H. (2014). The relation of flow-experience and physiological arousal under stress—can u shape it? *J. Exp. Soc. Psychol.* 53, 62–69. doi: 10.1016/j.jesp.2014.01.009

Pieper, S., Brosschot, J. F., van der Leeden, R., and Thayer, J. (2010). Prolonged cardiac effects of momentary assessed stressful events and worry episodes. *Psychosom. Med.* 72, 570–577. doi: 10.1097/PSY.0b013e3181dbc0e9

Porges, S. W. (2007). The polyvagal perspective. *Biol. Psychol.* 74, 116–143. doi: 10.1016/j.biopsycho.2006.06.009

Sawyer, R. K. (2006). Group creativity: musical performance and collaboration. *Psychol. Music* 34, 148–165. doi: 10.1177/0305735606061850

Sawyer, K. (2007). Group flow and group genius. *The NAMTA Journal* 40, 29–52.

Shaffer, F., McCraty, R., and Zerr, C. L. (2014). A healthy heart is not a metronome: an integrative review of the heart's anatomy and heart rate variability. *Front. Psychol.* 5:1040. doi: 10.3389/fpsyg.2014.01040

Task Force (1996). Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the north American Society of Pacing and Electrophysiology. *Circulation* 93, 1043–1065. doi: 10.1161/01.CIR.93.5.1043

Taylor, J. A., Carr, D. L., Myers, C. W., and Eckberg, D. L. (1998). Mechanisms underlying very-low-frequency RR-interval oscillations in humans. *Circulation* 98, 547–555. doi: 10.1161/01.cir.98.6.547

Theorell, T., and Bojner Horwitz, E. (2019). Emotional effects of live and recorded music in various audiences and listening situations. *Medicines* 6:16. doi: 10.3390/medicines6010016

Tozman, T., Magdas, E. S., MacDougall, H. G., and Vollmeyer, R. (2015). Understanding the psychophysiology of flow: a driving simulator experiment to investigate the relationship between flow and heart rate variability. *Comput. Hum. Behav.* 52, 408–418. doi: 10.1016/j.chb.2015.06.023

Tozman, T., and Peifer, C. (2016). "Experimental paradigms to investigate flow experience and its psychophysiology: inspired from stress theory and research" in *Flow experience: Empirical research and applications*. eds. L. Harmat, O. Anderssen, F. Ullén, J. Wright and G. Sadlo (Switzerland: Springer), 329–350.

Ullén, F., de Manzano, Ö., Theorell, T., and Harmat, L. (2010). "The physiology of effortless attention: correlates of state flow and flow proneness" in *Effortless attention*. ed. B. Bruya (Cambridge, MA, USA: The MIT Press), 205–218.

Ulrich-Lai, Y. M., and Herman, J. P. (2009). Neural regulation of endocrine and autonomic stress responses. *Nat. Rev. Neurosci.* 10, 397–409. doi: 10.1038/nrn2647

Wrigley, J. W., and Emmerson, S. B. (2013). The experience of the flow state in live music performance. *Psychol. Music* 23, 292–305. doi: 10.1177/0305735611425903

**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.

frontiers in Psychology

ORIGINAL RESEARCH
published: 06 May 2021
doi: 10.3389/fpsyg.2021.645993

# Typologies of Adolescent Musicians and Experiences of Performance Anxiety Among Instrumental Learners

*Ioulia Papageorgi**

*Department of Social Sciences, Faculty of Humanities and Social Sciences, University of Nicosia, Nicosia, Cyprus*

Literature suggests that music performance anxiety (MPA) is prevalent in adolescence, a developmental period with increased likelihood of experiencing anxiety under evaluative conditions. Evidence also indicate that individuals may respond to evaluative situations in distinct ways. Factors contributing to the individuality of responses in evaluative situations (such as test taking and musical performance) are not yet fully understood. This study investigated student typologies in adolescent instrumental learners. Participants included 410 learners who completed the Young Musicians' Performance Questionnaire. K-Means cluster analysis revealed three typologies: Cluster 1 – moderately anxious students evidencing lower levels of motivation and feeling ineffective but guarding their self-esteem; Cluster 2 – highly anxious students evidencing negative self-perceptions and being susceptible to experiencing maladaptive MPA; Cluster 3 – low anxious students evidencing high levels of motivation and confidence and inclined toward experiencing adaptive MPA. The 3-cluster solution effectiveness was validated with discriminant analysis. Significant associations between examination achievement and cluster membership revealed variations between clusters. Thematic analysis of qualitative data facilitated further understanding of their characteristics. This study adds to the body of MPA literature by exploring the different ways with which adolescent musicians interpret and respond to anxiety inducing situations. Findings have implications for clinical and educational practice.

**Keywords: music performance anxiety, adolescence, typology, K-means cluster analysis, thematic analysis**

OPEN ACCESS

***Edited by:***
*Gary Edward McPherson,*
*The University of Melbourne, Australia*

***Reviewed by:***
*Eckart Altenmüller,*
*Hanover University of Music, Drama and Media, Germany*
*Emery Schubert,*
*University of New South Wales, Australia*

****Correspondence:***
*Ioulia Papageorgi*
*papageorgi.i@unic.ac.cy*

***Specialty section:***
*This article was submitted to Performance Science, a section of the journal Frontiers in Psychology*

***Received:*** *24 December 2020*
***Accepted:*** *12 April 2021*
***Published:*** *06 May 2021*

***Citation:***
*Papageorgi I (2021) Typologies of Adolescent Musicians and Experiences of Performance Anxiety Among Instrumental Learners. Front. Psychol. 12:645993. doi: 10.3389/fpsyg.2021.645993*

## INTRODUCTION

Musical performance engages auditory and visual perception skills, attention, precise timing, extended control over movement, learning, memory, and emotion (Kanduri et al., 2015). Developing expertise in instrumental performance is a complex process requiring the development of aural, cognitive, technical, musical, communication, and performing skills (Hallam, 2006), and is categorised under complex skill learning. Music performance is an activity that is physically and psychologically demanding (Papageorgi and Kopiez, 2018) and can thus place physical and psychological stress on performers (Vervainioti and Alexopoulos, 2015). One of the most common psychological stressors reported in the literature is music performance anxiety

(MPA) (Barbar et al., 2014; Kenny et al., 2014; Burin and Osorio, 2017; Fernholz et al., 2019). Symptoms can be categorised under physiological (changes in the physiological state of the organism), mental (cognitive and emotional) and behavioural (visible changes in behaviour) and are experienced concurrently (Burin and Osorio, 2017). A major concern for musicians is that MPA can bring about practical complications and negatively affect the quality of performance, irrespective of the level of preparation and ability of the performer (Papageorgi and Welch, 2014a). MPA can vary in severity, to the extent of even forcing some musicians who are highly affected to put an end in their careers (Fernholz et al., 2019).

Contemporary conceptualisations of MPA view it as a multidimensional construct (Wilson, 2002; Osborne and Kenny, 2005; Papageorgi et al., 2007; Kenny et al., 2014). Contemporary theoretical models take into account the complexity of individual, task-related, and situational variables in the development and manifestation of MPA and view its development as a process that unfolds in a time sequence: pre-, during-, and post- performance (Papageorgi et al., 2007).

## Prevalence of MPA

Music performance anxiety prevalence in the literature has been mostly assessed through self-report psychometric tests and not by professionals on the basis of ICD or DSM criteria (Fernholz et al., 2019). Evidence suggests that MPA is highly prevalent in musicians (both children/adolescents and adults), but there is no consensus on prevalence rates. Between 15 and 70% of adult orchestral and choral musicians report the experience of MPA (Schulz, 1981; Gustafson and Rawson, 1983; Fishbein et al., 1988; Marchant-Haycox and Wilson, 1992; Steptoe, 2001; Brugués, 2009; Ryan and Andrews, 2009; Medeiros Barbar et al., 2014). A recent systematic review of MPA in professional musicians reported prevalence rates between 16.5 and 60% (Fernholz et al., 2019), noting that when focusing on reports indicating MPA to be a severe problem for musicians the evidence suggest that approximately 1/3 of musicians suffer from MPA.

Student musicians also report the presence of MPA to a significant extent. Wesner et al. (1990) reported that 21% of their sample of university students suffered from severe distress and 40% experienced moderate distress due to performance anxiety. American music students reported performance anxiety as one of the most disturbing and worrying issues in their lives (Dews and Williams, 1989). Around 50% of Swedish student and professional performers stated that they had often encountered performance anxiety, and 10% said that they always felt high distress before performing (Gustafson and Rawson, 1983). MPA was one of the most frequently reported problems for conservatoire students in the United Kingdom (Williamon and Thompson, 2006).

In younger musicians, Papageorgi (2020) reported that in a sample of 410 adolescent musicians, 11% reported high MPA levels. Fehm and Schmidt (2006) found that 33% of adolescent musicians experienced a negative impact of MPA, and about 10% reported that their musical career was negatively affected.

## MPA in Adolescence

Adolescence has been reported as a developmental period with increased vulnerability to MPA (Fehm and Schmidt, 2006; Kenny and Osborne, 2006). Younger musicians' experiences of MPA in terms of physiological, cognitive and behavioural symptoms are similar to those of adult musicians (Ryan, 1998, 2004; Fehm and Schmidt, 2006; Osborne and Kenny, 2008). Anxiety can also be evident in the early years. Boucher and Ryan (2011), using self-report, physiological and observational data found that children as young as 3–4 years old experience MPA. Recently, Zarza-Alzugaray et al. (2016) found that onset of musical training at or before the age of 7 may act as a protective barrier against MPA in music school pupils (mean age 12 years old).

A small number of cohort studies looked at the developmental trajectory of MPA in adolescence. Evidence suggest that that MPA increases with age as older adolescents appear to be more anxious (Osborne et al., 2005; Patston and Osborne, 2016). Dempsey and Comeau (2019) compared MPA in child (aged 7–12) and adolescent musicians (aged 13–17) and found that MPA levels increased with age, with adolescents experiencing significantly higher anxiety levels.

Papageorgi (2020) found that MPA in adolescents increased between the ages of 15 and 18, and decreased at the age of 19. The study also found evidence of cultural- and sex-specific variations in the manifestation of MPA. Girls experienced higher levels of MPA compared to boys, in agreement with the trend reported in the literature (Rae and McCambridge, 2004; Kenny and Osborne, 2006; Papageorgi, 2008; Papageorgi et al., 2013; Thomas and Nettelbeck, 2014; Patston and Osborne, 2016). In terms of the developmental trajectory, there was a steep increase in girls' levels of MPA between the ages of 15 and 18, a finding corroborating Patston and Osborne (2016) who also found that females experienced a steeper and more intense developmental trajectory.

The developmental trajectory may be affected by factors such as the personality and chosen musical genre of the performer, which may yield differentiations. Sarbescu and Dorgo (2014) in their study of 14–19 years old musicians reported that it was younger students who also had lower levels of emotional stability and lower performance frequency who reported higher levels of MPA. A further issue that must be taken into account is that the developmental trajectory of MPA may depend on the musical genre, with different developmental MPA profiles for classical and popular musicians. Nusseck et al. (2015), in a study of 7–20 years old music school students found that younger classical musicians (under age 16) were more anxious compared to their older counterparts, whereas the opposite trend was observed in popular musicians with older musicians (age 16+) evidencing higher MPA levels.

## Individual Differences in Responses to Performance Anxiety

Individuals do not respond to evaluative situations in the same way. Evidence suggests that individuals with different characteristics (e.g., age, gender, experience, and expertise level) can perceive and respond to conditions giving rise to performance anxiety or test anxiety in distinct ways (Losiak, 2005; Chow and Mercado, 2020).

Chow and Mercado (2020) looked at how task-specific expertise and past experiences moderate the degree to which individuals become anxious in a given performance context. Authors considered how individual differences arising from learning can influence the psychobiological, emotional, and cognitive processes that modulate anxious states during the performance of highly trained skills (such as in musical performance and sport performance). Their review concluded that experience dependent brain plasticity determines how a person will respond to each unique performance context. Authors suggested that in order to identify the processes that lead to maladaptive (problematic) or adaptive (advantageous) performance anxiety, theories of performance anxiety that clarify for how past experiences shape the emergence of emotional states that trigger cognitive appraisals are needed.

Nielsen et al. (2018) investigated the relationship between MPA, subjective performance quality and post-event rumination (PER) in university music students. Authors found that high-anxious music students showed more negative and less positive PER compared to the low-anxious music students after a 10-min solo performance. Furthermore, the development pattern for negative PER was related to MPA level. High-anxious musicians evidenced a slower decrease of negative PER compared to low- and moderate-anxious performers. These results suggest that the quality of musicians' responses to anxiety-inducing performances is dependent upon their MPA level, with those susceptible to higher MPA being more likely to need more time to recover after a performance.

## Profiles of Test Takers and Test/Performance Anxiety

The recognition that individuals can respond to anxiety-inducing evaluative situations in different ways has steered research to investigate the presence of profiles or types of test takers in relation to anxiety.

### Test Anxiety Literature

Cassady and Finch (2015) undertook a study with undergraduate students to explore whether it is possible to identify subcomponents of the cognitive test anxiety construct, hypothesising the presence of "types" of test anxious learners. They observed differences in the way low and high anxious individuals conceptualised cognitive test anxiety. Those with low levels of cognitive test anxiety understood cognitive test anxiety as a unidimensional construct. High test anxious students, however, represented it as a more complex construct and differentiated between two different aspects: Cognitive interference (both during the preparation stage and the performance phase) and perceived skill deficiencies in test taking. Their results suggest that the level of anxiety experienced is associated with how cognitive test anxiety is understood by learners.

Davis et al. (2008) employed cluster analysis in a sample of university students, exploring patterns of appraising tests. Authors were able to identify five subtypes (clusters) of test takers who varied their approach toward tests, their experience of anxiety, as well as how they coped with problems occurring during test taking: (1) Tests out of control; (2) Well-prepared for challenges; (3) Feeling hopeless; (4) Keeping tests in perspective; and (5) Bracing for the worst. These findings indicate that different profiles of test takers exist, differentiated by their overall approach toward tests, their experience of anxiety, and coping patterns utilised to manage test anxiety.

Putwain and Daly (2013) investigated whether secondary school students can be distinguished in distinct clusters on the basis of their test anxiety and academic buoyancy (resilience) scores, as well as their academic performance. Cluster analysis revealed the presence of five clusters with diverse combinations of test anxiety level and academic resilience. Cluster 1 represented high anxiety/low academic buoyancy and had the lowest academic performance. Cluster 2 corresponded to low anxiety/high academic buoyancy and had the highest academic performance. Cluster 3 represented moderate anxiety/moderate academic buoyancy. Cluster 4 represented moderate-high test anxiety/moderate buoyancy. Cluster 5 represented moderate anxiety/high buoyancy. These results highlight the distinctiveness of student responses to performance evaluative situations. Differences in academic performance between the clusters was also noted. Higher performance was observed in students exhibiting high academic buoyancy and either low (suggestive of lower threat appraisal) or moderate (signifying a buffering effect on performance) test anxiety levels.

In another study, Losiak (2005) investigated anxiety profiles in a non-clinical sample of undergraduate students. Cluster analysis suggested the presence of six clusters, which were categorised into two groups of profiles. The first group of profiles included three clusters in which the participants had average, moderate and low anxiety levels. These profiles were not differentiated, as no form of anxiety reactions or anxiety provoking situations dominated – they were rather unitary in characteristics and did not score extremely in anxiety measures. The second group of profiles included three clusters that were quite differentiated. One cluster included extremely anxious individuals who experienced motor anxiety in phobic situations. Another cluster included high anxious individuals characterized by physiological anxiety presented in daily life and evaluative situations (e.g., in exams, public speaking). The third cluster included participants who were very low anxious with a tendency to experience physiological anxiety in phobic and interpersonal situations (e.g., a date or social meetings). This study suggests that individuals can respond to anxiety-inducing situations in distinct ways.

Mammarella et al. (2018) conducted a study with primary school children to investigate the presence of subgroups expressing particular anxiety patterns on the basis of measures of general anxiety (GA), test anxiety (TA), and maths anxiety (MA). Authors also explored the role of personal factors protecting against anxiety (such as self-concept, resilience, academic self-concept, and academic buoyancy). Three clusters of students were identified: Low risk students (low scores in various forms of anxiety); Average risk students (average scores in anxiety measures); and High-risk students (high scores in anxiety measures) anxiety. The three clusters differed in their experiences of the different types of anxiety. For example, whereas the low risk group had low scores in all measures (GA, TA, and MA), the average risk students scored higher in GA and MA, and lower in TA. The high risk students evidenced higher levels of GA and TA, but lower levels of MA. Children with a low-risk profile reported a more positive academic self-concept and felt more competent overall compared to the other two groups.

Carey et al. (2017) measured math anxiety, test anxiety, general anxiety and mathematics and reading performance in primary school children in the United Kingdom. Latent Profile Analysis revealed different anxiety "profiles" in year 4 students and year 7 and 8 students. Year 4 students' data analysis produced four profiles: Low anxiety, Slight anxiety, Moderate anxiety and High anxiety. Scores in the various anxiety measures were homogenous within each profile. The analysis of year 7 and 8 students' scores revealed four groups with varying patterns of scores on each of the three anxiety measures. One group included participants with normative scores on all measures (Low anxiety). The second group evidenced high scores only on the measure of general anxiety (General anxiety). The third group scored highly on test anxiety and mathematics anxiety (Academic anxiety). The fourth group had high scores on all three anxiety measures (High anxiety). This study provides further evidence for the presence of differences in anxiety responses between students, and also suggests that there may be a developmental trajectory.

Stenlund et al. (2018) conducted a study aiming to identify and group test takers with similar patterns of test-taking behaviour and to investigate how these groups differ in relation to background characteristics and test performance in a high-stakes achievement test context (the Swedish SAT). Cluster analysis revealed three clusters of test takers with significantly different test-taking behaviour profiles: Moderate ($n$ = 741), Calm Risk Taker ($n$ = 637), and Test Anxious Risk Averse ($n$ = 513). The calm risk taker profile (high level of risk taking, low levels of test anxiety and motivation) had the higher performance. The test anxious risk averse profile (low degree of risk taking, high levels of test anxiety and motivation) had the lower performance.

In a recent study, Wang et al. (2020) looked at the developmental trajectory of mathematics anxiety (MA) in secondary school students and its association with cognitive, personality and environmental factors. Results identified four distinct growth trajectories of maths anxiety: (1) A non-anxious group (exhibited chronically low MA); (2) A highly anxious group (displayed moderately high MA over time); (3) A resilient group (exhibited high initial MA but it steadily decreased over time); and (4) A vulnerable group (reported low initial MA but it drastically increased over time). Researchers pointed that their findings highlight the presence of heterogeneity in the development of MA, and identified middle school as a critical period for MA development. They also called attention to the importance of examining developmental changes in cognitive, personality, and environmental factors in elucidating distinct MA trajectories in secondary school.

### MPA Literature

Wiedemann et al. (2019) investigated MPA and its anxiety correlates in undergraduate music students. They were able to identify, through cluster analysis, two groups of participants evidencing different patterns of anxiety. Cluster 2 showed elevated anxiety levels in nearly all DSM-5 anxiety measures used (e.g., generalised anxiety disorder and social anxiety disorder). On the contrary, Cluster 1 did not display pathological anxiety symptoms in the DSM-5 anxiety measures, but their MPA included both lower and higher levels. Findings support the differentiation between different MPA types. These may include MPA as an important disorder presented in an otherwise healthy musician, and MPA as part of a more complex psychopathology.

## The Current Study

The review of the literature suggests that MPA is prevalent in adolescence. Test anxiety literature also suggests that individuals may perceive and respond to evaluative situations in distinct ways, resulting in the presence of distinct profiles of test takers in relation to the experience of test anxiety. The factors that contribute to the individuality of responses in evaluative situations (such as test taking and musical performance) are not yet fully understood. Test anxiety literature has made progress in this area, by investigating through cluster analysis, the presence of profiles of test takers in relation to their test anxiety levels and achievement. With the exception of Wiedemann et al. (2019), MPA literature has not addressed this issue and so there is limited knowledge on the presence of typologies of instrumental learners in relation to performance anxiety.

The aim of the current study was to investigate the presence of student typologies in adolescent learners and explore associations with experiences of MPA and achievement in instrumental examinations. To the best knowledge of the author, no studies have investigated student typologies and MPA in adolescent learners and, as such, the current study has the potential to make a novel contribution to the literature. The current study focuses on classical musicians, who in the literature have been reported as more likely to experience higher MPA levels (Papageorgi et al., 2013; Papageorgi and Welch, 2014b; Perdomo-Guevara, 2014; Nusseck et al., 2015).

Findings from relevant studies in the field of test anxiety have guided the development of the research questions of the study. Davis et al. (2008) found different profiles of students based on test anxiety experienced and coping strategies they used. More recently, Putwain and Daly (2013) found different profiles of test takers with each profile evidencing different achievement levels. The investigation sought to explore whether similar phenomena occur with instrumental students and MPA. Literature on MPA has not yet explored whether distinguishable typologies/profiles of musicians can be detected on the basis of correlates of MPA and, in such case, whether different typologies are associated with different levels of achievement in instrumental examinations. To explore these issues, the following research questions were devised:

1. Is it possible, through cluster analysis, to identify distinct profiles (typologies) of adolescent musicians, based on correlates of performance anxiety and experiences of MPA?

2. If typologies of adolescent musicians are identified, how do the clusters relate to achievement in instrumental examinations?

# MATERIALS AND METHODS

## Participants

Four hundred and ten adolescent musicians took part in the study (ages 12–19 $M_{age}$ = 15.33). Data were collected in two geographical locations, Cyprus (51.5% of participants) and the United Kingdom (48.5% of participants). The decision to collect data from two geographical locations was driven by one of the main aims of the overall study, which was to investigate cultural variability in MPA, with relevant findings reported in an earlier publication (Papageorgi, 2020). All participants attended junior conservatoires where they undertook instrumental lessons, and/or played in youth orchestras. There were slightly more females (58%) compared to males (42%). Participants played a range of instruments, including piano/keyboard, string, woodwind, brass, percussion instruments, guitar, voice, and harp and identified themselves as classical musicians.

As participants came from two different geographical locations and nationalities, a comparison of the two nationality samples was initially conducted to assess their homogeneity and ascertain whether data should be analysed separately or collectively. Chi-square test analyses assessing the association between nationality and basic demographics of sex (males and females) and age group (younger adolescents 12–15, older adolescents 16–19) revealed no significant differences (sex $p = 0.554$; age group $p = 0.711$). Some statistically significant differences were observed in years playing first instrument and grade level, indicating that more British students played for longer [8+ years; $x^2(5) = 46.55$, $p < 0.001$] and were placed at advanced grades [grade 8; $x^2(8) = 75.20$, $p < 0.001$]. There was no significant association between the two nationalities in attainment as measured by the assessment grade received in the last exam taken ($p = 0.289$). Overall, the results suggested that the Cypriot and British samples were homogenous in basic characteristics and attainment, although the British students were slightly more advanced. Taking the evidence into consideration, it was decided to analyse data collectively.

## Materials

The Young Musicians' Performance Questionnaire (Papageorgi, 2007a, 2020) was used to collect data pertaining to learning, performance and experiences of musical performance anxiety, as part of a study investigating performance anxiety in adolescent musicians. Students completed the questionnaire on paper during a break from rehearsals or class. The items comprising the questionnaire were validated by two experts in the field during the development phase and prior to a pilot study (for further details, please see Papageorgi, 2020). Following necessary modifications after the pilot study, the questionnaire was finalised and then translated in Greek for the Cypriot participants using backward translation (for further details, see Papageorgi, 2020).

The questionnaire was broken down into four sections. Section A requested demographic information [sex, age, nationality, instrument(s) played, years playing, grade level, and mark (assessment grade) obtained in the last examination taken]. Section B included 54 statements (items) to which participants were requested to indicate their degree of agreement on a five-point Likert-type scale (1 = strongly disagree; 5 = strongly agree). The items comprised potential MPA predictor variables or correlates, identified on the basis of previous literature, with a focus on the three thematic areas suggested by Papageorgi et al. (2007) and Burin and Osorio (2017): the performer, the task and perceptions of the environment. Each of these variables was explored via items capturing its characteristics and related behaviours, as reported in the literature. They were grouped into 18 variables/mini scales (three items per variable/short scale), recoding any items necessary to ensure agreement in conceptual direction. Hair et al. (2010, p. 669) suggested that three items can ensure adequate coverage of a construct's theoretical domain. Furthermore, Gogol et al. (2014) suggested that short forms (three-item or single item) of domain-general and domain-specific academic anxiety measures can be recommended as psychometrically sound alternatives when study designs require brief measures. Internal reliability analyses of the short scales comprising each variable indicated acceptable internal consistency, with all Cronbach's α values >0.6. For the purposes of the current article, the variables indicated in **Table 1** were used (for further details see Papageorgi, 2020).

Section C included twenty items that together formed the Adolescent Musicians' Performance Anxiety Scale (AMPAS) (see Papageorgi, 2007a,b, 2020). Participants were asked to indicate the frequency of experiencing particular symptoms and cognitions on a five-point Likert type scale (1 = never; 5 = always). The possible range of scores ranged from 20 to 100, with higher scores indicating higher performance anxiety levels. The items were organised in subscales or variables/themes that reflected the multidimensional nature of MPA as exemplified in previous literature and, for the purposes of the current analysis, were combined into a new variable using a similar procedure as in Section B: Negative outcome expectancies (two items), negative experiences in performance (five items), evidence of pre-evaluation anxiety (three items), experience of physiological symptoms of anxiety (four items), concern about others' judgement (two items) and experience negative effects of anxiety (four items). Assessment of the psychometric properties of the AMPAS scale suggests that it is a promising psychometric instrument to measure performance anxiety in this population, with good internal reliability (Cronbach's α = 0.86) and 2 week test-retest reliability ($r$ = 0.91), and good construct validity ($r$ = 0.74 with KMPAI; Kenny, 2009 and $r$ = 0.80 with PAI, Nagel and Himle, 1989) (for further details see Papageorgi, 2020). Factor analysis revealed the presence of two higher order factors explaining 62% of the variance. The first factor explained 45% of the variance and had an internal focus with key variables being the anxiety symptoms, ruminations and effects of anxiety experienced (evidence of pre-evaluation anxiety, experience of physiological symptoms of anxiety, negative experiences in performance, negative outcome expectancies and experience

**TABLE 1 |** Variables included in section B of the questionnaire.

| Variable | Example item |
|---|---|
| Entity theory of ability | Talent is the most important component of success for a performer |
| Fear of evaluation | If I do poorly in a performing situation, others will question my ability in music |
| Incremental theory of ability | With lots of practice, one can accomplish anything. |
| Perception of critical parents with high expectations | The comments of my family about my performances are critical |
| Perception of supportive and encouraging parents | My parents are always supportive |
| Perception of receiving positive feedback from teacher | My teacher believes I could go on to be a professional musician |
| Perception of being under pressure to continue with music lessons | I am worried about how my parents would react if I decided to stop going to music lessons |
| Development of musical identity | My dream is to become a professional performer |
| Negative perception of anxiety | Anxiety can destroy the career of a musician |
| Positive self-concept in music | When I compare myself with other young performers I think that I am very good |
| Perfectionism | I am sad if I do not get a good result in an exam/audition |
| Effortful practice | I usually practice very hard |
| Low self-efficacy in music | I avoid working on pieces of music that look or sound difficult |
| Intrinsic motivation to learn a musical instrument | I enjoy learning to play a musical instrument and I want to be good at it |
| Experience of heightened anxiety in the presence of an audience | I am more nervous when I know someone is listening to me playing |
| Sensitivity to degree of self-exposure | Solo performances are scarier than group performances |
| Sensitivity to environmental conditions | I find that when I do not feel comfortable in the venue where I have to play, I am more anxious than usual |

negative effects of anxiety). The second factor explained 17% of the variance and had an external focus, the worry in relation to how others viewed the individual (concern about others' judgement). The six AMPAS variables/themes shown in **Table 2** have been used (for further details see Papageorgi, 2020).

Section D included six open ended questions that required students to write down their thoughts on a number of performance-related issues that explored (1) Self-concept in music (Q: Rate yourself in relation to other young performers – scale ranging from poor to excellent. Please explain your response), (2) Perfectionistic tendencies (Q: It is important to always get top marks – scale ranging from strongly disagree to strongly agree. Please explain your response), (3) Attributions of success (Q: Remember an occasion in which you performed and did really well. Please write down what you think led to that positive result), (4) Attributions of failure (Q: Remember an occasion in which you performed but did not go as well as you hoped. Please write down what you think led to that negative result), (5) Self-awareness regarding MPA level (Q: I would describe myself as a low, moderately or highly anxious performer (choose one option) and give reasons for your answer), and (6) Strategies used to cope with performance anxiety (Q: Write down any strategies you use to help you cope with performance anxiety and perform better).

## Ethical Considerations

The research was approved by the UCL Institute of Education Research Ethics Committee (the affiliation of the author at the time the study took place). Initially, contact was made with the directors of several conservatoires and youth orchestras, and only after receiving a positive response were they approached to discuss the data collection procedure. Participation to the study was voluntary, and informed consent from parents or caregivers was obtained for all participants. Participants were briefed about the rationale and aims of the study, both orally and in writing. The questionnaire included a cover page stating the aims of the study, reassuring that all information was confidential and that participants retained the right to withdraw at any time. The researcher's contact details were also included. Participants were assured that all information would be used for the purposes

**TABLE 2 |** Variables included in section C of the questionnaire.

| Variable name | Example item |
|---|---|
| Negative outcome expectancies | Sometimes, before an important I find myself thinking: "This is too difficult. I am not going to do well," even though I may have worked really hard in preparing for that event |
| Negative experiences in performance | During my recitals I feel great (RECODED) |
| Evidence of pre-evaluation anxiety | I worry a lot for several days before I take a recital examination in front of a jury |
| Experience of physiological symptoms of anxiety | During recitals my heart beats very fast |
| Concern about others' judgement | Especially if I score low in an exam or audition, I do not tell anyone exactly what my score was |
| Negative effects of anxiety | I believe that anxiety makes me forget parts of the music, makes it difficult to concentrate on my playing, and eventually has a negative result on my performance |

of the study and that they would remain anonymous in the dissemination of findings.

## Data Analysis

### Quantitative Data

Quantitative data were analysed with the IBM SPSS software programme. In order to identify typologies of adolescent performing musicians, K-Means cluster analysis, an optimisation clustering technique, was used. Cluster analysis aims at grouping individuals in such a way that those that are allocated into a particular group are similar (Bartholomew et al., 2002). The aim is to discover the arrangement of observations and clusters that maximises within-group homogeneity and, at the same time, between-group heterogeneity (Everitt et al., 2011). This approach enables grouping students into clusters based on the similarity of their responses, followed by researcher-generated descriptions of these on the basis of the most prominent variables in each cluster and can therefore provide information on the presence of different student typologies (also referred to as student profiles). K-Means cluster analysis was chosen as the aim was to investigate whether distinct profiles of adolescent musicians exist on the basis of the similarities of their responses concerning MPA experienced and the strategies used to cope with the demands of performance, as the literature on test anxiety suggests for other learning domains. This technique enabled to address a yet unanswered question in the MPA literature, whether instrumental learners perceive and respond to evaluative situations in consistent but distinct ways. Data were assessed for outliers, normality, linearity and homoscedasticity of independent variables and their suitability for K-Means cluster analysis was confirmed prior to the analysis. Variables were standardized into $z$-scores to enable their comparison and to minimize any bias in weighting which may have resulted from differing measurement scales and ranges. Validation of the cluster solution was conducted through Discriminant Analysis. Associations between the emerging clusters and achievement in instrumental examinations were investigated using Pearson's Chi-Square Test.

### Qualitative Data

Qualitative data, derived from responses to open-ended questions, were transcribed and entered into the NVivo software programme for thematic analysis. The analysis followed an iterative process of categorisation into themes according to a seven-stage process (Cooper and McIntyre, 1993):

1. Reading a random sample of scripts.
2. Identifying points of similarity and difference among these transcripts in relation to the research questions.
3. Generating theories, on the basis of 2, describing emergent answers to the research questions.
4. Testing theories against a new set of transcripts.
5. Testing new theories against transcripts that have already been dealt with.
6. Carrying all existing theories forward to new transcripts.
7. Repeating the above process until all data have been examined and all theories tested against all data.

The coding was conducted by the author, and was later validated by two experts through a process of checking the identified themes with a random selection of scripts. If disagreements were observed, these were discussed until an agreement was reached by all parties.

The purpose of the qualitative data was to enrich the quantitative data, providing further insight into the characteristics of the student typologies. The open questions also touched upon issues not covered by the quantitative data that are important in furthering our understanding of the MPA experience for each of the typologies, such as attributions of success and failure, and coping strategies for dealing with MPA. Qualitative data facilitated further understanding of the characteristics and behaviours of each type of student in relation to the development of self-concept, perfectionistic tendencies, attributions of success and failure, self-awareness regarding performance anxiety level and strategies used to cope with performance anxiety. Four participants from each cluster group, all cases with a small distance from the classification centre of each cluster, were selected to provide further insight into the typologies of adolescent performing musicians as they were illustrative of the characteristics of each type/profile of students. Effort was made to ensure that the cases chosen were representative of the sample in terms of the key demographic variables of nationality, gender and age. The presentation of the examples is carried out according to the theme of each open ended question.

# RESULTS

## Emerging Clusters – Typologies of Adolescent Musicians

The quantitative analysis initially explored various cluster solutions until the optimal was obtained. Cluster solutions of six, five, and four clusters were rejected in favour of a three cluster solution. The three cluster solution was optimal as the analysis showed that at the 9th iteration convergence was achieved due to no further change in cluster centres, signifying that the cluster model had stabilised. In the other cluster solutions iterations failed to converge at the default number of 10 iterations, suggesting that these models were not optimal. The three-cluster solution produced clusters broadly containing a similar number of students (Cluster 1: 144 students; Cluster 2: 121 students; Cluster 3: 145 students). This solution was successful in yielding three student groups with distinctively different characteristics as evidenced from the cluster centres' mean values (**Table 3** and **Figure 1**)[1]. One-way ANOVA with cluster membership as the independent variable confirmed the existence of statistically significant differences between the three groups in all dependent variables ($p < 0.05$) but one, which was nevertheless borderline significant ($p = 0.05$). Furthermore, the three cluster solution clearly differentiated students on the basis of reported MPA level, and grouped them into low (cluster 3), moderate (cluster 1) and high (cluster 2) MPA levels. This provided support for theoretical

[1] Variables are standardised, with a mean of 0 and a standard deviation of 1.

**TABLE 3 |** Final cluster centre mean values for each variable included in the analysis.

| | Cluster | | |
|---|---|---|---|
| | **1 Moderately anxious students evidencing lower levels of motivation and feeling ineffective but guarding their self-esteem** | **2 Highly anxious students evidencing negative self-perceptions and being susceptible to experiencing maladaptive MPA** | **3 Low anxious students evidencing high levels of motivation and confidence and inclined toward experiencing adaptive MPA** |
| Entity theory of ability | −0.13 | 0.12 | 0.03 |
| Fear of evaluation | −0.01 | 0.40 | −0.33 |
| Incremental theory of ability | −0.28 | 0.00 | 0.27 |
| Perception of supportive and encouraging parents | −0.48 | 0.22 | 0.29 |
| Perception of critical parents with high expectations | −0.27 | 0.20 | 0.10 |
| Perception of receiving positive feedback from teacher | −0.45 | −0.29 | 0.70 |
| Perception of being under pressure to continue with music lessons | −0.01 | −0.02 | 0.03 |
| Development of musical identity | −0.34 | −0.13 | 0.42 |
| Negative perception of anxiety | 0.17 | −0.01 | −0.15 |
| Positive self-concept in music | −0.27 | −0.51 | 0.70 |
| Perfectionism | −0.02 | 0.29 | −0.22 |
| Effortful practice | −0.45 | −0.27 | 0.66 |
| Low self-efficacy in music | 0.30 | 0.42 | −0.64 |
| Intrinsic motivation to learn a musical instrument | −0.65 | 0.14 | 0.52 |
| Experience of heightened anxiety in the presence of an audience | 0.02 | 0.87 | −0.74 |
| Sensitivity to degree of self-exposure | −0.09 | 0.74 | −0.54 |
| Sensitivity to environmental conditions | −0.21 | 0.38 | −0.11 |
| Negative outcome expectancies | 0.05 | 0.68 | −0.60 |
| Negative experiences in performance | −0.07 | 0.95 | −0.71 |
| Evidence of pre-evaluation anxiety | −0.16 | 0.83 | −0.52 |
| Experience of physiological symptoms of anxiety | −0.27 | 0.90 | −0.48 |
| Concern about others' judgement | 0.25 | −0.02 | −0.23 |
| Experience negative effects of anxiety | −0.14 | 0.66 | −0.40 |

conceptualisations of MPA in the literature (e.g., Papageorgi et al., 2007). The effectiveness of the proposed cluster solution was further validated using discriminant analysis (see relevant section below). For all the above reasons, the 3-cluster solution was considered to be optimal. The cluster membership of each case and the distance of each case from the classification cluster centre were saved for further analysis. The final cluster centre mean values of the variables in each of the clusters were then examined to identify the most prominent variables that guided the interpretation of the clusters.

## Cluster 1: Moderately Anxious Students Evidencing Lower Levels of Motivation and Feeling Ineffective but Guarding Their Self-Esteem

Cluster 1 students were strongly characterised by the fact that they reported low intrinsic motivation to learn a musical instrument. They perceived their parents not to be supportive/encouraging and they did not place great effort into practising. They also did not perceive receiving positive feedback from their teacher, suggesting that they felt inefficient and unsuccessful as performers. These students had not developed a strong sense of musical identity, had low-self-efficacy beliefs and a negative self-concept in music. They also held an entity theory of ability. Of the three groups of students, this group was the one with values closest to the mean (0) on the variables "evidence of pre-evaluation anxiety" and "experience of physiological symptoms of anxiety," suggesting that they experienced moderate arousal levels. Qualitative data (see below) suggested that this group may have tried to protect their self-esteem through their behaviour.

## Cluster 2: Highly Anxious Students Evidencing Negative Self-Perceptions and Being Susceptible to Experiencing Maladaptive MPA

Cluster 2 students were characterised by the fact that they had negative experiences in performance and experienced physiological arousal more intensely. They had the highest score on the variables "experience of physiological symptoms

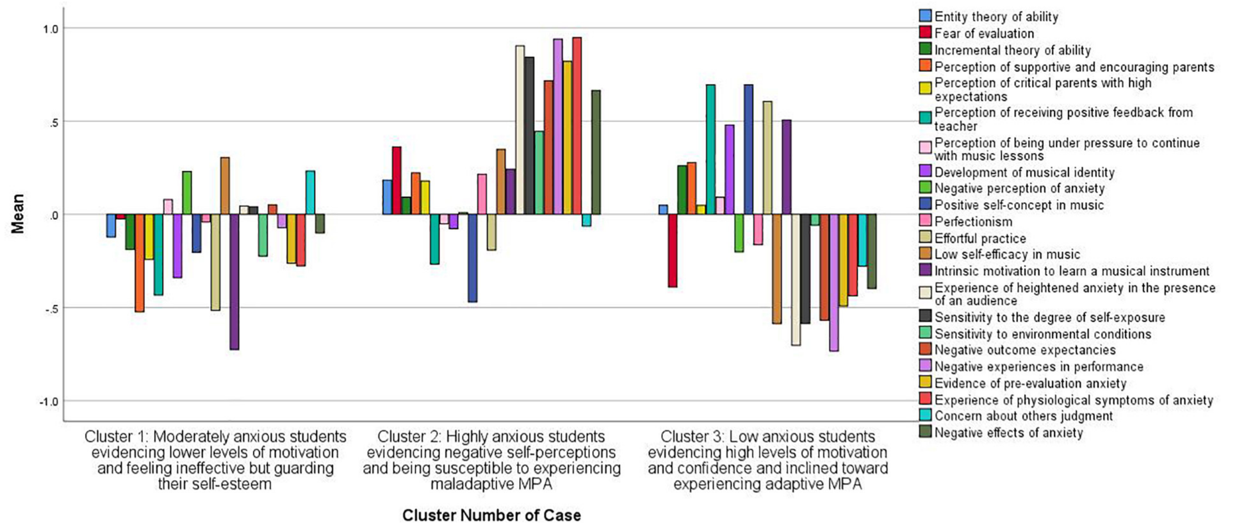



**FIGURE 1 |** The 3-cluster solution.

of anxiety" and "evidence of pre-evaluation anxiety." They were more anxious in the presence of audience and experienced high levels of pre-evaluation anxiety. In addition, their anxiety was influenced by how exposed they felt when performing (they felt more anxious when playing solo as opposed to playing in an orchestra, for example), they had negative outcome expectancies, experienced the negative effects of performance anxiety, held a negative self-concept in music, low self-efficacy beliefs and perfectionistic tendencies. All the above made these students less self-confident and more vulnerable to experiencing maladaptive performance anxiety. They did not perceive receiving positive feedback from their teacher, which may suggest that they were not successful or that they tended to interpret feedback in a negative way. They also indicated that they did not invest much effort in practising, suggesting an inverse association between motivation for practice and experience of musical performance anxiety.

### Cluster 3: Low Anxious Students Evidencing High Levels of Motivation and Confidence and Inclined Toward Experiencing Adaptive MPA

Cluster 3 students felt motivated in the presence of an audience, they had developed a positive self-concept in music and were confident players. They were intrinsically motivated, practised a lot and perceived that they received positive feedback from their teacher. They also had a strong sense of musical identity. They had some experience of physiological arousal and pre-evaluation anxiety, but their arousal and anxiety remained at low and manageable levels with values −0.52 and −0.48 on the variables "evidence of pre-evaluation anxiety" and "experience of physiological symptoms of anxiety," respectively. They reported that they had positive experiences in performance and that anxiety had positive effects on their performance (as indicated by the minus sign on the variables "negative experiences in performance" and "comment on the negative effects of anxiety"), indicative of anxiety having a facilitative or adaptive effect.

## Validation of Cluster Solution

Discriminant analysis was carried out to explore how efficient the cluster solution was. The three clusters of students formed the groups for the discriminant analysis and the twenty-three variables used in the cluster analysis (comprising sections B and C of the questionnaire) were placed as the independent variables. Two discriminant functions were calculated, which had a significant overall Wilk's lambda [$\Lambda = 0.15$, $\chi^2$ (46) = 548.64, $p < 0.001$]. After removal of the effects of the first function, the second discriminant function was also statistically significant [$\Lambda = 0.51$, $\chi^2$ (22) = 196.95, $p < 0.001$]. Both discriminant functions were successful at discriminating between the three groups of students. The first discriminant function accounted for 71% of the variance in the solution and the second one for 29%. The first function had an eigenvalue of 2.36 and a canonical correlation of 0.84. The $\eta^2$, obtained by squaring the canonical correlation was 0.71, indicating that 71% of the variability of the scores for the first discriminant function was accounted for by differences among the three groups of students. The second function had an eigenvalue of 0.972 and a canonical correlation of 0.70. The $\eta^2$ was 0.49, indicating that 49% of the variability of the scores for the second discriminant function was accounted for by differences among the three groups of students.

Function 1 was named "high anxiety and low effectiveness." It focused on issues that made students vulnerable to experiencing maladaptive performance anxiety, such as negative experiences in performance, experience of heightened anxiety in the presence of an audience, sensitivity to the degree of self-exposure, experience of physiological symptoms of anxiety, a negative perception of oneself, low self-confidence and a lack of effort. The perception of receiving negative feedback from the teacher was also an important variable in this function.

Function 2, which was orthogonal to the first function, was named "high motivation and success." The focus of this function was intrinsic motivation to learn a musical instrument.

Other important variables included supportive and encouraging parents, effortful practice, positive feedback and development of musical identity.

The canonical discriminant functions were evaluated at group means. This allowed the three cluster groups to be viewed in relation to the two discriminant functions. **Table 4** summarises the results.

Function 1 (high anxiety and low effectiveness) mostly distinguished clusters 2 and 3. These two groups of students were at the opposite end of the spectrum with respect to performance anxiety, as cluster 3 experienced adaptive performance anxiety that enhanced performance, whilst cluster 2 experienced maladaptive performance anxiety that obstructed their performance. Function 2 (high motivation and success) mostly distinguished clusters 1 and 2. Cluster 1 included students that evidenced low levels of motivation, moderate MPA and perceptions of being less successful, whilst cluster 2 included students that were moderately motivated but experienced high MPA, and perceived themselves to be less successful as they had negative performance experiences.

Cluster group membership prediction was also assessed, to investigate how well student group membership could be predicted by using a classification function. Results suggested that 88.3% of cluster 1, 92% of cluster 2 and 97.4% of cluster 3 students were predicted correctly. Overall, 92.8% of the sample was classified correctly. In order to estimate how well the classification functions could predict a new sample, the classification was estimated by selecting the leave-one-out option. Results showed that 81.6% of cluster 1, 85.1% of cluster 2 and 94.7% of cluster 3 students were correctly classified using the leave-one-out classification. Overall, 87.5% of the cross-validated cases were correctly classified.

## Associations Between Examination Achievement and Cluster Membership

Pearson's chi-square test was used to assess associations between achievement in instrumental exams and cluster membership in the data. A statistically significant association was observed between the mark obtained in the last instrumental examination students took and cluster membership [$\chi^2$ (6, $N$ = 321) = 14.61, $p$ = 0.024]. **Figure 2** illustrates that cluster 1 (moderately anxious students evidencing lower levels of motivation and feeling ineffective but guarding their self-esteem) were most likely to get a Pass/Grade C, as they formed the highest percentage from the three clusters in this category. Cluster 2 (highly anxious students evidencing negative self-perceptions and being susceptible to experiencing maladaptive MPA) were most likely to either Fail or get Merit/Grade B (they formed the largest percentage of the three groups in the Fail and Merit/B categories). Cluster 3 (low anxious students evidencing high levels of motivation and confidence and inclined toward experiencing adaptive MPA) were most likely to get a Distinction/Grade A, as they formed the largest percentage from the three clusters in this category. *Post hoc* tests revealed that it was mostly marks C and A that contributed to the significant association between clusters. Significantly more Cluster 1 students and significantly less Cluster 3 students than expected received mark C (Adj. residuals = 2.1 and −3.1, respectively). Furthermore, significantly more Cluster 3 students than expected received mark A (Adj. residual = 3.0).

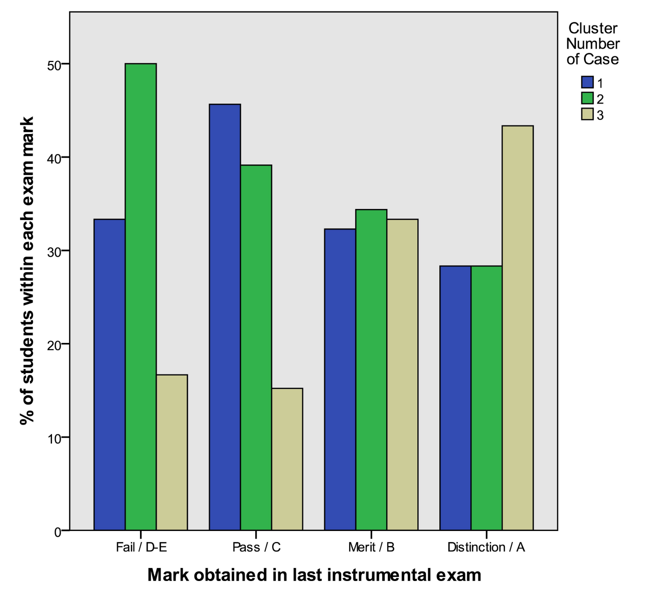



**FIGURE 2 |** Examination achievement and cluster membership.

**TABLE 4 |** Group centroids for discriminant functions.

| Cluster number of case | Function 1 | Function 2 |
|---|---|---|
| 1 | 0.237 | −1.362 |
| 2 | 2.018 | 0.851 |
| 3 | −1.755 | 0.581 |

## Qualitative Data Representing Each Student Type

### Cluster 1: Moderately Anxious Students Evidencing Lower Levels of Motivation and Feeling Ineffective but Guarding Their Self-Esteem

#### *Development of self-concept in music*

The score on the variable "perception of receiving positive feedback from teacher" (−0.45) for these students suggested that they felt the least successful of the three groups. Additionally, they had the lowest score (−0.45) on the variable "effortful practice." Their score on the variable "positive self-concept in music" (−0.27) suggested that they held a negative self-concept. However, when asked directly about their attainment, 62.5% reported that they considered themselves to be between good and excellent performers compared to other young musicians, indicating that they held a positive perception of themselves overall. This group of students perhaps made inaccurate self-assessment in relation to performance competence or was in denial in an effort to preserve their self-esteem. They explained their responses by saying that:

*"I think I am good because I have a lot of potential and if I put my mind to it, I can achieve a lot"*

(Male, Cypriot, 18 years old)

*"I think I am average because I am not talented, but I also don't practise enough because I don't have the time and so I do not perform as I well as I could. I have more potential. If I use it I believe I will reach a very high level"*

(Female, British, 16 years old)

These responses showed that they felt that they had potential as performers but both students implied that they did not work as much as they needed to in order to realise that potential. This might indicate an effort to maintain a positive self-concept in music and to guard self-esteem by attributing an unsatisfactory performance to lack of effort rather than lack of ability. Such students protected their self-esteem by thinking that if they actually practised enough, they would be successful. This defensive mechanism may have ensured that such students' self-esteem was not jeopardised.

***Perfectionistic tendencies***

The majority (60.4%) of students in cluster 1 agreed or strongly agreed with the statement "it is important for me to always get top marks in exams or auditions." Some explained this by saying that this would help them achieve their goals (being accepted at university).

*"I agree because although there is more to music than exams, exams show me that I am in fact progressing well – they will stand me in good stead for uni!"*

(Female, British, 16 years old)

Others appeared to be striving to achieve high marks because they saw it as a way of reinforcing their self-concept. High marks for this group of students might have therefore functioned as a way of ensuring they did not damage their self-esteem.

*"I agree because high marks make you feel better"*

(Male, Cypriot, 16 years old)

***Attributions of success***

When asked to remember an occasion in which they performed well and write down what they thought led to that positive result, students in cluster 1 responded in a variety of ways and evidenced both internal and external attributions of success. On the one hand, some mentioned the importance of their effort, feeling confident and being concentrated, which indicated an internal attribution of success, but also mentioned the venue (external attribution).

*"I knew the piece very well, I was confident in my own ability to play it. Not a threatening venue/atmosphere – small concert, sympathetic audience. Some nerves – helped me to focus"*

(Female, British, 18 years old)

Other students spoke about having a motive that helped them perform well and playing in a non-threatening venue/atmosphere (small audience, sympathetic audience), which pointed toward an external attribution of success. They also mentioned effort and concentration (internal attributions).

*"My effort, my concentration, the motive"*

(Male, Cypriot, 19 years old)

It is noteworthy that some students mentioned the presence of some nerves as facilitating concentration and focus during the performance.

***Attributions of failure***

In stating reasons that led to a past unsuccessful performance, students in cluster 1 mostly provided external attributions for failure.

*"I had not felt especially nervous beforehand, so when I went on stage I suddenly felt that I was not mentally prepared. Also, more critical audience – masterclass situation"*

(Female, British, 18 years old)

*"The atmosphere was cold and the audience was unfamiliar"*

(Male, Cypriot, 19 years old)

*"My school teacher and I fell out and she rushed the accompaniment. I was also quite upset"*

(Female, British, 16 years old)

These students strove to maintain a positive self-concept, and therefore attributed failures to situations beyond their control in an effort to protect their self-esteem. Some students also attributed their failure to low levels of anxiety prior to performing.

***Self-awareness regarding performance anxiety level***

The majority of students in cluster 1 described themselves as moderately anxious performers (22.7% – low anxious; 62.1% – moderately anxious; 15.2% – highly anxious). The qualitative data supported this:

*"I am anxious before the performance. When I start playing, not at all"*

(Male, Cypriot, 18 years old)

*"I do get anxious, but it's not uncontrollable panic"*

(Female, British, 18 years old)

The lack of intense nerves may have been related to a lack of caring about audience perceptions, although it could also indicate further protection of self-esteem.

*"I don't really care what other people think"*

(Female, British, 16 years old)

***Strategies for coping with performance anxiety***

Coping strategies used by students in cluster 1 seemed to focus on maintaining a positive approach to the performance and remaining confident. They believed that when adequate preparation was made, there was no reason to get anxious.

*"Breath and just think that most people are supportive and understand the ordeal of performance"*

(Female, British, 16 years old)

*"I think about my teacher's advice. In order to play a piece 100%, I must know it 200%. If I have practised, there is no reason for anxiety"*

(Male, Cypriot, 18 years old)

Maintaining a positive approach to the performance and ensuring that they practised sufficiently guarded against getting overly anxious.

Overall, students in cluster 1 appeared to be somewhat have low levels of motivation. Their comments indicated that they were driven by a desire to protect their self-concept. Some of their comments suggested one reason they might not have devoted the necessary effort in performance was in an effort to guard their self-esteem. Such behaviour allowed them to maintain a positive view of themselves despite placing low effort into practising, possibly thinking that if they had put the necessary effort they would have been successful. Reported strategies for coping with MPA were emotion-focused rather than problem-focused.

## Cluster 2: Highly Anxious Students Evidencing Negative Self-Perceptions and Being Susceptible to Experiencing Maladaptive MPA

### *Development of self-concept in music*

Almost half of the respondents in cluster 2 (48.3%) reported that they assessed themselves to be average performers and 14.2% reported to be below average or poor performers. This supported findings from the cluster analysis, which indicated negative self-concept in music and low self-efficacy beliefs in cluster 2.

> *"I think I am average because I don't believe that I have as much knowledge as I would like to in order to be a very good performer compared to others. This is mostly due to the fact that I don't play a musical instrument"* [This student was a singer].
> (Female, Cypriot, 18 years old)

> *"I think I am average because there are people who are both better and worse than me"*
> (Female, British, 18 years old)

These students seemed to construct their music self-concept by direct comparison with others. Whilst students in cluster 1 referred to their abilities and potentials and students in cluster 3 mentioned confidence and evidenced a positive approach to learning (see below), students in cluster 2 evaluated themselves on the basis of how they perceived the abilities of other young musicians.

### *Perfectionistic tendencies*

Most of the students in this group (79.8%) agreed or strongly agreed that it was important to receive top marks in exams or auditions. They indicated perfectionistic tendencies, this being one of the factors that possibly triggered their maladaptive response.

> *"If I do not get a high mark I feel as if I have let down myself, my teacher and my parents, as I know I am capable of achieving a high mark"*
> (male, British, 17 years old)

> *"I agree that it's important for me to get top marks in exams or auditions because others congratulate me and I feel better"*
> (Female, Cypriot, 12 years old)

Responses suggested that these students sought external approval and also felt external pressure to do well.

### *Attributions of success*

Responses to questions about the factors that led to a successful performance suggested that they attributed their successes to both internal and external factors. Some mentioned being well prepared and feeling comfortable in the performance environment.

> *"I was well prepared, I tried to lower my anxiety and I felt comfortable at the place I was playing"*
> (Female, Cypriot, 18 years old)

> *"I was well prepared. I was playing in front of people I knew"*
> (male, British, 17 years old)

Lowering anxiety levels seemed to produce positive results for these students, and an adequate preparation and a feeling of ease in the performing environment seemed to promote this.

### *Attributions of failure*

Students in cluster 2 mostly evidenced internal attributions to failure. This is an important difference from students in cluster 1, who attributed failure to external factors. Students in cluster 1 protected their self-esteem, whilst students in cluster 2, maintained and further jeopardised their low self-esteem.

> *"I hadn't practised as much as I should have, and I was very-very anxious"*
> (Female, Cypriot, 12 years old)

> *"Getting too worked up about the performance, several days beforehand"*
> (male, British, 17 years old)

Most students in this group attributed their failure to the fact that they had not practised enough. It is noteworthy that these students also mentioned that they felt highly anxious and experienced pre-performance anxiety for days before the event. There was an association between lack of adequate practice and the experience of high levels of performance anxiety.

### *Self-awareness regarding performance anxiety level*

Most students in cluster 2 considered themselves to be highly anxious (56.3%). In the comments they provided further details.

> *"I am generally very anxious before a performance or examination due to fear of something not going well or failing the examination"*
> (Female, Cypriot, 18 years old)

> *"I get very nervous generally"*
> (Female, British, 18 years old)

In justifying their response, students noted fear of failure and that they were generally anxious as individuals. The latter comment evidenced a link between trait anxiety and state anxiety, a well-documented association in MPA literature.

### *Strategies for coping with performance anxiety*

Responses of students in cluster 2 suggested that they saw effective practising as a means of relieving their anxiety about performance and feeling more secure about themselves.

> *"Practice – the right kind that focuses on the energy sections or especially difficult sections"*
> (Female, British, 18 years old)

They also rehearsed a lot, especially in the venue where the examination/concert was to take place.

> *"Rehearsing in the exam/concert venue so that we familiarise ourselves with the environment"*
>
> (Female, Cypriot, 12 years old)

Student comments suggested that students in cluster 2 had developed problem-focused coping strategies. They appeared to be more proactive in their approach to coping with anxiety compared to students in cluster 1.

### Cluster 3: Low Anxious Students Evidencing High Levels of Motivation and Confidence and Inclined Toward Experiencing Adaptive MPA

#### *Development of self-concept in music*

Students in cluster 3 demonstrated the most positive approach to performance and had a positive self-concept as musicians, as 90.8% of these students rated themselves between good and excellent performers. Confidence, in combination with low levels of performance anxiety, was considered to promote achievement.

> *"I am good because I generally feel confident and have little anxiety in my performances"*
>
> (Female, British, 14 years old)

Other students mentioned that they did not feel threatened by other musicians. They considered critical listening as an opportunity to take positive elements of other musicians and saw this as a positive learning experience.

> *"I am good because I want to feel confident and not get disappointed. I can take, by watching other young performers, the positive elements of their performance".*
>
> (male, Cypriot, 13 years old)

#### *Perfectionistic tendencies*

Most of the students in cluster 3 said that getting high marks in exams or auditions was important for them (72.1%). Some saw high marks to be a reward for the hard work they put into their preparation.

> *"I believe that getting high marks in exams or auditions is something like a reward for me, and so I pursue this"*
>
> (male, Cypriot, 15 years old)

Some students, however, disagreed with the importance given to high marks (14%). They acknowledged that an examination evaluates performance at one specific point in time, appreciated the volatility of an examination outcome due to external factors and considered that an examination result was not a reflection of ability or effort.

> *"Because examinations are not a total reflection of your talent or efforts, as your playing is assessed at a time where anything good or bad can happen"*
>
> (Female, British, 14 years old)

Overall, students in cluster 3 saw achievement as a reward for their hard work, but at the same time held a more pragmatic view, acknowledging that one specific examination result could not be a valid reflection of capability and effort.

#### *Attributions of success*

Students in cluster 3 attributed successful past performances both to internal and external factors. For instance, effort and teacher support.

> *"The effective and systematic practice and the support I received from my teacher"*
>
> (male, Cypriot, 13 years old)

Some students referred to their commitment to the music.

> *"I felt confident and I really loved the piece and I probably conveyed that to the judges and audience"*
>
> (Female, British, 14 years old)

#### *Attributions of failure*

Failure was attributed to both internal and external parameters by students in cluster 3. Some students attributed failure to anxiety because of lack of effort.

> *"Unfortunately, on some occasions I did not manage to give my best because of anxiety. I felt that I hadn't practised the pieces as much as I should have and that made me give an anxious performance"*
>
> (male, Cypriot, 15 years old)

Other factors relating to organisation issues and the venue were also mentioned.

> *"Poor organisation on competition's part, unsatisfying venue, leading to awkward feelings about playing"*
>
> (Female, British, 15 years old)

#### *Self-awareness regarding performance anxiety level*

Most students in cluster 3 perceived themselves to be moderately anxious performers (51.8%), although a significant percentage rated themselves as low anxious (39.7%). Students in cluster 3 experienced lower levels of physiological arousal compared to the other two clusters (see **Table 3**). Students raised a number of points in explaining their responses, such as enjoying performing rather than feeling threatened by it.

> *"I see music as something that pleases me and not as something that makes me anxious"*
>
> (male, Cypriot, 15 years old)

Other students spoke about being able to use performance anxiety to enhance the quality of their performance.

> *"Though I am usually quite anxious about performing, I usually manage to use this to my advantage to play well"*
>
> (Female, British, 15 years old)

#### *Strategies for coping with performance anxiety*

Coping strategies used by students in cluster 3 focused on having a positive attitude to the performance.

> *"I just think positively"*
>
> (male, Cypriot, 15 years old)

Other students also mentioned focusing on the positive things of their performance as opposed to focusing on the things that they did not like, as well as concentrating on communicating with the audience through their performance.

*"Have a really positive train of thought, think of the good things that you do in the piece and try to love the piece as much as possible because it will show in your performance and therefore will have a positive effect on the audience"*

(Female, British, 15 years old)

These students had developed active problem-focused strategies for coping with anxiety, not only before the performance, but also during the actual performance.

## DISCUSSION

The cluster analysis revealed the existence of three typologies or profiles of instrumental students: "moderately anxious students evidencing lower levels of motivation and feeling ineffective but guarding their self-esteem," "highly anxious students evidencing negative self-perceptions and being susceptible to experiencing maladaptive MPA," and "low anxious students evidencing high levels of motivation and confidence and inclined toward experiencing adaptive MPA." An inspection of both quantitative and qualitative data assisted in a more thorough understanding of the characteristics of each of these student types and their relationship with the experience of performance anxiety.

Cluster 1, moderately anxious students evidencing lower levels of motivation and feeling ineffective but guarding their self-esteem, were characterised by the fact that they reported low intrinsic motivation to engage in instrumental learning, placed low effort into practising and perceived their parents not to be supportive. They had a weak sense of musical identity and were troubled by low self-efficacy and negative self-concept beliefs. These students experienced moderate arousal levels and felt that they were not successful (they reported that the feedback they received from their teacher was negative). Inspection of the qualitative data suggested that these students strove to guard their self-esteem. High marks for these students served to reinforce self-concept. They attributed their failures to external factors that were beyond their personal control, a behaviour that might have served to protect self-esteem. With respect to their attitudes regarding achievement, there was indication that most had perfectionistic tendencies, and some saw success as a way of reinforcing their self-concept ("feeling better"). This group of students employed emotion-focused rather than problem-focused strategies. Most students focused on ensuring they felt positive and confident before the performance, something that they felt would guard them against feeling too anxious. These students were most likely to do moderately well (get a Pass) in instrumental examinations, which indicated that they did well enough to make sure they did not jeopardise their self-esteem. The coping strategies used, in combination with the attempt to guard self-esteem and lower levels of motivation correspond with some of the characteristics of cluster 1 (Tests out of control) in Davis et al. (2008).

Cluster 2, highly anxious students evidencing negative self-perceptions and being susceptible to experiencing maladaptive MPA, were characterised by the fact that they had negative experiences in performance, experienced physiological arousal more intensely and held negative views about themselves and their ability as performers. They were often apprehensive about an upcoming performance days before an event and usually had negative outcome expectancies. They commented strongly on the negative effects that anxiety may have on performance. Their personal characteristics included negative self-concept in music, low self-efficacy beliefs and perfectionism. They perceived the feedback that they received from their teacher to be negative and invested little effort in practice. This finding, in conjunction with the frequent negative outcome expectancies that they reported, supports earlier research which suggests that estimations of high probability of failure tend to result in the investment of little effort (Covington and Omelich, 1981; Carver and Scheier, 1986). Qualitative data suggested that these students tended to compare themselves with others and generally held negative self-views. Moreover, they felt the need for external approval and pressure to achieve in music, and perhaps this was the reason they considered achieving high marks to be so important. They tended to attribute their failures to internal factors, which may explain their negative self-concepts and low self-efficacy beliefs and perhaps contributed to maintaining their low self-esteem and promoted sensitivity to maladaptive performance anxiety. Weiner (1985) stressed the importance of achievement attributions, suggesting that when negative events are explained in terms of internal, stable, and controlled factors, students feel more pessimistic. They were self-aware regarding how anxious they were as performers, which may explain the use of proactive and problem-focused strategies for coping with performance anxiety, such as practising adequately and rehearsing in the performance venue in advance of an examination/concert. The importance of acclimatisation within the performance environment when students prepare for a performance has been highlighted by Hallam (2006), and is a view reinforced by the findings of the present study. These students were most likely to either Fail or do well (get Merit) in instrumental examinations. This might imply that when students were not in control of their nerves, these impeded their performance, whilst when they were able to cope with them they were able to perform well. This group had some similar characteristics to those described in cluster 3 (Feeling hopeless) by Davis et al. (2008) in terms of experiencing higher anxiety levels and good achievement (grade B). Nevertheless, in the current study some participants in this cluster of highly anxious students noted that they Failed in examinations, a finding consistent with cluster 1 (high anxiety/low academic buoyancy) in Putwain and Daly (2013) who had the lowest academic achievement. This latter finding supports literature linking physiological arousal with performance efficiency (Yerkes and Dodson, 1908; Wilson, 2002; Papageorgi et al., 2007), relating the experience of high levels of physiological arousal with maladaptive performance anxiety and the impairment of performance skills.

Cluster 3, low anxious students evidencing high levels of motivation and confidence and inclined toward experiencing adaptive MPA, were characterised by positive experiences in performances. When compared to the other two groups, they experienced lower arousal and anxiety levels in performance. Students in this group were generally confident, had a strong sense of musical identity, a positive self-concept in music and enjoyed their engagement with musical activities, without feeling threatened by the presence of an audience. They were intrinsically motivated and reported that they devoted a lot of time to practising. They perceived the feedback that they received to be positive. Although their physiological arousal levels were not at zero level (which would impede concentration on the task), these were maintained at a low and manageable level that eventually helped them perform well. The qualitative data reinforced these findings, showing that these students perceived themselves to be highly competent. They held a more pragmatic view on the importance of achievement. Achieving high marks in examinations and auditions was important for them, not because they sought external reassurance of their abilities, but because they saw this as a reward for their hard effort. They maintained a healthy attitude to such rewards, as they acknowledged that an examination result represented the quality of their performance at a specific point in time, and not their general ability in music. These students' attitudes in attributing failure in performance were a balanced combination of internal and external factors. They were self-aware in relation to their anxiety levels and stressed that they did not feel threatened by performance and that they engaged in this activity because it was something they inherently enjoyed doing. They were able to use the arousal they felt prior to performance in a positive way and believed that this enhanced the quality of their playing, supporting research stating that performance anxiety can lead to positive outcomes under certain conditions (Hamann, 1982; Gates and Montalbo, 1987; Wills and Cooper, 1988; Kemp, 1996; Papageorgi, 2008; Papageorgi et al., 2013). Overall, these students maintained a healthy, balanced and thoughtful approach to performance, which was also evident in the coping strategies they employed. They emphasised the importance of maintaining a positive approach to performance in dealing with performance anxiety (emotion-focused strategies) but also employed problem-focused strategies that were activated during performance, e.g., focusing on the music and communication. These students were most likely to do very well (get Distinction) in instrumental examinations, which further reinforced the fact that any anxiety experienced did not impede their performance but helped them perform at their best. The characteristics of this group are in agreement with cluster 2 (well-prepared for challenges) in Davis et al. (2008) and cluster 2 (low anxiety/high academic buoyancy) in Putwain and Daly (2013), where students with low anxiety levels evidenced high levels of achievement and higher levels of academic buoyancy. Evidence for cluster 3 suggest the development of resilience and successful coping with MPA, resulting in pre-performance arousal facilitating performance rather than hindering it.

Differences between the three clusters of students were observed in relation to the strategies they employed for coping with performance anxiety. Moderately anxious students evidencing lower levels of motivation and feeling ineffective but guarding their self-esteem (cluster 1) did not report any strategies other than being adequately prepared in order to minimise their anxiety, indicating that they were not proactive in developing problem-focused strategies specifically for performance, but their approach was emotion-focused. Highly anxious students evidencing negative self-perceptions and being susceptible to experiencing maladaptive MPA (cluster 2) mentioned effective practising, but also mentioned proactive problem-focused strategies, such as rehearsing in the venue prior to the performance. This indicates that these students planned their performance, perhaps in an effort to minimise the negative effects of anxiety. Whilst cluster 1 and 2 students' strategies mostly focused on strategies they could employ themselves, low anxious students evidencing high levels of motivation and confidence and inclined toward experiencing adaptive MPA (cluster 3) were able to shift their attention away from themselves. They mentioned concentrating on the enjoyment of music and focusing on communication with the audience in addition to maintaining a positive approach prior to performance. This indicated that cluster 3 students had devised active problem-focused coping strategies and that their thinking about performance had progressed to a more advanced level.

Overall, student responses are suggestive of differences in the way that each cluster approached instrumental learning and performance, as well as their experiences of MPA. Several of the concepts included in the study relate to concepts of motivation theory. Several findings from this study can be interpreted through the lens of the expectancy-value theory of achievement motivation, which states that individuals' choice, persistence, and performance on an activity can be explained by their expectations of success (beliefs about how well they will perform) and its incentive value [the extent to which they value the activity (Wigfield and Eccles, 2000)]. For example, highly anxious students' negative perceptions of their ability (on the basis of negative experiences in performance and negative feedback received from significant others) in combination with negative self-perceptions (low self-efficacy and low self-concept in music) and perfectionistic tendencies lead to negative outcome expectancies and investment of little effort in practice (low motivation for practice) in high-stakes situations (such as examinations). These conditions can increase MPA levels, resulting in additional negative experiences in performance and, in the long-term, they can contribute to more negative self-perceptions that further increase susceptibility to maladaptive anxiety in the future.

Current literature suggests that adolescence is a time of increased vulnerability to MPA (Fehm and Schmidt, 2006; Kenny and Osborne, 2006), and recent studies have revealed the presence of a developmental trajectory of MPA, with performance anxiety appearing to increase from childhood to adolescence and then throughout adolescent years (Osborne et al., 2005; Patston and Osborne, 2016; Papageorgi, 2020). Professional experience could act as a protective factor, as literature has suggested that the impact of anxiety on performance is mediated by musicians' performance experience when comparing higher education students with professional musicians (Papageorgi et al., 2013). It is important to note at this point that we must not ignore the possibility that some younger or still developing musicians suffering severely from MPA may have ceased their engagement with instrumental learning and so the adolescent, adult and professional musicians' samples presented in the literature may not include those who have experienced extreme MPA levels and severe negative effects as children.

## LIMITATIONS

Cluster analysis is a descriptive technique, which can result in solutions that are non-unique (Romesburg, 1984). A three-cluster solution was chosen as it was effective in distinguishing between different types of students and was supported by theoretical conceptualisations of MPA in the literature, but it should be acknowledged that the interpretation encompasses some degree of subjectivity. Although the cluster solution was validated through discriminant analysis, which provided support for the chosen interpretation, other researchers may have theoretical or statistical reasons to disagree. Another limitation is the use of self-report instruments, which may increase the likelihood of response bias (Bradburn et al., 2004). Future research could investigate the extent to which the three-cluster solution is replicated in other samples of instrumental musicians, not only children and adolescents but also adults. As the data were collected via a self-report questionnaire, assumptions about causality cannot be made. This study could, however, serve as a starting point for further research looking into cause-and-effect relationships between the variables that have been identified as being associated with the experience of MPA.

## CONCLUSION

This study adds to the body of MPA literature by exploring the different ways with which adolescent musicians interpret and respond to anxiety inducing situations. Findings have implications for clinical and educational practice. For example, interventions for students reporting low levels of motivation who feel less successful but, through their behaviours, aim to guard their self-esteem should focus on improving their sense of self and sense of achievement and reinforce that a person's self-worth should not depend on achievement grades. For students who experience high levels of performance anxiety, interventions should concentrate on techniques for managing physiological arousal and reducing the negative effects on the quality of performance. The aim should be to facilitate the experience of success to support the development of a positive sense of self and increase intrinsic motivation. Better understanding of how MPA is manifested and the different ways with which individuals experience and respond to evaluative situations is critical in the design of tailor-made intervention programs aiming at reducing MPA in adolescents, as well as in planning appropriate educational support. The ultimate aim should be to facilitate positive performance experiences so that students derive pleasure and benefit from their engagement with music.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the UCL Institute of Education Research Ethics Committee. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

IP conceptualised the study, collected and analysed the data, and authored the manuscript.

## ACKNOWLEDGMENTS

The information included in this manuscript is an extension of previous work, specifically a Ph.D. thesis completed at UCL Institute of Education (Papageorgi, 2007a).

## REFERENCES

Barbar, A. E. M., Crippa, J. A. S., and Osorio, F. L. (2014). Performance anxiety in Brazilian musicians: prevalence and association with psychopathology indicators. *J. Affect. Dis.* 152, 381–386. doi: 10.1016/j.jad.2013.09.041

Bartholomew, D. J., Steele, F., Moustaki, I., and Galbraith, J. I. (2002). *The Analysis and Interpretation of Multivariate Data for Social Scientists*. Boca Raton, FL: Chapman and Hall/CRC.

Boucher, H., and Ryan, C. A. (2011). Performance stress and the very young musician. *J. Res. Music Educ.* 58, 329–345. doi: 10.1177/0022429410386965

Bradburn, N. M., Sudman, S., Blair, E., and Stocking, C. (2004). "Question threat and response bias," in *Questionnaires*, Vol. 3, ed. M. Bulmer (London: SAGE Publications Ltd), 47–60.

Brugués, A. O. (2009). *Music Performance Anxiety—a Review of the Literature. Inaugural-Dissertation zur Erlagungdes Medizinischen Doktorgrades der Medizinischen Fakultätder Albert-Ludwigs-Universität*. Deutschland: Freiburg im Bresgau.

Burin, A. B., and Osorio, F. L. (2017). Music performance anxiety: a critical review of etiological aspects, perceived causes, coping strategies and treatment. *Arch. Clin. Psychiatry* 44, 127–133. doi: 10.1590/0101-60830000000136

Carey, E., Devine, A., Hill, F., and Szucs, D. (2017). Differentiating anxiety forms and their role in academic performance from primary to secondary school. *PLoS One* 12:e0174418. doi: 10.1371/journal.pone.0174418

Carver, C. S., and Scheier, M. F. (1986). "Functional and dysfunctional responses to anxiety: the interaction between expectancies and self-focused attention," in *Self-Related Cognitions in Anxiety and Motivation*, ed. R. Schwarzer Hillsdale, NJ: Erlbaum.111–141.

Cassady, J. C., and Finch, W. H. (2015). Using factor mixture modeling to identify dimensions of cognitive test anxiety. *Learn. Individ. Dif.* 41, 14–20. doi: 10.1016/j.lindif.2015.06.002

Chow, K., and Mercado, E. I. I. I. (2020). Performance anxiety and the plasticity of emotional responses. *Cogn. Emot.* 34, 1309–1325. doi: 10.1080/02699931.2020.1749568

Cooper, P., and McIntyre, D. (1993). Commonality in teachers' and pupils' perceptions of effective classroom learning. *Br. J. Educ. Psychol.* 63, 381–399. doi: 10.1111/j.2044-8279.1993.tb01066.x

Covington, M. V., and Omelich, C. L. (1981). As failures mount: affective and cognitive consequences of ability demotion in the classroom. *J. Educ. Psychol.* 73, 796–808. doi: 10.1037/0022-0663.73.6.796

Davis, H. A., DiStefano, C., and Schutz, P. A. (2008). Identifying patterns of appraising tests in first-year college students: implications for anxiety and emotion regulation during test taking. *J. Educ. Psychol.* 100, 942–960. doi: 10.1037/a0013096

Dempsey, E., and Comeau, G. (2019). Music performance anxiety and self-efficacy in young musicians: effects of gender and age. *Music Perform. Res.* 9, 60–79.

Dews, C. L. B., and Williams, M. S. (1989). Student musicians' personality styles, stresses, and coping patterns. *Psychol. Music* 17, 37–47. doi: 10.1177/0305735689171004

Everitt, B. S., Landuau, S., Leese, M. and Stahl, D. (2011). *Cluster Analysis*, 5th Edn. Newyork, NY: Wiley, 346.

Fehm, L., and Schmidt, K. (2006). Performance anxiety in gifted adolescent musicians. *Anxiety Dis.* 20, 98–109. doi: 10.1016/j.janxdis.2004.11.011

Fernholz, I., Mumm, J. L. M., Plag, J., Noeres, K., Rotter, G., Willich, S. N., et al. (2019). Performance anxiety in professional musicians: a systematic review on prevalence, risk factors and clinical treatment effects. *Psychol. Med.* 49, 2287–2306. doi: 10.1017/S0033291719001910

Fishbein, M., Middlestadt, S. E., Ottati, V., Strauss, S., and Ellis, A. (1988). Medical problems among ISCOM musicians: overview of a national survey. *Medi. Problems Performing Artists* 3, 1–8.

Gates, A. G., and Montalbo, P. J. (1987). The effect of low-dose beta blockade on performance anxiety in singers. *J. Voice* 1, 105–108. doi: 10.1016/s0892-1997(87)80032-2

Gogol, K., Brunner, M., Goetz, T., Martin, R., Ugen, S., Keller, U., et al. (2014). "My Questionnaire is Too Long!". the assessments of motivational-affective constructs with three-item and single-item measures. *Contemp. Educ. Psychol.* 39, 188–205. doi: 10.1016/j.cedpsych.2014.04.002

Gustafson, E., and Rawson, D. (1983). *Rambfeber Hos Strakmusiker (Stage fright in string players).* Unpublished Masters Thesis. Uppsala: Uppsala University.

Hair, J. F. Jr., Black, W. C., Babin, B. J., and Anderson, R. E. (2010). *Multivariate Data Analysis: A Global Perspective*. London: Pearson.

Hallam, S. (2006). *Music Psychology in Education*. London: Institute of Education.

Hamann, D. L. (1982). An assessment of anxiety in instrumental and vocal performances. *J. Res. Music Educ.* 30, 77–90. doi: 10.2307/3345040

Kanduri, C., Kuusi, T., Ahvenainen, M., Phlips, A. K., Lahdesmaki, H., and Jarvela, I. (2015). The effect of music performance on the transcriptome of professional musicians. *Sci. Rep.* 5:9506.

Kemp, A. E. (1996). *The Musical Temperament: Psychology and Personality of Musicians*. Oxford: Oxford University Press.

Kenny, D., and Osborne, M. S. (2006). Music performance anxiety: new insights from young musicians. *Adv. Cogn. Psychol.* 2, 103–112. doi: 10.2478/v10053-008-0049-5

Kenny, D. T. (2009). "The factor structure of the revised kenny music performance anxiety inventory," in *Paper presented at the International Symposium on Performance Science*, (Auckland).

Kenny, T., Driscoll, T., and Ackermann, B. (2014). Psychological well-being in professional orchestral musicians in Australia: a descriptive population study. *Psychol. Music* 42, 210–232. doi: 10.1177/0305735612463950

Losiak, W. (2005). Shapes of anxiety. analysis of anxiety profiles measured with S-R questionnaire. *Ansiedady Estres* 11, 157–162.

Mammarella, I. C., Donolato, E., Caviola, S., and Giofre, D. (2018). Anxiety profiles and protective factors: a latent profile analysis in children. *Pers. Individ. Dif.* 124, 201–208. doi: 10.1016/j.paid.2017.12.017

Marchant-Haycox, S. E., and Wilson, G. D. (1992). Personality and stress in performing artists. *Pers. Individ. Dif.* 13, 1061–1068. doi: 10.1016/0191-8869(92)90021-G

Medeiros Barbar, A. E., de Souza Crippa, J. A., and de Lima Osorio, F. (2014). Performance anxiety in Brazilian musicians: prevalence and association with psychopathology indicators. *J. Affect. Dis.* 15, 381–386. doi: 10.1016/j.jad.2013.09.041

Nagel, J. J., and Himle, D. P. (1989). Cognitive-behavioural treatment of music performance anxiety. *Psychol. Music* 17, 12–21. doi: 10.1177/0305735689171002

Nielsen, K., Studer, R. K., Hilderbrandt, H., Nater, U. M., Wild, P., Danuser, B., et al. (2018). The relationship between music performance anxiety, subjective performance quality and post-event rumination among music students. *Psychol. Music* 46, 136–152. doi: 10.1177/0305735617706539

Nusseck, M., Zander, M., and Spahn, C. (2015). Music performance anxiety in young musicians: comparison of playing classical or popular music. *Med. Probl. Perform. Art.* 30, 30–37. doi: 10.21091/mppa.2015.1005

Osborne, M. S., and Kenny, D. (2008). The role of sensitizing experiences in music performance anxiety in adolescent musicians. *Psychol. Music* 36, 447–462. doi: 10.1177/0305735607086051

Osborne, M. S., and Kenny, D. T. (2005). Development and validation of a music performance anxiety inventory for gifted adolescent musicians. *J. Anxiety Dis.* 19, 725–751. doi: 10.1016/j.janxdis.2004.09.002

Osborne, M. S., Kenny, D. T., and Holsomback, R. (2005). Assessment of music performance anxiety in late childhood: a validation study of the music performance anxiety inventory for adolescents (MPAI-A). *Int. J. Stress Manag.* 12, 312–333. doi: 10.1037/1072-5245.12.4.312

Papageorgi, I. (2007a). *Understanding Performance Anxiety in the Adolescent Musician.* Unpublished Doctoral Thesis. London: UCL Institute of Education.

Papageorgi, I. (2007b). "The influence of the wider context of learning, gender, age and individual differences on adolescent musicians' performance anxiety," in *Proceedings of the International Symposium on Performance Science*, eds A. Williamon and D. Coimbra (Utrecht: The Netherlands: European Association of Conservatoires (AEC)), 219–224.

Papageorgi, I. (2008). *Investigating musical performance: Performance Anxiety Across Musical Genres. Economic and Social Research Council, Teaching and Learning Research Programme (ESRC TLRP): Teaching and Learning Research Briefing (No. 57), November 2008*. Swindon: Economic and Social Research Council.

Papageorgi, I. (2020). Prevalence and predictors of musical performance anxiety in adolescent learners: contributions of individual, task-related and environmental factors. *Musicae Scientiae* doi: 10.1177/1029864920923128

Papageorgi, I., Creech, A., and Welch, G. (2013). Perceived performance anxiety in advanced musicians specializing in different musical genres. *Psychol. Music* 41, 18–41. doi: 10.1177/0305735611408995

Papageorgi, I., Hallam, S., and Welch, G. F. (2007). A conceptual framework for understanding musical performance anxiety. *Res. Stud. Music Educ.* 28, 83–107. doi: 10.1177/1321103x070280010207

Papageorgi, I., and Kopiez, R. (2018). "Psychological and physiological aspects of learning to perform," in *The Oxford Handbook of Music Education – Vocal, Instrumental, and Ensemble Learning and Teaching: An Oxford Handbook of Music Education*, Vol. 3, eds G. McPherson and G. Welch (Oxford: Oxford University Press), 184–208.

Papageorgi, I., and Welch, G. (2014b). "How do musicians develop their learning about performance?," in *Advanced Musical Performance: Investigations in Higher Education Learning SEMPRE Studies in the Psychology of Music*, eds I. Papageorgi and G. Welch (Aldershot: Ashgate Press), 170–186.

Papageorgi, I., and Welch, G. F. (eds) (2014a). *Advanced Musical Performance: Investigations in Higher Education Learning. SEMPRE Studies in the Psychology of Music*. Aldershot: Ashgate Press.

Patston, T., and Osborne, M. S. (2016). The developmental features of music performance anxiety and perfectionism in school age music students. *Perform. Enhancement Health* 4, 29–42.

Perdomo-Guevara, E. (2014). Is Music performance anxiety just an individual problem? exploring the impact of musical environments on performers'. *Approaches Perform. Emot.* 24, 66–74. doi: 10.1037/pmu0000028

Putwain, D. W., and Daly, A. L. (2013). Do clusters of test anxiety and academic buoyancy differentially predict academic performance? *Learn. Individ. Dif.* 27, 157–162. doi: 10.1016/j.lindif.2013.07.010

Rae, G., and McCambridge, K. (2004). Correlates of performance anxiety in practical music exams. *Psychol. Music* 32, 432–439. doi: 10.1177/0305735604046100

Romesburg, H. C. (1984). *Cluster Analysis for Researchers*. Belmont, CS: Lifetime Learning Publications.

Ryan, C. (1998). Exploring musical performance anxiety in children. *Med. Probl. Perform. Art.* 13, 83–88.

Ryan, C. (2004). Gender differences in children's experience of musical performance anxiety. *Psychol. Music* 32, 89–103. doi: 10.1177/0305735604039284

Ryan, C., and Andrews, N. (2009). An investigation into the choral singer's experience of music performance anxiety. *J. Res. Music Educ.* 57, 108–126. doi: 10.1177/0022429409336132

Sarbescu, P., and Dorgo, M. (2014). Frightened by the stage or by the public? Exploring the multidimensionality of music performance anxiety. *Psychol. Music* 42, 568–579. doi: 10.1177/0305735613483669

Schulz, W. (1981). "Analysis of a symphony orchestra," in *Stress and Music: Medical, Psychological, Sociological, and Legal STRAIN Factors in a Symphony Orchestra Musicians Profession*, ed. M. Piperek Vienna: Wilhelm Braumuller, 35–56.

Stenlund, T., Lyren, P.-E., and Eklof, H. (2018). The successful test taker: exploring test-taking behavior profiles through cluster analysis. *Eur. J. Psychol. Educ.* 33, 403–417. doi: 10.1007/s10212-017-0332-2

Steptoe, A. (2001). "Negative emotions in music making: the problem of performance anxiety," in *Music and Emotion: Theory and Research*, eds P. N. Juslin and J. A. Sloboda (Oxford: Oxford University Press), 291–307.

Thomas, J. P., and Nettelbeck, T. (2014). Performance anxiety in adolescent musicians. *Psychol. Music* 42, 624–634. doi: 10.1177/0305735613485151

Vervainioti, A., and Alexopoulos, E. C. (2015). Job related stressors of classical instrumental musicians: a systematic qualitative review. *Med. Probl. Perform. Art.* 30, 197–202. doi: 10.21091/mppa.2015.4037

Wang, Z., Wonjung, O., Malanchini, M., and Borriello, G. A. (2020). The developmental trajectories of mathematics anxiety: cognitive, personality, and environmental correlates. *Contemp. Educ. Psychol.* 61:101876. doi: 10.1016/j.cedpsych.2020.101876

Weiner, B. (1985). An attributional theory of achievement motivation. *Psychol. Rev*. 92, 548–573.

Wesner, R. B., Noyes, R., and Davis, T. L. (1990). The occurrence of performance anxiety among musicians. *J. Affect. Dis.* 18, 177–185. doi: 10.1016/0165-0327(90)90034-6

Wiedemann, A., Vogel, D., Voss, C., and Hoyer, J. (2019). Music performance anxiety and its anxiety correlates. *PsyArXiv* doi: 10.17605/OSF.IO/XPJYB

Wigfield, A., and Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemp. Educ. Psychol.* 25, 68–81. doi: 10.1006/ceps.1999.1015

Williamon, A., and Thompson, S. (2006). Awareness and incidence of health problems among conservatoire students. *Psychol. Music* 34, 411–430. doi: 10.1177/0305735606067150

Wills, G., and Cooper, C. L. (1988). *Pressure Sensitive: Popular Musicians Under Stress*. London: Sage.

Wilson, G. D. (2002). *Psychology for Performing Artists*, 2nd Edn. London: Whurr Publishers.

Yerkes, R. M., and Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *J. Comp. Neuro. Psychol*. 18, 459–482.

Zarza-Alzugaray, F. J., Orejudo, S., Casanova, O., and Mazas, B. (2016). Kenny music performance anxiety inventory: confirmatory factor analysis of the spanish version. *Psychol. Music* 44, 340–352. doi: 10.1177/0305735614567932

**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.

frontiers in Psychology

REVIEW
published: 10 June 2021
doi: 10.3389/fpsyg.2021.697726

# The State of Music Therapy Studies in the Past 20 Years: A Bibliometric Analysis

*Kailimi Li*[1†], *Linman Weng*[2†] and *Xueqiang Wang*[2,3*]

[1] School of Kinesiology, Shanghai University of Sport, Shanghai, China, [2] Department of Sport Rehabilitation, Shanghai University of Sport, Shanghai, China, [3] Department of Sport Rehabilitation Medicine, Shanghai Shangti Orthopedic Hospital, Shanghai, China

**Purpose:** Music therapy is increasingly being used to address physical, emotional, cognitive, and social needs of individuals. However, publications on the global trends of music therapy using bibliometric analysis are rare. The study aimed to use the CiteSpace software to provide global scientific research about music therapy from 2000 to 2019.

**Methods:** Publications between 2000 and 2019 related to music therapy were searched from the Web of Science (WoS) database. The CiteSpace V software was used to perform co-citation analysis about authors, and visualize the collaborations between countries or regions into a network map. Linear regression was applied to analyze the overall publication trend.

**Results:** In this study, a total of 1,004 studies met the inclusion criteria. These works were written by 2,531 authors from 1,219 institutions. The results revealed that music therapy publications had significant growth over time because the linear regression results revealed that the percentages had a notable increase from 2000 to 2019 ($t$ = 14.621, $P < 0.001$). The United States had the largest number of published studies (362 publications), along with the following outputs: citations on WoS (5,752), citations per study (15.89), and a high H-index value (37). The three keywords "efficacy," "health," and "older adults," emphasized the research trends in terms of the strongest citation bursts.

**Conclusions:** The overall trend in music therapy is positive. The findings provide useful information for music therapy researchers to identify new directions related to collaborators, popular issues, and research frontiers. The development prospects of music therapy could be expected, and future scholars could pay attention to the clinical significance of music therapy to improve the quality of life of people.

**Keywords: music therapy, aged, bibliometrics, health, web of science**

OPEN ACCESS

***Edited by:***
*Guadalupe López-Íñiguez, Sibelius Academy, Finland*

***Reviewed by:***
*Andy Wai Kan Yeung, The University of Hong Kong, Hong Kong, China*
*Carlos Mulet Forteza, University of the Balearic Islands, Spain*

***Correspondence:**
*Xueqiang Wang*
*wangxueqiang@sus.edu.cn*

*† These authors have contributed equally to this work and share first authorship*

***Specialty section:***
*This article was submitted to Performance Science, a section of the journal Frontiers in Psychology*

***Received:*** *20 April 2021*
***Accepted:*** *12 May 2021*
***Published:*** *10 June 2021*

***Citation:***
*Li K, Weng L and Wang X (2021) The State of Music Therapy Studies in the Past 20 Years: A Bibliometric Analysis. Front. Psychol. 12:697726. doi: 10.3389/fpsyg.2021.697726*

## INTRODUCTION

Music therapy is defined as the evidence-based use of music interventions to achieve the goals of clients with the help of music therapists who have completed a music therapy program (Association, 2018). In the United States, music therapists must complete 1,200 h of clinical training and pass the certification exam by the Certification Board for Music Therapists (Devlin et al., 2019). Music therapists use evidence-based music interventions to address the mental, physical, or emotional needs of an individual (Gooding and Langston, 2019). Also, music therapy is used as a solo standard treatment, as well as co-treatment with other disciplines, to address the needs in cognition, language, social integration, and psychological health and family support of an individual (Bronson et al., 2018). Additionally, music therapy has been used to improve various diseases in different research areas, such as rehabilitation, public health, clinical care, and psychology (Devlin et al., 2019). With neurorehabilitation, music therapy has been applied to increase motor activities in people with Parkinson's disease and other movement disorders (Bernatzky et al., 2004; Devlin et al., 2019). However, limited reviews about music therapy have utilized universal data and conducted massive retrospective studies using bibliometric techniques. Thus, this study demonstrates music therapy with a broad view and an in-depth analysis of the knowledge structure using bibliometric analysis of articles and publications.

Bibliometrics turns the major quantitative analytical tool that is used in conducting in-depth analyses of publications (Durieux and Gevenois, 2010; Gonzalez-Serrano et al., 2020). There are three types of bibliometric indices: (a) the quantity index is used to determine the number of relevant publications, (b) the quality index is employed to explore the characteristics of a scientific topic in terms of citations, and (c) the structural index is used to show the relationships among publications (Durieux and Gevenois, 2010; Gonzalez-Serrano et al., 2020). In this study, the three types of bibliometric indices will be applied to conduct an in-depth analysis of publications in this frontier.

While research about music therapy is extensively available worldwide, relatively limited studies use bibliometric methods to analyze the global research about this topic. The aim of this study is to use the CiteSpace software to perform a bibliometric analysis of music therapy research from 2000 to 2019. CiteSpace V is visual analytic software, which is often utilized to perform bibliometric analyses (Falagas et al., 2008; Ellegaard and Wallin, 2015). It is also a tool applied to detect trends in global scientific research. In this study, the global music therapy research includes publication outputs, distribution and collaborations between authors/countries or regions/institutions, intense issues, hot articles, common keywords, productive authors, and connections among such authors in the field. This study also provides helpful information for researchers in their endeavor to identify gaps in the existing literature.

## MATERIALS AND METHODS

### Search Strategy

The data used in this study were obtained from WoS, the most trusted international citation database in the world. This database, which is run by Thomson & Reuters Corporation (Falagas et al., 2008; Durieux and Gevenois, 2010; Chen C. et al., 2012; Ellegaard and Wallin, 2015; Miao et al., 2017; Gonzalez-Serrano et al., 2020), provides high-quality journals and detailed information about publications worldwide. In this study, publications were searched from the WoS Core Collection database, which included eight indices (Gonzalez-Serrano et al., 2020). This study searched the publications from two indices, namely, the Science Citation Index Expanded and the Social Sciences Citation Index. As the most updated publications about music therapy were published in the 21st century, publications from 2000 to 2019 were chosen for this study. We performed data acquisition on July 26, 2020 using the following search terms: title = ("music therapy") and time span = 2000–2019.

### Inclusion Criteria

**Figure 1** presents the inclusion criteria. The title field was music therapy (TI = music therapy), and only reviews and articles were chosen as document types in the advanced search. Other document types, such as letters, editorial materials, and book reviews, were excluded. Furthermore, there were no species limitations set. This advanced search process returned 718 articles. In the end, a total of 1,004 publications were obtained and were analyzed to obtain comprehensive perspectives on the data.

### Data Extraction

Author Lin-Man Weng extracted the publications and applied the EndNote software and Microsoft Excel 2016 to conduct analysis on the downloaded publications from the WoS database. Additionally, we extracted and recorded some information of the publications, such as citation frequency, institutions, authors' countries or regions, and journals as bibliometric indicators. The H-index is utilized as a measurement of the citation frequency of the studies for academic journals or researchers (Wang et al., 2019).

### Analysis Methods

The objective of bibliometrics can be described as the performance of studies that contributes to advancing the knowledge domain through inferences and explanations of relevant analyses (Castanha and Grácio, 2014; Merigó et al., 2019; Mulet-Forteza et al., 2021). CiteSpace V is a bibliometric software that generates information for better visualization of data. In this study, the CiteSpace V software was used to visualize six science maps about music therapy research from 2000 to 2019: the network of author co-citation, collaboration network among countries and regions, relationship of institutions interested in the field, network map of co-citation journals, network map of

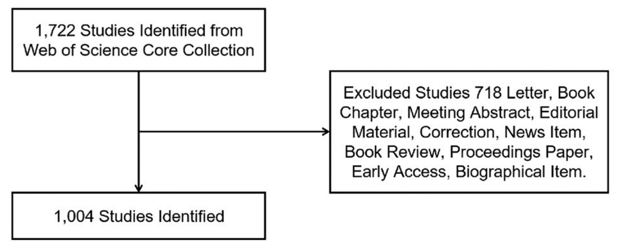



**FIGURE 1** | Flow chart of music therapy articles and reviews inclusion.

**Abbreviations:** WoS, Web of Science; ESI, essential science indicators; IF, impact factor; IMT, improvisational music therapy; ASD, autism spectrum disorder.

co-cited references, and the map (timeline view) of references with co-citation on top music therapy research. As noted, a co-citation is produced when two publications receive a citation from the same third study (Small, 1973; Merigó et al., 2019).

In addition, a science map typically features a set of points and lines to present collaborations among publications (Chen, 2006). A point is used to represent a country or region, author, institution, journal, reference, or keyword, whereas a line represents connections among them (Zheng and Wang, 2019), with stronger connections indicated by wider lines. Furthermore, the science map includes nodes, which represent the citation frequencies of certain themes. A burst node in the form of a red circle in the center indicates the number of co-occurrence or citation that increases over time. A purple node represents centrality, which indicates the significant knowledge presented by the data (Chen, 2006; Chen H. et al., 2012; Zheng and Wang, 2019). The science map represents the keywords and references with citation bursts. Occurrence bursts represent the frequency of a theme (Chen, 2006), whereas citation bursts represent the frequency of the reference. The citation bursts of keywords and references explore the trends and indicate whether the relevant authors have gained considerable attention in the field (Chen, 2006). Through this kind of map, scholars can better understand emerging trends and grasp the hot topics by burst detection analysis (Liang et al., 2017; Miao et al., 2017).

# RESULTS

## Publication Outputs and Time Trends

A total of 1,004 articles and reviews related to music therapy research met the criteria. The details of annual publications are presented in **Figure 2**. As can be seen, there were < 30

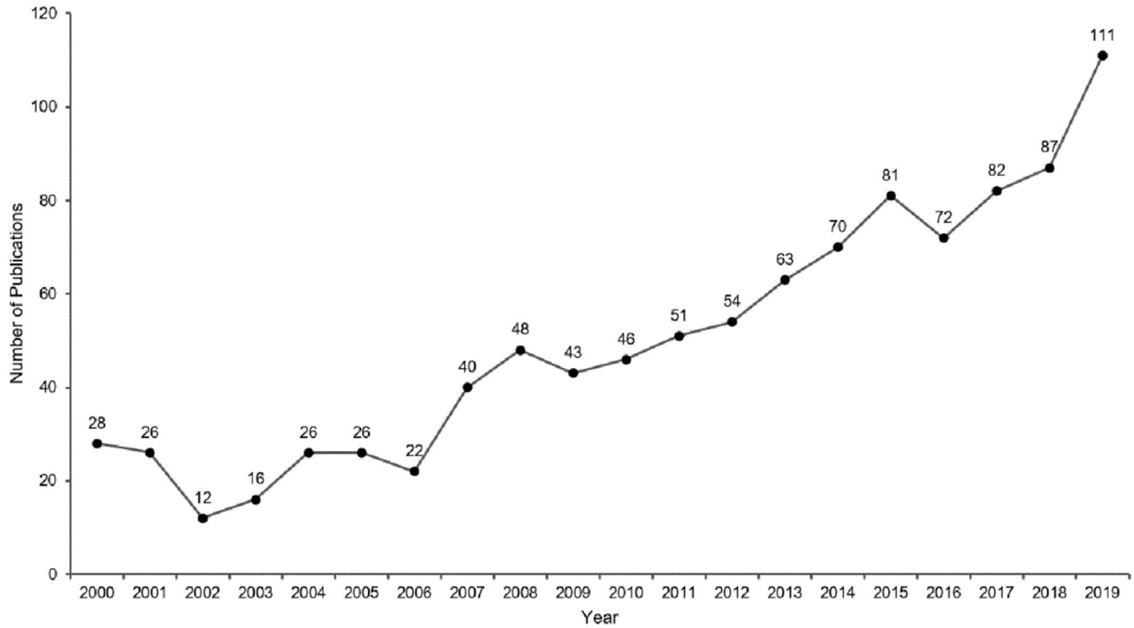



**FIGURE 2 |** Annual publication outputs of music therapy from 2000 to 2019.

**TABLE 1 |** Top 10 countries or regions of origin of study in the music therapy research field.

| Rank | Country or regions | Publications | Percentage (%) | Citations WoS | Citations per paper | H-index |
|---|---|---|---|---|---|---|
| 1 | USA | 362 | 36.056 | 5,752 | 15.89 | 37 |
| 2 | Germany | 96 | 9.562 | 1,343 | 13.99 | 20 |
| 3 | England | 95 | 9.462 | 1,841 | 19.38 | 25 |
| 4 | Australia | 88 | 8.765 | 1,492 | 16.95 | 21 |
| 5 | Norway | 72 | 7.171 | 1,957 | 27.18 | 25 |
| 6 | China | 53 | 5.279 | 767 | 14.47 | 17 |
| 7 | Denmark | 45 | 4.482 | 1,218 | 27.07 | 17 |
| 8 | Italy | 39 | 3.884 | 987 | 25.31 | 14 |
| 9 | Canada | 30 | 2.988 | 401 | 13.37 | 10 |
| 10 | Israel | 29 | 2.888 | 346 | 11.93 | 9 |

annual publications between 2000 and 2006. The number of publications increased steadily between 2007 and 2015. It was 2015, which marked the first time over 80 articles or reviews were published. The significant increase in publications between 2018 and 2019 indicated that a growing number of researchers became interested in this field. Linear regression can be used to analyze the trends in publication outputs. In this study, the linear regression results revealed that the percentages had a notable increase from 2000 to 2019 ($t = 14.621$, $P < 0.001$). Moreover, the $P < 0.05$, indicating statistical significance. Overall, the publication outputs increased from 2000 to 2019.

## Distribution by Country or Region and Institution

The 1,004 articles and reviews collected were published in 49 countries and regions. **Table 1** presents the top 10 countries or regions. **Figure 3** shows an intuitive comparison of the citations on WoS, citations per study, Hirsch index (H-index), and major essential science indicator (ESI) studies of the top five countries or regions. The H-index is a kind of index that is applied in measuring the wide impact of the scientific achievements of authors. The United States had the largest number of published studies (362 publications), along with the following outputs: citations on WoS (5,752), citations per study (15.89), and a high H-index value (37). Norway has the largest number of citations per study (27.18 citations). **Figure 4** presents the collaboration networks among countries or regions. The collaboration network map contained 32 nodes and 38 links. The largest node can be found in the United States, which meant that the United States had the largest number of publications in the field. Meanwhile, the deepest purple circle was located in Austria, which meant that Austria is the country with the most number of collaborations with other countries or regions in this research field. A total of 1,219 institutions contributed various music therapy-related publications. **Figure 5** presents the

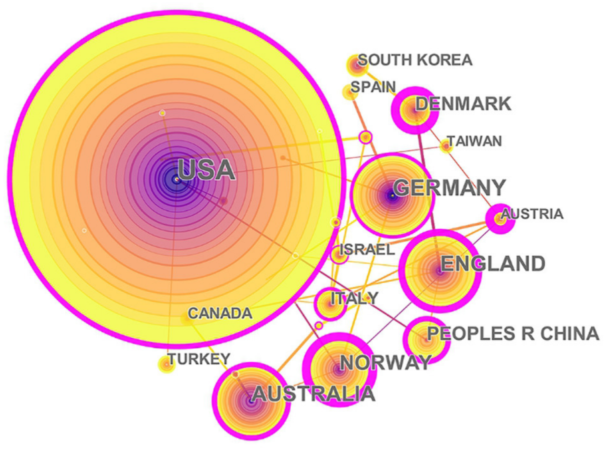



**FIGURE 4 |** The collaborations of countries or regions interested in the field. In this map, the node represents a country, and the link represents the cooperation relationship between two countries. A larger node represents more publications in the country. A thicker purple circle represents greater influence in this field.

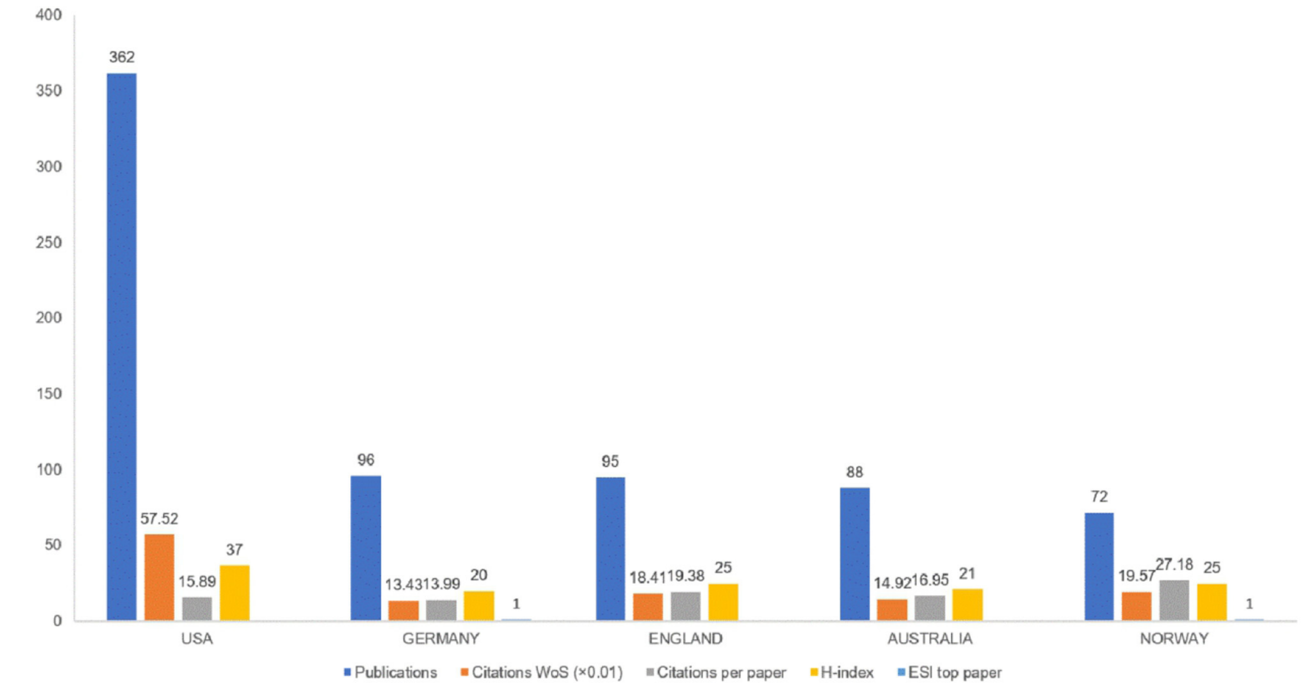



**FIGURE 3 |** Publications, citations on WoS (× 0.01), citations per study, H-index, and ESL top study among top five countries or regions.

collaborations among institutions. As can be seen, the University of Melbourne is the most productive institution in terms of the number of publications (45), followed by the University of Minnesota (43), and the University of Bergen (39). The top 10 institutions featured in **Table 2** contributed 28.884% of the total articles and reviews published. Among these, Aalborg University had the largest centrality (0.13). The top 10 productive institutions with details are shown in **Table 2**.

## Distribution by Journals

**Table 3** presents the top 10 journals that published articles or reviews in the music therapy field. The publications are mostly published in these journal fields, such as Therapy, Medical, Psychology, Neuroscience, Health and Clinical Care. The impact factors (IF) of these journals ranged between 0.913 and 7.89 (average IF: 2.568). Four journals had an impact factor >2, of which Cochrane Database of Systematic Reviews had the highest IF, 2019 = 7.89. In addition, the Journal of Music Therapy (IF: 2019 = 1.206) published 177 articles or reviews (17.629%) about music therapy in the past two decades, followed by the Nordic Journal of Music Therapy (121 publications, 12.052%, IF: 2019 = 0.913), and Arts in Psychotherapy (104 publications, 10.359%, IF: 2019 = 1.322). Furthermore, the map of the co-citation journal contained 393 nodes and 759 links (**Figure 6**). The high co-citation count identifies the journals with the greatest academic influence and key positions in the field. The Journal of Music Therapy had the maximum co-citation counts (658), followed by Cochrane Database of Systematic Reviews (281), and Arts in Psychotherapy (279). Therefore, according to the analysis of the publications and co-citation counts, the Journal of Music Therapy and Arts in Psychotherapy occupied key positions in this research field.

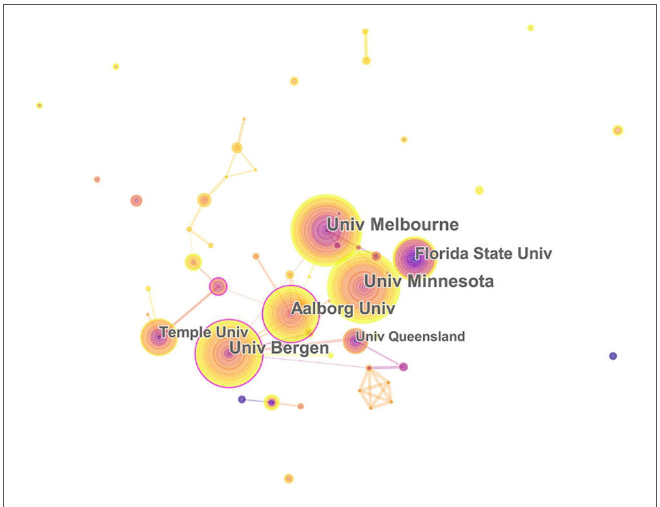



**FIGURE 5 |** The relationship of institutions interested in the field. University of Melbourne, Florida State University, University of Minnesota, Aalborg University, Temple University, University of Queensland, and University of Bergen. In this map, the node represents an institution, and the link represents the cooperation relationship between two institutions. A larger node represents more publications in the institution. A thicker purple circle represents greater influence in this field.

## Distribution by Authors

A total of 2,531 authors contributed to the research outputs related to music therapy. Author Silverman MJ published most of the studies (46) in terms of number of publications, followed by Gold C (41), Magee WL (19), O'Callaghan C (15), and Raglio A (15). According to co-citation counts, Bruscia KE (171 citations) was the most co-cited author, followed by Gold C (147 citations), Wigram T (121 citations), and Bradt J (117 citations), as presented in **Table 4**. In **Figure 7**, these nodes highlight the co-citation networks of the authors. The large-sized node represented author Bruscia KE, indicating that this author owned the most co-citations. Furthermore, the linear regression results revealed a remarkable increase in the percentages of multiple articles of authors ($t$ = 13.089, $P$ < 0.001). These also indicated that cooperation among authors had increased remarkably, which can be considered an important development in music therapy research.

## Analysis of Keywords

The results of keywords analysis indicated research hotspots and help scholars identify future research topics. **Table 5** highlights 20 keywords with the most frequencies, such as "music therapy," "anxiety," "intervention," "children," and

**TABLE 2 |** Top 10 institutions that contributed to publications in the music therapy field.

| Rank | Institution | Publications | Percentage (%) | Centrality |
|---|---|---|---|---|
| 1 | University of Melbourne | 45 | 4.482 | 0.08 |
| 2 | University of Minnesota | 43 | 4.283 | 0.00 |
| 3 | University of Bergen | 39 | 3.884 | 0.12 |
| 4 | Florida State University | 33 | 3.287 | 0.00 |
| 5 | Aalborg University | 32 | 3.187 | 0.13 |
| 6 | Temple University | 27 | 2.689 | 0.04 |
| 7 | University of Kansas | 20 | 1.992 | 0.00 |
| 8 | University of Queensland | 20 | 1.992 | 0.00 |
| 9 | Anglia Ruskin University | 16 | 1.594 | 0.08 |
| 10 | Bar Ilan University | 15 | 1.494 | 0.00 |

**TABLE 3 |** Top 10 journals that published articles in the music therapy field.

| Rank | Journal | Publications | Percentage (%) | IF (2019) |
|---|---|---|---|---|
| 1 | Journal of Music Therapy | 177 | 17.629 | 1.206 |
| 2 | Nordic Journal of Music Therapy | 121 | 12.052 | 0.913 |
| 3 | Arts in Psychotherapy | 104 | 10.359 | 1.322 |
| 4 | Analysis of the New York Academy of Sciences | 18 | 1.793 | 4.728 |
| 5 | Complementary Therapies in Medicine | 12 | 1.195 | 2.063 |
| 6 | Journal of Clinical Nursing | 10 | 0.996 | 1.972 |
| 7 | Journal of Palliative care | 10 | 0.996 | 1.200 |
| 8 | Cochrane Database of Systematic Reviews | 9 | 0.896 | 7.890 |
| 9 | Frontiers in Human Neuroscience | 9 | 0.896 | 2.673 |
| 10 | Psychology of Music | 9 | 0.896 | 1.712 |

*IF, impact factor.*

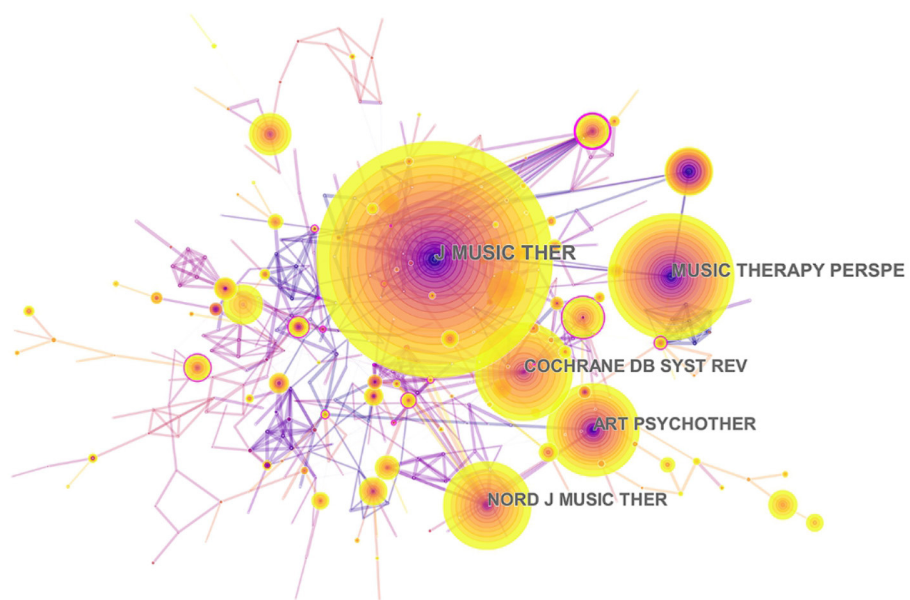



**FIGURE 6 |** Network map of co-citation journals engaged in music therapy from 2000 to 2019. Journal of Music Therapy, Arts in Psychotherapy, Nordic Journal of Music Therapy, Music Therapy Perspectives, Cochrane Database of Systematic Reviews. In this map, the node represents a journal, and the link represents the co-citation frequency between two journals. A larger node represents more publications in the journal. A thicker purple circle represents greater influence in this field.

"depression." The keyword "autism" has the highest centrality (0.42). **Figure 8** shows the top 17 keywords with the strongest citation bursts. By the end of 2019, keyword bursts were led by "hospice," which had the strongest burst (3.5071), followed by "efficacy" (3.1161), "health" (6.2109), and "older adult" (4.476).

## Analysis of Co-cited References

The analysis of co-cited references is a significant indicator in the bibliometric method (Chen, 2006). The top five co-cited references and their main findings are listed in **Table 6**. These are regarded as fundamental studies for the music therapy knowledge base. In terms of co-citation counts, "individual music therapy for depression: randomized controlled trial" was the key reference because it had the most co-citation counts. This study concludes that music therapy mixed with standard care is an effective way to treat working-age people with depression. The authors also explained that music therapy is a valuable enhancement to established treatment practices (Erkkilä et al., 2011). Meanwhile, the strongest citation burst of reference is regarded as the main knowledge of the trend (Fitzpatrick, 2005). **Figure 9** highlights the top 71 strongest citation bursts of references from 2000 to 2019. As can be seen, by the end of 2019, the reference burst was led by author Stige B, and the strongest burst was 4.3462.

**Figure 10A** presents the co-cited reference map containing 577 nodes and 1,331 links. The figure explains the empirical relevance of a considerable number of articles and reviews.

**TABLE 4 |** Top five authors of publications and top five authors of co-citation counts.

| Rank | Author | Publications | Percentage (%) | Centrality | Cited author | Co-citation counts |
|---|---|---|---|---|---|---|
| 1 | Silverman MJ | 46 | 4.582 | 0.00 | Bruscia KE | 171 |
| 2 | Gold C | 41 | 4.084 | 0.06 | Gold C | 147 |
| 3 | Magee WL | 19 | 1.892 | 0.01 | Wigram T | 121 |
| 4 | O'Callaghan C | 15 | 1.494 | 0.01 | Bradt J | 117 |
| 5 | Raglio A | 15 | 1.494 | 0.00 | Thaut MH | 116 |

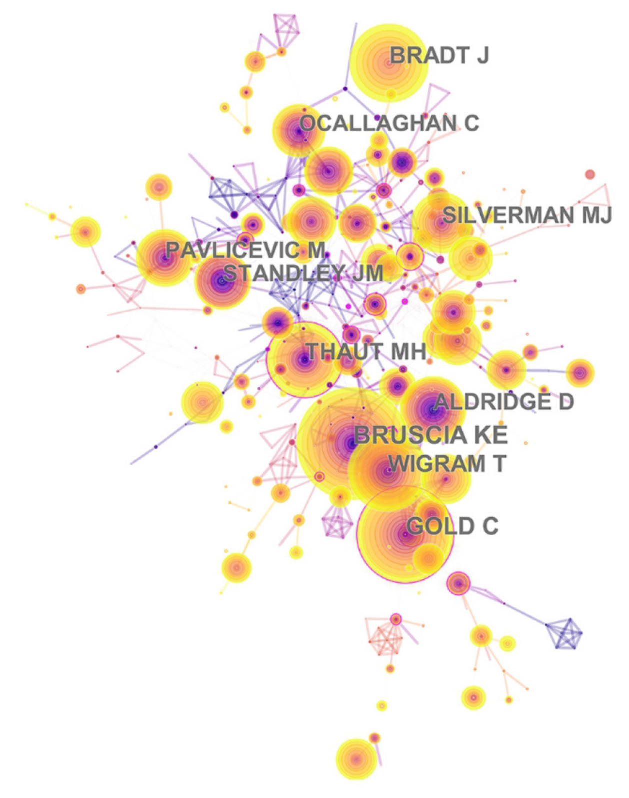



**FIGURE 7 |** The network of author co-citaion. In this map, the node represents an author, and the link represents the co-citation frequency between two authors. A larger node represents more publications of the author. A thicker purple circle represents greater influence in this field.

**Figure 10B** presents the co-citation map (timeline view) of reference from publications on top music therapy research. The timeline view of clusters shows the research progress of music therapy in a particular period of time and the thematic concentration of each cluster. "Psychosis" was labeled as the largest cluster (#0), followed by "improvisational music therapy" (#1) and "paranesthesia anxiety" (#2). These clusters have also remained hot topics in recent years. Furthermore, the result of the modularity Q score was 0.8258. That this value exceeded 0.5 indicated that the definitions of the subdomain and characters of clusters were distinct. In addition, the mean silhouette was 0.5802, which also exceeded 0.5. The high homogeneity of individual clusters indicated high concentration in different research areas.

## DISCUSSION

### Global Trends in Music Therapy Research

This study conducted a bibliometric analysis of music therapy research from the past two decades. The results, which reveal that music therapy studies have been conducted throughout the world, among others, can provide further research suggestions to scholars. In terms of the general analysis of the publications, the features of published articles and reviews, prolific countries or regions, and productive institutions are summarized below.

I. The distribution of publication year has been increasing in the past two decades. The annual publication outputs of music therapy from 2000 to 2019 were divided into three stages: beginning, second, and third. In the beginning stage, there were $< 30$ annual publications from 2000 to 2006. The second stage was between 2007 and 2014. The number of publications increased steadily. It was 2007, which marked the first time 40 articles or reviews were published. The third stage was between 2015 and 2019. The year 2015 was the key turning point because it was the first time 80 articles or reviews were published. The number of publications showed a downward trend in 2016 (72), but it was still higher than the average number of the previous years. Overall, music therapy-related research has received increasing attention among scholars from 2000 to 2020.

II. The articles and reviews covered about 49 countries or regions, and the prolific countries or regions were mainly located in the North American and European continents. According to citations on WoS, citations per study, and the H-index, music therapy publications from developed countries, such as United States and Norway, have greater influence than those from other countries. In addition, China, as a model of a developing country, had published 53 studies and ranked top six among productive countries.

III. In terms of the collaboration map of institutions, the most productive universities engaged in music therapy were located in the United States, namely, University of Minnesota (43 publications), Florida State University (33 publications), Temple University (27 publications), and University of Kansas (20 publications). It indicated that institutions in the US have significant impacts in this area.

IV. According to author co-citation counts, scholars can focus on the publications of such authors as Bruscia KE, Gold C, and Wigram T. These three authors come from the United States,

**TABLE 5 |** Top 20 keywords with the most frequency and centrality in music therapy study.

| Rank | Keyword | Frequency | Keyword | Centrality |
|---|---|---|---|---|
| 1 | Music therapy | 486 | Autism | 0.42 |
| 2 | Anxiety | 149 | People | 0.34 |
| 3 | Intervention | 116 | Brain | 0.32 |
| 4 | Children | 94 | Schizophrenia | 0.23 |
| 5 | Depression | 90 | Quality of life | 0.21 |
| 6 | Pain | 76 | Perception | 0.19 |
| 7 | Dementia | 71 | Plasticity | 0.17 |
| 8 | Music | 62 | Parent | 0.15 |
| 9 | Randomized controlled trial | 57 | Adolescent | 0.14 |
| 10 | Quality of life | 50 | Behavior | 0.12 |
| 11 | People | 48 | Mental health | 0.12 |
| 12 | Relaxation | 48 | Response | 0.12 |
| 13 | Recovery | 45 | Recovery | 0.11 |
| 14 | Stress | 45 | Stress | 0.11 |
| 15 | Care | 45 | Care | 0.10 |
| 16 | Cancer | 45 | Preterm infant | 0.10 |
| 17 | Behavior | 42 | Dementia | 0.09 |
| 18 | Symptom | 40 | Reliability | 0.09 |
| 19 | Rehabilitation | 39 | Mother | 0.09 |
| 20 | Adolescent | 38 | Self esteem | 0.09 |

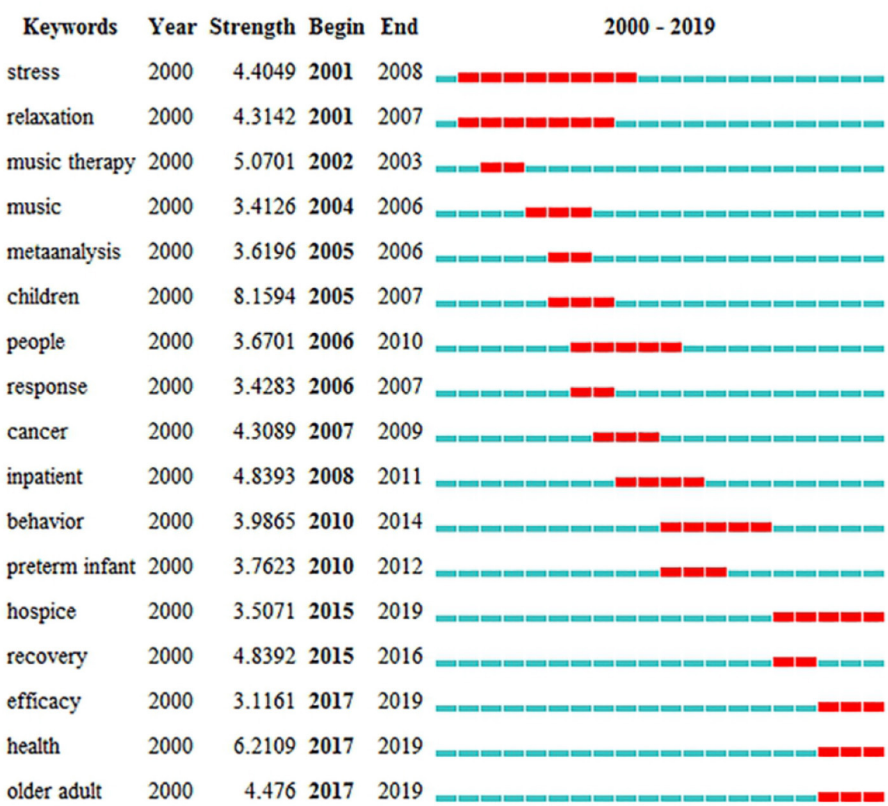



**FIGURE 8 |** The strongest citation bursts of the top 17 keywords. The red measures indicate frequent citation of keywords, and the green measures indicate infrequent citation of keywords.

**TABLE 6 |** Top five co-cited references with co-citation counts in the study of music therapy from 2000 to 2019.

| Rank | Cited reference | Co-citation counts | Publication year | Main findings |
|---|---|---|---|---|
| 1 | Individual music therapy for depression: randomized controlled trial | 43 | 2011 | Music therapy with its specific qualities is a valuable enhancement to working-age people with depression. |
| 2 | Dose-response relationship in music therapy for people with serious mental disorders: systematic review and meta-analysis | 39 | 2009 | Music therapy is an effective treatment which helps people with psychotic and non-psychotic mental disorders. |
| 3 | Music therapy for people with schizophrenia and schizophrenia-like disorders | 32 | 2011 | Music therapy can help people improve their emotional and relational competencies. |
| 4 | Music therapy for depression | 29 | 2008 | Music therapy is accepted by people with depression and is associated with improvements in mood. |
| 5 | Resource-oriented music therapy in mental health care | 29 | 2010 | An introduction to the resource-oriented approach to music therapyin mental health care. |

Norway, and Denmark, and it also reflected that these three countries are leading the research trend. Author Bruscia KE has the largest co-citation counts and is based at Temple University. He published many music therapy studies about assessment and clinical evaluation in music therapy, music therapy theories, and therapist experiences. These publications laid a foundation and facilitate the development of music therapy. In addition, in **Figure 11**, the multi-authored articles between 2000 and 2003 comprised 47.56% of the sample, whereas the publications of multi-authored articles increased significantly from 2016 to 2019 (85.51%). These indicated that cooperation is an effective factor in improving the quality of publications.

## Research Focus on the Research Frontier and Hot Topics

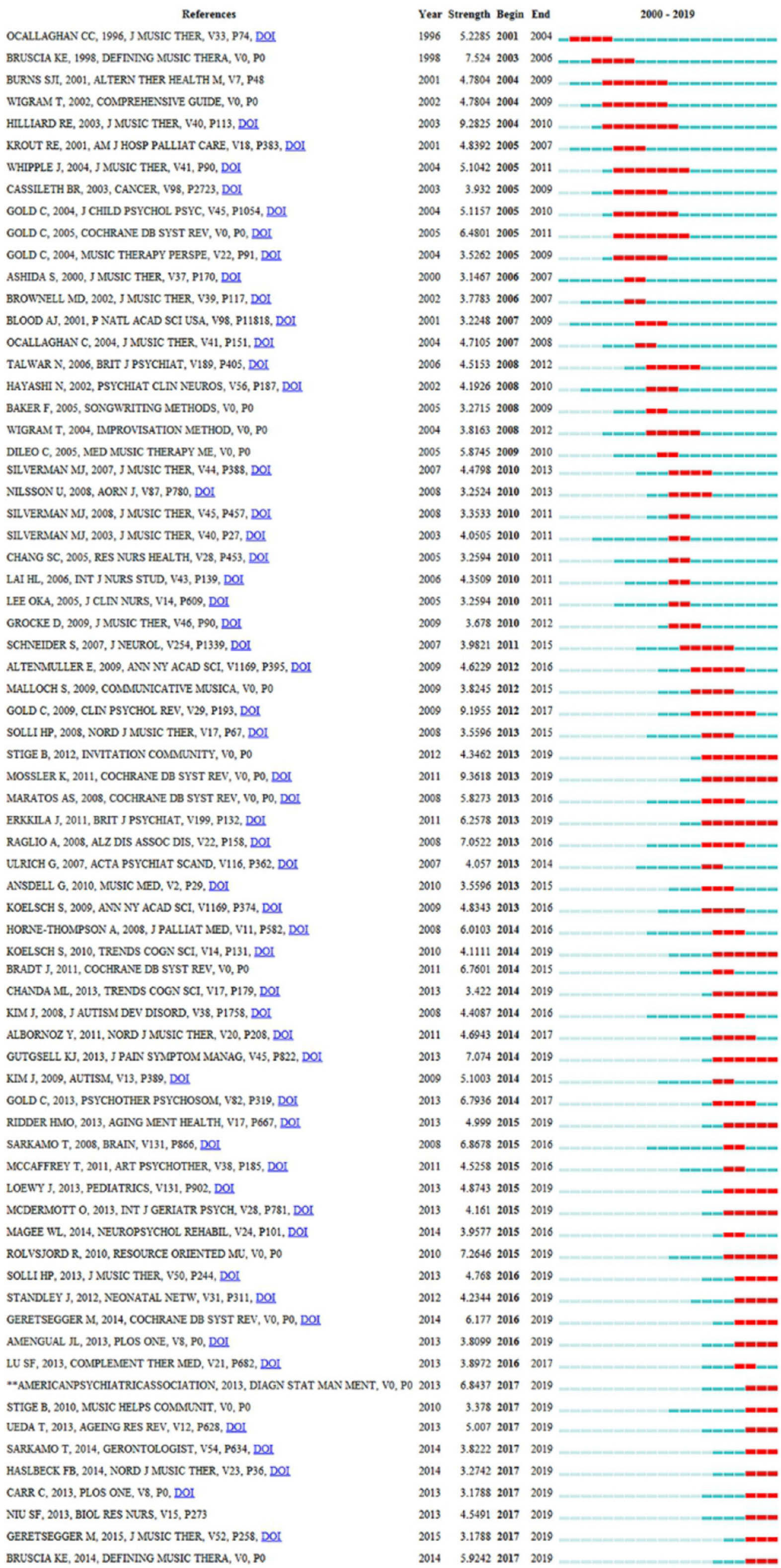
According to the science map analysis, hot music therapy topics among publications are discussed.

I. The cluster "#1 improvisational music therapy" (IMT) is the current research frontier in the music therapy research field. In general, music therapy has a long research tradition within autism spectrum disorders (ASD), and there have been more rigorous studies about it in recent years. IMT for children with autism is described as a child-centered method. Improvisational music-making may enhance social interaction and expression of emotions among children with autism, such as responding to communication acts (Geretsegger et al., 2012, 2015). In addition, IMT is an evidence-based treatment approach that may be helpful for people who abuse drugs or have cancer. A study applied improving as a primary music therapeutic practice, and the result indicated that IMT will be effective in treating depression accompanied by drug abuse among adults (Albornoz, 2011). By applying the interpretative phenomenological analysis and psychological perspectives, a study explained the significant role of music therapy as an innovative psychological intervention in cancer care settings (Pothoulaki et al., 2012). IMT may serve as an effective additional method for treating psychiatric disorders in the short and medium term, but it may need more studies to identify the long-term effects in clinical practice.

II. Based on the analysis of co-citation counts, the top three references all applied music therapy to improve the quality of life of clients. They highlight the fact that music therapy is an effective method that can cover a range of clinical skills, thus helping people with psychological disorders, chronic illnesses, and pain management issues. Furthermore, music therapy mixed with standard care can help individuals with schizophrenia improve their global state, mental state (including negative and general symptoms), social functioning, and quality of life (Gold et al., 2009; Erkkilä et al., 2011; Geretsegger et al., 2017).

III. By understanding the keywords with the strongest citation bursts, the research frontier can be predicted. Three keywords, "efficacy," "health," and "older adults," emphasized the research trends in terms of the strongest citation bursts.

a. Efficacy: This refers to measuring the effectiveness of music therapy in terms of clinical skills. Studies have found that a wide variety of psychological disorders can be effectively treated with music. In the study of Fukui, patients with Alzheimer's disease listened to music and verbally communicated with their music therapist. The results showed that problematic behaviors of the patients with Alzheimer's disease decreased (Fukui et al., 2012). The aim of the study of Erkkila was to determine the efficacy of music therapy when added to standard care. The result of this study also indicated that music therapy had specific qualities for non-verbal expression and communication when patients cannot verbally describe their inner experiences (Erkkilä et al., 2011). Additionally, as summarized by Ueda, music therapy reduced anxiety and depression in patients with dementia. However, his study cannot clarify what kinds of music therapy or patients have effectiveness. Thus, future studies should investigate music therapy with good methodology and evaluation methods (Ueda et al., 2013).
b. Health: Music therapy is a methodical intervention in clinical practice because it uses music experiences and relationships to promote health for adults and children (Bruscia, 1998). Also, music therapy is an effective means of achieving the optimal health and well-being of individuals and communities,



| References | Year | Strength | Begin | End | 2000 - 2019 |
|---|---|---|---|---|---|
| OCALLAGHAN CC, 1996, J MUSIC THER, V33, P74, DOI | 1996 | 5.2285 | 2001 | 2004 | |
| BRUSCIA KE, 1998, DEFINING MUSIC THERA, V0, P0 | 1998 | 7.524 | 2003 | 2006 | |
| BURNS SJI, 2001, ALTERN THER HEALTH M, V7, P48 | 2001 | 4.7804 | 2004 | 2009 | |
| WIGRAM T, 2002, COMPREHENSIVE GUIDE, V0, P0 | 2002 | 4.7804 | 2004 | 2009 | |
| HILLIARD RE, 2003, J MUSIC THER, V40, P113, DOI | 2003 | 9.2825 | 2004 | 2010 | |
| KROUT RE, 2001, AM J HOSP PALLIAT CARE, V18, P383, DOI | 2001 | 4.8392 | 2005 | 2007 | |
| WHIPPLE J, 2004, J MUSIC THER, V41, P90, DOI | 2004 | 5.1042 | 2005 | 2011 | |
| CASSILETH BR, 2003, CANCER, V98, P2723, DOI | 2003 | 3.932 | 2005 | 2009 | |
| GOLD C, 2004, J CHILD PSYCHOL PSYC, V45, P1054, DOI | 2004 | 5.1157 | 2005 | 2010 | |
| GOLD C, 2005, COCHRANE DB SYST REV, V0, P0, DOI | 2005 | 6.4801 | 2005 | 2011 | |
| GOLD C, 2004, MUSIC THERAPY PERSPE, V22, P91, DOI | 2004 | 3.5262 | 2005 | 2009 | |
| ASHIDA S, 2000, J MUSIC THER, V37, P170, DOI | 2000 | 3.1467 | 2006 | 2007 | |
| BROWNELL MD, 2002, J MUSIC THER, V39, P117, DOI | 2002 | 3.7783 | 2006 | 2007 | |
| BLOOD AJ, 2001, P NATL ACAD SCI USA, V98, P11818, DOI | 2001 | 3.2248 | 2007 | 2009 | |
| OCALLAGHAN C, 2004, J MUSIC THER, V41, P151, DOI | 2004 | 4.7105 | 2007 | 2008 | |
| TALWAR N, 2006, BRIT J PSYCHIAT, V189, P405, DOI | 2006 | 4.5153 | 2008 | 2012 | |
| HAYASHI N, 2002, PSYCHIAT CLIN NEUROS, V56, P187, DOI | 2002 | 4.1926 | 2008 | 2010 | |
| BAKER F, 2005, SONGWRITING METHODS, V0, P0 | 2005 | 3.2715 | 2008 | 2009 | |
| WIGRAM T, 2004, IMPROVISATION METHOD, V0, P0 | 2004 | 3.8163 | 2008 | 2012 | |
| DILEO C, 2005, MED MUSIC THERAPY ME, V0, P0 | 2005 | 5.8745 | 2009 | 2010 | |
| SILVERMAN MJ, 2007, J MUSIC THER, V44, P388, DOI | 2007 | 4.4798 | 2010 | 2013 | |
| NILSSON U, 2008, AORN J, V87, P780, DOI | 2008 | 3.2524 | 2010 | 2013 | |
| SILVERMAN MJ, 2008, J MUSIC THER, V45, P457, DOI | 2008 | 3.3533 | 2010 | 2011 | |
| SILVERMAN MJ, 2003, J MUSIC THER, V40, P27, DOI | 2003 | 4.0505 | 2010 | 2011 | |
| CHANG SC, 2005, RES NURS HEALTH, V28, P453, DOI | 2005 | 3.2594 | 2010 | 2011 | |
| LAI HL, 2006, INT J NURS STUD, V43, P139, DOI | 2006 | 4.3509 | 2010 | 2011 | |
| LEE OKA, 2005, J CLIN NURS, V14, P609, DOI | 2005 | 3.2594 | 2010 | 2011 | |
| GROCKE D, 2009, J MUSIC THER, V46, P90, DOI | 2009 | 3.678 | 2010 | 2012 | |
| SCHNEIDER S, 2007, J NEUROL, V254, P1339, DOI | 2007 | 3.9821 | 2011 | 2015 | |
| ALTENMULLER E, 2009, ANN NY ACAD SCI, V1169, P395, DOI | 2009 | 4.6229 | 2012 | 2016 | |
| MALLOCH S, 2009, COMMUNICATIVE MUSICA, V0, P0 | 2009 | 3.8245 | 2012 | 2015 | |
| GOLD C, 2009, CLIN PSYCHOL REV, V29, P193, DOI | 2009 | 9.1955 | 2012 | 2017 | |
| SOLLI HP, 2008, NORD J MUSIC THER, V17, P67, DOI | 2008 | 3.5596 | 2013 | 2015 | |
| STIGE B, 2012, INVITATION COMMUNITY, V0, P0 | 2012 | 4.3462 | 2013 | 2019 | |
| MOSSLER K, 2011, COCHRANE DB SYST REV, V0, P0, DOI | 2011 | 9.3618 | 2013 | 2019 | |
| MARATOS AS, 2008, COCHRANE DB SYST REV, V0, P0, DOI | 2008 | 5.8273 | 2013 | 2016 | |
| ERKKILA J, 2011, BRIT J PSYCHIAT, V199, P132, DOI | 2011 | 6.2578 | 2013 | 2019 | |
| RAGLIO A, 2008, ALZ DIS ASSOC DIS, V22, P158, DOI | 2008 | 7.0522 | 2013 | 2016 | |
| ULRICH G, 2007, ACTA PSYCHIAT SCAND, V116, P362, DOI | 2007 | 4.057 | 2013 | 2014 | |
| ANSDELL G, 2010, MUSIC MED, V2, P29, DOI | 2010 | 3.5596 | 2013 | 2015 | |
| KOELSCH S, 2009, ANN NY ACAD SCI, V1169, P374, DOI | 2009 | 4.8343 | 2013 | 2016 | |
| HORNE-THOMPSON A, 2008, J PALLIAT MED, V11, P582, DOI | 2008 | 6.0103 | 2014 | 2016 | |
| KOELSCH S, 2010, TRENDS COGN SCI, V14, P131, DOI | 2010 | 4.1111 | 2014 | 2019 | |
| BRADT J, 2011, COCHRANE DB SYST REV, V0, P0 | 2011 | 6.7601 | 2014 | 2015 | |
| CHANDA ML, 2013, TRENDS COGN SCI, V17, P179, DOI | 2013 | 3.422 | 2014 | 2019 | |
| KIM J, 2008, J AUTISM DEV DISORD, V38, P1758, DOI | 2008 | 4.4087 | 2014 | 2016 | |
| ALBORNOZ Y, 2011, NORD J MUSIC THER, V20, P208, DOI | 2011 | 4.6943 | 2014 | 2017 | |
| GUTGSELL KJ, 2013, J PAIN SYMPTOM MANAG, V45, P822, DOI | 2013 | 7.074 | 2014 | 2019 | |
| KIM J, 2009, AUTISM, V13, P389, DOI | 2009 | 5.1003 | 2014 | 2015 | |
| GOLD C, 2013, PSYCHOTHER PSYCHOSOM, V82, P319, DOI | 2013 | 6.7936 | 2014 | 2017 | |
| RIDDER HMO, 2013, AGING MENT HEALTH, V17, P667, DOI | 2013 | 4.999 | 2015 | 2019 | |
| SARKAMO T, 2008, BRAIN, V131, P866, DOI | 2008 | 6.8678 | 2015 | 2016 | |
| MCCAFFREY T, 2011, ART PSYCHOTHER, V38, P185, DOI | 2011 | 4.5258 | 2015 | 2016 | |
| LOEWY J, 2013, PEDIATRICS, V131, P902, DOI | 2013 | 4.8743 | 2015 | 2019 | |
| MCDERMOTT O, 2013, INT J GERIATR PSYCH, V28, P781, DOI | 2013 | 4.161 | 2015 | 2019 | |
| MAGEE WL, 2014, NEUROPSYCHOL REHABIL, V24, P101, DOI | 2014 | 3.9577 | 2015 | 2016 | |
| ROLVSJORD R, 2010, RESOURCE ORIENTED MU, V0, P0 | 2010 | 7.2646 | 2015 | 2019 | |
| SOLLI HP, 2013, J MUSIC THER, V50, P244, DOI | 2013 | 4.768 | 2016 | 2019 | |
| STANDLEY J, 2012, NEONATAL NETW, V31, P311, DOI | 2012 | 4.2344 | 2016 | 2019 | |
| GERETSEGGER M, 2014, COCHRANE DB SYST REV, V0, P0, DOI | 2014 | 6.177 | 2016 | 2019 | |
| AMENGUAL JL, 2013, PLOS ONE, V8, P0, DOI | 2013 | 3.8099 | 2016 | 2019 | |
| LU SF, 2013, COMPLEMENT THER MED, V21, P682, DOI | 2013 | 3.8972 | 2016 | 2017 | |
| **AMERICANPSYCHIATRICASSOCIATION, 2013, DIAGN STAT MAN MENT, V0, P0 | 2013 | 6.8437 | 2017 | 2019 | |
| STIGE B, 2010, MUSIC HELPS COMMUNIT, V0, P0 | 2010 | 3.378 | 2017 | 2019 | |
| UEDA T, 2013, AGEING RES REV, V12, P628, DOI | 2013 | 5.007 | 2017 | 2019 | |
| SARKAMO T, 2014, GERONTOLOGIST, V54, P634, DOI | 2014 | 3.8222 | 2017 | 2019 | |
| HASLBECK FB, 2014, NORD J MUSIC THER, V23, P36, DOI | 2014 | 3.2742 | 2017 | 2019 | |
| CARR C, 2013, PLOS ONE, V8, P0, DOI | 2013 | 3.1788 | 2017 | 2019 | |
| NIU SF, 2013, BIOL RES NURS, V15, P273 | 2013 | 4.5491 | 2017 | 2019 | |
| GERETSEGGER M, 2015, J MUSIC THER, V52, P258, DOI | 2015 | 3.1788 | 2017 | 2019 | |
| BRUSCIA KE, 2014, DEFINING MUSIC THERA, V0, P0 | 2014 | 5.9242 | 2017 | 2019 | |

**FIGURE 9 |** The strongest citation bursts among the top 71 references. The red measures indicate frequent citation of studies, and the green measures indicate infrequent citation of studies.

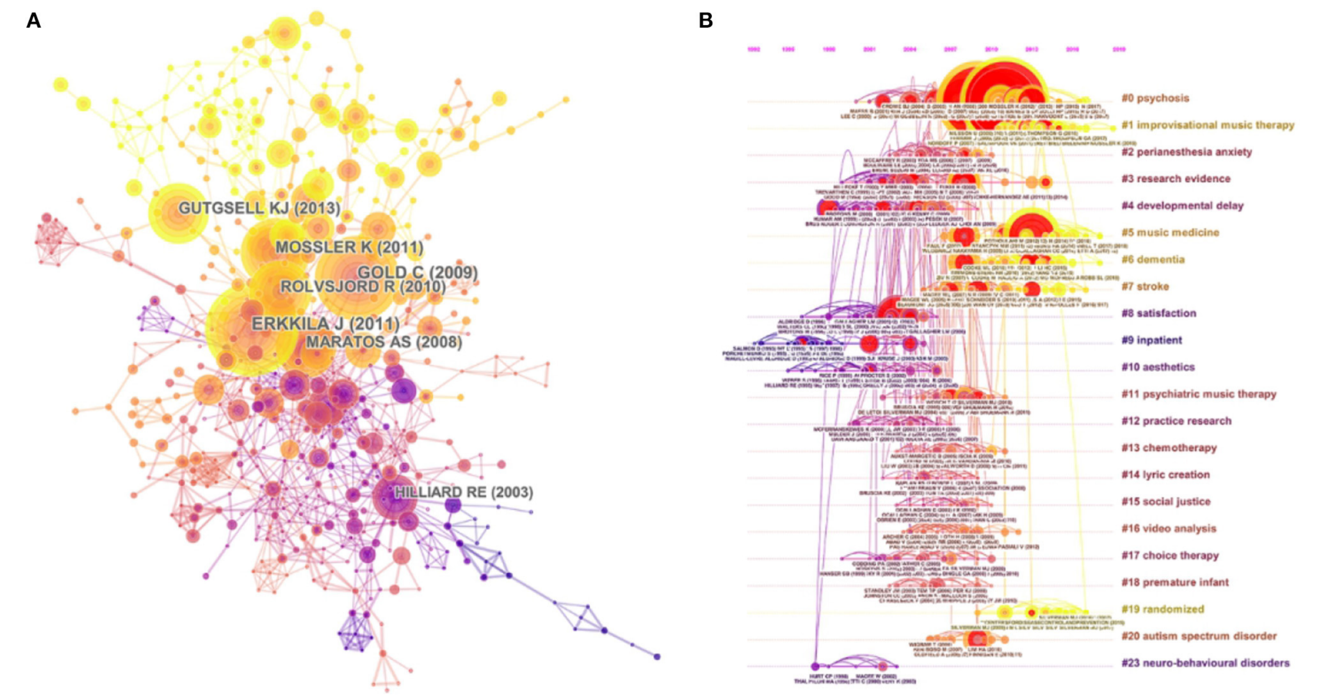



**FIGURE 10 | (A)** The network map of co-cited references and **(B)** the map (timeline view) of references with co-citation on top music therapy research. In these maps, the node represents a study, and the link represents the co-citation frequency between two studies. A larger node represents more publications of the author. A thicker purple circle represents greater influence in this field. **(A)** The nodes in the same color belong to the same cluster. **(B)** The nodes on the same line belong to the same cluster.

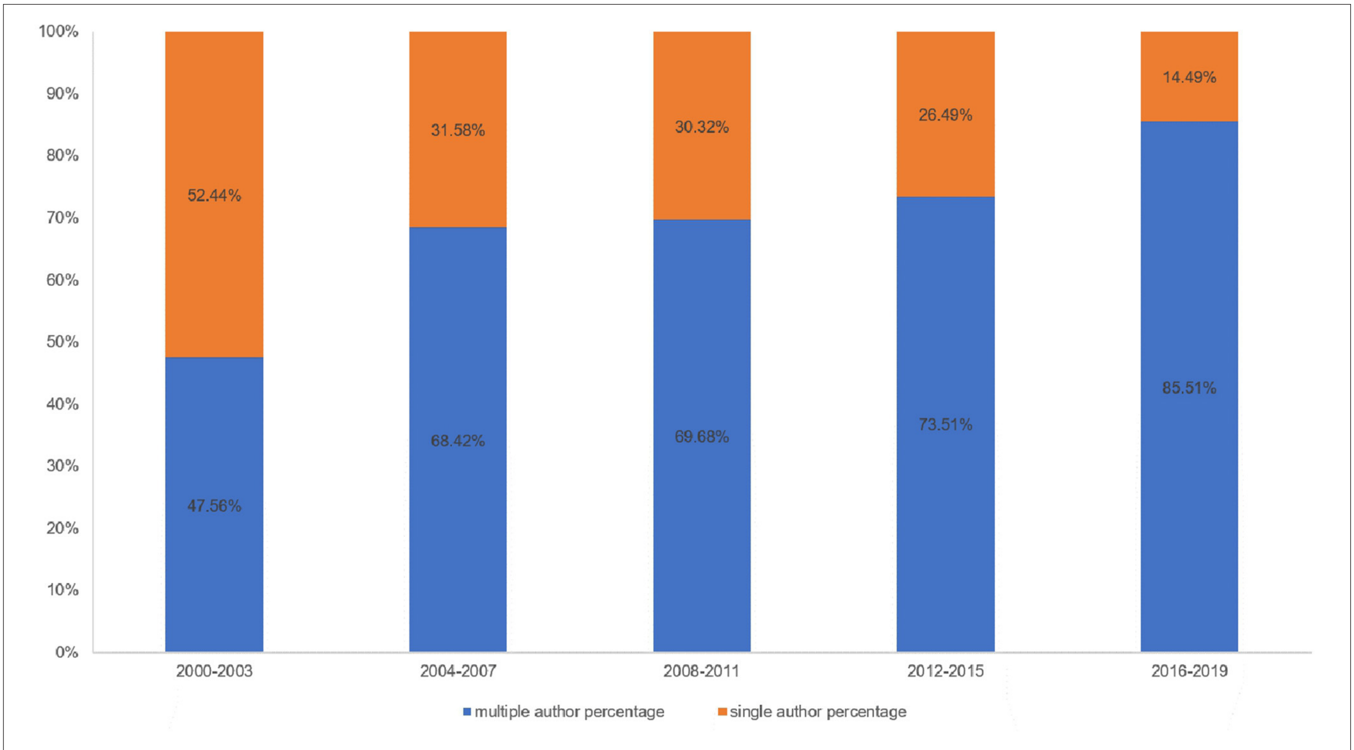



**FIGURE 11 |** The percentage of single- vs. multiple-authored articles. Blue bars mean multiple-author percentage; orange bars mean single-author percentage.

because it can be individualized or done as a group activity. The stimulation from music therapy can lead to conversations, recollection of memories, and expression. The study of Gold indicated that solo music therapy in routine practice is an effective addition to usual care for mental health care patients with low motivation (Gold et al., 2013). Porter summarized that music therapy contributes to improvement for both kids and teenagers with mental health conditions, such as depression and anxiety, and increases self-esteem in the short term (Porter et al., 2017).

c. Older adults: This refers to the use of music therapy as a treatment to maintain and slow down the symptoms observed in older adults (Mammarella et al., 2007; Deason et al., 2012). In terms of keywords with the strongest citation bursts, the most popular subjects of music therapy-related articles and reviews focused on children from 2005 to 2007. However, various researchers concentrated on older adults from 2017 to 2019. Music therapy was the treatment of choice for older adults with depression, Parkinson's disease, and Alzheimer's disorders (Brotons and Koger, 2000; Bernatzky et al., 2004; Johnson et al., 2011; Deason et al., 2012; McDermott et al., 2013; Sakamoto et al., 2013; Benoit et al., 2014; Pohl et al., 2020). In the study of Zhao, music therapy had positive effects on the reduction of depressive symptoms for older adults when added to standard therapies. These standard therapies could be standard care, standard drug treatment, standard rehabilitation, and health education (Zhao et al., 2016). The study of Shimizu demonstrated that multitask movement music therapy was an effective intervention to enhance neural activation in older adults with mild cognitive impairment (Shimizu et al., 2018). However, the findings of the study of Li explained that short-term music therapy intervention cannot improve the cognitive function of older adults. He also recommended that future researchers can apply a quality methodology with a long-term research design for the care needs of older adults (Li et al., 2015).

### Strengths and Limitations

To the best of our knowledge, this study was the first one to analyze large-scale data of music therapy publications from the past two decades through CiteSpace V. CiteSpace could detect more comprehensive results than simply reviewing articles and studies. In addition, the bibliometric method helped us to identify the emerging trend and collaboration among authors, institutions, and countries or regions.

This study is not without limitations. First, only articles and reviews published in the WoS Science Citation Index Expanded and Social Sciences Citation Index were analyzed. Future reviews could consider other databases, such as PubMed and Scopus. The document type labeled by publishers is not always accurate. For example, some publications labeled by WoS were not actually reviews (Harzing, 2013; Yeung, 2021). Second, the limitation may induce bias in frequency of reference. For example, some potential articles were published recently, and these studies could be not cited with frequent times. Also, in terms of obliteration by incorporation, some common knowledge or opinions become accepted that their contributors or authors are no longer cited (Merton, 1965; Yeung, 2021). Third, this review applied the quantitative analysis approach, and only limited qualitative analysis was performed in this study. In addition, we applied the CitesSpace software to conduct this bibliometric study, but the CiteSpace software did not allow us to complicate information under both full counting and fractional counting systems. Thus, future scholars can analyze the development of music therapy in some specific journals using both quantitative and qualitative indicators.

## CONCLUSIONS

This bibliometric study provides information regarding emerging trends in music therapy publications from 2000 to 2019. First, this study presents several theoretical implications related to publications that may assist future researchers to advance their research field. The results reveal that annual publications in music therapy research have significantly increased in the last two decades, and the overall trend in publications increased from 28 publications in 2000 to 111 publications in 2019. This analysis also furthers the comprehensive understanding of the global research structure in the field. Also, we have stated a high level of collaboration between different countries or regions and authors in the music therapy research. This collaboration has extremely expanded the knowledge of music therapy. Thus, future music therapy professionals can benefit from the most specialized research.

Second, this research represents several practical implications. IMT is the current research frontier in the field. IMT usually serves as an effective music therapy method for the health of people in clinical practice. Identifying the emerging trends in this field will help researchers prepare their studies on recent research issues (Mulet-Forteza et al., 2021). Likewise, it also indicates future studies to address these issues and update the existing literature. In terms of the strongest citation bursts, the three keywords, "efficacy," "health," and "older adults," highlight the fact that music therapy is an effective invention, and it can benefit the health of people. The development prospects of music therapy could be expected, and future scholars could pay attention to the clinical significance of music therapy to the health of people.

Finally, multiple researchers have indicated several health benefits of music therapy, and the music therapy mechanism perspective is necessary for future research to advance the field. Also, music therapy can benefit a wide range of individuals, such as those with autism spectrum, traumatic brain injury, or some physical disorders. Future researchers can develop music therapy standards to measure clinical practice.

## AUTHOR CONTRIBUTIONS

KL and LW: conceptualization, methodology, formal analysis, investigation, resources, writing—review, and editing. LW: software and data curation. KL: validation and writing—original draft preparation. XW: visualization, supervision, project administration, and funding acquisition. All authors contributed to the article and approved the submitted version.

## FUNDING

This study was supported by the Fok Ying-Tong Education Foundation of China (161092), the scientific and technological research program of the Shanghai Science and Technology Committee (19080503100), and the Shanghai Key Lab of Human Performance (Shanghai University of Sport) (11DZ2261100).


## REFERENCES


Albornoz, Y. (2011). The effects of group improvisational music therapy on depression in adolescents and adults with substance abuse: a randomized controlled trial. *Nord. J. Music Ther.* 20, 208–224. doi: 10.1080/08098131.2010.522717

Association, A. M. T. (2018). *History of Music Therapy*. Available online at: https://www.musictherapy.org/about/history/ (accessed November 10, 2020).

Benoit, C. E., Dalla Bella, S., Farrugia, N., Obrig, H., Mainka, S., and Kotz, S. A. (2014). Musically cued gait-training improves both perceptual and motor timing in Parkinson's disease. *Front. Hum. Neurosci.* 8:494. doi: 10.3389/fnhum.2014.00494

Bernatzky, G., Bernatzky, P., Hesse, H. P., Staffen, W., and Ladurner, G. (2004). Stimulating music increases motor coordination in patients afflicted with Morbus Parkinson. *Neurosci. Lett.* 361, 4–8. doi: 10.1016/j.neulet.2003.12.022

Bronson, H., Vaudreuil, R., and Bradt, J. (2018). Music therapy treatment of active duty military: an overview of intensive outpatient and longitudinal care programs. *Music Ther. Perspect.* 36, 195–206. doi: 10.1093/mtp/miy006

Brotons, M., and Koger, S. M. (2000). The impact of music therapy on language functioning in dementia. *J. Music Ther.* 37, 183–195. doi: 10.1093/jmt/37.3.183

Bruscia, K. (1998). *Defining Music Therapy 2nd Edition*. Gilsum: Barcelona publications.

Castanha, R. C. G., and Grácio, M. C. C. (2014). Bibliometrics contribution to the metatheoretical and domain analysis studies. *Knowl. Organiz.* 41, 171–174. doi: 10.5771/0943-7444-2014-2-171

Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *J. Am. Soc. Inform. Sci. Technol.* 57, 359–377. doi: 10.1002/asi.20317

Chen, C., Hu, Z., Liu, S., and Tseng, H. (2012). Emerging trends in regenerative medicine: a scientometric analysis in CiteSpace. *Expert Opin. Biol. Ther.* 12, 593–608. doi: 10.1517/14712598.2012.674507

Chen, H., Zhao, G., and Xu, N. (2012). "The analysis of research hotspots and fronts of knowledge visualization based on CiteSpace II," in *International Conference on Hybrid Learning*, Vol. 7411, eds S. K. S. Cheung, J. Fong, L. F. Kwok, K. Li, and R. Kwan (Berlin; Heidelberg: Springer), 57–68. doi: 10.1007/978-3-642-32018-_6

Deason, R., Simmons-Stern, N., Frustace, B., Ally, B., and Budson, A. (2012). Music as a memory enhancer: Differences between healthy older adults and patients with Alzheimer's disease. *Psychomusicol. Music Mind Brain* 22:175. doi: 10.1037/a0031118

Devlin, K., Alshaikh, J. T., and Pantelyat, A. (2019). Music therapy and music-based interventions for movement disorders. *Curr. Neurol. Neurosci. Rep.* 19:83. doi: 10.1007/s11910-019-1005-0

Durieux, V., and Gevenois, P. A. (2010). Bibliometric indicators: quality measurements of scientific publication. *Radiology* 255, 342–351. doi: 10.1148/radiol.09090626

Ellegaard, O., and Wallin, J. A. (2015). The bibliometric analysis of scholarly production: How great is the impact? *Scientometrics* 105, 1809–1831. doi: 10.1007/s11192-015-1645-z

Erkkilä, J., Punkanen, M., Fachner, J., Ala-Ruona, E., Pöntiö, I., Tervaniemi, M., et al. (2011). Individual music therapy for depression: randomised controlled trial. *Br. J. Psychiatry* 199, 132–139. doi: 10.1192/bjp.bp.110.085431

Falagas, M. E., Pitsouni, E. I., Malietzis, G. A., and Pappas, G. (2008). Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. *Faseb J.* 22, 338–342. doi: 10.1096/FJ.07-9492LSF

Fitzpatrick, R. B. (2005). Essential Science IndicatorsSM. *Med. Ref. Serv. Q.* 24, 67–78. doi: 10.1300/J115v24n04_05

Fukui, H., Arai, A., and Toyoshima, K. (2012). Efficacy of music therapy in treatment for the patients with Alzheimer's disease. *Int. J. Alzheimer's Dis.* 2012, 531646–531646. doi: 10.1155/2012/531646

Geretsegger, M., Holck, U., Carpente, J. A., Elefant, C., Kim, J., and Gold, C. (2015). Common characteristics of improvisational approaches in music therapy for children with autism spectrum disorder: developing treatment guidelines. *J. Music Ther.* 52, 258–281. doi: 10.1093/jmt/thv005

Geretsegger, M., Holck, U., and Gold, C. (2012). Randomised controlled trial of improvisational music therapy's effectiveness for children with autism spectrum disorders (TIME-A): study protocol. *BMC Pediatr.* 12:2. doi: 10.1186/1471-2431-12-2

Geretsegger, M., Mossler, K. A., Bieleninik, L., Chen, X. J., Heldal, T. O., and Gold, C. (2017). Music therapy for people with schizophrenia and schizophrenia-like disorders. *Cochrane Database Syst. Rev.* 5:CD004025. doi: 10.1002/14651858.CD004025.pub4

Gold, C., Mössler, K., Grocke, D., Heldal, T. O., Tjemsland, L., Aarre, T., et al. (2013). Individual music therapy for mental health care clients with low therapy motivation: multicentre randomised controlled trial. *Psychother. Psychosom.* 82, 319–331. doi: 10.1159/000348452

Gold, C., Solli, H. P., Krüger, V., and Lie, S. A. (2009). Dose-response relationship in music therapy for people with serious mental disorders: systematic review and meta-analysis. *Clin. Psychol. Rev.* 29, 193–207. doi: 10.1016/j.cpr.2009.01.001

Gonzalez-Serrano, M. H., Jones, P., and Llanos-Contrera, O. (2020). An overview of sport entrepreneurship field: a bibliometric analysis of the articles published in the Web of Science. *Sport Soc.* 23, 296–314. doi: 10.1080/17430437.2019.1607307

Gooding, L. F., and Langston, D. G. (2019). Music therapy with military populations: a scoping review. *J. Music Ther.* 56, 315–347. doi: 10.1093/jmt/thz010

Harzing, A.-W. (2013). Document categories in the ISI web of knowledge: misunderstanding the social sciences? *Scientometrics* 94, 23–34. doi: 10.1007/s11192-012-0738-1

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

Li, H. C., Wang, H. H., Chou, F. H., and Chen, K. M. (2015). The effect of music therapy on cognitive functioning among older adults: a systematic review and meta-analysis. *J. Am. Med. Dir. Assoc.* 16, 71–77. doi: 10.1016/j.jamda.2014.10.004

Liang, Y. D., Li, Y., Zhao, J., Wang, X. Y., Zhu, H. Z., and Chen, X. H. (2017). Study of acupuncture for low back pain in recent 20 years: a bibliometric analysis via CiteSpace. *J. Pain Res.* 10, 951–964. doi: 10.2147/jpr.S132808

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

McDermott, O., Crellin, N., Ridder, H. M., and Orrell, M. (2013). Music therapy in dementia: a narrative synthesis systematic review. *Int. J. Geriatr. Psychiatry* 28, 781–794. doi: 10.1002/gps.3895

Merigó, J. M., Mulet-Forteza, C., Valencia, C., and Lew, A. A. (2019). Twenty years of tourism Geographies: a bibliometric overview. *Tour. Geograph.* 21, 881–910. doi: 10.1080/14616688.2019.1666913

Merton, R. K. (1965). *On the Shoulders of Giants: A Shandean Postscript*. New York, NY: Free Press.

Miao, Y., Xu, S. Y., Chen, L. S., Liang, G. Y., Pu, Y. P., and Yin, L. H. (2017). Trends of long noncoding RNA research from 2007 to 2016: a bibliometric analysis. *Oncotarget* 8, 83114–83127. doi: 10.18632/oncotarget.20851

Mulet-Forteza, C., Lunn, E., Merigó, J. M., and Horrach, P. (2021). Research progress in tourism, leisure and hospitality in Europe (1969–2018). *Int. J. Contemp. Hospit. Manage.* 33, 48–74. doi: 10.1108/IJCHM-06-2020-0521

Pohl, P., Wressle, E., Lundin, F., Enthoven, P., and Dizdar, N. (2020). Group-based music intervention in Parkinson's disease - findings from a mixed-methods study. *Clin. Rehabil.* 34, 533–544. doi: 10.1177/0269215520907669

Porter, S., McConnell, T., McLaughlin, K., Lynn, F., Cardwell, C., Braiden, H. J., et al. (2017). Music therapy for children and adolescents with behavioural and emotional problems: a randomised controlled trial. *J. Child Psychol. Psychiatry* 58, 586–594. doi: 10.1111/jcpp.12656

Pothoulaki, M., MacDonald, R., and Flowers, P. (2012). An interpretative phenomenological analysis of an improvisational music therapy program for cancer patients. *J. Music Ther.* 49, 45–67. doi: 10.1093/jmt/49.1.45

Sakamoto, M., Ando, H., and Tsutou, A. (2013). Comparing the effects of different individualized music interventions for elderly individuals with severe dementia. *Int. Psychogeriatr.* 25, 775–784. doi: 10.1017/s1041610212002256

Shimizu, N., Umemura, T., Matsunaga, M., and Hirai, T. (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. Health* 22, 1614–1626. doi: 10.1080/13607863.2017.1379048

Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *J. Am. Soc. Inform. Sci.* 24, 265–269. doi: 10.1002/asi.4630240406

Ueda, T., Suzukamo, Y., Sato, M., and Izumi, S. (2013). Effects of music therapy on behavioral and psychological symptoms of dementia: a systematic review and meta-analysis. *Ageing Res. Rev.* 12, 628–641. doi: 10.1016/j.arr.2013.02.003

Wang, X. Q., Peng, M. S., Weng, L. M., Zheng, Y. L., Zhang, Z. J., and Chen, P. J. (2019). Bibliometric study of the comorbidity of pain and depression research. *Neural. Plast* 2019:1657498. doi: 10.1155/2019/1657498

Yeung, A. W. K. (2021). Is the influence of freud declining in psychology and psychiatry? A bibliometric analysis. *Front. Psychol.* 12:631516. doi: 10.3389/fpsyg.2021.631516

Zhao, K., Bai, Z. G., Bo, A., and Chi, I. (2016). A systematic review and meta-analysis of music therapy for the older adults with depression. *Int. J. Geriatr. Psychiatry* 31, 1188–1198. doi: 10.1002/gps.4494

Zheng, K., and Wang, X. (2019). Publications on the association between cognitive function and pain from 2000 to 2018: a bibliometric analysis using citespace. *Med. Sci. Monit.* 25, 8940–8951. doi: 10.12659/msm.917742

**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.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 22 July 2021
doi: 10.3389/fnins.2021.622053

# Sex Differences are Reflected in Microstructural White Matter Alterations of Musical Sophistication: A Diffusion MRI Study

*Mohammad-Mehdi Mehrabinejad[1], Parnian Rafei[2], Hossein Sanjari Moghaddam[1], Zeinab Sinaeifar[1] and Mohammad Hadi Aarabi[1,3]**

[1] Faculty of Medicine, Tehran University of Medical Sciences, Tehran, Iran, [2] Department of Psychology, Faculty of Psychology and Education, University of Tehran, Tehran, Iran, [3] Department of Neuroscience, Padova Neuroscience Center (PNC), University of Padova, Padova, Italy

**Background:** The human-specified ability to engage with different kinds of music in sophisticated ways is named "Musical Sophistication." Herein, we investigated specific white matter (WM) tracts that are associated with musical sophistication and musicality in both genders, separately, using Diffusion MRI connectometry approach. We specifically aimed to explore potential sex differences regarding WM alterations correlated with musical sophistication.

**Methods:** 123 healthy participants [70 (56.9%) were male, mean age = 36.80 ± 18.86 year], who were evaluated for musical sophistication using Goldsmiths Musical Sophistication Index (Gold-MSI) self-assessment instrument from the LEMON database, were recruited in this study. The WM correlates of two Gold-MSI subscales (active engagement and music training) were analyzed. Images were prepared and analyzed with diffusion connectometry to construct the local connectome. Multiple regression models were then fitted to address the correlation of local connectomes with Gold-MSI components with the covariates of age and handedness.

**Results:** a significant positive correlation between WM integrity in the corpus callosum (CC), right corticospinal tract (CST), cingulum, middle cerebellar peduncle (MCP), bilateral parieto-pontine tract, bilateral cerebellum, and left arcuate fasciculus (AF) and both active engagement [false discovery rate (FDR) = 0.008] and music training (FDR = 0.057) was detected in males. However, WM integrity in the body of CC, MCP, and cerebellum in females showed an inverse association with active engagement (FDR = 0.046) and music training (FDR = 0.032).

**Conclusion:** WM microstructures with functional connection with motor and somatosensory areas (CST, cortico-pontine tracts, CC, cerebellum, cingulum, and MCP) and language processing area (AF) have significant correlation with music engagement and training. Our findings show that these associations are different between males and females, which could potentially account for distinctive mechanisms related to musical perception and musical abilities across genders.

**Keywords: musicality, musical sophistication, diffusion magnetic resonance imaging, connectometry, white matter, white matter microstructural changes, music perception and cognition**

OPEN ACCESS

***Edited by:***
*Francisco Javier Zarza-Alzugaray, University of Zaragoza, Spain*

***Reviewed by:***
*McNeel Gordon Jantzen, Western Washington University, United States*
*Desirée García Gil, Complutense University of Madrid, Spain*
*Oscar Casanova, University of Zaragoza, Spain*

***Correspondence:***
*Mohammad Hadi Aarabi*
*mohammadhadiarabi@gmail.com*

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

***Received:*** *27 October 2020*
***Accepted:*** *28 June 2021*
***Published:*** *22 July 2021*

***Citation:***
*Mehrabinejad M-M, Rafei P, Sanjari Moghaddam H, Sinaeifar Z and Aarabi MH (2021) Sex Differences are Reflected in Microstructural White Matter Alterations of Musical Sophistication: A Diffusion MRI Study. Front. Neurosci. 15:622053. doi: 10.3389/fnins.2021.622053*

## INTRODUCTION

The human-specified ability and skills to engage with different kinds of music in ways with different sophistication degrees is named "Musical Sophistication" (Merriam and Merriam, 1964; Müllensiefen et al., 2014). There is a long scientific history for the assessment of music abilities and behavior and music perception in humankind (Boyle and Radocy, 1987; Rickard et al., 2015), and measuring musicality and musical sophistication, and their correlates had always been a challenge in different research paradigms.

The history of assessing musical sophistication goes back to a hundred years ago since the first measurement of musical talent was developed by Carl Seashore and colleagues in 1919 (Seashore et al., 1956). However, the majority of assessments employed for assessing musical ability (i.e., musicality) of individuals have been carried out in specific populations such as professionals in the field of music (e.g., musicians) or those who suffer from a pathological musical condition (e.g., amusia) (Brockmeier et al., 2011; Law and Zentner, 2012; Sato et al., 2015; Larrouy-Maestri et al., 2019). This vast omission in musicality testing paradigms led to the development of the recently developed batteries such as the Goldsmiths Musical Sophistication Index (Gold-MSI) self-assessment instrument (Müllensiefen et al., 2014).

The Gold-MSI is a self-report measure consisting of five subscales and two subjective auditory listening tests (a melodic memory task and beat perception task), as well as an exercise regarding a sound similarity judgmental task (Müllensiefen et al., 2014; Baker et al., 2020). Musical sophistication in the Gold-MSI paradigm is conceptualized as a psychometric construct that is involved in developing musical skills, achievements, and expertise and defines a multi-faceted musical behavior in individuals (Müllensiefen et al., 2014). As a result of the aforementioned attributes, the Gold-MSI questionnaire has been shown to be suitable for assessing musicality in the non-musician population in a comprehensive way (Baker et al., 2020). Moreover, this instrument is capable of assessing a broad range of musical-related abilities, consisting of the individual's performance on a musical instrument, their listening expertise, the ability to engage with music in functional settings or to communicate about music (Müllensiefen et al., 2014).

There are a number of neuroimaging studies that have investigated the neural basis, functional neural connectivity, and neuro-anatomical evidences of musical perception and expertise in grey matter (GM) of the humankind brain (Parsons, 2001; Schmithorst and Wilke, 2002; Elmer et al., 2013; Wu et al., 2013; Oechslin et al., 2018; van Vugt et al., 2021). In general, they have identified significant pieces of evidence on differences in the brain GM in brain areas such as the cerebellum (Gaser and Schlaug, 2003), Broca's area (Sluming et al., 2002), and Planum temporal (Keenan et al., 2001; Burkhard et al., 2020) in musicians and non-musicians. Several studies also demonstrate that brain structures with a close connection with the motor system have significant association with musicality (Schlaffke et al., 2020; Møller et al., 2021).

However, the mainstream studies in this field are majorly focused on the structural and functional characteristics of the GM in professional musicians, and the WM correlates of musicality is hitherto considered as a relatively understudied phenomenon in this field. In general, empirical pieces of evidence show that musicians exhibit changes in the white matter structure of their brain (Levitin, 2012). For instance, the Corticospinal Tract (CST) has been shown to have reduced fractional anisotropy (a measure of the directionality of water diffusion) in professional musicians, which indicates increased radial diffusivity (Imfeld et al., 2009). This notion is, however, discrepant among published studies investigating the diffusivity measures between higher and lower fractional anisotropy values of known tracts in response to musical training (Schlaug, 2015). These variations and inconsistencies in such findings have been reported to be potentially influenced by factors such as fiber density, cell membrane density, axon collateral sprouting, axon diameter, myelination, and fiber coherence. Higher fractional anisotropy values has been thought to reflect more aligned fibers in a specific tract, whereas lower fractional anisotropy values indicate less alignment of fibers in addition to more axonal sprouting and more branching of axons close to the cortical target region (Wan et al., 2014; Rüber et al., 2015; Schlaug, 2015). Moreover, the macro and microstructural organization of the Arcuate Fasciculus (AF)—a prominent WM tract connecting temporal and frontal brain regions— has been shown to have a predictive role in learning rate and learning speed in musical tasks related to rhythm and melody training (Vaquero et al., 2018). Other studies suggest that musical training is associated with microstructural adaptations in the AF, which appear as increased tract volume in the right AF of musicians compared to non-musicians (Halwani et al., 2011). A more recent study found that the microstructural organization of WM tracts that connect auditory and frontal motor regions in both hemispheres of the brain may serve as a neural foundation of the musicality or musicians' advantage (Li et al., 2021). Nevertheless, the most frequently reported WM microstructural differences in musicians compared with non-musicians appears to be in the cross-hemispheric connections (i.e., Corpus Callosum CC) (for a review, see Moore et al., 2017).

In light of the small number of studies investigating the WM alterations related to musicality and given the conflicting results that former studies presented, we aimed to investigate the brain white matter (WM) integrity to identify the microstructural patterns associated with musicality in general population.

Diffusion MRI Connectometry approach, the analysis method that we used in this study, is a novel approach, is reported to be more sensitive than diffusion tensor imaging (DTI) metrics (Yeh et al., 2016), and gives additional spatial resolution to track the WM fibers (Rahmani et al., 2017; Moghaddam et al., 2020; Mehrabinejad et al., 2021). Therefore, in the present study, we investigated specific WM tracts associated with musical sophistication and musicality utilizing diffusion MRI connectometry as an exploratory approach and looked for any variations in WM microstructural patterns related to demographic characteristics such as gender, a predictor of self-reported musical sophistication (Gold-MSI) subscales (Greenberg et al., 2015).

We hypothesized that those WM tracts that were formerly found to be associated with auditory and motor regions, such as

AF, may be significantly correlated with two Gold-MSI subscales (active engagement and music training). To the best of our knowledge, no previous study has specifically examined the WM correlates of musical sophistication using the diffusion MRI connectometry approach. In addition, we were interested in exploring potential sex differences related to the association between WM trajectories and each Gold-MSI subscale. Given the empirical evidence showing significant within-hemisphere and inter-hemisphere differences between males and females' brain structural and functional connectomes using DTI (Ingalhalikar et al., 2014), and the gap in the literature in terms of sex differences in musical sophistication and its neural correlates, we sought potential differences by running discrete analyses for male and female subjects of this study.

## MATERIALS AND METHODS

### Overview

To recognize the similarity in local connectivity patterns and identify the pathways of WM tracts, diffusion MRI connectometry approach measures the density of water diffusion through different directions of a voxel. Thus, water diffusivity measurement which is the speed of water diffusion in different directions, and is the primary concern of conventional DTI analysis, is replaced by measuring the density of water diffusion in diffusion MRI connectometry; which leads us to the identification of the local connectivity of fibers and tracking the subcomponents of the tract pathways which are significantly associated with our study variables (i.e., musical sophistication).

### Study Data

In the current study, we obtained all the required data from the "Leipzig Study for Mind-Body-Emotion Interactions" (LEMON) dataset (Babayan et al., 2019)[1]. The LEMON study was carried out in four series from September 2013 to September 2015. After prescreening via telephone interview, participants who met the eligibility criteria were invited to Max Planck Institute for Human Cognitive and Brain Sciences (MPI-CBS) for further evaluations. Exclusion criteria were present or past history of any cardiovascular (hypertension, congenital heart disease, or heart attacks), psychiatric [conditions needing more than 2 weeks therapy within last 10 years, post-traumatic stress disorder (PTSD), psychosis, or suicidal attempts], neurological [e.g., multiple sclerosis (MS), epilepsy, stroke], or malignant diseases, and also some particular medication usage (e.g., centrally active drugs, cortisol, alpha- or beta-blocker, extensive alcohol, benzodiazepine, cocaine, amphetamines, cannabis, or opiates) as well as any MRI contraindications.

A total of 227 eligible German-speaking participants who were screened via a telephone interview in day 0, participated in a 5-day survey. All participants were examined at the Day Clinic for Cognitive Neurology of the University Clinic Leipzig and the MPI-CBS in Leipzig, Germany. Briefly, all enrolled participants were asked to complete: (1) four fMRI and one structural scan in one session; (2) a battery of personality questionnaire; and (3) a set of cognitive, attention, and creativity related tasks. The Gold-MSI questionnaire was given in the first day of the study.

### Participants

Of 227 participated individuals in the LEMON study, 123 healthy participants, who were evaluated for musical sophistication using Gold-MSI from the LEMON database, were recruited in this study. The present study was carried out in accordance with the World Medical Association Declaration of Helsinki revised in 1989 and approved by the Ethics Committee of the University of Leipzig (reference number 154/13-ff).

### Goldsmiths Musical Sophistication Index (Gold-MSI)

The Gold-MSI evaluates self-reported musical abilities and behaviors on multiple aspects in the general population (Müllensiefen et al., 2014). Five subscales of musical sophistication consist of active musical engagement, self-reported perceptual abilities, musical training, self-reported singing abilities, and sophisticated emotional engagement with music (Müllensiefen et al., 2014). Of five main subscales (active engagement, musical training, perceptual abilities, singing abilities, and emotions) only first two subscales were included in the LEMON study.

Active engagement is defined as the level of music engagement including reading, writing, and listening to music as well as the time and income spent on music and music events attendance (Müllensiefen et al., 2014). Musical training reveals the musical dedication according to the time (peak hour per day and amount of training years) spent on training and number of instruments played (Müllensiefen et al., 2014). A subset of 16 items, scoring on a seven-point Likert scale (1: "absolutely disagree" to 7: "absolutely agree") were measured. Higher scores are attributed to higher musical sophistication. The German version of this scale was used in this study (Schaal et al., 2014).

### Image Acquisition

Magnetic resonance images were acquired with a 3 Tesla scanner (MAGNETOM Verio, Siemens Healthcare GmbH, Erlangen, Germany) and a 32-channel head coil in addition to a multi-band accelerated sequence combined with an in-plane GRAPPA (TR = 7,000 ms, TE = 80 ms, GRAPPA acceleration factor = 2, bandwidth = 1,502 Hz/Px, field of view = $220 \times 220$ mm$^2$, and voxel size = $1.7 \times 1.7 \times 1.7$ mm$^3$) aiming for diffusion MRI data collection. Consequently, seven b0 and sixty diffusion MRI images were recorded.

### Imaging Data and Statistical Analysis

Preprocessing steps, i.e., head motion, eddy current distortions, and susceptibility artifacts because of the magnetic field inhomogeneity correction, were carried out using the ExploreDTI toolbox[2] (Leemans et al., 2009). Diffusion data were reconstructed within Montreal Neuroimaging Initiative

[1] http://doi.org/10.15387/fcp_indi.mpi_lemon

[2] http://www.exploredti.com

(MNI) space, using q-space diffeomorphic reconstruction (Yeh et al., 2011) to obtain the Spin Diffusion Function (SDF; main component of diffusion connectometry). Subsequently, a diffusion sampling length ratio of 1.25 was used.

Diffusion metrics association with two subscales of Gold-MSI (active engagement and musical training) were analyzed through dMRI connectometry (Yeh et al., 2016). A multiple regression model was performed to consider these subscales in female and male participants, separately. Age and handedness were taken into account in all the analyses to adjust for possible confounding effects. T-score threshold of 2.5 was defined to delineate local connectomes. Deterministic fiber tracking algorithm was used to estimate WM tracts (Yeh et al., 2013). After normalization, topology-informed pruning was undertaken to prevent false positive tracking. Tracks were generated from bootstrap resampling and a length threshold of 20 voxel distance was utilized to highlight tracks. The seeding number for each permutation testing was set to 100,000. A total of 2,000 randomized permutations were employed to the group label to obtain the null distribution of track lengths in order to estimate the false discovery rate (FDR).

## RESULTS

dMRI of 123 healthy participants underwent analysis to investigate the association between musical sophistication subscales with WM connectivity in different genders. Of those, 70 (56.9%) were male, 120 (97.6%) were right-handed, and were on average middle-aged adults (mean age $\pm$ standard deviation (SD): 36.80 $\pm$ 18.86 year) (**Table 1**).

The mean score of active engagement and musical training were 23.95 $\pm$ 8.36 (range: 7–44) and 19.65 $\pm$ 10.09 (range: 7–40), respectively. Males and females did not differ significantly in their active engagement and musical training scores ($p = 0.45$ and 0.80, respectively) (**Table 2**).

dMRI connectometry analysis revealed significant correlations (either positive or negative) between quantitative anisotropy (QA) of some specific WM trajectories with each Gold-MSI subscales in different genders (**Table 3**).

Active engagement was significantly and positively correlated with WM integrity in genu, body, and splenium of CC, bilateral parieto-pontine tract, right cortico-spinal tract (CST), middle cerebellar peduncle (MCP), bilateral cortico-thalamic pathway, right fronto-pontine tract, right cingulum, bilateral cerebellum, and left arcuate fasciculus (AF) in male participants (FDR = 0.008, **Figure 1**). Besides, it was significantly, but negatively, correlated with QA values in body and splenium of the CC in females, as well (FDR = 0.046) (**Table 3** and **Figure 2**).

We also observed direct yet marginally significant association between musical training and QA values of genu, body and splenium of CC, bilateral cingulum, right CST, bilateral parieto-pontine tract, and left AF in males (FDR = 0.057, **Figure 3**). Besides, inverse association was detected in genu, body and splenium of CC in females (FDR = 0.032) (**Table 3** and **Figure 4**).

## DISCUSSION

Alterations in WM architecture in individuals with high musical abilities is a common report across studies probing into the association between WM characteristics and musical abilities. However, the connection of musical sophistication, which is a broad concept of musical-related abilities in the population of non-professional musicals, with the WM microstructures has been overlooked. Thus, in the current study, we explored the association of musical sophistication and abilities related to musicality and the WM microstructures in the brain of a population with a diverse musical training and engagement background. In order to do this, we assessed the WM correlates of active engagement and musical training as the two Gold-MSI subscales. In general, we showed that the WM microstructure connectivity pattern correlates with the Gold-MSI subscales of musical sophistication and gender. To elaborate on, our findings support our hypothesis regarding the sex differences in terms of WM microstructural alterations related to musical active engagement and musical training. Within the male participants, results exhibited higher WM coherence in CC, CST, MCP, AF, cingulum, and parietopontine tract with a higher score in both investigated subscales of Gold-MSI (i.e., active engagement and musical training) and higher integrity in the cortico-thalamic tract, fronto-pontine tract, and cerebellum with higher active engagement, whereas findings related to the female participants showed a negative correlation between WM connectivity in CC and both investigated Gold-MSI subscales.

Multiple efforts have been taken to underpin the neural basis of musical perception in both WM and GM (Parsons, 2001; Schmithorst and Wilke, 2002; Elmer et al., 2013; Wu et al., 2013; Oechslin et al., 2018). For instance, Heschl's gyrus in the primary auditory cortex has indicated to be a possible marker of musicality due to a MEG study comparing professional and amateur musicians with non-musicians which showed a significant difference in MEG activity between the groups in the primary auditory cortex, moreover, a brain volumetric difference has also been found between groups (Schneider et al., 2002). Results have shown that there are significant pieces of evidence on differences in the brain GM in brain areas such as the cerebellum (Gaser and Schlaug, 2003), Broca's area (Sluming et al., 2002), and Planum temporal (Keenan et al., 2001; Burkhard et al., 2020) in musicians and non-musicians. Nevertheless, the direct findings related to WM are limited, and the interpretation of our results requires discussing our findings based on previously known functions of the structures, as follows:

Active engagement and musical training are defined as the engagement and dedication level with musical activities and musical events (Müllensiefen et al., 2014). Both these two subscales have a close association with motor and somatosensory activities. For instance, active musical engagement has been defined as the amount of time and effort spent on music, whereas the musical training refers to the formal amount of musical training received. Both these functions are involved with motor activities regardless of the specific musical activity or instrument that one may employ. As described further, our findings are in accordance with the previously known functions

**TABLE 1 |** Demographic characteristics of study participants ($n = 123$).

| | |
|---|---|
| **Age** | 36.80 ± 18.86 (years) |
| **Gender** | Male: 70 (56.9%) |
| | Female: 53 (43.1%) |
| **Handedness** | Right-handed: 120 (97.6%) |
| | Left-handed: 0 (0%) |
| | Ambidextrous: 3 (2.4%) |

**TABLE 2 |** Gold-MSI results for active engagement and musical training subscales in all participants and each gender separately.

| Variables | Total (Mean (*SD*)) | Male (Mean (*SD*)) | Female (Mean (*SD*)) | *p*-value |
|---|---|---|---|---|
| **Active engagement** | 23.95 (8.36) | 24.44 (8.41) | 23.30 (8.33) | 0.45 |
| **Musical training** | 19.65 (10.09) | 19.46 (10.28) | 19.91 (9.92) | 0.80 |

**TABLE 3 |** WM tracts with significant association with Gold-MSI subscales in each gender.

| Variables | WM tracts | |
|---|---|---|
| | **Positive** | **Negative** |
| **Active engagement** | | |
| **Male** | Genu, body, and splenium of CC, bilateral parieto-pontine tract, right CST, MCP, bilateral cortico-thalamic pathway, right fronto-pontine tract, right cingulum, bilateral cerebellum, and left AF | None |
| **Female** | None | Body and splenium of CC |
| **Musical training** | | |
| **Male** | Genu, body and splenium of CC, bilateral cingulum, right CST, bilateral parieto-pontine tract, and left AF | None |
| **Female** | None | Genu, body, and splenium of CC |

of the motor cortex, which regulates the execution and control of movements. The motor cortex is mainly divided into five cortical regions in two main brain lobes, including the frontal lobe (primary motor cortex, premotor cortex, and supplementary motor area) and parietal lobe (posterior parietal cortex and primary somatosensory cortex) (Campbell, 1905). The CST, which was associated with a higher Gold-MSI score in our study, is one of the main pyramidal tracts and projects from the motor cortex to lower motor neurons in the spinal cord, which regulates the movements of limbs and trunk (**Figure 5**). Cortico-pontine tracts are bundles that arise from each lobe cortex [e.g., fronto-pontine (associated with higher active engagement in our study), parieto-pontine, and etc.] and terminate in pontine nuclei (**Figure 5**; Rea, 2015). The cortico-pontine tracts allow the coordination of motor functions by communicating with the opposite cerebellum which was associated with the higher active engagement scores of the Gold-MSI in our study. This communication takes places through the MCP which appeared to be associated with higher Gold-MSI scores in our study (Rea, 2015). This result is consistent with previous studies depicting higher cerebellar volume in musicians relative to non-musicias (Hutchinson et al., 2003). The same study also reported a positive correlation between relative cerebellar volume and lifelong intensity of musical practice which represents structural adaptation to long-term motor and cognitive functional demands in the cerebellum. A circuitry model that explains the motor system interconnections mentions that the cortico-basal ganglia-thalamo-cortical loop is a neural circuit system with both inhibitory and excitatory fibers (**Figure 5**; Silkis, 2001): Cortical inputs into the basal ganglia and thalamic inputs into the cortex are excitatory, whereas the basal ganglia outputs to the thalamus are inhibitory. Additionally, the basal ganglia and cerebellum, which their microstructural alterations were associated with higher scores of active engagement subscale in our study, modulate the output of the CST. CST which was also shown to be associated with higher Gold-MSI score. Along with other descending motor neurons and they receive inputs from the motor and somatosensory cortex, brain stem, and spinal cord, and consecutively project back to the motor cortex through thalamus (**Figure 5**).

Our findings show a significant association between the CC with higher Gold-MSI scores in our male participants, and conversely, lower scores in female participants. The CC is the main commissural bundle in brain which interconnects the two contralateral lobes. The AF that was shown to be associated with higher Gold-MSI score in our study, is traditionally known for its critical role in language processing functions music development, interconnects the two main language processing

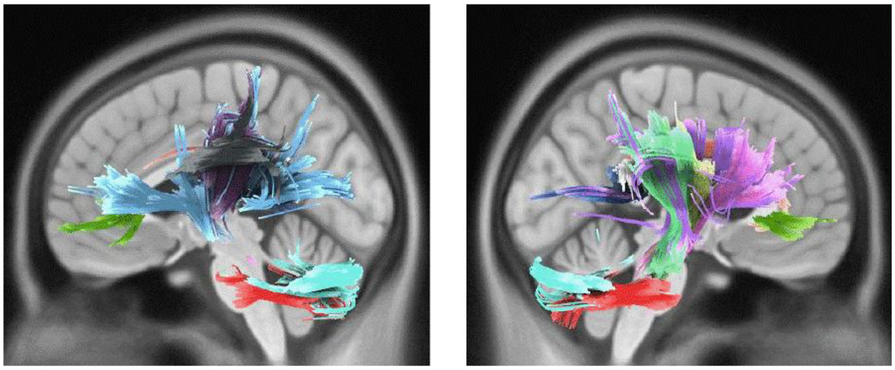


**FIGURE 1 |** White matter pathways with significantly positive association with Active engagement in male participants (FDR = 0.008).

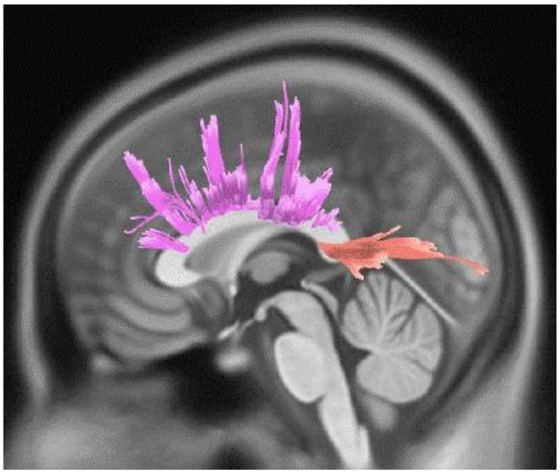


**FIGURE 2 |** White matter pathways with significantly negative association with Active engagement in female participants (FDR = 0.046).

areas Broca's area in the inferior frontal gyrus and Wernicke's area in the posterior-superior temporal gyrus (Eichert et al., 2019). The AF is confirmed to be responsible for some clinical disorders, including conduction aphasia (Acharya and Maani, 2020), tone-deafness (Loui et al., 2009), and stuttering (Cieslak et al., 2015). The cingulum, that was shown to be associated with a higher Gold-MSI score in our study, is the main component of the limbic system. The cingulum is a fiber bundle beneath the cingulate cortex interconnecting the frontal lobe with the temporal and parietal lobes closely above the CC (Bruni and Montemurro, 2009). Ultimately, it seems that fiber bundles with close functional connection with the motor system might play a significant role in both active engagement and musical training subscales of the Gold-MSI. Previous studies have also followed our findings and reported the increased GM volume in the primary motor cortex, premotor cortex, somatosensory areas, parietal cortex, prefrontal cortex, and cerebellum in musicians (Gaser and Schlaug, 2003; Han et al., 2009; Lai et al., 2012; Acer et al., 2018; Oechslin et al., 2018). Apart from similar studies that investigated GM alterations, previous efforts have been taken to address the WM differences mainly between musicians and non-musicians using DTI. For instance, through the investigation of DTI human studies on musical perception, the AF was recognized as a bundle of WM that is responsible for language and music functions (Loui and Schlaug, 2009), and a lower WM integrity in AF was observed in musically tone-deaf individuals (Loui et al., 2009). Besides, autistic children were reported to have a lower FA

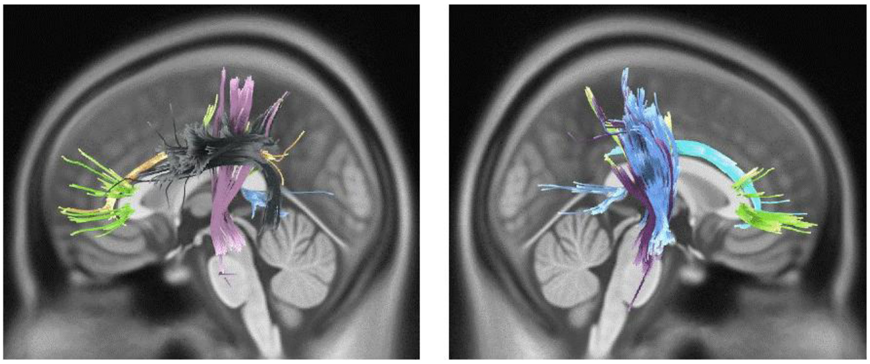


**FIGURE 3 |** White matter pathways with marginally significantly positive association with musical training in male participants (FDR = 0.057).

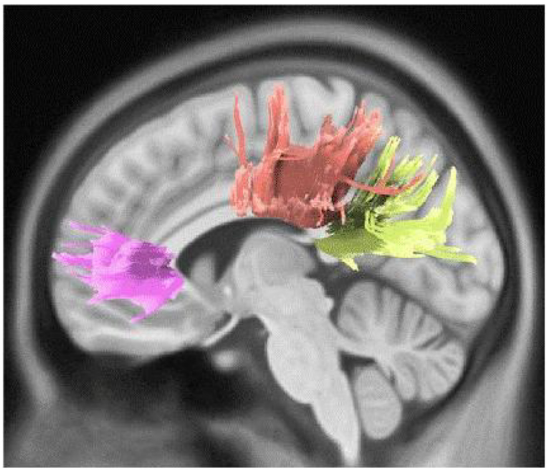


**FIGURE 4 |** White matter pathways with significantly negative association with musical training in female participants (FDR = 0.032).

value in left AF compared to controls (Lai et al., 2012). In a study done in a Mandarin population with amusia, DTI-derived results showed the higher WM integrity in the right posterior AF as well as lower WM integrity in the right anterior AF in amusics (Chen et al., 2018). Moreover, increased FA value in the right AF was observed as the effect of musical training on the WM changes in a short time (20 min, three times per week, for 4 weeks), compared to controls (Moore et al., 2017). Above all, the higher integrity and volume of right AF are reported to have an association with music learning speed and rate (Vaquero et al., 2018). Therefore, on top of the aforementioned findings, our study also suggests the higher WM coherence in AF to have a significant connection with musical sophistication.

In line with our findings, CC changes have been shown to be different among individuals with different musical abilities. Importantly, our results found a marginally significant correlation between musical training and QA values of CC in males. Even so, our findings notably demonstrated that WM integrity changes in CC were significantly different between genders, being positive in males and negative in females. This difference could potentially arise from the impact of "age of onset" of musical training, which may have been different across our groups of participants. Nonetheless, finding from several DTI studies support our results related to CC changes associated with musical training. For instance, seminal study which investigated the differences in musicians and non-musicians' brain structure, found the anterior part of the CC to be larger in musicians (Schlaug et al., 1995). Another study also deduced that intensive and professional musical training since childhood leads to significant changes in WM architecture (Schmithorst and Wilke, 2002). A more recent study examined the association between musical perceptional abilities, assessed by the Profile of Music Perception Skills (PROMS) and WM microstructure using DTI, and reported specific parts of the CC to be significantly involved with musicianship (Rajan et al., 2019). WM plasticity in early trained musicians was also detected as the higher WM integrity in posterior mid-body and isthmus of CC in them (Steele et al., 2013). Another longitudinal study compared the children with more than 2 years of musical training with non-trained controls and reported the highest WM integrity (FA value) in the CC in trained children, particularly in the crossing bundles interconnecting superior frontal, sensory, and motor areas (Habibi et al., 2018). Moreover, professional drummers showed to have higher microstructural diffusion properties in the CC than non-musical controls (Schlaffke et al., 2020). These microstructural changes in the CC among the musicians are probably owing to the bimanual coordination –the ability to simultaneously control multiple movements– that playing a musical instrument demands (Swinnen, 2002; Palomar-García et al., 2017).

Our results also show a positive association between the CST microstructural alterations with higher scores in both GOLD-MSI scores in males, but not in females. The maturation of CST fibers has been shown to correspond to the improvement of fine finger movements (Paus et al., 1999). The differences of diffusion parameters in the CST of professional musicians and non-musicians investigations suggested a lower WM integrity in the CST and plastic changes in WM in professional musicians compared to non-musicians (Imfeld et al., 2009). Besides, a higher WM integrity in the CST, superior longitudinal fasciculus and the CC was reported in dancers comparing to musicians in a previous DTI survey which investigated the WM alteration between dancers and musicians based on their different required motor functions (Giacosa et al., 2016). Moreover, the tract volume and number of streamlines of superior and middle cerebellar peduncles, which were associated with higher Gold-MSI score in our study, were previously reported to be higher in musicians compared to non-musicians (Abdul-Kareem et al., 2011). Another recent study evaluated the GM and WM alterations between professional musicians and non-musicians and found the lower WM coherence in CC, superior longitudinal fasciculus, forceps major and minor, and right AF but higher FA value in right CST as well as increased GM volume in bilateral cerebellar hemispheres, supramarginal and angular gyrus, left parietal lobule, and left temporal lobe in the professional musicians compared to non-musicians control group (Acer et al., 2018). Thus, given the significant role of CST in the motor system, and sufficient pieces of evidence suggesting the significant CST WM differences in musicians, we can conclude that our results showing WM changes in CST and its correlation with both musical sophistication subscales is consistent with formerly existing evidence.

Although this study followed an exploratory approach for investigating WM microstructural alterations links with musicality, we hypothesized that there could be potential sex differences in terms of these alterations and associations with active musical engagement and training. In light of the ample evidence of the gender differences in brain structures using traditional or voxel-based morphometry (Amunts et al., 2000; Nopoulos et al., 2000; Good et al., 2001), and with specific attention to various cognitive domains (Hyde, 2016), we aimed to investigate the potential gender differences in musical sophistication. In line with our hypothesis, we found

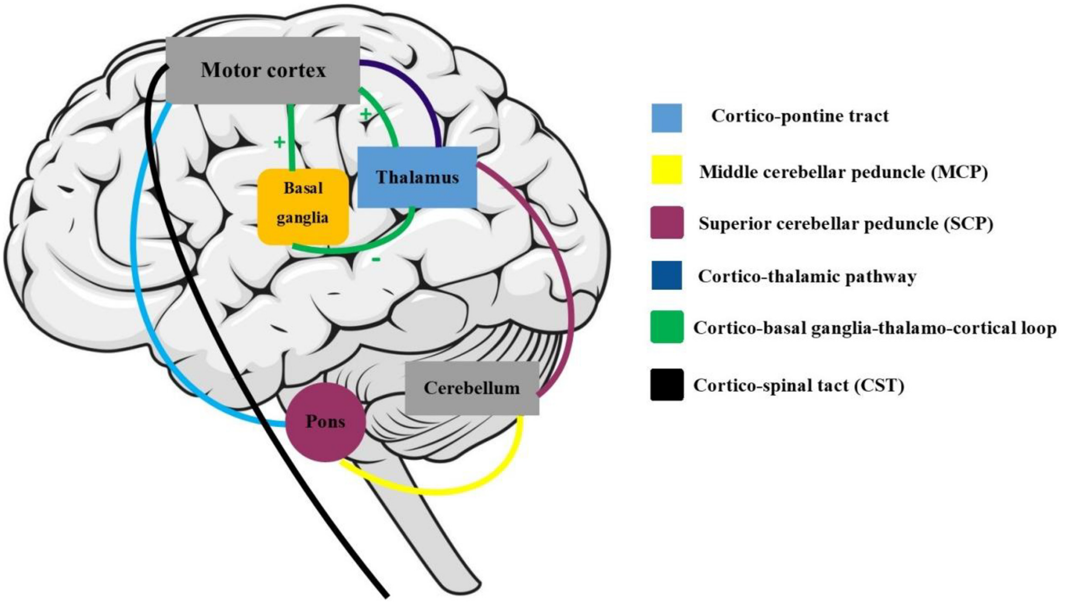



**FIGURE 5 |** Motor and somatosensory system: the maturation of cortico-spinal tract (CST) is responsible for smooth and fine movements of limbs. The basal ganglia and cerebellum modulate the CST functions through the afferent (cortex, brain stem, and spinal cord) and efferent (through the thalamus) fibers. The cortico-basal ganglia-thalamo-cortical circuit is a modulator motor system with inhibitory and excitatory inputs.

WM microstructural differences in males compared to females in both investigated GOLD-MSI subscales. Notably, limited studies have addressed the gender differences in the field of music. For instance, auditory processing was reported to be different in males comparing to females in animal studies (Yoder et al., 2015). Besides, the anterior CC was observed to be significantly larger in male musicians in comparison to non-musicians; however, this structural change was not observed in females equivalent groups (Lee et al., 2003). Furthermore, relatively similar to our findings in both genders, music processing was reported to be conducted bilaterally, with right dominance in females' and males' hemispheres (Koelsch et al., 2003). These pieces of evidence derived from neural data are also consistent with findings from a study that shows female advantages at recognizing familiar melodies stemming from their superiority in declarative memory and behavioral sex differences in higher-level aspects of musical cognition (Miles et al., 2016).

Our findings should be interpreted in light of some limitations. These findings only specify a number of WM tracts with significant association with the variable of interest, yet the causal association remains unclear. Thus, it remains unanswered whether the musical engagement or training leads to differences in WM tracts or vice versa; and if some genetic or non-genetic etiologic factors result in these WM microstructural alterations and make them predisposing factors for musical sophistication. With respect to our findings about sex differences in WM alterations related to musical sophistication, our study could have benefitted from more measurements of musical training for detecting potential confounding factors such as age of onset of musical training or musical training intensity. Additionally, it is noteworthy to mention that dMRI connectometry is a novel WM structural analytic technique to address the WM connectivity of specific bundle fibers, but the functional features of those trajectories are failed to address. Hence, further multimodal imaging modalities are expected to rectify this crucial knowledge gap (Burunat et al., 2015).

This study not only provides a basis for the investigation of differences regarding musical skills and talent, but also serves as a potential for further studies in the etiology, prevention, and management of clinical conditions associated with problems in comprehension and processing of information in the aforementioned brain structures. Future studies could take into account the factors such as age of onset of musical training, musical training intensity, and different aspects of auditory processing and their association with WM microstructural alterations. In addition, our sample was precisely controlled for potential neurological, psychological, and neurocognitive confounding factors; however, future studies may explore ethnical and/or cultural aspects of musical sophistication and their associated neural alterations, which may exhibit contradictory results relative to our findings that took place on a population from a WEIRD society (Hoffman et al., 2011). Finally, it is worth mentioning that we employed two specific subscales of the GOLD-MSI, including active musical engagement and musical training, to explore their correlations with WM microstructural alterations and potential sex differences in this regard; but the above mentioned association with three other subscales of GOLD-MSI (i.e., self-reported perceptual abilities, self-reported singing abilities, and sophisticated emotional engagement with music) could be further investigated to unravel which WM tracts and their connectivity are associated with those aspects of musical sophistication and potential sex differences.

## CONCLUSION

In conclusion, our study is the first dMRI study investigation that explored the brain microstructural alterations related to musical sophistication in a healthy population. Our findings pointed to the notion that WM microstructures with functional connection with motor and somatosensory areas (such as the corpus callosum, corticospinal tract, cingulum, cerebellar peduncle, parieto-pontine tract, and cerebellum), and language processing areas (such as the arcuate fasciculus) have a significant correlation with active musical engagement and training. Although some caution is warranted since this is the first study that investigated sex differences in brain's white matter microstructural alterations related to musical sophistication using a novel approach, significant sex differences were observed indicating the corpus callosum to be the only WM tract associated with musical sophistication in females, while a wide range of WM microstructures were shown to be linked with this ability in males. Our results are consistent with the idea that the coordination between auditory and motor systems is necessary for music performance, particularly musical active engagement and training.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

## ETHICS STATEMENT

The present study was carried out in accordance with the World Medical Association Declaration of Helsinki revised in 1989 and approved by the Ethics Committee of the University of Leipzig (reference number 154/13-ff).

## AUTHOR CONTRIBUTIONS

MA and PR conceived of the presented idea. MA performed the computations and data analysis. M-MM, PR, HS, and ZS wrote the first draft of the manuscript. All authors discussed the results and contributed to the final manuscript.

## REFERENCES


Abdul-Kareem, I. A., Stancak, A., Parkes, L. M., and Sluming, V. (2011). Increased gray matter volume of left pars opercularis in male orchestral musicians correlate positively with years of musical performance. *J. Magn. Resonance Imag.* 33, 24–32. doi: 10.1002/jmri.22391

Acer, N., Bastepe-Gray, S., Sagiroglu, A., Gumus, K. Z., Degirmencioglu, L., Zararsiz, G., et al. (2018). Diffusion tensor and volumetric magnetic resonance imaging findings in the brains of professional musicians. *J. Chem. Neuroanatomy* 88, 33–40. doi: 10.1016/j.jchemneu.2017.11.003

Acharya, A. B., and Maani, C. V. (2020). *Conduction Aphasia. in StatPearls [Internet]*. Treasure Island, FL: StatPearls.

Amunts, K., Jäncke, L., Mohlberg, H., Steinmetz, H., and Zilles, K. (2000). Interhemispheric asymmetry of the human motor cortex related to handedness and gender. *Neuropsychologia* 38, 304–312. doi: 10.1016/s0028-3932(99)00075-5

Babayan, A., Erbey, M., Kumral, D., Reinelt, J. D., Reiter, A. M., Röbbig, J., et al. (2019). A mind-brain-body dataset of MRI, EEG, cognition, emotion, and peripheral physiology in young and old adults. *Sci. Data* 6, 1–21.

Baker, D. J., Ventura, J., Calamia, M., Shanahan, D., and Elliott, E. M. (2020). Examining musical sophistication: a replication and theoretical commentary on the Goldsmiths Musical Sophistication Index. *Musicae Sci.* 24, 411–429. doi: 10.1177/1029864918811879

Boyle, J. D., and Radocy, R. E. (1987). *Measurement and Evaluation of Musical Experiences*. New Yoerk, NY: Schirmer Books.

Brockmeier, S., Fitzgerald, D., Searle, O., Fitzgerald, H., Grasmeder, M., Hilbig, S., et al. (2011). The MuSIC perception test: a novel battery for testing music perception of cochlear implant users. *Cochlear Implants Int.* 12, 10–20. doi: 10.1179/146701010x12677899497236

Bruni, J. E., and Montemurro, D. G. (2009). *Human Neuroanatomy: A Text, Brain Atlas, and Laboratory Dissection Guide*. Oxford: Oxford University Press.

Burkhard, A., Hänggi, J., Elmer, S., and Jäncke, L. (2020). The importance of the fibre tracts connecting the planum temporale in absolute pitch possessors. *NeuroImage* 211:116590. doi: 10.1016/j.neuroimage.2020.116590

Burunat, I., Brattico, E., Puoliväli, T., Ristaniemi, T., Sams, M., and Toiviainen, P. (2015). Action in perception: prominent visuo-motor functional symmetry in musicians during music listening. *PLoS One* 10:e0138238. doi: 10.1371/journal.pone.0138238

Campbell, A. W. (1905). *Histological Studies on the Localisation of Cerebral Function*. Cambridge: Cambridge University Press.

Chen, X., Zhao, Y., Zhong, S., Cui, Z., Li, J., Gong, G., et al. (2018). The lateralized arcuate fasciculus in developmental pitch disorders among mandarin amusics: left for speech and right for music. *Brain Struct. Funct.* 223, 2013–2024.

Cieslak, M., Ingham, R. J., Ingham, J. C., and Grafton, S. T. (2015). Anomalous white matter morphology in adults who stutter. *J. Speech Lang. Hearing Res.* 58, 268–277. doi: 10.1044/2015_jslhr-s-14-0193

Eichert, N., Verhagen, L., Folloni, D., Jbabdi, S., Khrapitchev, A. A., Sibson, N. R., et al. (2019). What is special about the human arcuate fasciculus? Lateralization, projections, and expansion. *Cortex* 118, 107–115. doi: 10.1016/j.cortex.2018.05.005

Elmer, S., Hänggi, J., Meyer, M., and Jäncke, L. (2013). Increased cortical surface area of the left planum temporale in musicians facilitates the categorization of phonetic and temporal speech sounds. *Cortex* 49, 2812–2821. doi: 10.1016/j.cortex.2013.03.007

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

Giacosa, C., Karpati, F. J., Foster, N. E., Penhune, V. B., and Hyde, K. L. (2016). Dance and music training have different effects on white matter diffusivity in sensorimotor pathways. *Neuroimage* 135, 273–286. doi: 10.1016/j.neuroimage.2016.04.048

Good, C. D., Johnsrude, I., Ashburner, J., Henson, R. N., Friston, K. J., and Frackowiak, R. S. (2001). Cerebral asymmetry and the effects of sex and handedness on brain structure: a voxel-based morphometric analysis of 465 normal adult human brains. *Neuroimage* 14, 685–700. doi: 10.1006/nimg.2001.0857

Greenberg, D. M., Müllensiefen, D., Lamb, M. E., and Rentfrow, P. J. (2015). Personality predicts musical sophistication. *J. Res. Pers.* 58, 154–158. doi: 10.1016/j.jrp.2015.06.002

Habibi, A., Damasio, A., Ilari, B., Veiga, R., Joshi, A. A., Leahy, R. M., et al. (2018). Childhood music training induces change in micro and macroscopic brain structure: results from a longitudinal study. *Cereb. Cortex* 28, 4336–4347. doi: 10.1093/cercor/bhx286

Halwani, G. F., Loui, P., Rüber, T., and Schlaug, G. (2011). Effects of practice and experience on the arcuate fasciculus: comparing singers, instrumentalists, and non-musicians. *Front. Psychol.* 2:156. doi: 10.3389/fpsyg.2011.00156

Han, Y., Yang, H., Lv, Y.-T., Zhu, C.-Z., He, Y., 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

Hoffman, M., Gneezy, U., and List, J. A. (2011). Nurture affects gender differences in spatial abilities. *Proc. Natl. Acad. Sci. U.S.A.* 108:14786. doi: 10.1073/pnas.1015182108

Hutchinson, S., Lee, L. H., Gaab, N., and Schlaug, G. (2003). Cerebellar volume of musicians. *Cereb. Cortex* 13, 943–949. doi: 10.1093/cercor/13.9.943

Hyde, J. S. (2016). Sex and cognition: gender and cognitive functions. *Curr. Opin. Neurobiol.* 38, 53–56. doi: 10.1016/j.conb.2016.02.007

Imfeld, A., Oechslin, M. S., Meyer, M., Loenneker, T., and Jancke, L. (2009). White matter plasticity in the corticospinal tract of musicians: a diffusion tensor imaging study. *Neuroimage* 46, 600–607. doi: 10.1016/j.neuroimage.2009.02.025

Ingalhalikar, M., Smith, A., Parker, D., Satterthwaite, T. D., Elliott, M. A., Ruparel, K., et al. (2014). Sex differences in the structural connectome of the human brain. *Proc. Natl. Acad. Sci. U.S.A.* 111:823.

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

Koelsch, S., Maess, B., Grossmann, T., and Friederici, A. D. (2003). Electric brain responses reveal gender differences in music processing. *Neuroreport* 14, 709–713. doi: 10.1097/00001756-200304150-00010

Lai, G., Pantazatos, S. P., Schneider, H., and Hirsch, J. (2012). Neural systems for speech and song in autism. *Brain* 135, 961–975. doi: 10.1093/brain/awr335

Larrouy-Maestri, P., Harrison, P. M., and Müllensiefen, D. (2019). The mistuning perception test: a new measurement instrument. *Behav. Res. Methods* 51, 663–675. doi: 10.3758/s13428-019-01225-1

Law, L. N., and Zentner, M. (2012). Assessing musical abilities objectively: construction and validation of the Profile of Music Perception Skills. *PLoS One* 7:e52508. doi: 10.1371/journal.pone.0052508

Lee, D. J., Chen, Y., and Schlaug, G. (2003). Corpus callosum: musician and gender effects. *Neuroreport* 14, 205–209. doi: 10.1097/00001756-200302100-00009

Leemans, A., Jeurissen, B., Sijbers, J., and Jones, D. (2009). "ExploreDTI: a graphical toolbox for processing, analyzing, and visualizing diffusion MR data," in *Proceedings of the International Society for Magnetic Resonance in Medicine*, (Honolulu).

Levitin, D. J. (2012). What does it mean to be musical? *Neuron* 73, 633–637. doi: 10.1016/j.neuron.2012.01.017

Li, X., Zatorre, R. J., and Du, Y. (2021). The microstructural plasticity of the arcuate fasciculus undergirds improved speech in noise perception in musicians. *Cereb. Cortex* bhab063. [Online ahead of print]. doi: 10.1093/cercor/bhab063

Loui, P., Alsop, D., and Schlaug, G. (2009). Tone deafness: a new disconnection syndrome? *J. Neurosci.* 29, 10215–10220. doi: 10.1523/jneurosci.1701-09.2009

Loui, P., and Schlaug, G. (2009). Investigating musical disorders with diffusion tensor imaging: a comparison of imaging parameters. *Ann. N. Y. Acad. Sci.* 1169:121. doi: 10.1111/j.1749-6632.2009.04781.x

Mehrabinejad, M.-M., Sanjari Moghaddam, H., Mohammadi, E., Hajighadery, A., Sinaeifar, Z., and Aarabi, M. H. (2021). Sex differences in microstructural white matter alterations of mathematics anxiety based on diffusion MRI connectometry. *Neuropsychology* 35:197. doi: 10.1037/neu0000684

Merriam, A. P., and Merriam, V. (1964). *The Anthropology of Music*. Evanston, IL: Northwestern University Press.

Miles, S. A., Miranda, R. A., and Ullman, M. T. (2016). Sex differences in music: a female advantage at recognizing familiar melodies. *Front. Psychol.* 7:278. doi: 10.3389/fpsyg.2016.00278

Moghaddam, H. S., Mehrabinejad, M.-M., Mohebi, F., Hajighadery, A., Maroufi, S. F., Rahimi, R., et al. (2020). Microstructural white matter alterations and personality traits: a diffusion MRI study. *J. Res. Personal.* 88:104010. doi: 10.1016/j.jrp.2020.104010

Møller, C., Garza-Villarreal, E. A., Hansen, N. C., Højlund, A., Bærentsen, K. B., Chakravarty, M. M., et al. (2021). Audiovisual structural connectivity in musicians and non-musicians: a cortical thickness and diffusion tensor imaging study. *Sci. Rep.* 11, 1–14.

Moore, E., Schaefer, R. S., Bastin, M. E., Roberts, N., and Overy, K. (2017). Diffusion tensor MRI tractography reveals increased fractional anisotropy (FA) in arcuate fasciculus following music-cued motor training. *Brain Cogn.* 116, 40–46. doi: 10.1016/j.bandc.2017.05.001

Müllensiefen, D., Gingras, B., Musil, J., and Stewart, L. (2014). The musicality of non-musicians: an index for assessing musical sophistication in the general population. *PLoS One* 9:e89642. doi: 10.1371/journal.pone.0089642

Nopoulos, P., Flaum, M., O'leary, D., and Andreasen, N. C. (2000). Sexual dimorphism in the human brain: evaluation of tissue volume, tissue composition and surface anatomy using magnetic resonance imaging. *Psychiatry Res. Neuroimag.* 98, 1–13. doi: 10.1016/s0925-4927(99)00044-x

Oechslin, M. S., Gschwind, M., and James, C. E. (2018). Tracking training-related plasticity by combining fMRI and DTI: the right hemisphere ventral stream mediates musical syntax processing. *Cereb. Cortex* 28, 1209–1218. doi: 10.1093/cercor/bhx033

Palomar-García, M., 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.

Parsons, L. M. (2001). Exploring the functional neuroanatomy of music performance, perception, and comprehension. *Ann. N. Y. Acad. Sci.* 930, 211–231. doi: 10.1111/j.1749-6632.2001.tb05735.x

Paus, T., Zijdenbos, A., Worsley, K., Collins, D. L., Blumenthal, J., Giedd, J. N., et al. (1999). Structural maturation of neural pathways in children and adolescents: *in vivo* study. *Science* 283, 1908–1911. doi: 10.1126/science.283.5409.1908

Rahmani, F., Sobhani, S., and Aarabi, M. H. (2017). Sequential language learning and language immersion in bilingualism: diffusion MRI connectometry reveals microstructural evidence. *Exp. Brain Res.* 235, 2935–2945. doi: 10.1007/s00221-017-5029-x

Rajan, A., Valla, J. M., Alappatt, J. A., Sharda, M., Shah, A., Ingalhalikar, M., et al. (2019). Wired for musical rhythm? A diffusion MRI-based study of individual differences in music perception. *Brain Struct. Funct.* 224, 1711–1722. doi: 10.1007/s00429-019-01868-y

Rea, P. (2015). *Essential Clinical Anatomy of the Nervous System*. Cambridge, MA: Academic Press.

Rickard, N. S., Chin, T., Coutinho, E., and Scherer, K. R. (2015). "The MUSEBAQ: a comprehensive and modular instrument for assessing musical engagement," in *Proceedings of the 4th International Conference on Music and Emotion (ICME'4)*, (Geneva).

Rüber, T., Lindenberg, R., and Schlaug, G. (2015). Differential adaptation of descending motor tracts in musicians. *Cereb. Cortex* 25, 1490–1498. doi: 10.1093/cercor/bht331

Sato, K., Kirino, E., and Tanaka, S. (2015). A voxel-based morphometry study of the brain of university students majoring in music and nonmusic disciplines. *Behav. Neurol.* 2015:274919.

Schaal, N. K., Bauer, A.-K. R., and Müllensiefen, D. (2014). Der Gold-MSI: replikation und validierung eines fragebogeninstrumentes zur messung musikalischer erfahrenheit anhand einer deutschen stichprobe. *Musicae Sci.* 18, 423–447. doi: 10.1177/1029864914541851

Schlaffke, L., Friedrich, S., Tegenthoff, M., Güntürkün, O., Genç, E., and Ocklenburg, S. (2020). Boom Chack Boom—A multimethod investigation of motor inhibition in professional drummers. *Brain Behav.* 10:e01490.

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., 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

Schmithorst, V. J., and Wilke, M. (2002). Differences in white matter architecture between musicians and non-musicians: a diffusion tensor imaging study. *Neurosci. Lett.* 321, 57–60. doi: 10.1016/s0304-3940(02)00054-x

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

Seashore, C. E., Lewis, D., and Saetveit, J. G. (1956). *Seashore Measures of Musical Talents.* Oxford: Psychological Corp.

Silkis, I. (2001). The cortico-basal ganglia-thalamocortical circuit with synaptic plasticity. II. Mechanism of synergistic modulation of thalamic activity via the direct and indirect pathways through the basal ganglia. *Biosystems* 59, 7–14. doi: 10.1016/s0303-2647(00)00135-0

Sluming, V., Barrick, T., Howard, M., Cezayirli, E., Mayes, A., and Roberts, N. (2002). Voxel-based morphometry reveals increased gray matter density in

Broca's area in male symphony orchestra musicians. *Neuroimage* 17, 1613–1622. doi: 10.1006/nimg.2002.1288

Steele, C. J., Bailey, J. A., Zatorre, R. J., and Penhune, V. B. (2013). Early musical training and white-matter plasticity in the corpus callosum: evidence for a sensitive period. *J. Neurosci.* 33, 1282–1290. doi: 10.1523/jneurosci.3578-12.2013

Swinnen, S. P. (2002). Intermanual coordination: from behavioural principles to neural-network interactions. *Nat. Rev. Neurosci.* 3, 348–359. doi: 10.1038/nrn807

van Vugt, F. T., Hartmann, K., Altenmüller, E., Mohammadi, B., and Margulies, D. S. (2021). The impact of early musical training on striatal functional connectivity. *NeuroImage* 238:118251. doi: 10.1016/j.neuroimage.2021.118251

Vaquero, L., Ramos-Escobar, N., François, C., Penhune, V., and Rodríguez-Fornells, A. (2018). White-matter structural connectivity predicts short-term melody and rhythm learning in non-musicians. *NeuroImage* 181, 252–262. doi: 10.1016/j.neuroimage.2018.06.054

Wan, C. Y., Zheng, X., Marchina, S., Norton, A., and Schlaug, G. (2014). Intensive therapy induces contralateral white matter changes in chronic stroke patients with Broca's aphasia. *Brain Lang.* 136, 1–7. doi: 10.1016/j.bandl.2014.03.011

Wu, J., Zhang, J., Ding, X., Li, R., and Zhou, C. (2013). The effects of music on brain functional networks: a network analysis. *Neuroscience* 250, 49–59. doi: 10.1016/j.neuroscience.2013.06.021

Yeh, F.-C., Badre, D., and Verstynen, T. (2016). Connectometry: a statistical approach harnessing the analytical potential of the local connectome. *Neuroimage* 125, 162–171. doi: 10.1016/j.neuroimage.2015.10.053

Yeh, F.-C., Verstynen, T. D., Wang, Y., Fernández-Miranda, J. C., and Tseng, W.-Y. I. (2013). Deterministic diffusion fiber tracking improved by quantitative anisotropy. *PLoS One* 8:e80713. doi: 10.1371/journal.pone.0080713

Yeh, F.-C., Wedeen, V. J., and Tseng, W.-Y. I. (2011). Estimation of fiber orientation and spin density distribution by diffusion deconvolution. *Neuroimage* 55, 1054–1062. doi: 10.1016/j.neuroimage.2010.11.087

Yoder, K. M., Phan, M. L., Lu, K., and Vicario, D. S. (2015). He hears, she hears: are there sex differences in auditory processing? *Dev. Neurobiol.* 75, 302–314. doi: 10.1002/dneu.22231

**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.

frontiers in Psychology

ORIGINAL RESEARCH
published: 21 September 2021
doi: 10.3389/fpsyg.2021.717842

# The Communication of Timbral Intentions Between Pianists and Listeners and Its Dependence on Auditory-Visual Conditions

Shen Li[1,2], Renee Timmers[3] and Weijun Wang[1,2]*

[1] School of Psychology, Central China Normal University, Wuhan, China, [2] Key Laboratory of Adolescent Cyberpsychology and Behavior, Ministry of Education, Wuhan, China, [3] Department of Music, The University of Sheffield, Sheffield, United Kingdom

The perceptual experiment reported in this article explored whether the communication of five pairs of timbral intentions (bright/dark, heavy/light, round/sharp, tense/relaxed, and dry/velvety) between pianists and listeners is reliable and the extent to which performers' gestures provide visual cues that influence the perceived timbre. Three pianists played three musical excerpts with 10 different timbral intentions ($3 \times 10 = 30$ music stimuli) and 21 piano students were asked to rate perceived timbral qualities on both unipolar Likert scales and non-verbal sensory scales (shape, size, and brightness) under three modes (vision-alone, audio-alone, and audio-visual). The results revealed that nine of the timbral intentions were reliably communicated between the pianists and the listeners, except for the dark timbre. The communication of tense and relaxed timbres was improved by the visual conditions regardless of who is performing; for the rest, we found the individuality in each pianist's preference for using visual cues. The results also revealed a strong cross-modal association between timbre and shape. This study implies that the communication of piano timbre is not based on acoustic cues alone but relates to a shared understanding of sensorimotor experiences between the performers and the listeners.

**Keywords: music communication, timbre perception, performer-listener chain, auditory-visual condition, cross-modal correspondence**

OPEN ACCESS

**Edited by:**
Gary Edward McPherson,
The University of Melbourne, Australia

**Reviewed by:**
Laura Bishop,
University of Oslo, Norway
Zachary Wallmark,
University of Oregon, United States

***Correspondence:**
Weijun Wang
wangwj@mail.ccnu.edu.cn

**Specialty section:**
This article was submitted to Performance Science, a section of the journal Frontiers in Psychology

**Received:** 31 May 2021
**Accepted:** 18 August 2021
**Published:** 21 September 2021

**Citation:**
Li S, Timmers R and Wang W (2021) The Communication of Timbral Intentions Between Pianists and Listeners and Its Dependence on Auditory-Visual Conditions. Front. Psychol. 12:717842. doi: 10.3389/fpsyg.2021.717842

## INTRODUCTION

Compared to the communication of emotions or expressiveness in musical performance, research conducted into the communication of timbral intention between performers and listeners is rare. From the performers' perspective, they intend to produce explicit timbres in their performances, which may lead to their self-satisfaction and sense of achievement (Holmes, 2011), relate to their holistic perception of expressive elements and music structure (Li et al., 2020), or be relevant to emotional expression (Juslin, 2000). In addition, piano teachers and students usually work on explicit timbral intentions in piano lessons using metaphors, gestures and modeling etc. to improve their communication of timbre-related performance goals (Li and Timmers, 2021). However, whether pianists can communicate timbral intentions reliably to the listeners under different auditory-visual presentation conditions remains unknown. An interview study (Li et al., 2020) on the conceptualization of piano timbre revealed pianists' extensive utilization of timbral intentions in piano performance; the findings also suggested that a pianist's concept of timbre is enriched by embodied experience, such as bodily preparations, indicating the relevance of visual cues. As a follow-up study, a perceptual experiment was conducted to examine the accuracy of communication of timbral intentions to listeners and its dependence on the visual and aural components of musical performance. It aims to explore the core question of how piano timbre is communicated, by using several sub-questions: Do pianists communicate timbral intentions to the audience and is the communication reliable? And what is the relevance of the auditory and/or visual components of a musical performance in timbre communication?

## THE RESEARCH BACKGROUND OF PIANO TIMBRE

To better understand the research background and research purpose, it is necessary to clarify the meaning and scope of timbre specified in this study, which we distinguished as "micro-perspective" and "macro-perspective." In general, timbre studies have usually adopted a macro-perspective, which considers the differences in timbre produced from various sound objects (i.e., different sound sources). The "macro-perspective" focuses on the characterization of timbre of specific instruments or instrument groups. This could also concern differences within an instrument group (i.e., the timbre of piano A is different from that of piano B). An important aim may be to uncover the timbre space through which listeners categorize or distinguish instrument timbres (e.g., timbre intervals: McAdams, 2013). However, this study investigates piano timbre from a "micro-perspective"—which focuses on the timbral nuances produced from one instrument, from the point of view that a specific instrument can still have a variety of timbres depending on how the instrument is played (e.g., piano A has contrasting timbres when the key is either gently pressed or quickly struck).

One difficulty in the study of piano timbre lies in the interweaving of timbre and other performance parameters (e.g., intensity, articulation, and tempo) in the pianists' conceptions of piano timbre. It might be difficult for pianists to believe that, by controlling performed intensity, different ways of touching the keys have little effects on produced timbre—what they've changed are the attack noises (i.e., finger-key noise, key-keyframe noise; Goebl et al., 2014). Timbre may also be phrased as tone quality, tone color in the discourses of pianists.[1] Sometimes we found these terminologies were used interchangeably in the writings and discourses of pianists, with the combined effect of timbre and other musical features being mentioned. Ortmann (1935) referred to "tone quality" and suggested that the perception of tone quality is subjective and results from our unified reaction to three variants: pitch, intensity, and duration. Bernays and Traube (2014) put forward the notion of *composite timbre*, referring to the complexity of piano timbre interwoven with other performance parameters when considering piano timbre in a musical and polyphonic context. Their study selected five verbal descriptors (dry, bright, round, velvety, and dark) and the pianists conceptualized and performed the given music pieces in accordance with these timbral nuances. For these reasons, this study investigates the beliefs and utilizations of piano timbre within the conventions of the pianistic community, rather than from an acoustic perspective. In other words, we will examine timbral communication in the context that pianists conceptualize their entire performance as expressing a timbral feature (i.e., mental conception: Kochevitsky, 1967) and rely on the coordination and adjustment of other musical parameters to align with the timbral intention.

## MUSIC COMMUNICATION BETWEEN PERFORMER AND LISTENERS

### The Role of Visual Information

Previous studies that investigated the role of visual information provided by the performer in communicating to the listeners mainly consider the communication of expressivity (Davidson, 1993, 1995; Broughton and Stevens, 2009) and emotional intentions (Dahl and Friberg, 2007). In Davidson's study (1993), a pianist was instructed to play a piece of music in three differently expressive ways (deadpan, projected, and exaggerated) and listeners were asked to rate the *expressivity* in the condition of seeing only, hearing only, or both seeing and hearing. This study found that once the pianist was playing with a certain degree of expressivity (i.e., in a projected and exaggerated manner), listeners could only differentiate between different degrees when the visual information was present and not in the audio-only presentation. Employing a similar experimental paradigm, Broughton and Stevens (2009) verified this result in the communication of expressivity in a marimba performance in which the audio-visual condition helped the listeners' differentiation of performances played with different expressive intentions compared to the audio-only condition. These studies suggested that the bodily movements and gestures of the performers provide identifiable information through which listeners can detect musical expressivity.

Performers express and communicate emotions to audiences [see a review by Juslin and Timmers (2010)] and visual information plays a role in the performer-audience communication chain (Camurri et al., 2003; Timmers et al., 2006; Dahl and Friberg, 2007) and for a duo partner (Wöllner, 2020). Dahl and Friberg (2007) examined whether the specific emotions (happy, sad, angry, and fear) expressed by the performer could be recognized by listeners using visual information only. This study confirmed that listeners (not necessarily musically trained) could easily recognize emotions such as happy and sad and were roughly accurate with anger, but failed to perceive fear. In a dancing-related study, Camurri et al. (2003) indicated that listeners can successfully detect a dancer's emotional intentions (e.g., joy, fear, grief, and anger) through movement cues only. These studies implied that even without the aural cues, visual cues are informative for listeners to recognize the emotional intentions of performers.

However, to the author's current knowledge, there is no existing research that investigates the communication of timbre in music performance and considers the visual effects of performers' gestures in this process. A possible exception may

[1] See Li and Timmers (2021) and Li et al. (2020) for a detailed explanation of the similarities and differences in these terminologies.

be a research project by Wapnick et al. (2004) who examined the visual perception of tone quality. The results indicated that the ratings of perceived tone quality together with the other five performance items (note accuracy, rhythmic accuracy, expressivity, adherence to style, and overall impression) were higher in the audio-visual presentation mode than the audio-only presentation. The present study aims to examine whether the communication of timbral intention between the pianist and the listeners is reliable and to what extent visual information influences judgment.

## Experimental Paradigm

### Changing Audio-Video Presentation Modes

The first common approach in empirical studies of visual communication in music performance is to: (a) instruct the performer to play the same piece in a different manner (expressive vs. non-expressive, or with different emotional intentions) and (b) vary the mode of audio-visual presentation of the recordings and investigate its influence on the listeners' evaluation of the performance. Relevant research has been conducted to investigate the role of vision in the evaluation of emotional engagement (Timmers et al., 2006), performance quality (Wapnick et al., 2000, 2004), and expressive intention (Davidson, 1993, 1995). Similarly, visual information from a singer's facial expression (e.g., eyebrow raised or not, eyes widened or not, lip movement) affected a listener's judgment of the emotional valence of a music piece (Thompson et al., 2005).

### Employing Congruent/Incongruent Pairings

Another approach employed by researchers is to make congruent or incongruent pairings of audio-video stimuli and examine their impact on listeners' responses. This approach helps to understand the impact of auditory-visual integration on participants' processing of new music stimuli, as the "McGurk" effect demonstrated in the studies of speech communication (McGurk and MacDonald, 1976). Relevant research has been conducted to investigate the extent to which vision influences the processing of tone duration (Schutz and Lipscomb, 2007), pitch interval (Thompson et al., 2005; Thompson and Russo, 2007), and timbre (Saldaña and Rosenblum, 1993). For example, Thompson et al. (2005, Experiment 3) examined the influence of (in)congruent pairings of singers' facial expressions with singing clips on the perception of interval size and found that seeing an incongruent pairing (e.g., a small melodic interval accompanied by images of singing a larger interval) resulted in ratings of a larger interval than either of the original pairings. The perceptual experiment conducted by Schutz and Lipscomb (2007) demonstrated that the visual perception of stroke action (seeing either longer or shorter physical gestures) in a marimba performance influences listeners' perception of tone duration. In addition, Saldaña and Rosenblum (1993) examined the perception of cello timbre influenced by the visual information of stroke action (plucking/bowing) in cello performances. Their studies found that watching a video performance enabled the listeners to distinguish between plucking and bowing and that seeing a bowing movement led to higher ratings of bowing timbre and seeing a plucking action resulted in a larger plucking timbre response.

The above studies suggest that we could either manipulate the presentation mode of one stimulus, or combine two stimuli that convey different timbres. This study will use the first approach (changing aural-visual presentation modes) to examine the visual communication of timbral intentions between pianists and listeners, due to the ease of operation and the study purpose. Playing with different timbral intentions leads to a variety of performance parameters e.g., timing and performance speed, which result in the difficulty of synchronizing audio with unmatched video in the "artificial" audio-visual stimuli (i.e., incongruent pairing). Instead, the second approach will be more effective in examining the perception of single piano tones or chords, but the focus of this research is the polyphonic musical background.

There are increasing numbers of scholars (Dahl and Friberg, 2007; Behne and Wöllner, 2011) investigating in detail which visual aspects and which specific regions of bodily movements have an impact on listeners' music perception, instead of looking performers' bodily movements as an integrated part. For example, ancillary gestures—those that are not required for producing sounds (Wanderley et al., 2005; Wanderley and Vines, 2006), have been found to have an impact on audiences' perception of musical tention and phrasing (Vines et al., 2006). Sound-producing gestures—those movements that are effectively produce sounds (e.g., the hands of a marimba player, or the lips of a singer; Jensenius et al., 2010), can influence listeners' judgement of musical notes at a perceptual level (see section Employing Congruent/Incongruent Pairings). Dahl and Friberg (2007) tested the influence of viewing condition (full, nohands, nohead, and only-head) on the detection of emotions. They found that happy emotion was better detected when seeing upper body movement while sad emotion was better recognized when the head movement information was available. Seeing the whole body led to the highest accuracy of recognized emotions. These studies suggested that we could combine two angles in the video recordings—one focused on the movement of the upper body and the head, and the other one focused on the sound-producing gestures (i.e., finger movements).

## AN EMBODIED PERSPECTIVE ON THE COMMUNICATION OF TIMBRAL INTENTIONS

### From an Information-Processing Model to an Embodied Model

The rationale of communicating a timbral intention from the performer to the listeners can be explained by several well-known models, for instance, the information-processing model (Kendall and Carterette, 1990) and the lens model (Juslin, 1998, 2000; Juslin and Lindström, 2010). According to Kendall and Carterette (1990), the information-processing model of music communication addresses how the composer and performer convey their expressive intentions and how the listener perceives the intended message from the composer/performer *via* message recoding. The composer's encoding process goes from ideation to notation; the performer then recodes the notation into acoustic signals and, finally, these acoustic signals are recognized by listeners and result in ideation. In this context, timbral intentions work as a message that the performer intends to communicate to their audience and the encoding and recoding of this timbral message relies on a successful recognition of acoustic cues. For acoustic cues, Juslin and Lindström (2010) adapted the lens model from seminal work Brunswik's(1956) and applied it to the musical context. The modified lens model explains the utilization of cues (i.e., tempo, loudness, articulation etc.) in the communicative process of specific emotions between the performer and the listeners, which suggests that a successful communicative process happens when the performer's cue utilization matches that of the listeners. Both models have acknowledged the feasibility of communicating implicit messages (i.e., emotions) within the musical communication process; however, these models seem to regard musical communication purely as a sonic art, as can be seen in Juslin and Lindström's model of acoustic cues as a lens of musical expression, as well as Kendall and Carterette's model which regards the acoustic signal as transmitting the composer/performer's thoughts to the listener's ideation.

The theory of embodied cognition has prompted the understanding of listeners' musical experience, as revealed by the following research. For example, Davidson and Correia (2002) highlighted the role of bodily movements in shaping the listeners' perception of musical expressiveness. More importantly, the common knowledge of bodily experience i.e., balance, scale, force, and cycles (Johnson, 1987) forms the common ground on which performers and audiences can communicate and appreciate the musical work (Davidson and Correia, 2002). Leman (2008) extended the previous research and proposed a model of embodied music communication based on the decoding and encoding of motor control. The performer realizes the musical goal/idea through corporeal articulations, while the instrument is the mediation which may form a closed loop with the performer with haptic, sonic and visual feedback whilst also transmitting the sonic and visual energy to the listener. Listeners can make sense of the communicative process through a mirror process i.e., a corporeal and cerebral understanding of the intended actions. Leman's (2008) model suggests multimodal sensing of music communication and extends the understanding of musical communication from the decoding/encoding of sonic forms to include those of motor control.

It is worth noting that listeners' embodied listening experience does not merely incorporate the visual perception of musical performance, it could involve more covert imitation of sound-producing actions. Godøy (2006) related the notion of gestural-sonorous object to musical sounds, referring to the phenomenon that humans would mentally trace sound as a continuous process (i.e., the onsets, contours, textures and envelops involving hands, fingers, arms etc.) in the perception or imagination of musical sounds. Cox (2016) proposed the mimetic hypothesis and paid attention to the covert mimetic behavior (MMI, mimetic motor imagery) besides the overt imitation of performance actions (MMA, mimetic motor action). He suggested that, MMI plays important roles in the comprehension of musical sounds, for instance the instrumental timbre—despite the limitations of human voice to overtly imitate other instrumental timbres, "what matters is the attempt to emulate the sound, to feel something of what it would be like to make such sounds, and to thereby feel something of what it would be like to be an entity capable of making such sounds" (p. 32).

## Sensorimotor Perceptions in Timbral Communication

The embodied model of music communication implies that timbral communication may be associated with sensorimotor perceptions, as indicated by several studies (De Poli et al., 1998, 2017; Baraldi et al., 2006). For example, De Poli et al. (2017) claimed that sensorimotor expressivity is the part of musical expressivity that is not covered by musical and emotional expressivity and that it reflects certain cross-modal correspondence (CMC) features. They further developed an FEI (friction, elasticity, and inertia) metaphor that characterizes the sensorimotor expressivity in a kinetic-energy two-dimensional space (De Poli et al., 2009) and demonstrates that participants can describe perceived expressivity *via* FEI metaphors (e.g., hard-soft, light-heavy categories).

Timbre communication may also involve CMC features (Wallmark and Kendall, 2018; Wallmark, 2019a,b). For example, timbre metaphors such as bright/dark, rough/smooth reflected cross-modal correspondence with vision and touch (Wallmark, 2019b). Timbre metaphors in CMC categories encompass "an embodied conceptual transfer process by which an auditory target domain (timbre) is understood in reference to a non-auditory source domain (vision, touch, taste, and smell)" (2019a, p. 594). Therefore, we predict that the listeners in our study will describe perceived timbral qualities in piano performances by referencing multimodal sensations.

## Communicating Timbre: Emotion and Musical Expressivity

In the field of music emotion studies, timbre is closely related to the perceived emotions of listeners (Gabrielsson and Juslin, 1996; Balkwill and Thompson, 1999; Hailstone et al., 2009) as well as the experienced emotions of performers during music performance (Holmes, 2011; Van Zijl and Sloboda, 2011). For example, the selection of instrumental timbre effectively influences listeners' judgment of emotion conveyed by the music (Balkwill and Thompson, 1999). With a systematic control on other performance parameters (loudness, tempo, and melody), there is still a robust effect induced by timbral feature (e.g., spectral content, attack) on the ratings of listeners' perceived emotion (Hailstone et al., 2009). Juslin and Timmers (2010) suggested that: a bright timbre is used for expressing happiness; a soft timbre is suitable for expressing tenderness; a dull timbre is good for expressing sadness; a sharp timbre is appropriate for expressing anger. From a performer's perspective, producing timbral nuances is not only about playing techniques, but is inseparable from musicians' own emotions experienced during music practice/performance. Holmes (2011) suggested that the employment of specific timbre is the key element of musicians' internal motivation to make a successful performance during which process their felt emotions are mainly feelings of satisfaction and elation. Van Zijl and Sloboda (2011) found that focusing on 'sound/tone color' is one of the key elements of the inner technique toward an expressive music interpretation, and the process of focusing helped them to bring their own emotions in line with the musical emotions.

The above research seems to suggest that the expression of timbral nuances, driven by an affective intention of the performer, appears to be an accessory of emotional communication in musical performance. However, this study would rather consider timbral communication not under the umbrella of emotional communication, but within the framework of musical expressivity (as shown in **Figure 1**). Our previous research (Li and Timmers, 2021) found that playing a short musical phrase or an entire section with specific timbral intention without indicating any relevance to affective intention is naturalistic and feasible. The study results revealed that the piano teachers and students usually worked on an entire musical phrase with a consistent timbral intention (e.g., "lute-like timbre," "timbre of princess and earl") and employed multimodal communication strategies (e.g., modeling, verbal explanation, physical touch, gesture, etc.) to achieve these timbre-related goals. According to Bernays and Traube (2014), verbal instruction of basic emotions and timbral intentions are not comparable, and the vocabularies of describing piano timbre are consensual and meaningful to pianists (Faure, 2000; Bellemare and Traube, 2005). The study of timbre semantics (Wallmark, 2019a) suggests that the description of timbre impression can include many possibilities beside emotion, including acoustics, cross-modal correspondence (CMC), mimesis, action, etc.

Timbre, as an important musical message which conveys musical expressivity between performers and listeners, has received scholarly attention (Juslin and Laukka, 2003; Barthet et al., 2010). Barthet et al. (2010) found that changes in timbral features across expressive levels in a clarinet performance did not happen at every note but were specific to some notes or groups of notes or specific musical passages. When cellists were asked to make either expressive performance or physically constrained performance (Rozé et al., 2017), the timbre was modified in the constrained postural condition. The above studies imply that the communication of timbral intention in piano performance might involve not only affective purposes but working as an important musical message when conveying musical expressivity from performers to listeners. As mentioned before, sensorimotor expressivity is the core component of the embodied communication of musical expressivity.

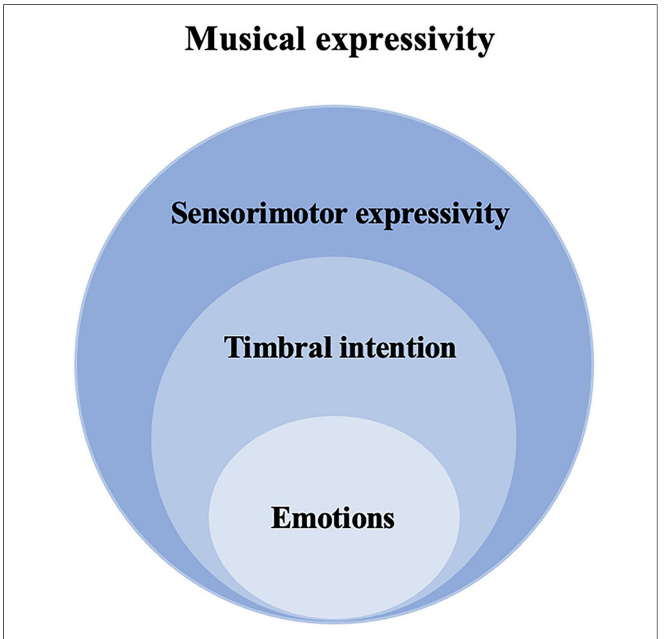



**FIGURE 1** | Logic relationship between timbre, emotion, and musical expressivity.

## The Current Research

This study aims to examine the communication of piano timbre between pianists and listeners and the influence of audio-visual presentation modes on the communication outcome. The focus on piano timbre originates from the touch-tone debate in piano performance, corresponds the trend of increasing attention on semantic studies of piano timbre (Bellemare and Traube, 2005; Bernays, 2013; Kojucharov and Rodà, 2015), and sheds light on the understanding of the relationship between timbre and musical expressivity (Barthet et al., 2010; Ystad et al., 2019). This research expects that, like expressivity and emotions, pianists can express different timbral intentions in their performances and listeners can detect those intentions from either the variations in sound or the visual cues in performative gestures (H1). We also expect that the listeners' ratings of perceived timbre in response to performances with contrasting timbres will be significantly different (H2). Concerning the visual influence, we expect that audio-visual presentation modes of the music performances will influence the communication of timbral intentions (H3).

Hypothesis 1 (H1): the communication of timbral intentions is reliable and successful in the performer-listener communication chain.

Hypothesis 2 (H2): the main effect of heard timbre on timbre ratings: The listeners can successfully differentiate between performances with contrasting timbral intentions.

Hypothesis 3 (H3): the interaction effect of AV condition and heard timbre: the differentiation between contrasting timbral intentions will be influenced by the auditory-visual presentation modes.

We also developed 10 timbre descriptors (see **Table 1** below) as the perceptual scales for participants in the current perceptual experiment. Bright, dark, round, velvety, and dry were selected from the study of Bernays and Traube (2014),[2] whilst the

[2]The validity of using the five timbre descriptors as performance instructions has been tested in their study with sound analysis and measurements of performance strategies (e.g. attack depth, speed, decay, etc.) by five pianists. Their study

**TABLE 1 |** Ten timbre descriptors used in the perceptual experiment.

| | Timbre descriptors | |
|---|---|---|
| Pair 1 | Bright* | Dark* |
| Pair 2 | Round* | Sharp |
| Pair 3 | Light | Heavy |
| Pair 4 | Relaxed | Tensed |
| Pair 5 | Velvety* | Dry* |

*refers to the timbre descriptors borrowed from Bernays and Traube's (2014) study.*

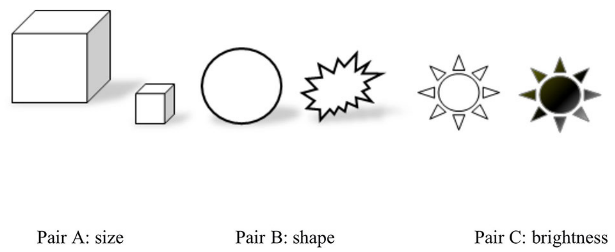



**FIGURE 2 |** The contrasting images employed in this study to examine cross-modal correspondence (CMC) related to piano timbre perception.

others are either their antonyms (sharp) or new descriptors in pairs (heavy/light, relaxed/tense). These 10 timbre metaphors aim to examine the multidimensionality of timbre perception. The heavy and light descriptors represent a kinetic experience associated with piano timbre perception; relaxed and tense relate to the muscular sensations when responding to piano timbre. It is noteworthy that the contrast between "velvety" and "dry" was not noticeable; however, they described the characteristics of an object's surface and represent different types of tactile feelings.

In addition to the 10 metaphors, we employed non-verbal perceptual scales (i.e., image choice) to directly test CMC features with musical timbre i.e., three pairs of images that show contrasting brightness, shape, and size (see **Figure 2**). Timbre-shape and timbre-brightness associations have been found: a soft timbre (e.g., piano) is strongly associated with round shapes together with blue, green, or lighter grayscales, whereas a harsh timbre (e.g., crashing cymbals) is associated with angular shapes together with red, yellow, or darker grayscales (Adeli et al., 2014). Additionally, timbre can be associated with visual textures that vary in terms of sharpness, compactness, and sensory dissonance (Giannakis, 2006). Therefore, this study proposes

> Hypothesis 4: the timbre expressed by the performer will significantly influence the listeners' choice between contrasting images.

indicated that there was no difficulty in pianists interpreting the same musical piece with distinct timbral intentions, and subtle changes in expressing different timbres were found by sound and keyboard motion analysis.

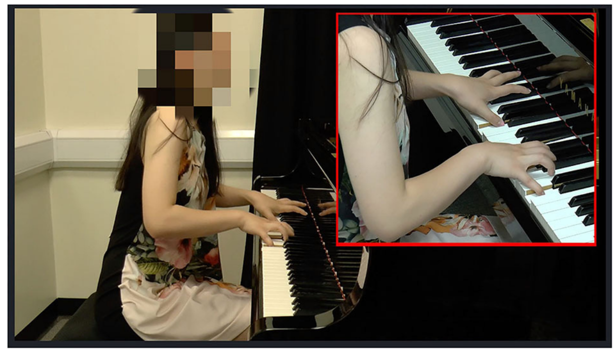


**FIGURE 3 |** The video presentation with two camera angles.

# EXPERIMENT

## Participants

*Performers.* Three female Chinese pianists (age 23–27) from the Department of Music at the University of Sheffield were asked to give performances to create the stimuli used in the listening experiment. They were all majoring in piano performance studies (one Ph.D. student, two Master's students) and regularly performed classical, baroque and contemporary music as part of their performance repertoire.

*Listeners.* Twenty-one music students (19 females and two males; Mean age = 21.89, SD = 2.03) from the music department at the University of Henan in China participated in the listening experiment. Fourteen of them were undergraduates and seven of them were postgraduates. All participants were majoring in piano performance studies. Based on large effect sizes found in the three-way interaction (Partial Eta Squared 0.15 is considered large, Cohen, 1988), we conducted a post hoc power analysis using G power 3.1 (Faul et al., 2007), that determined we had 99% power to detect large effects in a three-way interaction with 5% Type I error rate.

## Stimuli

In total, there were three musical pieces x 3 AV presentations x 10 instructed timbres totaling 90 music excerpts. The three musical pieces was performed by three different pianists (piece 1 pianist 1, piece 2 pianist 2, piece 3 pianist 3), and each musical piece was played with 10 types of timbral intentions, which were audio- and video-recorded. The video recordings were captured from two angles: a global view that showed the sitting pianists from the side (i.e., the viewing of ancillary gestures), and a more local view that was focused on the hands and finger (i.e., the viewing of sound-producing gestures; see **Figure 3**).

In the preparation of experimental stimuli, the three pianists practiced and recorded all three musical pieces with 10 timbres, but only one musical piece played by each pianist was selected for inclusion in the listening experiment to avoid boredom and the effect of repetitive listening on listeners' responses (Morimoto and Timmers, 2012). Otherwise, the listeners had to listen to the same musical piece 30 times (three performers playing one musical piece repetitively using 10 different timbres). When matching the performer and the piece, we considered the criteria of fewer mistakes and more expression in the performances of each pianist to optimize the quality of the musical material presented to listeners.

The three pianists' performances were recorded on a Yamaha grand piano (Disklavier Pro S6) in the Sound House Studio of the Department of Music at the University of Sheffield. Panasonic HC-V770k HD Camcorder and Tascam DR-05 Audio Recorder Kit were used to obtain higher recording quality.

Three pianists were given the pieces [selected from Bernays and Traube (2014), see **Appendix 1**] 1 month before the recording and each of them experienced a recording trial before the actual recording to enable them to get used to the cameras, the procedure, and the piano. The practice time also allowed them to arrive at the appropriate performance manner for each special timbre. They were told to play the piece with a consistent timbral intention and performance instructions were explained in both written (see **Appendix 3**) and oral forms. To ensure that pianists played in response to timbral intentions rather than emotional intentions, the objectives of recording their performances were clearly explained to them. The three pianists were told that their performances would be used as stimuli in a perceptual experiment and the listeners' focus would be perceived timbral quality. All three pianists have participated in the previous interview study on piano timbre, which gave them and the researchers a common understanding of the research objectives. The original musical pieces borrowed from Bernays' study do not have any expression, dynamic, articulation, phrasing, and accent markings, which gave the performers scope to vary performance parameters to achieve a desired tone quality.

## Procedure

The experiment tested participants individually by presenting the recordings on a computer monitor. After the stimuli presented, 21 participants were asked to rate the presence of the 10 timbres on a 1–9 unipolar Likert scales (task 1) and to choose from three pairs of non-verbal sensory scales (task 2). We considered using unipolar scales instead of categorical responses, allowing for ambiguous and relative open responses (e.g., both bright and light might be perceived to certain extent, or none of the timbres might be strongly perceived). Each music excerpt lasted around 15 s and the participants were required to complete two tasks for each excerpt intuitively.

Participants gave their responses on a paper-based response sheet (see **Appendix 2**, two tasks in responses to each excerpt). Response sheet was written in both English and Chinese. Although there might be potential differences between Chinese and English timbre descriptors (Namba et al., 1991), the participants reported no difficulty in understanding 10 timbre descriptors which were also commonly used in Chinese context. The ordering of musical excerpts was randomized to ensure that the same timbre could not occur more than twice consecutively, and we prepared two different random orderings for counterbalancing. The monitor had a 15″ screen and was placed in front of the participants within a comfortable viewing distance (80–100 cm). The experiment contains two practice trials and 90 experiment trials. Recordings were presented over headphones at a comfortable level. Participants were encouraged to respond intuitively after experiencing the stimuli for the first time, but they could play the recording repeatedly if they asked, or they were hesitant about the answer. The majority of them played the stimuli just once and responded quickly, but two out of 21 participants took half hour longer than others due to unsure answers and replay of stimuli. The duration of the entire experiment normally took between 45 min to an hour. In the middle of the experiment, there was a short break.

# RESULTS

To examine how well the performers communicated the timbre to the participants, we calculated the percentage of correct answers in the ratings (perceived timbre) for each target timbre (heard timbre). A one-sample t-test was conducted to test whether the number of correct answers was above the chance level (10%). Timbre ratings were then entered into a three-way ANOVA with repeated measures for heard timbre (two levels: target, opposite), AV condition (three levels: AO, AV, and VO), and pianist/piece (three levels: P1, P2, and P3). To simplify the analysis and assure statistical power, the ability to communicate a particular timbre was tested per timbre pair rather than by comparing across all 10 timbre levels individually. Finally, PCA was conducted to explore which of the 10 timbres capture the same variance, with the inclusion of acoustic and visual information for further comparison.

## Percent Correct Score

We calculated the percentage of correct answers for each target timbre by re-coding data as either 1 (correct) or 0 (incorrect). The percent correct scores were then further divided into an absolute percentage and a relative percentage.[3] For example, in the perception of the dark timbre, only 15.34% of listeners gave the highest ratings for darkness when exposed to an excerpt performed with *dark*[4] timbre, and the score was even below chance level when seen without presentation of audio (VO: 9.52%). However, using relative percentage correct, the success of communication of darkness is above chance (19.15%): listeners' rating of darkness was relatively high for this target timbre compared to when asked to perform the same music with a different timbre.

The results of the percent correct answers for each target timbre are summarized in **Table 2** below, including the specific scores in each audio-visual condition and the average score. These results indicate that percent correct responses for *dark*, *round*, and *bright* timbres were the lowest (dark<round<bright), and percent correct responses for *light*, *relaxed*, and *sharp* timbres were the highest (sharp>relaxed>light). It is noteworthy that the relative percent correct answers regarding round timbre and sharp timbre was considerably higher than the absolute

[3]The method of calculating absolute and relative scores of percent correct answers is borrowed from Timmers and Ashley (2007, Experiment 2).

[4]Timbre metaphors will now be italicized for better readability and display of the results.

**TABLE 2 |** Percentage of listeners' correct answers for a target timbre (values in each AV condition and their average).

| Rated timbre | Absolute percentage[a] | | | | Relative percentage[b] | | | |
|---|---|---|---|---|---|---|---|---|
| | Average | AO | AV | VO | Average | AO | AV | VO |
| Dark | **15.34** | 12.70 | 23.81 | 9.52 | **19.15** | 17.49 | 23.33 | 17.46 |
| Round | **19.15** | 13.33 | 20.63 | 22.22 | **32.98** | 28.33 | 36.51 | 31.75 |
| Bright | **23.4** | 28.57 | 20.63 | 21.67 | **29.26** | 23.81 | 31.75 | 31.67 |
| Velvety | **24.34** | 23.81 | 22.22 | 26.98 | **29.63** | 26.98 | 30.16 | 31.75 |
| Tense | **26.98** | 19.05 | 30.16 | 31.75 | **25.4** | 20.63 | 25.40 | 30.16 |
| Dry | **27.13** | 31.75 | 26.98 | 21.67 | **23.94** | 25.40 | 22.22 | 25.00 |
| Heavy | **31.91** | 36.67 | 28.57 | 30.16 | **30.32** | 38.33 | 28.57 | 25.40 |
| Light | **34.92** | 30.16 | 36.51 | 38.10 | **39.15** | 39.68 | 42.86 | 34.92 |
| Relaxed | **35.45** | 31.75 | 34.92 | 39.68 | **34.92** | 28.57 | 38.10 | 38.10 |
| Sharp | **39.68** | 47.62 | 42.86 | 28.57 | **52.38** | 61.90 | 55.56 | 39.68 |
| Average | **27.83** | 27.54 | 28.73 | 27.03 | **31.71** | 31.11 | 33.45 | 30.59 |

[a] *When the rating of the target timbre was higher than ratings of the other nine timbres when the listener was presented with the performance of a target timbre.*
[b] *When the rating of target timbre was relatively higher for the performance with target timbre than in the other performances of a piece.*

percentage. This implies that the performed piece did not sound very round or sharp, but performers were able to change the degree to which those excerpts were perceived as round or sharp. When considering the influence of AV stimuli on the absolute/relative percent correct score, *tense* and *relaxed* timbre were more likely to be rated as high when visual information was presented—as can be seen by the higher percent correct score in both absolute and relative percentages in the audio-visual and the visual-only condition compared to the audio-only condition. The impact of the AV stimuli on other timbre evaluations was more variable, the details of which will be investigated in the next analysis of variance of the timbre ratings.

To examine whether the evaluation of each timbre was above chance level, a one-sample $t$-test was conducted, in which the test value was 0.1. This is a chance probability that a target timbre is the highest rated timbre. The values of t, df, and mean score for the one-sample t-test are summarized below (**Table 3**). This analysis showed that the mean values of all variables were significantly higher than the chance level of 10%, which gives the first indication of a reliable communication of the target instruction. Thus, the first hypothesis (H1) was supported. Interestingly, for nine timbres the communication was even successful in the visual-only condition, except for the *dark* timbre (9.52% in VO condition).

**TABLE 3 |** $T$-value, $p$-value, and mean score of absolute percent correct in the one-sample $t$-test.

| Rated timbres | Mean | $t$-value | df | $p$-value |
|---|---|---|---|---|
| *Dark* percent correct | 0.15 | 2.03 | 188 | <0.05 |
| *Round* percent correct | 0.19 | 3.18 | 187 | <0.05 |
| *Bright* percent correct | 0.23 | 4.33 | 187 | <0.01 |
| *Velvety* percent correct | 0.24 | 4.58 | 188 | <0.01 |
| *Tense* percent correct | 0.27 | 5.25 | 188 | <0.01 |
| *Dry* percent correct | 0.27 | 5.27 | 187 | <0.01 |
| *Heavy* percent correct | 0.32 | 6.43 | 187 | <0.01 |
| *Light* percent correct | 0.35 | 7.17 | 188 | <0.01 |
| *Relaxed* percent correct | 0.35 | 7.30 | 188 | <0.01 |
| *Sharp* percent correct | 0.40 | 8.32 | 188 | <0.01 |

## Three-Way ANOVA Results

In the repeated measures three-way ANOVA, the main effect of heard timbre (two levels: target timbre, opposite timbre) is a key indicator of successful timbral communication, as the score indicates the differences in timbre ratings across two performances with two contrasting timbres (target, opposite). In other words, significant differences suggest that listeners can successfully recognize performers with target timbres, otherwise performances with target timbre and opposite timbre are confounded. The overall ANOVA results are in **Table 4**.

### Main Effects

We found the main effect of heard timbre for all timbre ratings except for the *dark* timbre, indicating that the listeners failed to recognize performances with *dark* timbre regardless of who was performing and in which condition. This result is in line with the results in section Participants, which show that the percentage of correct answers for *dark* evaluation is the lowest (15%) among the 10 timbres. Therefore, H2 was mostly supported - the listeners can successfully differentiate between performances with contrasting timbral intentions except when the music was played with dark timbre.

### Two-Way and Three-Way Interaction

*Heard Timbre* × *condition*. The two-way interaction between heard timbre and condition (three level: AO, AV, and VO) suggests that the recognition of a target timbre is influenced by auditory-visual presentation modes. We found a significant interaction effect between heard timbre and AV condition for the communication of three timbres: *relaxed*, *tense*, and *velvety*. For *relaxed* timbre evaluation, planned contrast analysis between

**TABLE 4 |** Three-way repeated measures ANOVA main effects and interactions for listeners' evaluation of 10 timbres.

| | | Evaluation | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Effects** | | **Brightness** | **Darkness** | **Heaviness** | **Lightness** | **Roundness** | **Sharpness** | **Relaxed** | **Tense** | **Dry** | **Velvety** |
| Timbre | **F** | **22.88**** | 3.92 | **59.02**** | **52.44**** | **32.38**** | **45.38**** | **39.05**** | **45.30**** | **31.18**** | **70.25**** |
| | $\eta p^2$ | 0.77** | 0.2 | 0.76** | 0.73** | 0.66** | 0.73** | 0.66** | 0.70** | 0.62** | 0.79** |
| | df | (1, 16) | (1, 16) | (1, 19) | (1, 19) | (1, 17) | (1, 17) | (1, 20) | (1, 20) | (1, 19) | (1, 19) |
| Pianist/ Piece | **F** | **11.90**** | **31.48**** | **73.12**** | **57.28**** | **6.80*** | **10.25**** | **8.42**** | **12.83**** | **8.93**** | **17.22**** |
| | $\eta p^2$ | 0.43 | 0.66** | 0.79** | 0.75 | 0.53* | 0.61 | 0.30** | 0.39** | 0.32** | 0.48** |
| | df | (2, 32) | (2, 32) | (2, 38) | (2, 38) | (2, 34) | (2, 34) | (2, 40) | (2, 40) | (1, 19) | (1, 19) |
| AV stimuli | **F** | 0.78 | **4.32*** | **3.81*** | 0.91 | **11.56**** | **14.75**** | 0.48 | 1.39 | 1.02 | **4.37*** |
| | $\eta p^2$ | 0.05 | 0.21* | 0.17* | 0.05 | 0.64** | 0.68** | 0.02 | 0.07 | 0.05 | 0.19* |
| | df | (2, 32) | (2, 32) | (2, 38) | (2, 38) | (2, 34) | (2, 34) | (2, 40) | (2, 40) | (2. 38) | (2. 38) |
| Timbre * Piece | **F** | 0.93 | 0.86 | 1.28 | **7.96**** | **11.31**** | **8.21**** | **8.11**** | 3.13 | 2.85 | **3.84*** |
| | $\eta p^2$ | 0.06 | 0.05 | 0.06 | 0.30** | 0.63** | 0.57** | 0.29** | 0.14 | 0.13 | 0.17* |
| | df | (2, 32) | (2, 32) | (2, 38) | (2, 38) | (2, 34) | (2, 34) | (2, 40) | (2, 40) | (2. 38) | (2. 38) |
| Timbre * AV | **F** | 0.28 | 0.83 | 0.17 | 0.11 | 0.11 | 1.09 | **3.81*** | **6.95**** | 2.29 | **3.80*** |
| | $\eta p^2$ | 0.02 | 0.05 | 0.01 | 0.01 | 0.01 | 0.06 | 0.16* | 0.26** | 0.11 | 0.17* |
| | df | (2, 32) | (2, 32) | (2, 38) | (2, 38) | (2, 34) | (2, 34) | (2, 40) | (2, 40) | (2. 38) | (2. 38) |
| Piece * AV | **F** | 1.86 | 1.36 | 2.21 | 0.91 | 2.22 | 1.55 | **3.10*** | 0.53 | 0.62 | 0.41 |
| | $\eta p^2$ | 0.10 | 0.08 | 0.10 | 0.05 | 0.12 | 0.08 | 0.13* | 0.03 | 0.03 | 0.02 |
| | df | (4, 64) | (4, 64) | (4, 76) | (4, 76) | (4, 68) | (4, 68) | (4, 80) | (4, 80) | (4, 76) | (4, 76) |
| Timbre* Piece * AV | **F** | 0.31 | 1.16 | **4.80**** | 1.20 | 1.52 | **4.04*** | 1.83 | 0.73 | 0.75 | **3.31*** |
| | $\eta p^2$ | 0.02 | 0.07 | 0.45** | 0.06 | 0.08 | 0.19* | 0.08 | 0.04 | 0.04 | 0.15* |
| | df | (4, 64) | (4, 64) | (4, 76) | (4, 76) | (4, 68) | (4, 68) | (4, 80) | (4, 80) | (4, 76) | (4, 76) |

*$^{**}p < 0.01$; $^{*}p < 0.05$.*

*Bold values indicate the significant results in the main effects and interaction effects.*

conditions contains vision and sound-only condition suggested that the communication was better when both audio and vision contained than the sound-only condition, $F_{(1, 20)} = 15.56$, $p < 0.01$, $\eta p^2 = 0.44$. For *tense* timbre evaluation, planned contrast results indicated that tension was differentiated more clearly in the AV conditions than in the audio-only condition, $F_{(1, 20)} = 5.00$, $p < 0.05$, $\eta p^2 = 0.20$. **Figure 4** displays the evaluation for relaxed timbre (left) and tense timbre respectively, where the interval (i.e., differences in ratings across two performance with target timbre and opposite timbre) was larger in visual conditions. However, the participants differentiated the *velvety* timbre more clearly in the audio-only condition than in the audio-visual condition, $F_{(1, 20)} = 9.10$, $p < 0.01$, $\eta p^2 = 0.31$ (**Figure 5**).

*Heard Timbre × condition × piece/performer*. The three-way interaction between heard timbre, condition, and piece/performer suggests the extent to which a successful communication of a target timbre in particular audio-visual conditions is reliant on the performer/piece. We found significant interaction results for the communication of *heavy* timbre (**Figure 6**), *sharp* timbre (**Figure 7**), and *velvety* timbre (**Figure 8**). The results of planned contrasts were labeled in these figures where significant three-way interaction was found. More specifically, P1 communicated *heavy* timbre better than P3 in aural conditions while P3 communicated better in the visual-only condition; P1 communicated *sharp* timbre better in the visual-only condition while P3 communicated better in aural conditions. In addition, P1 communicated *velvety* timbre more effectively in aural conditions than P2 whereas P2 was better in the visual-only condition.

To sum up, H3 was partially supported: the above results suggest that only the communication of *relaxed* and *tense* timbres is not reliant on the performer's differences. In other cases, there is an individual difference in the utilization of the audio-visual condition to communicate the timbral intention to the listeners. Each performer may be specialized in communicating a particular timbre with particular aural or visual cues.

## Non-verbal Sensory Judgment Analysis

Judgment of the size, shape, and brightness was collected as categorical data in the form of either: A (bigger size, rounder shape, and brighter version), or B (smaller size, sharper shape, and darker version) in the questionnaire. This data was replaced with a score of either 0 or 1, to give a method of calculating the mean across different participants ($N = 21$). The measure of each type of judgment (e.g., size evaluation while hearing *bright* timbre) for each participant was calculated *via* an average across AV stimuli and three music pieces (Mean = SUM divided by 9). **Table 5** displays the value of the size, shape, and brightness evaluations in response to five pairs of timbre intentions. A paired sample $T$-test was conducted to compare the mean difference in non-verbal judgments in response to

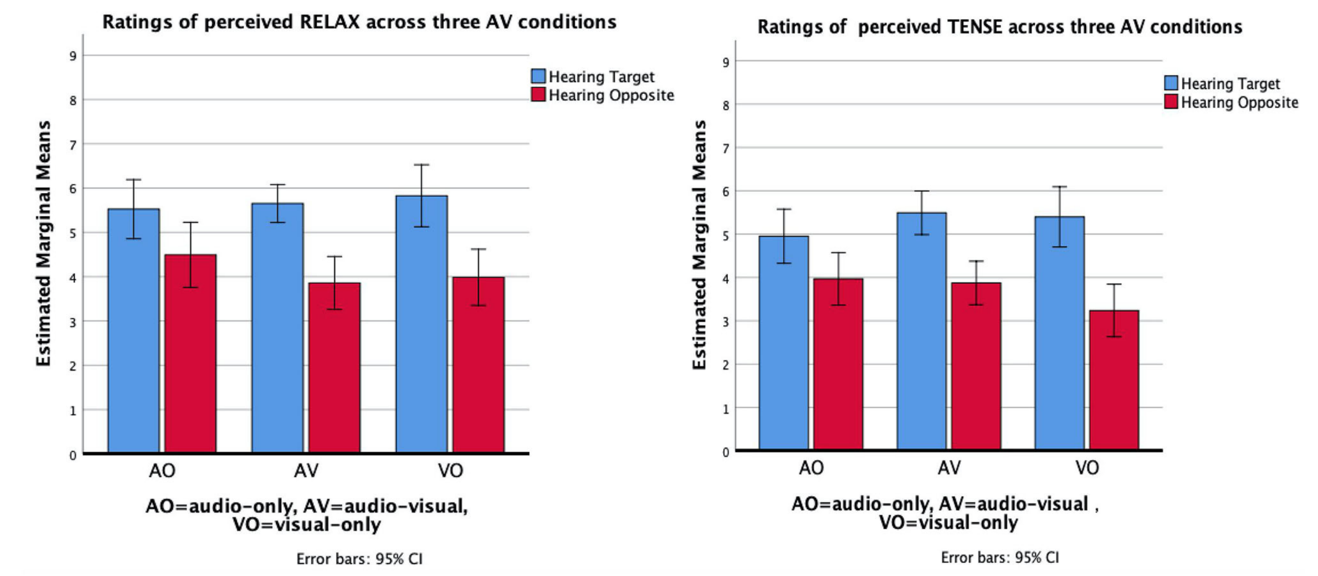



**FIGURE 4 |** Mean values of the evaluation of *Relaxed* timbre **(left)** and *Tense* timbre **(right)** across three AV stimuli found by averaging the pianist**/**piece factor. Error bars represent confidence intervals.

contrasting performances—e.g., the mean difference in size evaluation between hearing *bright* timbre and hearing *dark* timbre. Significant differences are shown with *.

The results indicate a strong timbre-shape association in all five pairs of timbral intentions, hence H4 was supported. Participants tended to choose the round object when hearing/seeing performances with *dark*, *round*, *light*, *relaxed,* and *velvety* timbral intentions and chose the sharp object when hearing/seeing performances with *bright*, *sharp*, *heavy*, *tense,* and *dry* timbral intentions. In contrast, size was associated with the perception of two pairs of timbres (*bright/dark* timbre, *heavy/light* timbre), and brightness was associated with only one timbre perception pair (*relaxed/tense* timbre).

## Principal Components Analysis (PCA) and Acoustic/Visual Information

The PCA revealed the presence of three components with eigenvalues >1, which account for a total of 69.6% of the variance in evaluations of the 10 timbres, explaining 38.71, 19.2, and 11.68% of the variance respectively. To aid the interpretation of the three components, varimax rotation was used and the results are displayed in **Table 6**. Loading values that were <0.3 were excluded from the table, so blanks in the table indicate where low loading exists.

As a result, the first dimension (named as Round-dimension, eigenvalue: 38.7%) includes the timbre evaluation of round, velvety, relaxed (positive scores) and non-dry (negative score). This dimension is related to touch and movement qualities and is most strongly associated with Round. It is a combination of tactile feelings of roundness and velvetiness, and low dryness. The second component (named as Heavy-dimension) relates to the evaluation of heavy, dark, tense, and sharp. This is a combination of high intensity and negative valence and is associated with haptic sensations of sharpness and tactile

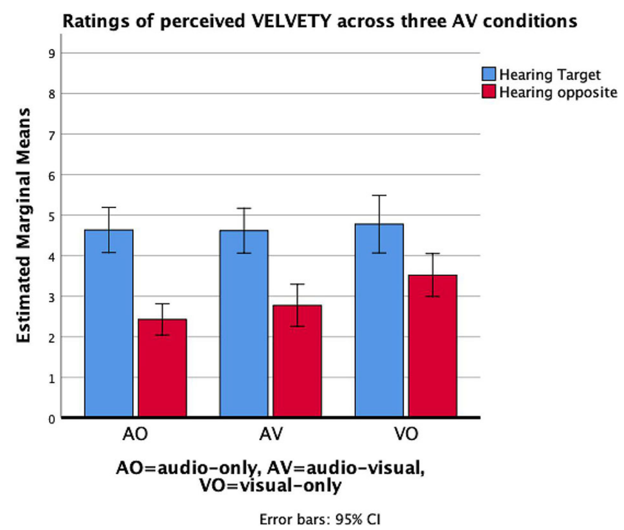



**FIGURE 5 |** Mean values of the evaluation of *Velvety* timbre across three AV stimuli found by averaging in piece/performer factor. Error bars represent confidence intervals.

sensations of heaviness. It is most strongly associated with Heavy and relates to experiences of weight and negative intensity. The last component (named as Luminance-dimension) is most strongly associated with Brightness followed by lightness. It seems that the Luminance-dimension is both positive in valence, high in space and lightweight.

The **Appendix 4** displays the audio waves and video snapshot[5] of the three pairs of performances (round vs. velvety, heavy

[5]Video snapshots were captured at the same time slot to obtain a relatively fair comparison. However, video shortcuts failed to provide dynamic and continuous gestural changes e.g. how fingers are articulated (smooth or detached).

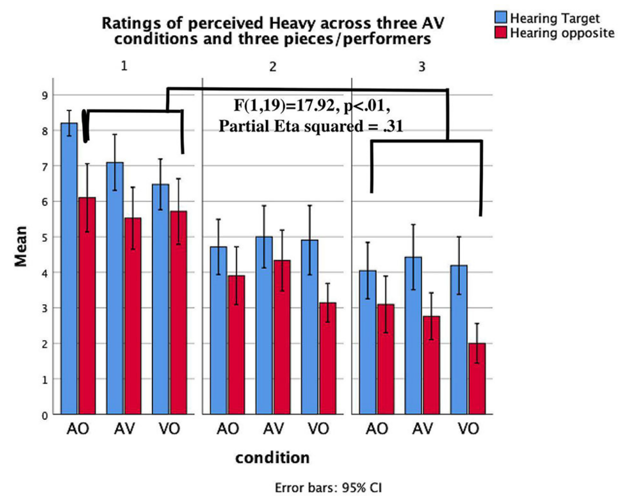



**FIGURE 6 |** Mean values of the evaluation of *Heavy* timbre across three pianists/pieces and AV stimuli. Error lines represent confidence intervals.

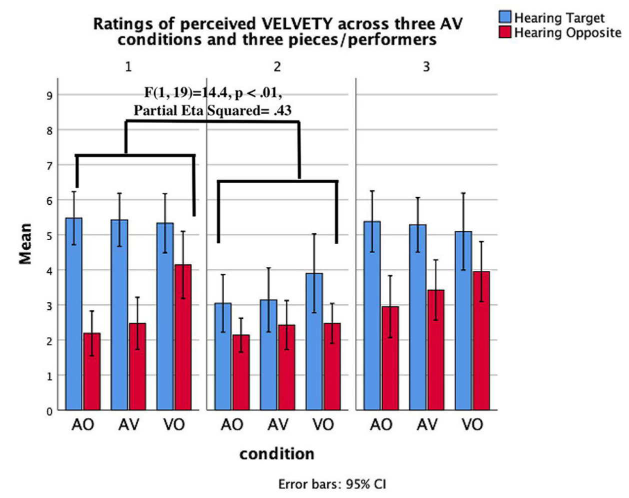



**FIGURE 8 |** Mean value of the evaluation of *Velvety* timbre across three pianists/pieces and AV stimuli. Error lines represent confidence intervals.

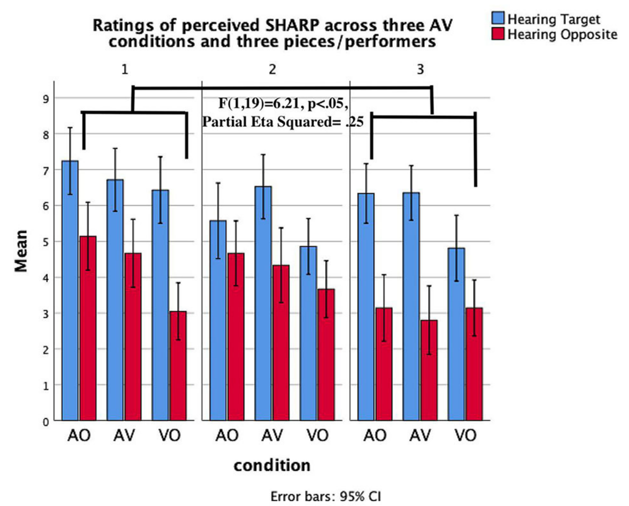



**FIGURE 7 |** Mean values of the evaluation of *Sharp* timbre across three pianists/pieces and AV stimuli. Error lines represent confidence intervals.

vs. dark, bright vs. light) by three performers. We compared the aural/visual information in pairs, to explore whether the pianists were defaulting to the same performance style/gestures or not when playing these highly correlated timbres. Observable differences were found in these pairs:

Round vs. Velvety: P1 differentiated these two timbres using intensity (*round* is louder than *velvety*); while P2 used performance tempo to make contrast (*round* is quicker than *velvety*). In the visual cues, P1 and P2 changed sound-producing gestures, with higher hand position and curved finger to produce *round* timbre while using lower hand position and flatter finger to play *velvety* timbre. The difference in P3 was little.

Heavy vs. Dark: For P1 and P3, *heavy* timbre was interpreted as much louder than *dark* timbre. P2 played *slower* in dark timbre compared to *heavy* timbre. In general, all the three pianists were shown more solemn facial expression, and P3 showed more intensive feelings with inward upper-body movements toward the piano.

Bright vs. Light: In general, these two timbres were played much softer for the three performers. Differences were found in the facial expression and posture for P1 and P3, that *bright* timbre was played with more delighted and cheerful expression and movements.

# DISCUSSION

## The Reliability of Communicating Timbre

The first research question examined in this study is whether the communication of timbral intention between pianists and listeners is reliable. The study results indicate that all the 10 timbres are communicated with an accuracy above chance level, although the average percentage correct was considerably higher for some timbres (*sharp, relaxed, and light*) than for others. Furthermore, the repeated-measures ANOVA results showed that nine of 10 timbres were reliably communicated, except for the *dark* timbre. This suggests that timbral intention can be communicated between the performer and the listeners like other abstract intentions including expressivity (Davidson, 1993), emotions (Juslin, 2000), and sensorimotor feelings (De Poli et al., 2017). There is already evidence demonstrating that instrument-expertise influences the perception of musical expressivity due to the activation of sensory representation in the observation of motor plan (Broughton and Davidson, 2014).

The difficulty of communicating the 10 timbres varies one from the other. *Bright* timbre was the easiest to communicate and not influenced by the AV condition factor or performer/piece, while *dark* timbre was the most difficult one and the

**TABLE 5 |** Mean judgment scores for size, shape, and brightness in the perception of 10 timbres.

| | Size | Shape | Brightness |
|---|---|---|---|
| | (small) 0 – 1 (big) | (sharp) 0 – 1 (round) | (dark) 0 – 1 (bright) |
| Bright | 0.47* | 0.45** | 0.53 |
| Dark | 0.35* | 0.62** | 0.57 |
| Round | 0.49 | 0.66** | 0.54 |
| Sharp | 0.59 | 0.17** | 0.44 |
| Heavy | 0.58** | 0.41** | 0.53 |
| Light | 0.31** | 0.72** | 0.60 |
| Tense | 0.41 | 0.40** | 0.48* |
| Relaxed | 0.38 | 0.68** | 0.63* |
| Dry | 0.39 | 0.38** | 0.54 |
| Velvety | 0.46 | 0.76** | 0.57 |

***$p < 0.01$; *$p < 0.05$.*

**TABLE 6 |** Factor loading after varimax rotation.

| Evaluation | Component 1 Round (Touch-movement) | Component 2 Heavy (Weight-Negative intensity) | Component 3 Bright (Luminance) |
|---|---|---|---|
| Round | **0.851** | | |
| Velvety | **0.827** | | |
| Dry | −0.626 | 0.333 | 0.340 |
| Relaxed | 0.595 | | 0.486 |
| Heavy | | **0.811** | −0.328 |
| Dark | | **0.765** | −0.320 |
| Tense | −0.473 | 0.650 | |
| Sharp | −0.595 | 0.614 | |
| Bright | | | **0.829** |
| Light | | −0.344 | **0.744** |

*Bold values highlight the most two representative timbres in each dimension.*
*The highlighted red mark indicates the three clusters in PCA. They were framed for readability and highlighting.*

communication even failed in the visual-alone condition. One of the potential reasons could be that this experiment uses a Yamaha piano which usually has a brighter sound, resulting in the superior communication of brightness.

The differences in the communication outcome imply that music performers may consider suitable strategies (i.e., acoustic cues or visual cues) for the expression and communication of different timbres and that music educators are encouraged to focus on the teaching and learning of more difficult ones (i.e., *dark, round, and bright*) in piano lessons. Facial expression can be considered to differentiate *bright* and *dark* timbre, which has been demonstrated to be an effective way of differentiating the effects conveyed in the vocal performance (Thompson et al., 2005). For example, **Figure 9** below shows that two performers in this research used different visual communication strategies to differentiate *bright* and *dark* timbre: one used facial expression while the other one used forward/inward upper-body movements.

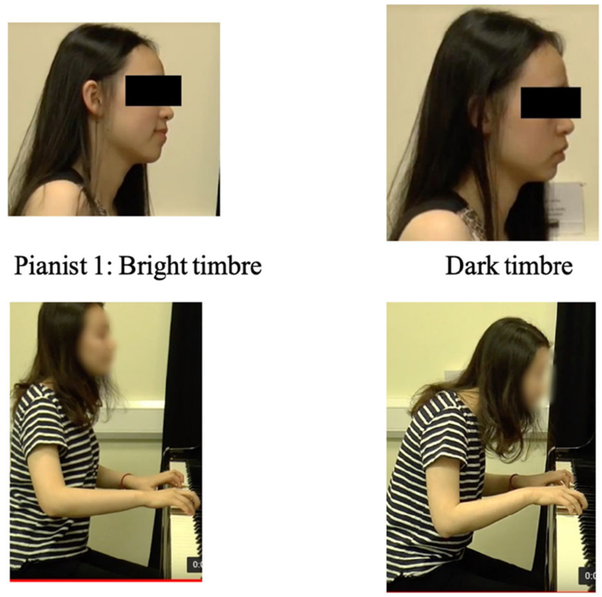



**FIGURE 9 |** The facial expression and posture of two performers when expressing bright and dark timbres.

## Hearing and Seeing Piano Timbre

The subsequent research question we investigated concerns the relevance of auditory and/or visual cues in the communication of timbral intention. Firstly, the comparison of sound waves across five pairs of timbres indicated that the inconsistency in the choice of other performance parameters (either using intensity or performance tempo) among the three pianists; however, the reliability of communicative process implied that the participants had a common, consensual understanding of the meaningful use of performance parameters to achieve certain timbre. This result is in line with Bernays and Traube's (2014) study that showed an individuality in the playing style of five pianists in the expression of the same timbral intention, with a personalized choice of musical features (articulation, pedal, attack, dynamics, etc.).

Secondly, the importance of visual cues in music performance is verified for the paring of *relaxed* and *tense* timbres, where visual information determined the effectiveness of timbral communication independent of the pianist/piece. For the remaining timbres (*heavy, light, round, sharp, dry, and velvety*), the influence of the AV condition was in most cases accompanied by an interaction involving the effect of pianist/piece (e.g., when communicating *sharp* timbre, P1 is better in visions while P3 is better in sounds), implying the difference in the preference of using visual cues among the pianists. Interestingly, increasing numbers of scholars (Thompson et al., 2005; Schutz and Lipscomb, 2007; Schutz, 2008; Davidson, 2012) have started to claim the benefits of using bodily communication in a music performance for the performers, including enhancing auditory experience, shaping the perception of musical notes, or facilitating the comprehension of lyrics. According to Schutz (2008), percussive instruments such as the piano and marimba could be most successful in benefiting from visual communication due to limited timbral nuances and percussive/short note length. We therefore encourage pianists to use bodily communication (i.e., appropriate facial expressions, gestures, and postures) in the communication of timbral intention, in addition to the auditory component of the music performance.

We integrated two angles including sound-producing gestures and ancillary gestures in the video stimuli. Although it is unknown whether the visual communication of 10 timbres differs in the reliance on specific aspects of bodily movements, we anticipate that the visual information of sound-producing gestures underlies the perception of 10 timbres—seeing how the keyboard is touched (i.e., percussive/non-percussive, curved/flat, hard/soft, etc.) influenced the listeners' perception of piano timbre, as suggested by the McGurk effect in music perception (Saldaña and Rosenblum, 1993; Schutz and Lipscomb, 2007). Further study needs to be taken to investigate the impact of ancillary gestures on timbre perception.

Last but not the least, the successful communication of timbral intention even in the vision-alone condition suggests an embodied perspective in music communication. One possible explanation is that understanding a pianist's silent movements possibly involves the simulation of a performative action in the listener's head, which leads to the triggering of an auditory image of the timbral effect (Keller, 2012). This is in line with previous studies (Camurri et al., 2003; Dahl and Friberg, 2007) showing that listeners can still detect the emotional intention of performers even when the sound information is unavailable. Brain imaging studies supported this and found that there are activations of certain motor-related areas in the brain when imagining music with the "inner ear" (Zatorre and Halpern, 2005) and in the perception of musical sounds (Haueisen and Knösche, 2001). This study extends the view of embodied music listening, by showing the successful communication of timbral intentions even without aural cues.

In this study, the listeners can "stand in someone else's shoes" to understand the actions, intentions, and sounds of the pianists by simulating the motor activity. According to Molnar-Szakacs and Overy (2006), a "similar or equivalent motor network is engaged by someone listening to singing/drumming as the motor network engaged by the actual singer/drummer" (p. 236) and therefore the communication of timbre was also successful in audio-alone conditions due to a simulation of a motor activity similar or equivalent to that associated with the timbral effect. The loop between sensory experience and motor commands has been explained by the "internal model" in the research into sensorimotor perceptions (Keller, 2012; Maes et al., 2014), which suggests its benefits for musicians themselves including action planning and self-monitoring (Novembre and Keller, 2014) and the benefits for co-players such as synchronization and coordination (Keller et al., 2007). This study demonstrated the benefits of an internal model in performer-listener communication, which is a shared understanding of actions and sounds that occur in the pianists' communicative process of timbral intentions to the listeners. We, therefore, suggest that the listeners are also performers because a mirror process happens in the perception of piano timbre through the encoding of expressive gestures into sounds (timbres) and the decoding of sounds (timbres) into expressive gestures (cf. Leman and Maes, 2014).

## Timbre-Related Cross-Modal Correspondence

The results of non-verbal sensory judgement indicated that a cross-modal timbre-shape association was found in all 10 timbres. When music was played with *dark*, *round*, *light*, *relaxed*, and *velvety* timbral intentions, it was felt to be rounder than when played with *bright*, *sharp*, *heavy*, *tense,* and *dry* intentions. This timbre-shape association is in line with several previous studies which have found a sharpness/color association with either *soft* or *harsh* timbres (Adeli et al., 2014) as well as visual textural associations (Giannakis, 2006). In addition to a timbre-shape association, this study also found a timbre-size association in two pairs of timbres (*bright/dark*, *heavy/light*).

Wallmark and Kendall (2018) applied the theory of conceptual metaphor (Lakoff and Johnson, 1980) to understand timbre semantics. They explained that when timbre is described with light contrast (e.g., bright/dark) and textural feature (e.g., rough, smooth), it reflects the conceptual metaphors that SOUND IS LIGHT and SOUND IS TEXTURE as people use the source domain of vision and tactile sensations as a reference to drive the meaning of the target domain (i.e., timbre). This statement reflects how abstract domains of human experience can be understood by concrete, embodied accounts. The embodied accounts to explain timbre-related CMC can also be found in Spence (2011) who provided three for the underlying reasoning of CMC: statistical, semantic-mediated, and embodied accounts. Statistical and embodied accounts explained the results of repeated exposure to, and physical interaction with, instruments and the extent to which the weak synaesthesia experience related to piano timbre can be established as a part of musical training (Spence, 2011). Studies have indicated that repeated exposure to statistical co-current pairs of stimuli can help to shape the coupling pairs, even in circumstances where the stimuli are unrelated [e.g., the association between a feeling of stiffness and visual luminance, Ernst (2007)]. Semantic-mediated accounts help to understand the linguistic descriptions related to piano timbre and their possible impact in generating a cross-modal coupling experience [cf. Dolscheid et al. (2013)]. For example, in the production of round timbre, a teacher's verbal description (round), modeling, and the embodiments (round handshape) may help the student to establish an association between the sound outcome and the round shape.

## Piano Timbre as a Component of Expressivity

For a long time, the definition of musical expression/expressiveness has been deeply influenced by Seashore's 1938 statement on "deviation from the score"[6] [see a discussion in Clarke and Doffman (2014)]. The shortcomings of

[6]which defines expression as a deviation from the notated values (i.e. pure tone, true pitch, even dynamics, rigid rhythm etc.) on the musical score.

this definition and the over-emphasis on the sonic properties of performance have been noted by several scholars, for instance leading to a trend of the musical score being the primary ontological focus of music (Dogantan-Dack, 2014) and a disembodied, ahistorical account of musical performance by conceptualizing the score as the piece (Clarke, 2004). This research would like to assure that, in the pursue of timbral nuances, musical expressiveness is imprinted in the acoustic variance and the subtle control of gestures and touch of pianists, as if "touch is the expressive skeleton on which the pianist enfolds the expressive flesh" (Dogantan-Dack, 2014, p. 7). Touch, gestures, bodily movements in piano playing actively shape a listener's multimodal perception of piano timbre, resulting in sensorimotor perception and cross-modal correspondences.

## IMPLICATIONS AND CONCLUSIONS

There are some implications for instrumental music pedagogy and performance. This study offers interesting insights into the question of "what do listeners perceive in timbre in music performance." From the listener's perspective, the perception of timbre is not merely variations in acoustic information but related to multimodal perception including kinaesthetic and muscular sensations (heavy-light, tensed-relax) and cross-modal correspondences (physical size and shape). This is consistent with the studies of the semantics of timbre (Bellemare and Traube, 2005; Bernays, 2013; Kojucharov and Rodà, 2015; Saitis and Weinzierl, 2019) that have shown rich subjective experiences and multidimensionality in response to timbre. Therefore, music teachers and students are encouraged to use cross-domain metaphors and multimodal communications (gestures, modeling, touch, etc.) in the teaching and learning of timbre-related performance goals in music lessons (Li and Timmers, 2021).

This study suggests that timbral intention can work well as a bridge that connects the performer and the listener, due to the resonance of sensorimotor knowledge induced by musical sounds. It also has implications for music education that teachers should encourage students to mobilize timbre as a deeper motivation for expressing intention/message and communicating it to their listeners. As Dogantan-Dack (2017) pointed out, classical musicians are facing pressure due to the sustainability of the profession itself and the diversity of today's musical genres and practices. In the process of working on timbral intentions and communicating them to the listeners, student-pianists may increase their musical competency and sense of autonomy [i.e., expressive freedom extending beyond the composer's intention, Dogantan-Dack (2017)], which may help to fulfill their psychological needs at this stage of their musical education and maintain the motivation to persist and engage with music (Wise et al., 2017).

One of the limitations in this study is the unsystematic control of performance parameters such as dynamics and tempo, and we acknowledge that these elements can influence musical perception. This enables a more open musical interpretation for the performers to express different timbral intentions, which probably work as co-variables to influence the listeners' judgments. Another limitation is the limited sample size used in the present study; thus, the results of the variety in communication outcomes of 10 timbres must be taken as tentative. Future research on the visual communication of piano timbre may consider using less timbres, for instance relaxed/tense and heavy/light, to expand our understanding of touch qualities (e.g., soft/hard, percussive/non-percussive) with deeper insights into bodily feelings in piano playing. Alternatively, future research may also consider using non-pianists as the control group (i.e., the level of musical training as the between-group factor). There is already evidence demonstrating that instrument-expertise influences the perception of musical expressivity due to the activation of sensory representation in the observation of motor plan (Broughton and Davidson, 2014). Future research can consider more restricted control of performance parameters using a single piano tone or chord. Furthermore, future research may be able to use point-light displays to replace bodily movement in the videos, to avoid the influence of familiarity of performers or their facial expressions. This method could also help to create congruent or incongruent stimuli, by synchronizing the movement features of one stimulus with the performance data of another in terms of onset/offset time, duration, dynamics, and pedaling [cf. Vuoskoski et al. (2014)]. For example, a new artificial video can be generated *via* synchronizing movement features in a "tense" performance with the sound signal of a "relaxed" performance thereby allowing an examination of the extent to which visual information influences or modifies listeners' judgments.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the University of Sheffield. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the participants for the publication of any identifiable images or data included in this article.

## AUTHOR CONTRIBUTIONS

SL and RT designed the experiment. SL was responsible for data collection, data analysis, and the draft of manuscript. RT and WW contributed to refine arguments, the presentation of results, and to improve the readability. All authors contributed to the article and approved the submitted version.

## FUNDING

This research was funded by Chinese Postdoctoral International Exchange Program [grant number 273060] and Research on adolescent internet adaptation-oriented optimization method for personalized information service [NSFC Grant No. 71974072].

## ACKNOWLEDGMENTS

We would like to thank the Department of Music of The University of Sheffield (UK) and the Department of Music of the Henan University (China) for providing the experiment equipment and venue. Our special thanks go to Dr. Mengjiao Yan, Yiyun Liu, and Yikun Huang for making the recordings and the pianist-students who participated in our experiment. We would like to thank the two composers Frederic Chiasson and Ana Dall'Ara-Majek who composed the music, and Michel Bernays and Caroline Traube for their permissions of using the musical score.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2021.717842/full#supplementary-material

**Supplementary Video 1 |** P1_bright.

**Supplementary Video 2 |** P1_dark.

**Supplementary Video 3 |** P1_heavy.

**Supplementary Video 4 |** P1_light.

**Supplementary Video 5 |** P1_relax.

**Supplementary Video 6 |** P1_tense.

**Supplementary Video 7 |** P1_round.

**Supplementary Video 8 |** P1_sharp.

**Supplementary Video 9 |** P1_dry.

**Supplementary Video 10 |** P1_velvety.

## REFERENCES

Adeli, M., Rouat, J., and Molotchnikoff, S. (2014). Audiovisual correspondence between musical timbre and visual shapes. *Front. Hum. Neurosci*. 8:352. doi: 10.3389/fnhum.2014.00352

Balkwill, L. L., and Thompson, W. F. (1999). A cross-cultural investigation of the perception of emotion in music: psychophysical and cultural cues. *Music Percept*. 17, 43–64. doi: 10.2307/40285811

Baraldi, F. B., Poli, G. D., and Rodá, A. (2006). Communicating expressive intentions with a single piano note. *J. New Music Res.* 35, 197–210. doi: 10.1080/09298210601045575

Barthet, M., Depalle, P., Kronland-Martinet, R., and Ystad, S. (2010). Acoustical correlates of timbre and expressiveness in clarinet performance. *Music Percept*. 28, 135–154. doi: 10.1525/mp.2010.28.2.135

Behne, K. E., and Wöllner, C. (2011). Seeing or hearing the pianists? A synopsis of an early audiovisual perception experiment and a replication. *Music. Sci.* 15, 324–342.

Bellemare, M., and Traube, C. (2005). "Verbal description of piano timbre: exploring performer-dependent dimensions," in *Proceedings of the 2nd Conference on Interdisciplinary Musicology (CIM05)*, Montreal, QC: CIM.

Bernays, M. (2013). *The Expression and Production of Piano Timbre: Gestural Control and Technique, Perception and Verbalisation in the Context of Piano Performance and Practice*. (Doctoral dissertation), Université de Montréal, Montreal, QC, Canada.

Bernays, M., and Traube, C. (2014). Investigating pianists' individuality in the performance of five timbral nuances through patterns of articulation, touch, dynamics, and pedaling. *Front. Psychol.* 2014:157. doi: 10.3389/fpsyg.2014.00157

Broughton, M., and Stevens, C. (2009). Music, movement and marimba: an investigation of the role of movement and gesture in communicating musical expression to an audience. *Psychol. Music* 37, 137–153. doi: 10.1177/0305735608094511

Broughton, M. C., and Davidson, J. W. (2014). Action and familiarity effects on self and other expert musicians' Laban effort-shape analyses of expressive bodily behaviors in instrumental music performance: a case study approach. *Front. Psychol.* 5:1201. doi: 10.3389/fpsyg.2014.01201

Brunswik, E. (1956). *Perception and the Representative Design of Psychological Experiments*. Berkeley, Los Angeles, CA: University of California Press. doi: 10.1525/9780520350519

Camurri, A., Lagerlöf, I., and Volpe, G. (2003). Recognizing emotion from dance movement: comparison of spectator recognition and automated techniques. *Int. J. Hum. Comp. Stud.* 59, 213–225. doi: 10.1016/S1071-5819(03)00050-8

Clarke, E. (2004). "Empirical methods in the study of performance," in *Empirical Musicology: Aims, Methods, Prospects*, eds E. Clarke and N. Cook (Oxford: Oxford University Press), 77–102. doi: 10.1093/acprof:oso/9780195167498.003.0005

Clarke, E., and Doffman, M. (2014). "Expressive performance in contemporary concert music," in *Expressiveness in Music Performance: Empirical Approaches across Styles and Cultures,* eds D. Fabian, R. Timmers and E. Schubert (Oxford: Oxford University Press), 98–114. doi: 10.1093/acprof:oso/9780199659647.003.0006

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences (2nd Edn.)*. Hillsdale, NJ: Erlbaum.

Cox, A. (2016). *Music and Embodied Cognition: Listening, Moving, Feeling, and Thinking*. Bloomington, IN; Indianapolis, IN: Indiana University Press. doi: 10.2307/j.ctt200610s

Dahl, S., and Friberg, A. (2007). Visual perception of expressiveness in musicians' body movements. *Music Percept.* 24, 433–454. doi: 10.1525/mp.2007.24.5.433

Davidson, J. W. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychol. Music* 21, 103–113. doi: 10.1177/030573569302100201

Davidson, J. W. (1995). "What does the visual information contained in music performances offer the observer? Some preliminary thoughts," in *Music and the Mind Machine*, ed R. Steinberg (Berlin, Heidelberg: Springer), 105–113. doi: 10.1007/978-3-642-79327-1_11

Davidson, J. W. (2012). Bodily movement and facial actions in expressive musical performance by solo and duo instrumentalists: two distinctive case studies. *Psychol. Music* 40, 595–633. doi: 10.1177/0305735612449896

Davidson, J. W., and Correia, J. S. (2002). "Body movement," in *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning*, eds R. Parncutt and G. E. McPherson (Oxford; New York, NY: Oxford University Press), 237–252.

De Poli, G., Mion, L., and Rodá, A. (2009). Toward an action based metaphor for gestural interaction with musical contents. *J. New Music Res.* 38, 295–307. doi: 10.1080/09298210902773941

De Poli, G., Murari, M., Canazza, S., Rodà, A., and Schubert, E. (2017). "Beyond emotion: multi-sensory responses to musical expression," in *The Routledge Companion to Embodied Music Interaction*, eds M. Lesaffre, P. J. Maes, and M. Leman (New York, NY; London: Routledge), 78–86. doi: 10.4324/9781315621364-9

De Poli, G., Rodà, A., and Vidolin, A. (1998). Note-by-note analysis of the influence of expressive intentions and musical structure in violin performance. *J. New Music Res.* 27, 293–321. doi: 10.1080/09298219808570750

Dogantan-Dack, M. (2014). "Philosophical reflections on expressive music performance," in *Expressiveness in Music Performance: Empirical Approaches Across Styles and Cultures*, eds D. Fabian, R.

Timmers and E. Schubert (Oxford: Oxford University Press), 3–21. doi: 10.1093/acprof:oso/9780199659647.003.0001
Dogantan-Dack, M. (2017). "Expressive freedom in classical performance: insights from a pianist-research," in *Musicians in the Making: Pathways to Creative Performance*, eds J. Rink, H. Gaunt, and A. Williamon (Oxford: Oxford University Press), 131–135.
Dolscheid, S., Shayan, S., Majid, A., and Casasanto, D. (2013). The thickness of musical pitch: psychophysical evidence for linguistic relativity. *Psychol. Sci.* 24, 613–621. doi: 10.1177/0956797612457374
Ernst, M. O. (2007). Learning to integrate arbitrary signals from vision and touch. *J. Vis.* 7, 7–7. doi: 10.1167/7.5.7
Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Method.* 39, 175–191. doi: 10.3758/BF03193146
Faure, A. (2000). *Des Sons Aux Mots, Comment Parle-t-on du Timbre Musical?* (Doctoral dissertation), Ecole des Hautes Etudes en Sciences Sociales, Paris, France.
Gabrielsson, A., and Juslin, P. N. (1996). Emotional expression in music performance: between the performer's intention and the listener's experience. *Psychol. Music* 24, 68–91. doi: 10.1177/0305735696241007
Giannakis, K. (2006). A comparative evaluation of auditory-visual mappings for sound visualisation. *Org. Sound* 11, 297–307. doi: 10.1017/S1355771806001531
Godøy, R. I. (2006). Gestural-sonorous objects: embodied extensions of Schaeffer's conceptual apparatus. *Org. Sound* 11, 149–157. doi: 10.1017/S1355771806001439
Goebl, W., Bresin, R., and Fujinaga, I. (2014). Perception of touch quality in piano *tones. J. Acoust. Soc. Am.* 136, 2839–2850. doi: 10.1121/1.4896461
Hailstone, J. C., Omar, R., Henley, S. M., Frost, C., Kenward, M. G., and Warren, J. D. (2009). It's not what you play, it's how you play it: timbre affects perception of emotion in music. *Quart. J. Exp. Psychol.* 62, 2141–2155. doi: 10.1080/17470210902765957
Haueisen, J., and Knösche, T. R. (2001). Involuntary motor activity in pianists evoked by music perception. *J. Cogn. Neurosci.* 13, 786–792. doi: 10.1162/08989290152541449
Holmes, P. A. (2011). An exploration of musical communication through expressive use of timbre: the performer's perspective. *Psychol. Music* 40, 301–323. doi: 10.1177/0305735610388898
Jensenius, A. R., Wanderley, M. M., Godøy, R. I., and Leman, M. (2010). "Concepts and methods in research on music-related gestures," in *Musical Gestures: Sound, Movement, and Meaning*, eds R. I. Godøy and M. Leman (New York, NY; London: Routledge), 12–35.
Johnson, M. (1987). *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. London: University of Chicago Press. doi: 10.7208/chicago/9780226177847.001.0001
Juslin, P. N. (1998). *A Functionalist Perspective on Emotional Communication in Music Performance. Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences, 78*. Uppsala: Uppsala University Library.
Juslin, P. N. (2000). Cue utilization in communication of emotion in music performance: relating performance to perception. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 1797–1813. doi: 10.1037/0096-1523.26.6.1797
Juslin, P. N., and Laukka, P. (2003). Communication of emotions in vocal expression and music performance: different channels, same code? *Psychol. Bullet*. 129:770. doi: 10.1037/0033-2909.129.5.770
Juslin, P. N., and Lindström, E. (2010). Musical expression of emotions: modelling listeners' judgements of composed and performed features. *Music Anal.* 29, 334–364. doi: 10.1111/j.1468-2249.2011.00323.x
Juslin, P. N., and Timmers, R. (2010). "Expression and communication of emotion in music performance," in *Handbook of Music and Emotion: Theory, Research, Applications*, eds P. N. Juslin and S. A. John (New York, NY: Oxford University Press), 453–492.
Keller, P. E. (2012). Mental imagery in music performance: underlying mechanisms and potential benefits. *Ann. N. Y. Acad. Sci.* 1252, 206–213. doi: 10.1111/j.1749-6632.2011.06439.x
Keller, P. E., Knoblich, G., and Repp, B. H. (2007). Pianists duet better when they play with themselves: on the possible role of action simulation in synchronization. *Consciousness Cogn.* 16, 102–111. doi: 10.1016/j.concog.2005.12.004
Kendall, R. A., and Carterette, E. C. (1990). The communication of musical expression. *Music Percept.* 8, 129–163. doi: 10.2307/40285493
Kochevitsky, G. (1967). *The Art of Piano Playing: A Scientific Approach*. Secaucus, NJ: Summy-Birchard Music.
Kojucharov, E., and Rodà, A. (2015). "Towards an empirically based definition of piano touch: touch/timbre relationship and key motion measurements," in *Proceedings of the 9th Triennial Conference of the European Society for the Cognitive Sciences of Music (ESCOM*2015*)*, eds J. Ginsborg, A. Lamont, and S. Bramley. (Manchester, England), 516–521.
Lakoff, G., and Johnson, M. (1980). *Metaphors We Live By*. Chicago: The University of Chicago Press.
Leman, M. (2008). *Embodied Music Cognition and Mediation Technology*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/7476.001.0001
Leman, M., and Maes, P. J. (2014). The role of embodiment in the perception of music. *Empir. Musicol. Rev.* 9, 236–246. doi: 10.18061/emr.v9i3-4.4498
Li, S., and Timmers, R. (2020). Exploring pianists' embodied concepts of piano timbre: an interview study. *J. New Music Res.* 49, 477–492. doi: 10.1080/09298215.2020.1826532
Li, S., and Timmers, R. (2021). Teaching and learning of piano timbre through teacher-student interactions in lessons. *Front. Psychol.* 2021:576056. doi: 10.3389/fpsyg.2021.576056
Maes, P. J., Leman, M., Palmer, C., and Wanderley, M. (2014). Action-based effects on music perception. *Front. Psychol.* 4:1008. doi: 10.3389/fpsyg.2013.01008
McAdams, S. (2013). "Musical timbre perception," in *The Psychology of Music, 3rd Edn*, ed D. Deutsch (London, Waltham, and San Diego: Academic Press), 35–67. doi: 10.1016/B978-0-12-381460-9.00002-X
McGurk, H., and MacDonald, J. (1976). Hearing lips and seeing voices. *Nature* 264, 746–748. doi: 10.1038/264746a0
Molnar-Szakacs, I., and Overy, K. (2006). Music and mirror neurons: from motion to'e'motion. *Soc. Cogn. Affect. Neurosci.* 1, 235–241. doi: 10.1093/scan/nsl029
Morimoto, Y., and Timmers, R. (2012). "The effect of repeated listening on pleasure and boredom responses," in *Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music,* eds E. Cambouropoulos, C. Tsougras, P. Mavromatis, and K. Pastiadis (Thessaloniki), 693–697.
Namba, S., Kuwano, S., Hashimoto, T., Berglund, B., Da Rui, Z., Schick, A., et al. (1991). Verbal expression of emotional impression of sound: A cross-cultural study. *J. Acoust. Soc. Jpn. (E)* 12, 19–29.
Novembre, G., and Keller, P. E. (2014). A conceptual review on action-perception coupling in the musicians' brain: what is it good for? *Front. Hum. Neurosci.* 8:603. doi: 10.3389/fnhum.2014.00603
Ortmann, O. (1935). What is tone-quality? *Musical Quart.* 21, 442–450. doi: 10.1093/mq/XXI.4.442
Rozé, J., Aramaki, M., Kronland-Martinet, R., and Ystad, S. (2017). Exploring the perceived harshness of cello sounds by morphing and synthesis techniques. *J. Acoust. Soc. Am.* 141, 2121–2136. doi: 10.1121/1.4978522
Saitis, C., and Weinzierl, S. (2019). "The semantics of timbre," in *Timbre: Acoustics, Perception, and Cognition*, eds K. Siedenburg, C. Saitis, S. McAdams, A. N. Popper, and R. R. Fay (Cham: Springer), 119–150. doi: 10.1007/978-3-030-14832-4_5
Saldaña, H. M., and Rosenblum, L. D. (1993). Visual influences on auditory pluck and bow judgments. *Percept. Psychophys.* 54, 406–416. doi: 10.3758/BF03205276
Schutz, M. (2008). Seeing music? What musicians need to know about vision. *Empir. Musicol. Rev.* 3, 83–108. doi: 10.18061/1811/34098
Schutz, M., and Lipscomb, S. (2007). Hearing gestures, seeing music: vision influences perceived tone duration. *Perception* 36, 888–897. doi: 10.1068/p5635
Seashore, C. E. (1938). *The Psychology of Music*. New York, NY: McGraw Hill. doi: 10.2307/3385515
Spence, C. (2011). Crossmodal correspondences: a tutorial review. *Attent. Percept. Psychophys.* 73, 971–995. doi: 10.3758/s13414-010-0073-7
Thompson, W. F., Graham, P., and Russo, F. A. (2005). Seeing music performance: visual in?uences on perception and experience. *Semiotica* 2005, 203–227. doi: 10.1515/semi.2005.2005.156.203

Thompson, W. F., and Russo, F. A. (2007). Facing the music. *Psychol. Sci*. 18, 756–757. doi: 10.1111/j.1467-9280.2007.01973.x

Timmers, R., and Ashley, R. (2007). Emotional ornamentation in performances of a *Handel* sonata. *Music Percept. Interdiscipl. J*. 25, 117–134. doi: 10.1525/mp.2007.25.2.117

Timmers, R., Marolt, M., Camurri, A., and Volpe, G. (2006). Listeners' emotional engagement with performances of a Scriabin étude: an explorative case study. *Psychol. Music* 34, 481–510. doi: 10.1177/0305735606067165

Van Zijl, A. G., and Sloboda, J. (2011). Performers' experienced emotions in the construction of expressive musical performance: an exploratory investigation. *Psychol. Music* 39, 196–219. doi: 10.1177/0305735610373563

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., and Levitin, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition* 101, 80–113.

Vuoskoski, J. K., Thompson, M. R., Clarke, E. F., and Spence, C. (2014). Crossmodal interactions in the perception of expressivity in musical performance. *Attent. Percept. Psychophys.* 76, 591–604. doi: 10.3758/s13414-013-0582-2

Wallmark, Z. (2019a). A corpus analysis of timbre semantics in orchestration treatises. *Psychol. Music* 47, 585–605. doi: 10.1177/0305735618768102

Wallmark, Z. (2019b). Semantic crosstalk in timbre perception. *Music Sci*. 2, 1–18. doi: 10.1177/2059204319846617

Wallmark, Z., and Kendall, R. A. (2018). "Describing sound: the cognitive linguistics of timbre," in *The Oxford Handbook of Timbre*, eds E. I. Dolan and A. Rehding (New York, NY: Oxford University Press), 9780190637224.013.14. doi: 10.1093/oxfordhb/9780190637224.013.14

Wanderley, M., and Vines, B. (2006). "Origins and functions of clarinettists' ancillary gestures," in *Music and Gesture,* eds A. Gritten and E. King (Aldershot, UK: Ashgate),165–191.

Wanderley, M. M., Vines, B. W., Middleton, N., McKay, C., and Hatch, W. (2005). The musical significance of clarinetists' ancillary gestures: An exploration of the field. *J. New Music Res.* 34, 97–113.

Wapnick, J., Mazza, J. K., and Darrow, A. A. (2000). Effects of performer attractiveness, stage behavior, and dress on evaluation of children's piano performances. *J. Res. Music Educ*. 48, 323–335. doi: 10.2307/3345367

Wapnick, J., Ryan, C., Lacaille, N., and Darrow, A. A. (2004). Effects of selected variables on musicians' ratings of high-level piano performances. *Int. J. Music Educ.* 22, 7–20. doi: 10.1177/0255761404042371

Wise, K., James, M., and Rink, J. (2017). "Performers in the practice room," in *Musicians in the Making: Pathways to Creative Performance*, eds J. Rink, H. Gaunt, and A. Williamon (Oxford: Oxford University Press), 143–163. doi: 10.1093/acprof:oso/9780199346677.003.0011

Wöllner, C. (2020). Call and response: musical and bodily interactions in jazz improvisation duos. *Musicae Scientiae* 24, 44–59. doi: 10.1177/1029864918772004

Ystad, S., Aramaki, M., and Kronland-Martinet, R. (2019). "Timbre from sound synthesis and high-level control perspectives," in *Timbre: Acoustics, Perception, and Cognition*, eds K. Siedenburg, C. Saitis, S. McAdams, A. N. Popper, and R. R. Fay (Cham: Springer), 361–389. doi: 10.1007/978-3-030-14832-4_13

Zatorre, R. J., and Halpern, A. R. (2005). Mental concerts: musical imagery and auditory cortex. *Neuron* 47, 9–12. doi: 10.1016/j.neuron.2005.06.013

**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.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Psychology

ORIGINAL RESEARCH
published: 04 October 2021
doi: 10.3389/fpsyg.2021.701761

# EEG Correlates of Middle Eastern Music Improvisations on the *Ney* Instrument

**Mohammad Yaghmour[1], Padmakumari Sarada[1], Sarah Roach[1], Ibrahim Kadar[2], Zhivka Pesheva[2], Ali Chaari[1] and Ghizlane Bendriss[1]***

[1] Premedical Division, Weill Cornell Medicine Qatar, Doha, Qatar, [2] Qatar Music Academy, Doha, Qatar

The cognitive sciences have witnessed a growing interest in cognitive and neural basis of human creativity. Music improvisations constitute an ideal paradigm to study creativity, but the underlying cognitive processes remain poorly understood. In addition, studies on music improvisations using scales other than the major and minor chords are scarce. Middle Eastern Music is characterized by the additional use of microtones, resulting in a tonal–spatial system called *Maqam*. No EEG correlates have been proposed yet for the eight most commonly used *maqams*. The *Ney*, an end-blown flute that is popular and widely used in the Middle East was used by a professional musician to perform 24 improvisations at low, medium, and high tempos. Using the EMOTIV EPOC+, a 14-channel wireless EEG headset, brainwaves were recorded and quantified before and during improvisations. Pairwise comparisons were calculated using IBM-SPSS and a principal component analysis was used to evaluate the variability between the *maqams*. A significant increase of low frequency bands theta power and alpha power were observed at the frontal left and temporal left area as well as a significant increase in higher frequency bands beta-high bands and gamma at the right temporal and left parietal area. This study reveals the first EEG observations of the eight most commonly used *maqam* and is proposing EEG signatures for various *maqams*.

**Keywords: improvisation, EEG, *Ney*, *Maqam*, prefrontal, cognition**



**Edited by:**
Oscar Casanova,
University of Zaragoza, Spain

**Reviewed by:**
Kai Yang,
PLA Information Engineering University, China
Annika Susann Wienke,
University of Bremen, Germany

***Correspondence:**
Ghizlane Bendriss
ghb2002@qatar-med.cornell.edu

**Specialty section:**
This article was submitted to Performance Science, a section of the journal Frontiers in Psychology

**Received:** 28 April 2021
**Accepted:** 14 September 2021
**Published:** 04 October 2021

**Citation:**
Yaghmour M, Sarada P, Roach S, Kadar I, Pesheva Z, Chaari A and Bendriss G (2021) EEG Correlates of Middle Eastern Music Improvisations on the Ney Instrument. Front. Psychol. 12:701761. doi: 10.3389/fpsyg.2021.701761

## INTRODUCTION

Human creativity has been the focus of thousands of studies and is still a topic of debate with considerable heterogeneity of evidence in brain research. Often, creativity is placed in the right hemisphere (Joseph, 1988; Hoppe, 1998; Demarin et al., 2016), but contradicting theories have been proposed to describe the underlying processes: (1) the dominance of right hemisphere activity (Joseph, 1988), (2) the low cortex activity (Carlsson et al., 2000; Starchenko et al., 2014), (3) the high neural connectivity (Razumnikova and Yashanina, 2014; Mayseless and Shamay-Tsoory, 2015), and (4) the prefrontal and frontal brain activation (Arden et al., 2010). Music improvisation refers to both the process and product of spontaneous creativity of music and constitutes a good model to study neural correlates of creative processes (Liu et al., 2012). Music improvisations, although spontaneous, are constructions resulting from successive decision-making processes. Nevertheless, research studies exploring the neural activity that underlies the creative process of music improvising remain scarce. Moreover, no study has explored the neural correlates of creating Middle Eastern Music, which involves the use of more than the major and minor scales.

Indeed, Middle Eastern Music is characterized by the use of *maqams* (literally "place" and "position"), which are recognized as the system of melodic tunes used in traditional Middle Eastern Music from Turkey, Azerbaijan, Israel, Iran, and all Arab countries from Middle East to North Africa. These *maqams* are not exactly the equivalent of "scales" in Western Music and are characterized by defined tonal–spatial factors, while the rhythmic-temporal features are free (Touma, 1971). On the contrary, a waltz would have a fixed rhythmic-temporal organization, while the tonal–spatial component (the melody) is free. As a result, *maqams* differ in their intervals between the first few notes and are characterized by combinations of phrases that can include microtones such as quartertones, for example. Therefore, *maqams* resonate as various melodic tunes, that can be classified based on their characteristics into more than 50 families and subfamilies. The names given to each of these *maqams* are not subject to consensus, thus creating confusion in the literature. For this reason, it is always preferable to define the *maqam* by its set of intervals signatures as described in **Table 1**.

The music improviser builds the scales by using combinations of these families and subfamilies of *maqams*. In addition, associations of *maqams* to different emotional status are well rooted in Middle Eastern cultures (Kligman, 2001; Powers, 2005). Indeed, since the 9th century, philosophers and scientists such Al-Farab (870–950) have associated different *maqams* to different sets of emotions (Naroditskaya, 2009; Yöre, 2012). For example, the *maqam* called *Rast* is believed to trigger happiness, the *maqam* called *Saba*, is associated to sadness. Kligman (1993) has published a first trial of canonization of the *maqams* in the prayers of Syrian Jews in Brooklyn, New York, during their Shabbat morning service, where he described the association of *maqams* to various prayers. In his research, Kligman (2001) details how each maqam is said to convey a unique emotion and therefore preferable to use in specific prayers. In Turkey, research and clinical trials involving the use of *maqams* have been flourishing in the past years, and *maqams* are still believed to be associated to specific emotions (Özaslan et al., 2012; Bekiroğlu et al., 2013; Tumata, n.d.). Studies that support association of these maqams to specific emotions remain extremely scarce and none of them has explored the possibility of EEG-based emotion recognition. While Self-Assessment Mannequins have become a standard for emotion detection, a growing number of studies are using the development of machine learning algorithms and brain computer interfaces to propose novel methods for emotional assessments using brain signals, that are not be dependent on individual's ability to express themselves or grasp of their mental state (Suhaimi et al., 2020; Torres et al., 2020). These studies have benefited from previous EEG explorations of neural correlates on Western Music and instruments. Lopata et al. (2017) compared improvisation-trained vs. non-improvisation-trained western musicians and showed an increase in right frontal upper alpha-band activity during more creative tasks such as improvisation, and their results suggested that creativity is probably a trainable mental state. Sasaki et al. (2019) showed in a study involving 14 male guitarists that improvisation over scale is characterized by an increase in power of theta, alpha, and beta bands in prefrontal and motor regions. Dikaya and Skirtach (2015) reported a distinguished EEG pattern in professional musicians during improvisation which was the predominant activation of the left-hemispheric cortical regions simultaneously with high interhemispheric integration in the high-frequency band.

**TABLE 1 |** Interval signatures of the eight *Maqam*s studied.

| | Ascending scale intervals |
|---|---|
| Kurd | 1/2 - 1- 1-1- 1/2- 1- 1 |
| Saba | 3/4 - 3/4 - 1/2 - 1 1/2 - 1/2-1 - 1 |
| Ajam (major) | 1 - 1- 1/2 - 1- 1-1- 1/2 |
| Nahawand (minor) | 1- 1/2- 1- 1 -1/2 - 1 1/2 - 1/2 |
| Hijaz | 1/2 - 1 1/2 - 1/2 -1- 3/4 - 3/4- 1 |
| Huzam | 3/4 - 1- 1/2 - 1 1/2 - 1/2 -1- 3/4 |
| Bayati | 3/4 - 3/4 - 1- 1- 1/2 - 1- 1 |
| Rast | 1- 3/4 - 3/4 - 1- 1- 3/4 - 3/4 |

*By convention, the first few intervals, here noted in blue, give the name to the maqam.*

Unfortunately, evidence on how humans perceive and process other modes of music such as the tonal–spatial system of *Maqams* is almost inexistent. Despite a recent study published in scientific reports by Teixeira Borges et al. (2019) that supported the role of scaling behavior of music in determining the emotions elicited, no study has yet explored and compared the neural correlates of different *maqams.* The present single subject study is the first exploration of EEG correlates to the eight most commonly used *maqams* and is believed to allow the expansion of this area of research to include Middle Eastern Music. Prior to this study, no research has focused on how *Ney* playing translates into EEG correlates.

One of the specificities of Middle Eastern Music is that it can include microtones such as quartertones, used in many *maqams,* but not all instrument can play those tones. This case study exploration relied on a performance using the *Ney* instrument, an end-blown flute, that has been played for more than 4500 years in the Middle East and is still an important component of today's Middle Eastern musical ensemble. Playing the *Ney* requires a close coordination of mouth and jaw embouchure, lip contraction, diaphragm, and breathing control, in addition to fingering control. Moreover, similarly to the Clarinet, the *Ney* is known to be one of the most difficult Middle Eastern instruments to play, not only because it does not overblow in the octave, so almost every note has its own fingering and embouchure, but also because the *Ney* comes in thirteen different sizes, each one allowing the musician to play different ranges of octaves. These characteristics of the *Ney* instrument result in the necessity to coordinate multiple cognitive processes, including integration of sensory feedback, attention, working memory, decision making, movement, etc. (Ninaus, 2011).

*Ney* players and teachers often mention that playing the *Ney* is similar to singing and refer to the *Ney* as a continuum of their body into which they blow not only air, but also meaning and "words." Although no research has previously focused on the cognitive demand of performing with this instrument, evidence supports the parallel of woodwind instruments and speech. Because woodwinds are hold at the continuity of the body and playing such instruments involves several parts of the body involved in singing (jaw, lips, tongue, mouth muscles, and breathing), it is expected that similarities are found between the cognitive demand of *Ney* and speech or singing (Zarate, 2013). Indeed, the musician's lips function as a valve might for a woodwind instrument, just as the vocal folds need to be controlled to produce various resonances that are important for timbre, loudness, shape, and sharpness (Wolfe et al., 2009).

In addition to investigating *Ney* improvisations, this single-subject study is the first to explore EEG commonalities and differences among various music modes or *maqams* of Middle Eastern Music. The EEG is an excellent method to run a first exploration on this original topic, as it gives an overview of several cognitive processes in real time. Those cognitive processes include: working memory, retrieval, focus, concentration, relaxation, planning, decision making, motor planning, and arousal. Therefore, the aims of this short study are as following: (a) explore the EEG commonalities and specificities of *Ney* performance of various modes or *maqams*; (b) identify the EEG correlates of improvisations on *Ney* and comparing it with singing or spontaneous speech; (c) identify interesting patterns or signatures of *maqams* to explore further. This EEG-based study will set the ground for further explorations finding applications in musicology, music psychology, music performance, neurofeedback training, and music-based therapeutic interventions.

## MATERIALS AND METHODS

### Materials

The EEG was recorded using the EMOTIV EPOCx headset (EMOTIV, San Francisco, CA, United States), a wireless headset that consists of 14 saline-based electrodes, recording at 14 sites according to the international 10–20 locations (AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4) and two reference electrodes CMS/DRL at P3 and P4 (**Figure 1**). The headset also includes 9-axis motion sensors and detects head movements. The EMOTIV measurement system recorded in a sequential sampling at a sampling rate of 128 Hz, for a band width of 0.16–43 Hz. EEG signals were collected at a resting state with eyes closed, then relaxed with eyes opened. Signals includes digital notch filters at 50 and 60 Hz and built-in digital fifth order Sinc filter. Although the device is able to record performance metrics and facial expressions through its sensors, only EEG signals are analyzed and presented in this present paper. The data were wirelessly recorded using EMOTIV-PRO software running on the experimental computer.

### Experimental Set-Up and Recording

The subject is a healthy right-handed professional *Ney* player in the age range 30–40 years. Subject was comfortably sitting on a chair in a classroom with dim light and no other electronic than the recording computer. Subject was asked to play a total of 24 improvisations in a row, in the following melodic tunes: *Kurd, Saba, Bayati, Hijaz, Huzam, Ajam, Nahawand*, and *Rast*. Each melodic tune or *maqam is* defined by the intervals depicted in **Table 1**. The 24 tasks were listed in front of the performer. Each improvisation was played for 1 min at three different tempos: 60, 100, and 120 bpm, before moving on to the following *maqam*.

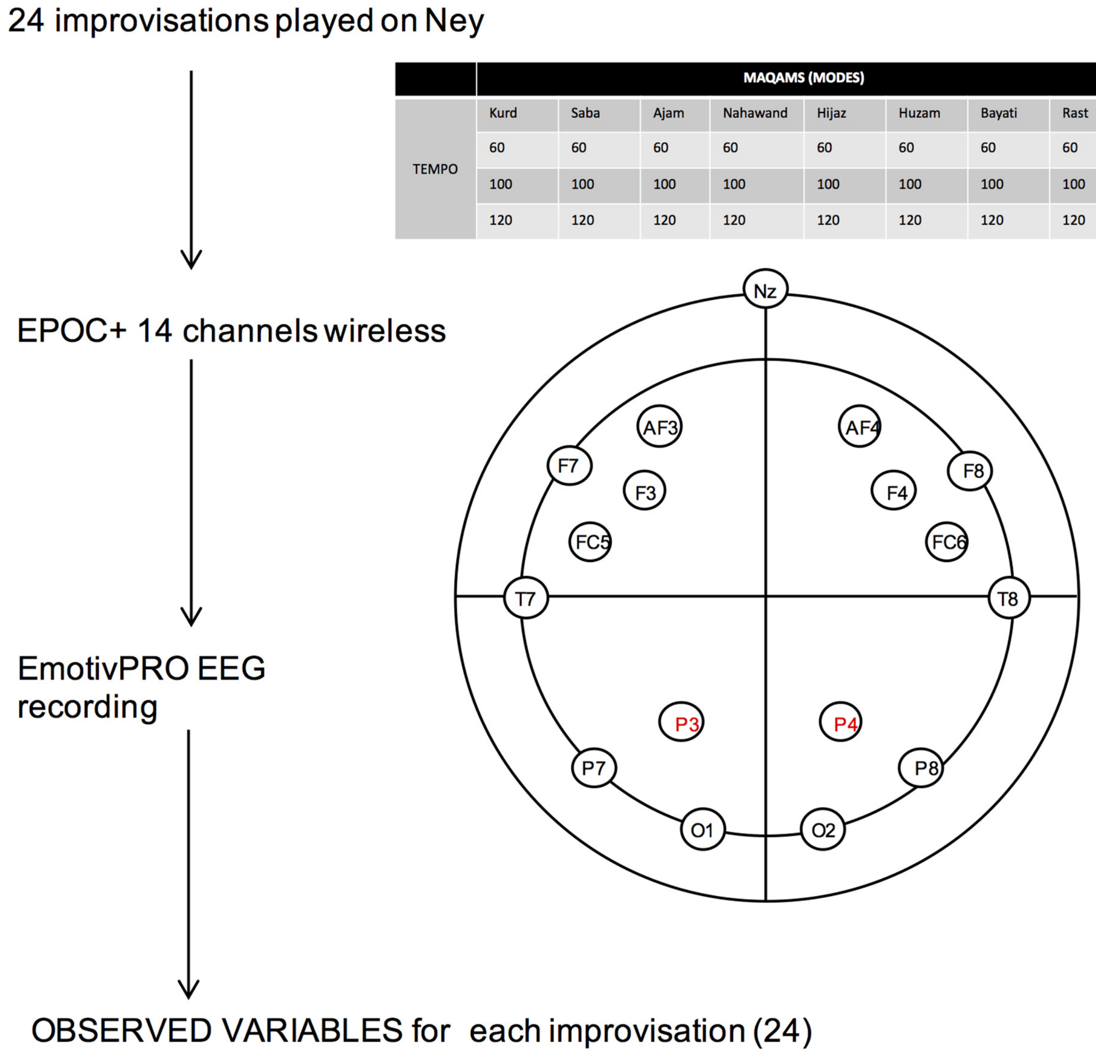
Another recording, referred to as the baseline (eyes opened, relaxed, and non-blowing status) was performed before starting the series of improvisations. The EEG was continuously recording throughout the experiment, which lasted 34 min and 32 s, including the setting and baseline recording as well as the time in between improvisations. To identify timestamps of beginning and end of each improvisation, markers were added by the experimenter using the keystroke marker function of the EMOTIV-PRO as described in their guidelines (Emotiv-Pro, 2013) and confirmed post-experimentally by visual observation of the timeline on the screen recording video of the entire experiment. A total of 25 min was included in the present analysis.

Improvisations were recorded using Audacity software and the EEG was recorded as described below. **Figure 1** shows a summary of the design. The 24 improvisational audio clips can be found in **Supplementary Files**.

### Data Processing and Analysis

Using a band pass filtering system, the EMOTIV-PRO pre-processes and extracts the power spectra of the following frequency bands: theta (4–8 Hz); Alpha (8–12 Hz); low beta (12–16 Hz); high beta (16–25 Hz); and gamma (25–45 Hz). The EMOTIV-PRO provides the power spectra for all timestamps and electrode location in a csv file. All data recorded between improvisations were eliminated, and data were organized in an excel file based on the following variables: *maqam*, tempo, frequency band, and electrodes. Markers on timestamps were used to define beginning and end of each improvisation during the recording, allowing the calculation of mean power spectra in excel for the entire duration (1 min) of each improvisation. Data were uploaded into IBM-SPSS (version 26.0) to perform descriptive and statistical analysis. For the descriptive analysis, the mean powers of each frequency band per *maqam*, regardless of the tempo, as well as the mean powers of each frequency band per tempo, regardless of *maqam*, were compared. In addition, the mean powers of each electrode separately were compared between each *maqam* regardless of tempo, and between each tempo regardless of *maqam*.

This multivariate study used a two-way analysis of variance (ANOVA) to test the hypotheses. The outcome variable was the power, which is quantitative, while all the independent variables were qualitative such as *maqam* (8 levels), frequency bands (5 levels), tempo (4 levels), and electrode (14 levels). The statistical analysis was done using SPSS to see the main effects as well as the interaction effects of the independent variables on the outcome variable. The interaction term reveals whether the effect of one independent variable on the outcome variable is the same for all values of the other independent variable. *Post hoc* tests were




| | MAQAMS (MODES) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Kurd | Saba | Ajam | Nahawand | Hijaz | Huzam | Bayati | Rast |
| TEMPO | 60 | 60 | 60 | 60 | 60 | 60 | 60 | 60 |
| | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | 120 | 120 | 120 | 120 | 120 | 120 | 120 | 120 |

| | ELECTRODES | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | AF3 | F7 | F3 | FC5 | T7 | P7 | O1 | O2 | P8 | T8 | FC6 | F4 | F8 | AF4 |
| WAVE POWER | Theta | Theta | Theta | Theta | Theta | Theta | Theta | Theta | Theta | Theta | Theta | Theta | Theta | Theta |
| | Alpha | Alpha | Alpha | Alpha | Alpha | Alpha | Alpha | Alpha | Alpha | Alpha | Alpha | Alpha | Alpha | Alpha |
| | BetaL | BetaL | BetaL | BetaL | BetaL | BetaL | BetaL | BetaL | BetaL | BetaL | BetaL | BetaL | BetaL | BetaL |
| | BetaH | BetaH | BetaH | BetaH | BetaH | BetaH | BetaH | BetaH | BetaH | BetaH | BetaH | BetaH | BetaH | BetaH |
| | Gamma | Gamma | Gamma | Gamma | Gamma | Gamma | Gamma | Gamma | Gamma | Gamma | Gamma | Gamma | Gamma | Gamma |

**FIGURE 1 |** Study design. Detail of the 24 improvisations played and the 14 electrode positions, as well as the variables observed for each improvisation.

administered for the pairwise comparisons. Separate plots were prepared for main effects and interaction effects. The assumption of homogeneity was tested using Levene's test of homogeneity. The confidence level was set at 95%. Research questions were answered here using the two-way ANOVA technique. Heatmaps produced on excel were computed for each frequency band separately and included minimum and maximum values from the mean power spectra for all *maqams*. To better visualize eventually asymmetry, cells were then reorganized into right/left and frontal/occipital for each *maqam*.

Using IBM-SPSS, Kaiser–Meyer–Olkin, and Bartlett's tests were used to measure sampling adequacy for principal component analysis (PCA) analysis. Eigen values and scree plots were computed to find the number of components that can be used. Four components were used for PCA analysis of the eight *maqams*. Detailed output of the analysis is available in **Supplementary File 3**.

# RESULTS

## Changes in Slow Oscillations at the Left Frontal and Left Temporal Area

For each *maqam*, a different improvisation was performed at 60, 100, and 120 bpm, which allowed the calculation of a mean of spectral powers for each electrode and for each *maqam*, compared to the baseline (non-blowing and non-improvising resting state). **Figure 2** suggests a highly significant increase of theta bands at the frontal left area F7 for almost all maqams ($p$-value $< 0.001$), except for the maqam Saba. The details of the means and $p$-values for all electrodes are available in **Supplementary File 1**.

Because these means were performed between different tempos, they reflect the effect of the maqam regardless of the tempo. There seem to be some topographical differences as well as power differences as the tempo increases. However, the effect of the tempo was not tested statistically, since only one improvisation per maqam and per tempo was performed. A **Supplementary Figure** that details the mean power of bands obtained at each tempo is available in **Supplementary File 3**.

Observations suggest some specificities of improvising on certain maqams. A highly significant increase of theta bands compared to the baseline is observed for the *maqam Ajam* in the frontal left area F3 ($p$-value = 0.000) (**Supplementary Table 1**) and the *maqam Hijaz* at the left temporal area T7 ($p$-value = 0.003) (**Supplementary Table 1**). In addition, we also observe at F7 a significant increase of alpha bands for the *maqam Bayati* at the left frontal area F7 ($p$-value = 0.039) and the *maqam Rast* ($p$-value = 0.05) (**Table 1**).

## Changes in Fast Oscillations at the Left Parietal and Right Temporal Area

At the left parietal area P7, the two *maqams Saba and Huzam* showed significant increase of beta-high bands ($p$-value = 0.034, $p$-value = 0.050, respectively) and gamma bands ($p$-value = 0.002, $p$-value = 0.009, respectively) (**Supplementary Table 1**). Two other maqams, *Nahawand and Hijaz*, showed a significant increase at P7, but only for gamma bands ($p$-value = 0.043, $p$-value = 0.018) (**Supplementary Table 1**).

At the right temporal area T8, almost all *maqams* were characterized by a highly significant increase in gamma bands ($p$-value = 0.000). The only maqam that did not show a significant increase in gamma at T8 was the *maqam Kurd*. For beta-high bands, the increase was significant at T8 for the maqam Saba ($p$-value = 0.001), Nawahand ($p$-value = 0.023), Hijaz ($p$-value = 0.014), Huzam ($p$-value = 0.004), and Bayati ($p$-value = 0.028), but not significant for Kurd, Ajam, and Rast (**Supplementary Table 1**).

Data suggest that, at the contrary to beta-high bands, the activity of beta-low bands is increased not only in T8 and P7, but also in F7. However, the pairwise comparisons for beta-low were not significant (**Supplementary File 4**).

## Improvisations Done on Different Maqams Induce Different EEG Signatures

Three different improvisations were played for each maqam, at three different tempos and the mean power spectra for each maqam were used to perform pairwise comparisons with the baseline and between all *maqams*. **Table 2** summarizes the significant changes observed in all maqams as compared to the baseline and suggests the existence of *maqam* EEG signatures. Improvisations on the *maqams saba* and *huzam* both showed significant increase of beta high and gamma bands at P7 and T8. However, the maqam *huzam* differed by a significant increase of theta at F7. In addition, improvisations made on maqam *saba* produced significantly higher power spectra of gamma bands than all other *maqams* ($p$-value $< 0.001$) (**Table 2**) not only at locations P7 and T8, but also at locations T7, P8, and O1 (**Supplementary Figure 4**).

The *maqam Hijaz* is the only *maqam* studied that includes an interval of 1.5 tone and also happens to be the only maqam on which improvisations induced a significant increase of theta at the left temporal region T7 (**Table 2**). Improvisations done on *Hijaz* showed the highest theta activity at almost all brain locations (**Supplementary Figure 4**). The maqam *hijaz* depicts significant changes at similar topographic area to improvisations done on maqam *nahawand*, with significant increase of theta bands at F7, beta-high at T8, and gamma at P7 and T8.

The improvisations on *Kurd* showed the lowest power of gamma across all locations and whatever the tempo, did not increase in T8 as in all other *maqams* (**Supplementary Figure 4**). The pairwise comparisons confirmed that the gamma activity of improvisations on *Kurd* in T8 are not significantly different from the baseline and significantly lower than on all other *maqams* ($p$-value $< 0.001$). The only location where significant changes were observed for improvisations on *maqam kurd* was F7 with an increase in theta bands (**Table 2**).

**Figure 3** depicts heatmaps for all the *maqams* tested. The heatmaps are organized in such a way to simulate scalp maps and show the asymmetry of brain activation for each band type at the frontal, temporal, parietal, and occipital regions of the brain. Although these do not account for the significance of the changes, these maps propose a hypothetical patterns of brainwave activation per *maqam*, to be explored further.

In addition, a principal component analysis was performed on the eight maqams to explore any eventual proximity between the *maqams*. The Bartlett's test was significant ($p$-value $< 0.05$) and concluded that our data have enough variance to be partitioned using factor analysis. **Figure 4** depicts the component plot in rotated space and suggests the presence of three clusters, with the *maqam kurd* being the closest to zero. **Figure 4** suggests the presence of three clusters of improvisations around the *maqam kurd*: (1) *bayati* and ajam, (2) *saba*, *rast*, and *hijaz*, and (3) *nahawand* and

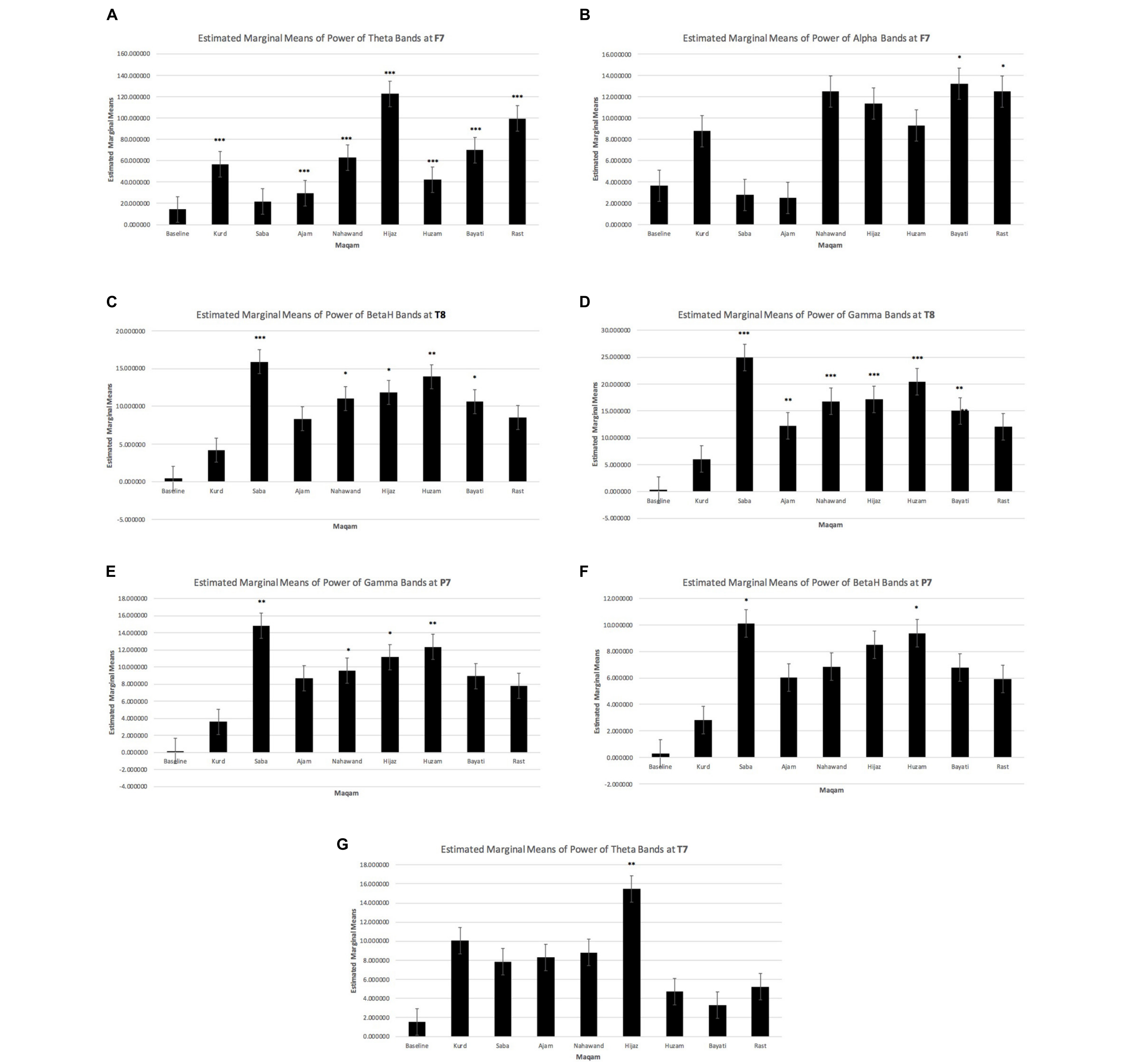



**FIGURE 2 |** Left fronto-temporal increase of low frequency bands and parietal left and temporal right increase of high frequency bands during improvisations for each maqam as compared to the baseline. **(A)** Theta bands at F7. **(B)** Alpha bands at F7. **(C)** Beta high bands at T8. **(D)** Gamma bands at T8. **(E)** Gamma bands at P7. **(F)** Beta high bands at P7. **(G)** Theta bands at T7.

*huzam*. These clusters will be interpreted and discussed in the following section.

# DISCUSSION

## EEG Correlates of Improvisations on the *Ney* Instrument

Musical improvisation is a complex musical behavior that captured attention of a growing number of scientists. This single subject case aimed at exploring human creativity by exploring improvisations *via* the prism of Middle Eastern Music and using the *Ney* flute, which were not explored before. In this single case study, we observed significant changes in the powers of the low frequency bands (theta and alpha) in the left frontal and left temporal areas —F7, F3, and T7. In addition, significant increases in the powers of the higher frequency bands (beta-high and gamma) were observed in the left parietal and right temporal areas, P7 and T8. These results align with previous studies done on western music improvisation and using other

**TABLE 2 |** Proposed EEG signatures for improvisations at eight commonly used *maqams*.

| Improvisation mode | Significant *p*-values (compared to baseline) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Theta (4–8 Hz) | | | Alpha (8–12 Hz) | Beta high (16–25 Hz) | | Gamma (25–45 Hz) | |
| | F7 | F3 | T7 | F7 | P7 | T8 | P7 | T8 |
| Kurd | 0.000 | – | – | – | – | – | – | – |
| Saba | – | – | – | – | 0.034 | 0.001 | 0.002 | 0.000 |
| Ajam | 0.001 | 0.000 | – | – | – | – | – | 0.010 |
| Nahawand | 0.000 | – | – | – | – | 0.023 | 0.043 | 0.000 |
| Hijaz | 0.000 | – | 0.003 | – | – | 0.014 | 0.018 | 0.000 |
| Huzam | 0.000 | – | – | – | 0.050 | 0.004 | 0.009 | 0.000 |
| Bayati | 0.000 | – | – | 0.039 | – | 0.028 | – | 0.002 |
| Rast | 0.000 | – | – | 0.053 | – | – | – | 0.010 |

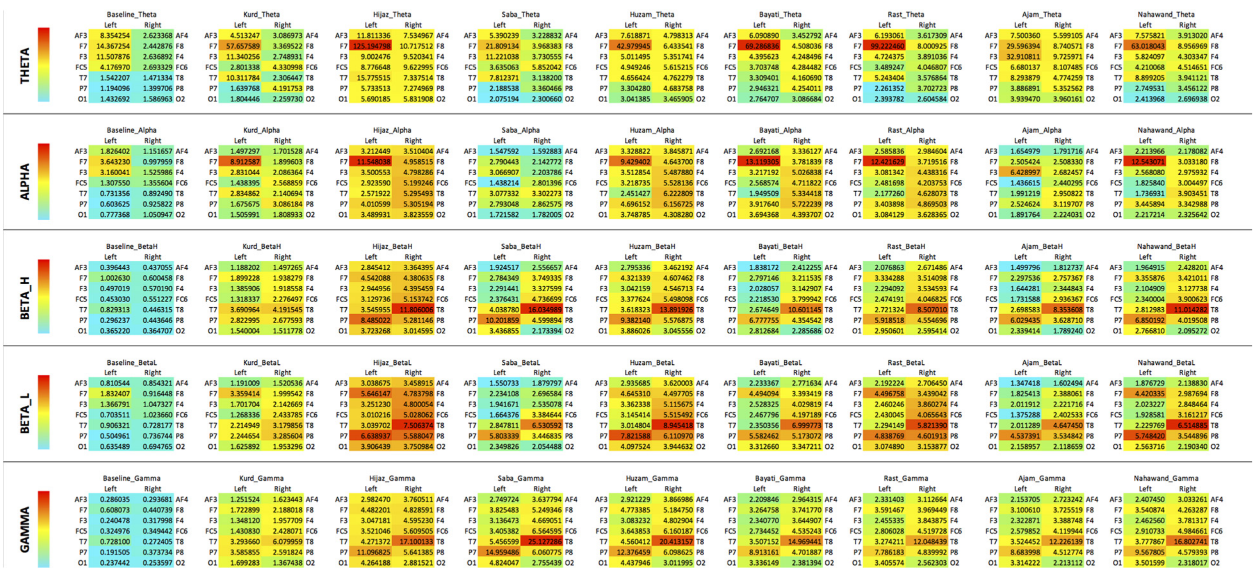



**FIGURE 3 |** Heatmaps of power spectra of frequency bands at 14 electrode sites for each *maqam*. Heatmap were computed for each frequency band type across all *maqams*. Power spectra are organized in a way to simulate scalp map (frontal on top, occipital on bottom, left, and right hemispheres).

instruments. Indeed, Dikaya and Skirtach (2015) showed in a cohort of 136 musicians, amateur, and professionals, that professional musicians were distinguished by a predominant activation of the left hemisphere, with a simultaneous integration between both hemispheres in the higher frequency bands, which is similar to what we have observed in this experiment done on a professional musician. Other studies showed this increase of the EEG spectral power at the prefrontal cortical area, when playing guitar (Sasaki et al., 2019), rock music (Tachibana et al., 2019), and piano (Pinho et al., 2014). The specific increase of frontal theta is facilitated by emotions, concentration, and mental tasks (Aftanas and Golocheikine, 2001; Marcuse et al., 2016; Katahira et al., 2018). While EEG allows an extremely good temporal resolution, it does not provide a good spatial resolution and as a result, studies using EEG do not inform on accurate location of the observed data. Some have used other techniques to explore improvisation cognitive demand such as Limb and Braun (2008), who used fMRI on jazz pianists improvising novel melodies using pre-existing chord patterns (equivalent to our experiment with the *maqams*) and highlighted an activation in lateral and prefrontal area. In a study of Faber (2014), an exploration involving 32 channels during improvisations in various settings, a strong activity of the frontal and central–temporal regions were also observed suggesting improvised music is a communicative medium. Itthipuripat et al. (2013) showed that frontal theta constitutes a signature of a successful working memory manipulation, as theta bands are also known to facilitate the encoding of temporary episodic memories into long-term memory (Baddeley, 2003; Itthipuripat et al., 2013). Interestingly, at the contrary to all other *maqams*, improvisations done on the maqam *Saba* did not induce significant increase of theta bands at the F7 frontal area, suggesting the possible use of different cognitive processes for improvisations done on this *maqam*.

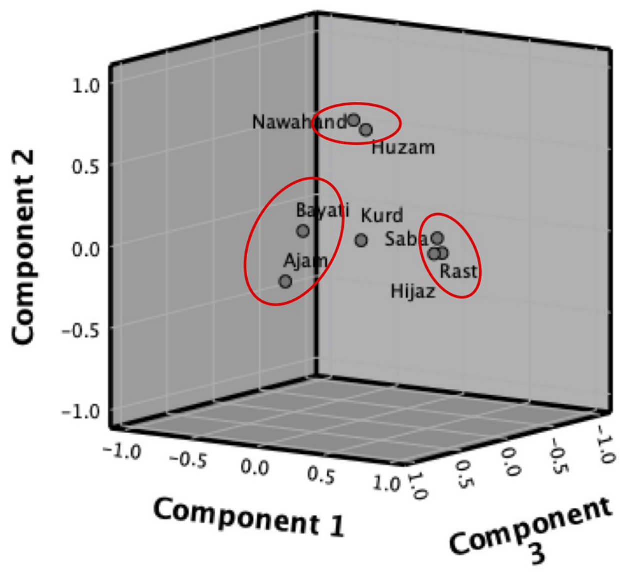



**FIGURE 4 |** Principal component analysis plot of the eight *maqam* commonly used in Middle Eastern Music. Detailed output available in **Supplementary File 5**.

The ability to improvise is one of the highest levels of musical achievement, as it requires from the improviser to master the music language necessary to spontaneously compose original music. The cognitive processes underlying improvisations have been compared to spontaneous speech (Peretz et al., 2015). Therefore, it is thought to be a powerful mean to express oneself and communicate with others. Nevertheless, these processes are still poorly understood. The data obtained through this single subject study corroborate the existing studies showing an increase in the EEG spectral power at the pre-frontal area (Sasaki et al., 2019), suggesting a strong role for the region F7 or left frontal area of the brain, involved in controlling language-related movement and executive functions such as planning, organizing, and self-monitoring. We have found that most improvisations in this present study resulted in significant increase of beta-high and gamma activity at parietal left P7 and temporal right T8 area. Wan et al. (2014) explored the EEG signals of a pianist on improvisations using Western Music scales have also suggested that frontal, parietal, and temporal regions play a key role in differentiating improvisations from playing composed music. Interestingly, they noted the strong involvement of T8 but not P7 (Wan et al., 2014). While T8 represents the right temporal lobe, close to the amygdala and hippocampus and involved in auditory processing and music appreciation (Stewart et al., 2006), P7 is involved in logical or verbal understanding, word recognition during auditory processing. We emitted the hypothesis that EEG signals on woodwind instrument such as the *Ney* would be similar to those during spontaneous speech. The strong involvement of F7, T7, and P7 that we noted for *Ney* improvisations were also spotted by Chengaiyan et al. (2020) in a study of speech imagery (or imagining of speaking), where they observed that left frontal and left temporal electrodes (where T5 correspond to our P7) were activated for speech and speech imagery processes (Chengaiyan et al., 2020).

## EEG Signatures of *maqams* in Middle Eastern Music

Musicology research studies on *maqams* are extremely limited and almost exclusively run in Turkey and Israel. However, no study has explored yet the neural correlates of these *maqams* in the same way studies have explored the EEG correlates of major and minor chords (Petsche et al., 1996; Virtala et al., 2013). When having a closer look at the EEG signals elicited by improvisations at each maqam, we observed that each *maqam* was characterized by a topographically unique combination of significant electroencephalographic changes, suggesting the existence of what we would call *maqam* EEG signatures, as presented in **Table 2**. Jenni et al. (2017) explored processing of western tonal music major and minor EEG signals and observed an increased activity of higher frequency bands for the minor scale. Interestingly, in the present study, we observed the same pattern between improvisations on *Nahawand* (minor) and the those on *Ajam* (major). But because Middle Eastern Music is also using other tones than the whole tone and semi-tone, a greater number of scales is possible yielding the maqams families studied. It is therefore of great importance that we understand the interactions between the *maqams*, in order to build future studies on emotional correlates of these *maqams*. Therefore, we have opted for a principal component analysis to extract eventual clusters that could help us understand which feature in these improvisations seem to play a key role. The principal component analysis suggested the presence of three clusters of improvisations around the *maqam Kurd*: (1) *Bayati* and *Ajam*, (2) *Saba*, *Rast*, and *Hijaz*, and (3) *Nahawand* and *Huzam*. The main challenge in studying the *maqam* system resides in the terminology used, as these names actually refer to the first sets of intervals described in **Table 1**.

The term *jins* (or plural *ajnas*) refers to the building blocks of a *maqam* scale, which always has a lower and an upper *jins*. By convention, *maqams* are classified based on their lower *jins*, and the first note of the second *jins* is called the dominant and is the second most important note after the tonic.

By comparing **Table 1** and the PCA plot, we understand that the clusters we see could correspond to the sets of intervals they have in common. Indeed, *Ajam* – characterized by the intervals 1-1-1/2 – is also present in within the unfolding of the *Bayati* scale 3/4 - 3/4 - 1- 1- 1/2 - 1- 1. The maqam *Saba* includes some *Kurd* (1/2 - 1- 1) and some *Hijaz* (1/2 - 1 1/2 - 1/2), while the *maqam Hijaz* includes some *Rast (*1- 3/4 - 3/4), possibly explaining the proximity of these three *maqams* on the PCA. As for the *Kurd* set of intervals 1/2 - 1- 1 is retrieved in *Kurd*, *Saba*, *Ajam*, *Bayati*, *Nahawand* and could explain why this particular *maqam* is not fitting in any of the clusters.

These interactions between *maqams* are well known by professional musicians and structural proximity of these scales are used to create improvisations that increase in complexity by mixing closely related *maqams*. Dikaya and Skirtach (2015) suggested that high-frequency coherent connections increased with the level of difficulty of the musical improvisation.

This result is important because it highlights the importance of considering intervals, tones, and microtones, in studies on processing of music and emotional correlates.

## Implications for Studies on EEG-Based Detection of Emotions

Gu (2014) in her book, "Cultural history of Arabic Language," mentioned the emotions that are commonly associated to each of the presently studied *maqam*. *Kurd* evoking freedom, romance, and gentleness; *Saba* evoking sadness or pain; *Ajam* evoking strength; *Nahawand* evoking drama and emotional extremes; *Hijaz* evoking desert, solitude, and enchantment; *Huzam* evoking old days; *Bayati* evoking femineity, joy, and vitality; and *Rast* evoking pride and power. Although there is no consensus among musicologists on what the mood each *maqam* is associated with, it is surprising to see to which extent these various mood-*maqam* associations have been passed on throughout history without any scientific methodology or validated emotional assessment to support these claims (Kligman, 2001; Powers, 2005; Gu, 2014; Marks, 2019).

EEG signals play an important role in research on human emotions, which in turn are involved in cognitive processes such as memory, learning, and decision-making (Zhang et al., 2020). There are converging evidence from the literature that gamma bands, in addition to being associated with focus and concentration, are associated with negative emotions such as sadness and worry (Oathes et al., 2008; Yang et al., 2020). A recent study using functional network analysis by Yang et al. (2020) exposed native Chinese individuals to 180 pictures selected from the Chinese Affective Picture System (CAPS) (Lu et al., 2005) and recorded their EEG responses as well as their self-assessment Manikin rating scales (Morris, 1995). While no significant difference in brain network were found at low frequency bands, significant differences were observed between positive and negative emotions In the high gamma bands. They concluded on the existence of neural signatures for emotional states in the high gamma bands, particularly against negative stimuli (Yang et al., 2020).

Interestingly, the *Saba maqam is* consistently called by musicians of the Middle East the "*sad maqam*" and is also depicting the highest gamma spectral power in this case study. This result represents the first piece of EEG evidence that could eventually support the historical claim that *Saba maqam* is associated to negative emotions. However, since we did not perform any emotional assessment during the experiment, further studies are needed to conclude on this. In addition, as this results from a single subject study, more explorations on larger and culturally diverse samples need to be performed.

*The Hijaz maqam* showed the highest theta activity across all brain locations. The theta bands have been shown to increase during sleep, deep meditation, and spiritual awareness. Some have described the *Hijaz maqam* as being "snake charming music." Powers (2005, p. 9) writes that "this *Maqam* is associated with the lonely treks of the camel caravans and with fascination and enchantment." Lee et al. (2018) have reviewed the neural signals underlying several meditation practices including focused attention, open-monitoring, transcendental attention and loving kindness meditation. It is only during focused attention that a significant increase of theta was observed across both anterior and posterior parts of the brain. While further characterization of the oscillatory activities during improvisation on *Hijaz* are necessary, the present data suggest similarities between improvisations on *Hijaz* and focused attention practices.

Music and perception have been substantially researched in the field of music psychology. Through existing neuronal measuring methodologies (EEG, fMRI, and MEG), studies shed light on brain functionality and mechanisms involved when passively listening or playing an instrument (Limb and Braun, 2008; Virtala et al., 2013; Almudena et al., 2021). It is generally agreed that music stimulates a combination of different processes including short-term memory, the nature of different emotions produced by music, concentrations, pleasure and non-pleasure, and self-reflection (Almudena et al., 2021). Studies of Almudena et al. (2021) and Virtala et al. (2013) support that different mechanisms are involved in the perception of major vs. minor, and consonant vs. dissonance chords in infants, adults and school-age children, which correlate with findings in the present study. Furthermore, Limb's study (Limb and Braun, 2008) tries to break the code of spontaneous music performance, with the assumption that this creative music process is predicative on novel combination of ordinary mental processes. It was hypothesized that short term memory would be associated with hierarchical top-down subtle changes in other systems, such as sensorimotor area and limbic structures used to regulate memory and emotional tone. The study suggests that the prefrontal cortex is of critical importance for processes which include self-reflection, and sensory processes as integral component.

To conclude, this present case explored the power of low and high frequency bands across 14 cortical locations during Middle Eastern Music improvisations played using the *Ney* instrument using the tonal–spatial system of *Maqams*. This case provides further support to the already published studies on the important role during musical improvisations of the left hemisphere with the significant increase of the low frequency bands at the frontal and temporal left area, as well as the more integrated activity in both hemispheres at higher frequency bands.

In addition, this case introduces, for the first time in neuromusicology, the question of EEG *Maqam* signatures, where signatures found seems to follow the maqam's intervals signatures, supporting the necessity of referring to *Maqams* by their intervals rather than their names.

Single case studies are being used by many across multiple sessions to obtain consistent results using brain activity measurements (Farrugia et al., 2021).

Finally, this case's results can be used as ground study to design further studies, including: (1) establishing the cognitive demand for each mode or *maqam* on professional's vs. amateur musicians or improvisation vs. composed music, (2) exploring the listener's and the performer's perception of intended emotions by using the present recordings of various *maqams* and combination of self-assessment mannequin and EEG-based emotion detection, and

(3) increasing the sample size in order to confirm the proposed correlates and explore the possibilities of neurofeedback training to improve performance on every *maqam*.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

GB worked on the conceptualization, methodology, analysis, investigation, interpretation and writing, and supervised all the work. MY worked on the methodology, ran the investigation, managed software processing of data, did the analysis, and wrote parts of the manuscript. PS worked on the methodology, managed software processing of data, did the analysis, and wrote parts of the manuscript. SR participated in running the investigation, writing, reviewing, and editing. IK ran the investigation, participated in interpretation, writing, reviewing, and editing. ZP and AC participated in writing, reviewing, and editing. All authors have read and agreed to the published version of the manuscript.

## FUNDING

The publication of this article was funded by the Weill Cornell Medicine – Qatar Distributed eLibrary.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2021.701761/full#supplementary-material

**Supplementary File 1 |** Mean power spectra of theta, alpha, beta-low, beta-high, and gamma at each of the 14 electrodes: pairwise comparisons between baseline and each *maqam*.

**Supplementary File 2 |** Mean power spectra of theta, alpha, beta-low, beta-high, and gamma bands at the 14 electrode sites and for all maqams.

**Supplementary File 3 |** Detail of the power spectra of theta, alpha, beta-low, beta-high, and gamma bands at the 3 different tempos tested, at the 14 electrode sites and for all maqams.

**Supplementary File 4 |** Mean power spectra of theta, alpha, beta-low, beta-high, and gamma bands at the 14 electrodes organized by *maqam*.

**Supplementary File 5 |** Detailed output of principal component analysis.

**Supplementary Audios |** Improvisations played are available in **Supplementary Files**.

## REFERENCES

Aftanas, L. I., and Golocheikine, S. A. (2001). Human anterior and frontal midline theta and lower alpha reflect emotionally positive state and internalized attention: high-resolution EEG investigation of meditation. *Neurosci. Lett.* 310, 57–60. doi: 10.1016/S0304-3940(01)02094-8

Almudena, G., Manuel, S., and Julián Jesús, G. (2021). "EEG analysis during music perception," in *Electroencephalography-From Basic Research to Clinical Applications*, ed. H. Nakano (London: IntechOpen), doi: 10.5772/intechopen.94574

Arden, R., Chavez, R. S., Grazioplene, R., and Jung, R. E. (2010). Neuroimaging creativity: a psychometric view. *Behav. Brain Res.* 214, 143–156. doi: 10.1016/j.bbr.2010.05.015

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4, 829–839. doi: 10.1038/nrn1201

Bekiroğlu, T., Ovayolu, N., Ergün, Y., and Ekerbiçer, H. Ç (2013). Effect of Turkish classical music on blood pressure: a randomized controlled trial in hypertensive elderly patients. *Complement. Ther. Med.* 21, 147–154. doi: 10.1016/j.ctim.2013.03.005

Carlsson, I., Wendt, P. E., and Risberg, J. (2000). On the neurobiology of creativity. Differences in frontal activity between high and low creative subjects. *Neuropsychologia* 38, 873–885. doi: 10.1016/S0028-3932(99)00128-1

Chengaiyan, S., Retnapandian, A. S., and Anandan, K. (2020). Identification of vowels in consonant–vowel–consonant words from speech imagery based EEG signals. *Cogn. Neurodyn.* 14, 1–19. doi: 10.1007/S11571-019-09558-5

Demarin, V., Bedekoviæ, M. R., Puretiæ, M. B., and Pašiæ, M. B. (2016). Arts, brain and cognition. *Psychiatr. Danub.* 28, 343–348.

Dikaya, L. A., and Skirtach, I. A. (2015). Neuropsychology neurophysiological correlates of musical creativity: the example of improvisation. *Psychol. Russia* 8, 84–97. doi: 10.11621/pir.2015.0307

Emotiv-Pro (2013). *Emotiv Gitbook*. San Francisco, CA: Emotiv-Pro. Available online at: https://emotiv.gitbook.io/emotivpro/event_markers/keystroke_markers (accessed August 3, 2021).

Faber, S. (2014). *The Communicative Processes Of Musicians Engaged in Synchronous Play*. Finland: University of Jyväskylä.

Farrugia, N., Lamouroux, A., Rocher, C., Bouvet, J., and Lioi, G. (2021). Beta and theta oscillations correlate with subjective time during musical improvisation in ecological and controlled settings: a single subject study. *Front. Neurosci.* 15:626723. doi: 10.3389/FNINS.2021.626723

Gu, S. (2014). *A Cultural History of the Arabic Language*. Jefferson: McFarland & Company.

Hoppe, K. D. (1998). Hemispheric specialization and creativity. *Psychiatr. Clin. North Am.* 11, 303–315.

Itthipuripat, S., Wessel, J. R., and Aron, A. R. (2013). Frontal theta is a signature of successful working memory manipulation. *Exp. Brain Res.* 224, 255–262. doi: 10.1007/s00221-012-3305-3

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

Joseph, R. (1988). The right cerebral hemisphere: emotion, music, visual-spatial skills, body-image, dreams, and awareness. *J. Clin. Psychol.* 44, 630–673. doi: 10.1002/1097-4679(198809)44:5

Katahira, K., Yamazaki, Y., Yamaoka, C., Ozaki, H., Nakagawa, S., and Nagata, N. (2018). EEG correlates of the flow state: a combination of increased frontal theta and moderate frontocentral alpha rhythm in the mental arithmetic task. *Front. Psychol.* 9:300. doi: 10.3389/fpsyg.2018.00300

Kligman, M. (1993). Modes of prayer: the canonization of the maqamat in the prayers of the syrian jews in brooklyn, New York on jstor. *JSTRO Proc. World Congr. Jew. Stud.* 2, 259–266., Kligman, M. (2001). The bible, prayer, and maqam: extra-musical associations of syrian jews. *Ethnomusicology* 45:443. doi: 10.2307/852866

Lee, D. J., Kulubya, E., Goldin, P., Goodarzi, A., and Girgis, F. (2018). Review of the neural oscillations underlying meditation. *Front. Neurosci.* 12:178. doi: 10.3389/fnins.2018.00178

Limb, C. J., and Braun, A. R. (2008). Neural substrates of spontaneous musical performance: an 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

Lopata, J. A., Nowicki, E. A., and Joanisse, M. F. (2017). Creativity as a distinct trainable mental state: an EEG study of musical improvisation. *Neuropsychologia* 99, 246–258. doi: 10.1016/j.neuropsychologia.2017.03.020

Lu, B., Hui, M., and Yu-Xia, H. (2005). The development of native chinese affective picture system–a pretest in 46 college Students. *Chin. Ment. Health J.* 19, 719–722.

Marcuse, L. V., Fields, M. C., and Yoo, J. (2016). "The normal adult EEG," in *Rowan's Primer of EEG*, eds L. V. Marcuse, M. C. Fields, and J. Yoo (London: Elsevier), 39–66. doi: 10.1016/b978-0-323-35387-8.00002-0

Marks, E. (2019). Arab musical culture in a jewish liturgy. *Musica Judaica* 22, 66–86.

Mayseless, N., and Shamay-Tsoory, S. G. (2015). Enhancing verbal creativity: modulating creativity by altering the balance between right and left inferior frontal gyrus with tDCS. *Neuroscience* 291, 167–176. doi: 10.1016/j.neuroscience.2015.01.061

Morris, J. D. (1995). Observations: SAM: the self-assessment manikin an efficient cross-cultural measurement of emotional response 1. *J. adv. Res.* 35:63.

Naroditskaya, I. (2009). The philosophy of music by abu nasr muhammad al-farabi (review). *Asian Music* 40. 133–137. doi: 10.1353/amu.0.0030

Ninaus, P. (2011). *The Fingering Logic and Performing of Woodwind Instruments*. Sacramento, CA: GRIN.

Oathes, D. J., Ray, W. J., Yamasaki, A. S., Borkovec, T. D., Castonguay, L. G., Newman, M. G., et al. (2008). Worry, generalized anxiety disorder, and emotion: evidence from the EEG gamma band. *Biol. Psychol.* 79, 165–170. doi: 10.1016/j.biopsycho.2008.04.005

Özaslan, T. H., Serra, X., and Arcos, J. L. (2012). "Characterization of embellishments in ney performances of makam music in Turkey". in *Proceedings of the 13th International Society for Music Information Retrieval Conference, ISMIR 2012*, *November* 2014, Porto, 13–18.

Peretz, I., Vuvan, D., Lagrois, M. É, and Armony, J. L. (2015). Neural overlap in processing music and speech. *Philos. Trans. R. Soc. B* 370:20140090. doi: 10.1098/rstb.2014.0090

Petsche, H., Von Stein, A., and Filz, O. (1996). EEG aspects of mentally playing an instrument. *Cogn. Brain Res.* 3, 115–123. doi: 10.1016/0926-6410(95)00036-4

Pinho, A. L., de Manzano, Ö, Fransson, P., Eriksson, H., and Ullén, 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

Powers, C. (2005). *Arabic Musical Scales: Basic Maqam Notation*. Toronto, ON: G. L. Design.

Razumnikova, O. M., and Yashanina, A. A. (2014). Personality specific differences in EEG reactivity on convergent and divergent thinking. *Int. J. Psychophysiol.* 94:160. doi: 10.1016/j.ijpsycho.2014.08.702

Sasaki, M., Iversen, J., and Callan, D. E. (2019). Music improvisation is characterized by increase EEG spectral power in prefrontal and perceptual motor cortical sources and can be reliably classified from non-improvisatory performance. *Front. Hum. Neurosci.* 13:435. doi: 10.3389/fnhum.2019.00435

Starchenko, M. G., Kireev, M. V., and Medvedev, S. V. (2014). Brain organization in creative thinking. *Int. J. Psychophysiol.* 94:160. doi: 10.1016/j.ijpsycho.2014.08.703

Stewart, L., Von Kriegstein, K., Warren, J. D., Griffiths, T. D., and Griffiths, T. (2006). Music and the brain: disorders of musical listening. *Brain* 129, 2533–2553. doi: 10.1093/brain/awl171

Suhaimi, N. S., Mountstephens, J., and Teo, J. (2020). EEG-based emotion recognition: a state-of-the-art review of current trends and opportunities. *Comput. Intell. Neurosci.* 2020:8875426. doi: 10.1155/2020/8875426

Tachibana, A., Noah, J. A., Ono, Y., Taguchi, D., and Ueda, S. (2019). Prefrontal activation related to spontaneous creativity with rock music improvisation: a functional near-infrared spectroscopy study. *Sci. Rep.* 9, 1–13. doi: 10.1038/s41598-019-52348-6

Teixeira Borges, A. F., Irrmischer, M., Brockmeier, T., Smit, D. J. A., Mansvelder, H. D., and Linkenkaer-Hansen, K. (2019). Scaling behaviour in music and cortical dynamics interplay to mediate music listening pleasure. *Sci. Rep.* 9:17700 doi: 10.1038/s41598-019-54060-x

Torres, E. P., Torres, E. A., Hernández-Álvarez, M., and Yoo, S. G. (2020). EEG-based BCI emotion recognition: a survey. *Sensors* 20, 1–36. doi: 10.3390/S20185083

Touma, H. H. (1971). The maqam phenomenon: an improvisation technique in the music of the middle east. *Ethnomusicology* 15:38. doi: 10.2307/850386

Tumata (n.d.). *Turkish Music Tonalities (Makam) and their Effects on Human Beings | Tumata*. Available online at: https://tumata.com/en/turkish-music-therapy/turkish-tonalities-and-their-effects (accessed April 18, 2021).

Virtala, P., Huotilainen, M., Partanen, E., Fellman, V., and Tervaniemi, M. (2013). Newborn infants' auditory system is sensitive to Western music chord categories. *Front. Psychol.* 4:492. doi: 10.3389/fpsyg.2013.00492

Wan, X., Crüts, B., and Jensen, H. J. (2014). The causal inference of cortical neural networks during music improvisations. *PLoS One* 9:e112776. doi: 10.1371/JOURNAL.PONE.0112776

Wolfe, J., Garnier, M., and Smith, J. (2009). Vocal tract resonances in speech, singing, and playing musical instruments. *HFSP J.* 3, 6–23. doi: 10.2976/1.2998482

Yang, K., Tong, L., Shu, J., Zhuang, N., Yan, B., and Zeng, Y. (2020). High gamma band EEG closely related to emotion: evidence from functional network. *Front. Hum. Neurosci.* 14:89. doi: 10.3389/fnhum.2020.00089

Yöre, S. (2012). Maqam in music as a concept, scale and phenomenon. *Zeitschrift Für Die Welt Der Türken* 4, 267–286.

Zarate, J. M. (2013). The neural control of singing. *Front. Hum. Neurosci.* 7:237. doi: 10.3389/fnhum.2013.00237

Zhang, L., Gan, J. Q., Zhu, Y., Wang, J., and Wang, H. (2020). EEG source-space synchrostate transitions and Markov modeling in the math-gifted brain during a long-chain reasoning task. *Hum. Brain Mapp.* 41, 3620–3636. doi: 10.1002/hbm.25035

**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.

frontiers in Psychology

ORIGINAL RESEARCH
published: 08 October 2021
doi: 10.3389/fpsyg.2021.722082

# Social Support as a Facilitator of Musical Self-Efficacy

*Santos Orejudo[1], Francisco Javier Zarza-Alzugaray[1]*, Oscar Casanova[1]* and Gary Edward McPherson[2]*

*[1] Faculty of Education, University of Zaragoza, Zaragoza, Spain, [2] Melbourne Conservatorium of Music, University of Melbourne, Melbourne, VIC, Australia*

OPEN ACCESS

**Edited by:**
*Gunter Kreutz,*
*Carl von Ossietzky Universität Oldenburg, Germany*

**Reviewed by:**
*Graham Frederick Welch,*
*University College London, United Kingdom*
*Clemens Wöllner,*
*University of Hamburg, Germany*
*Esther H. S. Mang,*
*Hong Kong Baptist University, Hong Kong, SAR China*

***Correspondence:**
*Francisco Javier Zarza-Alzugaray*
*fjzarza@unizar.es*
*Oscar Casanova*
*ocasanov@unizar.es*

**Specialty section:**
*This article was submitted to Performance Science, a section of the journal Frontiers in Psychology*

**Received:** *08 June 2021*
**Accepted:** *17 September 2021*
**Published:** *08 October 2021*

**Citation:**
*Orejudo S, Zarza-Alzugaray FJ, Casanova O and McPherson GE (2021) Social Support as a Facilitator of Musical Self-Efficacy. Front. Psychol. 12:722082. doi: 10.3389/fpsyg.2021.722082*

Previous research has shown that musical self-efficacy is one of the predictors of academic achievement, but few studies have analyzed the function of social support in the construction of musical self-efficacy. In this study we analyze the relationship between three sources of support perceived by music students – parents, teachers, and peers – and their influence on levels of self-efficacy for learning and for public performance. We analyze three groups of students under the hypothesis that relationships among those variables can vary with age and the level of education. A total of 444 students enrolled in six Spanish music schools, two music universities, and four advanced music schools, completed the Social Support Scale for Music Students, as well as the General Musical Self-Efficacy Scale. Results reveal significant relationships among the aforementioned variables, with considerable variation according to academic level. For the youngest students enrolled in advanced music schools (*conservatorios profesionales*), the role of parents and teachers was crucial, especially for predicting self-efficacy for learning, which, in turn, is the best predictor of self-efficacy for public performance. For the 16–18-year-olds enrolled in the same advanced music schools, their peers play a particularly relevant role in reinforcing their self-efficacy for learning. Social support had a negligible influence on the self-efficacy of university-level students, but they did experience a strong relationship between self-efficacy for learning, on the one hand, and public performance, on the other. We interpret these results in view of potential long-term careers in music, relating them with a series of different agents.

**Keywords: social support, musical self-efficacy, music students, academic level, age, parents, teachers**

## INTRODUCTION

The theory of self-efficacy is one of the most relevant theoretical contributions to the study of human behavior. Originally defined by Bandura (1997, 2006), in his Socio-Cognitive Theory, self-efficacy is seen as "the conviction that one can successfully execute the behavior required to produce the outcome" (Bandura, 1997, p. 79). In other words, self-efficacy refers to the beliefs people hold about the extent to which they can complete a task in a particular situation: for example, in the area of music, which is the focus in this study (McPherson and McCormick, 2006). This approach to self-efficacy lends importance to the situational context and the specific domain in which we are analyzing a subject's behavior, although other models can likewise be applied. Thus, for example, the Theory of General Self-Efficacy (Baessler and Schwarcer, 1996; Schwarzer and Jerusalem, 2010) postulates general self-efficacy as an attitude that can be adopted to face a series of stressors in a variety of different environments. Orejudo et al. (2017) have used that approach to define a profile of personal vulnerability in the face of performance anxiety within Barlow's anxiety model (Barlow, 2000).

In both cases, the general approach and the specific focus are both relevant, since they each have the potential for explaining behaviors, cognitions, and emotional responses. A situational approach helps approximate performance in educational contexts (Zimmerman, 2000; Pajares and Schunk, 2001), leading to the postulation of the self-efficacy hypothesis (SEH), which has been used to help explain students' choice of goals, the efforts they invest along with the strategies they employ to reach them, and the validation processes that serve as feedback for their study progress (Panadero and Alonso-Tapia, 2014). Music has been one of the concrete areas in which this theoretical framework has been developed and applied, as we expound below.

## Musical Self-Efficacy

Since the year 2000, McPherson and his collaborators started conducting studies on self-efficacy in the area of musical activity (McPherson and McCormick, 2000, 2006; McCormick and McPherson, 2003). Their first studies attempted to prove the relationship between self-efficacy and various levels of musical achievement. Subsequently, Papageorgi et al. (2007) postulated that self-efficacy is a key component in helping us to understand the training process undergone by music students, who, as they learn, need to develop skills to help them face a performance situation in front of an audience, and manage their performance anxiety. Along similar lines, Upitis et al. (2017c, p. 413) have described self-efficacy as "one's beliefs in one's abilities to achieve goals and complete tasks." The tasks music students are required to accomplish are typically associated with performing in front of an audience: either in examinations, or in concerts. To be successful, one must have acquired the technical skills needed to prepare for and master repertoire to be performed. But apart from requiring the gradual mastery of those skills, musical training likewise necessitates the development of motivational abilities that enable the individual to persist in the task, especially when coping with difficulties and setbacks. In this way, self-efficacy is situated within models of self-regulated learning (Zimmerman, 2000; Varela et al., 2016).

Based on the above, it is evident that researchers who have studied the role played by self-efficacy in music training seek to understand: (1) the essential relationship between self-efficacy and musical performance; (2) the inclusion of self-efficacy in a theoretical framework that organizes data and allows researchers to structure their findings, making self-efficacy easy to evaluate; and (3) the particular types of factors that exert an influence on musical self-efficacy.

As mentioned above, one of the most relevant findings that can help us grasp the usefulness of self-efficacy in musical training is its relationship with public performance. Studying results from a sample of 332 music students, McCormick and McPherson (2003) reported that musical self-efficacy was the best predictor of their externally evaluated excellence in a public performance. Another study by the same team (McPherson and McCormick, 2006) replicated that result using an entirely different sample of 686 Australian students. Again, self-efficacy was shown to have a high predictive value for performance, thus confirming the significant role played by musical practice and by the cognitive strategies applied by the students. Subsequent studies, including Hewitt (2015) and Miksza and Tan (2015), have confirmed this relationship between self-efficacy and public performance.

Since self-efficacy is defined as a domain-specific construct, it can be adapted to other areas, such as music education (Hendricks, 2014). Several research teams have adapted a series of tools to two basic processes in that field: the study phase (preparation of repertoire during the phase of technical competency development), and the specific situation of performing in public in front of an audience. To examine both dimensions, Ritchie and Williamon (2007, 2011) devised the *General Musical Self-Efficacy Scale*, which was based on the *Self-Efficacy Scale* developed by Sherer and Adams (1983). By restricting itself to these two components, self-efficacy for learning, and self-efficacy for public performance, this model parsimoniously represents the most relevant tasks a student or a musician must accomplish to achieve a successful public performance.

Other authors have proposed additional areas of research in music. Thus, for instance, in a series of transcultural studies, Randles (Randles, 2011; Randles and Smith, 2012; Randles and Muhonen, 2015; Randles and Ballantyne, 2018; Randles and Tan, 2019) has analyzed aspects associated with creativity, composition, improvisation, and the act of playing different types of music. He considers that these dimensions of music serve as areas of development for musicians; areas in which to acquire a sense of competency in order to achieve satisfactory professional development. The results of these transcultural studies confirm the existence of the above-mentioned types of self-efficacy, while explaining certain differences among the levels achieved by students, based on their previous training experiences. Similarly, Watson (2010) has studied self-efficacy for improvising in jazz contexts, Egilmez (2015) for the handling of anxiety in one's own perception of a public musical performance, and Miksza and Tan (2015) for self-regulated music learning. Further concrete areas of musical training can be studied with the construct of self-efficacy, since it defines itself as domain-specific. Kurtuldu and Bulut (2017) designed a scale to evaluate students' self-efficacy toward piano lessons, and Watson (2010), adapted this scale for jazz. Other authors (Papageorgi et al., 2013; Girgin, 2020) have highlighted the lack of self-efficacy in facing a series of challenges and demands: a slump in motivation, on the one hand, or increased levels of burnout, on the other.

As mentioned above, studies in self-efficacy have benefited from the implementation of a wider theoretical framework regarding the subject of learning: concretely, self-regulated learning models (Zimmerman, 2000; Panadero and Alonso-Tapia, 2014). In studies in this area focusing on music, self-efficacy has also established itself as a key concept within self-regulated learning (McPherson and McCormick, 2000; Varela et al., 2016). An important way of explaining the relationship of self-efficacy with public music performance is by highlighting its role as a component of the self-regulated earning three-phase model (forethought, performance, and reflection) (McPherson, 2022). In this context, self-efficacy is seen as a component of self-motivation beliefs related to feelings of confidence, outcome expectations, interest, and passion which are based on the *want* ("I want to do this") and *can* ("I can do this") parts of motivation (McPherson, 2022; see also Varela et al., 2016). Self-efficacy includes the beliefs we hold about our own capacity to perform at an expected level of achievement (self-efficacy for performance), or the beliefs we hold about our own capacity to implement or learn the types of processes that will allow us to master a musical challenge within a practice session (self-efficacy for learning). We achieve our best level of performance when we feel confident, which is why self-efficacy beliefs are critical in expert performance. The emphasis is on believing we *can* do something (rather than *will* do something). Research shows that we tend to overestimate our capacity to achieve and perform. This is not necessarily bad, but our evaluations do need to be realistic because of this. Importantly, when our personal self-efficacy is high, we are more likely to set challenging goals for ourselves, and search for strategies to achieve them (McPherson, 2022).

The recent uptake in studies regarding self-regulated learning, the relative ease in evaluating it, and the existence of a well-defined theoretical corpus have given rise to a series of studies of the construct's applicability (Hewitt, 2015; Miksza and Tan, 2015; Varela et al., 2016; Waters, 2020). Research has sought to improve students' self-regulation through interventions (Watson, 2010; Mieder and Bugos, 2017; Miksza et al., 2018; MacAfee and Comeau, 2020), and studied the important role played by contextual factors in its development and modification, occasionally over brief periods of time (Hendricks, 2014), and using well-defined evaluation methods and validated measurement tools (Ritchie and Williamon, 2007, 2011; Watson, 2010; Randles and Muhonen, 2015; Kurtuldu and Bulut, 2017; McPherson et al., 2019; Osborne et al., 2020).

Although the number of studies on musical self-efficacy has increased in recent years, few studies have attempted to explore how it develops. In Bandura's model, four different sources of self-efficacy are postulated: mastery experiences, vicarious experiences, verbal persuasion, and emotional states (Hendricks, 2014; Zelenak, 2015). The greatest amount of evidence has accumulated in relation with the importance of practice, especially the types of mastery experiences that are regarded as the basic condition for musical performance (McCormick and McPherson, 2003; McPherson and McCormick, 2006). However, not all approaches to practice are effective, since students need to simultaneously associate their practice routine with programing elements that are autonomous, strategic, and self-regulated; what researchers refer to as "deliberate practice" (Hallam et al., 2012; McPherson et al., 2012; Upitis et al., 2017c). This kind of practice approach increases with the passage of time under the influence of academic requirements that progressively become more demanding.

Further information has been gathered about the effect of emotional states, mostly associated with performance anxiety (Papageorgi et al., 2010). Hardly any evidence has been gathered, however, regarding the eventual influence of other sources. Although clear evidence has been found for the influence exerted by parents, teachers and peers on the development of a musical career (Orejudo et al., 2020), practically no evidence has been gathered to ascertain whether the role played by these agents leads to an improvement in musical self-efficacy. Previous studies have shown the role they can play by strengthening other elements associated with success in a musical career, such as providing the trainee with a series of resources to help them handle performance anxiety (Zarza-Alzugaray et al., 2020), but no direct data has been provided regarding the relationship between the support coming from those agents and the development of the musical self-efficacy. Nevertheless, before we continue to address this study's goals, we proceed to analyze in further detail the role of social support in musical training.

## Social Support in Music

As shown by Gruber et al. (2008), and Lehmann and Kristensen (2014) parents, teachers, and peers provide an important source of information for learners because of their role as "persons in the shadow." Successful musical careers always rely on the existence of certain people who are relevant and key in supporting the artist's training, in early as well as in advanced stages. Here we are dealing with the psychological concept of "social support" (Caplan, 1974; Cobb, 1976; Sarason et al., 1990). Social support involves different forms of psychological support and resources provided by significant people in the learner's environment who help satisfy their basic needs in interacting with others, act as a source that can be trusted, and are valued and loved because they are able to maintain open communication channels that are based on mutual feelings of responsibility and commitment.

Growing evidence demonstrates the presence of social support as a key factor that determines musical success (Davidson et al., 1996; McPherson and Davidson, 2002; Creech and Hallam, 2003; Sichivitsa, 2007; Margiotta, 2011; Nogaj and Ossowski, 2015). Moore et al. (2003) related social support to progress in an artist's musical career; Nogaj and Ossowski (2015) related it to achievement, Sichivitsa (2007) found that parental support is a basic factor in music students' self-concept, whereas Howe and Sloboda (1991) highlight the role played by parents and siblings in the initial stages of musical training. Social support is defined as a multidimensional construct that involves different types of support: instrumental and emotional support, as well as what a series of personal agents can provide – in the case of music, those agents are the family, teachers, and peers (Orejudo et al., 2020). Creech (2009) has specified that parents support their children by three different types of means: behavioral support, cognitive support, and personal support. Through these types of support, parents can enhance the teacher's educational task by helping the student organize their study, providing them with opportunities to interact with music, and helping them establish expectations and goals. As can be readily observed, such parental activities come to form part of a series of conditions that encourage the student's development of positive perceptions about their own personal value, thereby generating self-regulated learning skills.

Although parents undoubtedly fulfill a basic role in their children's musical training, this does not occur at the margins of what is achieved by teachers and peers. All three groups are jointly regarded as the main support sources for music students, responsible for generating the motivational and emotional processes they need in order to pursue their training career (Ryan et al., 2000; Lehmann and Kristensen, 2014; Nogaj and Ossowski, 2015). Indeed, these three “source groups” are not regarded as mutually independent. For instance, parents who have enjoyed musical training and have a direct relationship with music are generally perceived as better sources of support than those for whom it is not the case (Sichivitsa, 2007; Ritchie and Williamon, 2013; Orejudo et al., 2020). They effectively encourage the student to persevere with their musical training (Jeppsson and Lindgren, 2018). Upitis et al. (2017a) found that a family relative or a custodian who plays an instrument provides a positive contribution to self-efficacy. The student has a greater enjoyment of the training situation, and this, in turn, can reinforce intrinsic motivation that might otherwise be lacking.

Teachers are further key agents in music student training. By establishing a direct relationship with self-efficacy sources, teachers can play a key role in the learning process, thereby providing an essential contribution to student motivation (Upitis et al., 2017c). This activity involves a number of aspects such as the establishment of short-term and long-term goals, monitoring the latter process, choosing repertoire, providing feedback *via* exams or a general evaluation of the student, teaching coping strategies to face performance anxiety, and acting as a social and emotional support agent in collaboration with the family. Regarding this important role played by teachers, there is a certain amount of evidence. Upitis et al. (2017a) report that teacher quality is an important factor enabling students to enjoy their achievements, particularly their public performances. Waters (2020) examined which factors have a decisive influence on the effectiveness of orientation provided by the teachers: students put their teachers’ advice to best use when they approach the learning context with a proactive attitude. Conversely, when students are not sufficiently autonomous in this sense, those who manage to adopt the strategies suggested by the teachers nevertheless have very little perception of how to control them and thereby do not succeed in transforming them into tools that improve their learning. Such students eventually perceive that they have less control in shaping their learning, with the result that their self-efficacy declines.

Peers are likewise regarded as a source of social support for musicians, but little evidence has been found of their relationship with musical development. Hendricks (2014) ascertained that when girls, in particular, feel that they are receiving a substantial amount of social support, they experience greater levels of self-efficacy. This is more likely to occur if the context is not perceived as being highly competitive. Siblings can also be a source of motivation for music students. Howe and Sloboda (1991) ascertained that elder brothers and sisters play an important role in the musical practice activities of their younger siblings.

## Aims, Research Questions, and Objectives

As mentioned above, in recent years a considerable number of studies have been published, ascertaining that self-efficacy is a useful construct for the analysis of musical training. A series of studies have proved its relationship with performance in different contexts, and/or have developed tools to evaluate it in such contexts, giving rise, on occasion, to full-fledged intervention programs. But few studies exist on the role of social support in maintaining musical self-efficacy. Age, sex, and type of instrument (solo or orchestral) are relevant variables analyzing musical self-efficacy as well as performance anxiety (Casanova et al., 2018; Zarza-Alzugaray et al., 2020). In this study, however, we have focused on the variables of age and formative level, since they have been less analyzed until now. To our knowledge, no study has been published with an attempt to compare the sole exerted in social support by different sources – parents, teachers, and peers – at different levels of education.

Our study’s purpose was thus to analyze the types and amount of support perceived by music students at two academic levels: those enrolled in university-level music academies (*conservatorios superiores*), and those enrolled in advanced music schools (*conservatorios profesionales*). *Conservatorios superiores* are institutions of musical learning for students who want to embark on a professional music-related career. At the same age at which they would start studying at university, they gain access to *conservatorios superiores* after having concluded studies in a *conservatorio professional.* This is usually when they are 18 years old. At the other educational level, *conservatorios profesionales* enroll students with more heterogeneous profiles: some are studying to gain access to a *conservatorio superior*, while others are learning music without necessarily holding a long-term professional perspective in mind. For purposes of analysis, we differentiate two age groups in *conservatorios profesionales*: younger students, on the one hand, and 16–18-year-olds, on the other. We chose to apply this division based on evidence (Orejudo et al., 2020) that students in the latter age group are combining musical training with secondary education which allows them access to university, and are going through a decision-making process regarding their professional future, which can affect the level of their commitment to musical activities.

We expected to find a relationship between social support and self-efficacy for learning and for public performance (Ritchie and Williamon, 2007, 2011; Upitis et al., 2017c), given the possibility that such relationships can vary in terms of age and academic level. We therefore carried out the analysis on three separate groups of students: university-level music students (*conservatorios superiores*), 16–18-year-old students enrolled in *conservatorios profesionales*, and younger students (11–15-year-old) enrolled in the latter kind of institution. In support of this assumption, certain authors postulate that family support should be more relevant in early stages (Davidson et al., 1996; Margiotta, 2011), and that teachers in initial musical training stages need to have a series of competencies that differ from those required for more advanced stages (Moore et al., 2003).

A second assumption refers to the relationship among the different types of self-efficacy. We speculated that self-efficacy for learning will be a strong predictor of self-efficacy for public performance. This seems plausible within the theoretical framework of self-regulated learning, in which the preparation of repertoire prior to performance and the management of performance anxiety serves as relevant factors in the development of musical competency. We thus expected to find a direct relationship between these, although, admittedly, learning situations do not necessarily imply performance experiences; the predictors of the two types of self-efficacy might thus eventually be different.

To test these assumptions and to ascertain whether the relationships among these variables can differ in function of age, we analyzed our data using structural equation modeling (SEM) with the sources of support (parents, teachers, and peers) as exogenous variables and the sources of self-efficacy as endogenous variables. We assumed that the relationship among them could be direct, and that self-efficacy would have a relationship of partial mediation with the sources of support. This analysis technique also allowed us to compare equality of regression weights in different groups (Byrne, 2010).

# MATERIALS AND METHODS

## Participants

Our sample comprised 415 music students, 296 of whom (71.3%) were enrolled in *conservatorios profesionales* (advanced, pre-university music schools), whereas 119 (28.7%) were enrolled in *conservatorios superiores* (university-level music academies). We established three large age groups: 141 [34%, *Mean age* ($M$) = 13.69, SD = 1.17, range: 11–15] participants were 15 years old or younger (enrolled in *conservatorios profesionales*); 158 were ages 16–18 (38%, $M$ = 16.91, SD = 0.80), all enrolled in *conservatorios profesionales*; the last group, age 19 and older (28%, $M$ = 22.41, SD = 4.45, 19–43), were enrolled in *conservatorios superiores* (university-level music academies). Regarding distribution by gender, 44.6% ($n$ = 185) were male, whereas 55.4% ($n$ = 230) were female, without any significant association ($\chi^2$ = 4.194; $p$ = 0.123) between a student's gender and their age group.

## Measures

The *Social Support Scale* proposed by Ryan et al. (2000) is designed to evaluate the level of social support perceived by music students. It measures that support through a series of independent scales corresponding to each of the social agents: parents, teachers, and peers, associated with 12, 9, and 10 items respectively, measured on a 7-point Likert-type scale (from 1, "not very much," to 7, "a lot"). A Spanish version of this scale (Orejudo et al., 2020) has been validated for the academic levels featured in this study. The authors found one factorial structure for parent support (9 items, $\alpha$ = 0.866) and teacher support (10 items, $\alpha$ = 0.866), but two different factors for peer support: one related to musical training activities (5 items, $\alpha$ = 0.785), and the other related to facing taunts (3 items, $\alpha$ = 0.935). For this reason, we used the same four social support subscales in our study.

The *General Musical Self-Efficacy Scale* proposed by Ritchie and Williamon (2007, 2011) is a 1–7-point Likert scale (completely disagree-completely agree) made up of 22 items grouped into two subscales: musical self-efficacy for learning, and musical self-efficacy for performing. Six items in each of the subscales are reverse-coded: items 2, 4, 5, 8, 10, and 11 in the learning factor, and items 2, 3, 4, 6, 7, and 8 in the performance factor. In its English-language version, this scale has good psychometric properties as applied to different age groups, including young students (Ritchie and Williamon, 2013). A Spanish version (Orejudo et al., 2020) confirms the scale's reliability and validity. Internal consistency was good both in the "self-efficacy for learning" scale (10 items, $\alpha$ = 0.773) as well as in the "self-efficacy for performance" scale (10 items, $\alpha$ = 0.773). There was also good temporal stability (correlation ranging from 0.515 to 0.539 after a period of 1 month).

## Procedure

After having received an affirmative response from the above-cited institutions of musical learning, we proceeded to gather the data in person, on the premises of each institution. The research team or a local professor visited the academies in order to operate *in situ*, with the task of administering and gathering the questionnaires (this lasted approximately 30 min per session). Students participated on a voluntary, anonymous basis, and they had no external incentive to participate in the study.

## Statistical Procedure

Analysis of results was conducted in two phases. In a first phase, we applied descriptive analysis of the means of the scales that were used, differentiating by age group to conduct an initial exploration of results and to test whether there were differences. Correlations were obtained among all the factor scores of the variables in the three age-groups. In view of the high number of correlations in each group, we adjusted the level of significance by applying the Bonferroni correction ($p$ = 0.0033). In a second phase, we tested the hypothetical model of causal structure by applying *SEM*. In this model we posited the Social Support Scales of parents, teachers, and peers as exogenous variables, and the two self-efficacy subscales as endogenous variables. The model is displayed in **Figures 1–3**, and was tested with IBM-SPSS software and its AMOS extension (v. 17). The estimation method chosen to test the measurement model was maximum likelihood whenever multivariate normal distribution criteria were met. We initially obtained correlations among all exogenous and endogenous variables in each of the subsamples we analyzed. Then a comparison was made between the two subsamples by applying Fisher's Z transformation of the correlation coefficient. The model's goodness of fit was tested using the $\chi^2$ test, as well as the normal and the $\chi^2$ degrees of freedom ratio (*DCIM/GL* in Amos), by *RMSEA* and *GFI* indicators, and by their critical levels as indicated by authors such as Schermelleh-Engel et al. (2003) and Byrne (2010). We applied multi-group analysis to verify whether the interviewees of different age groups displayed significant differences in terms of influencing relationships. To make this distinction between models, we compared a series of nested models, the results of which are described in section "Results." To contrast differences between groups, the models were compared by calculating differences in $\chi^2$ and the *AIC* index (Byrne, 2010).

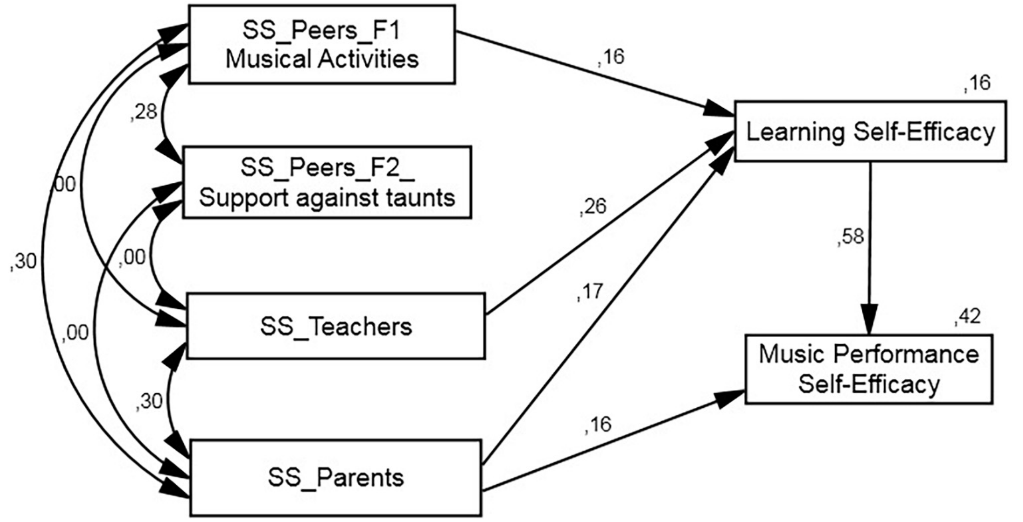



**FIGURE 1 |** Up to 15 years old graph path.

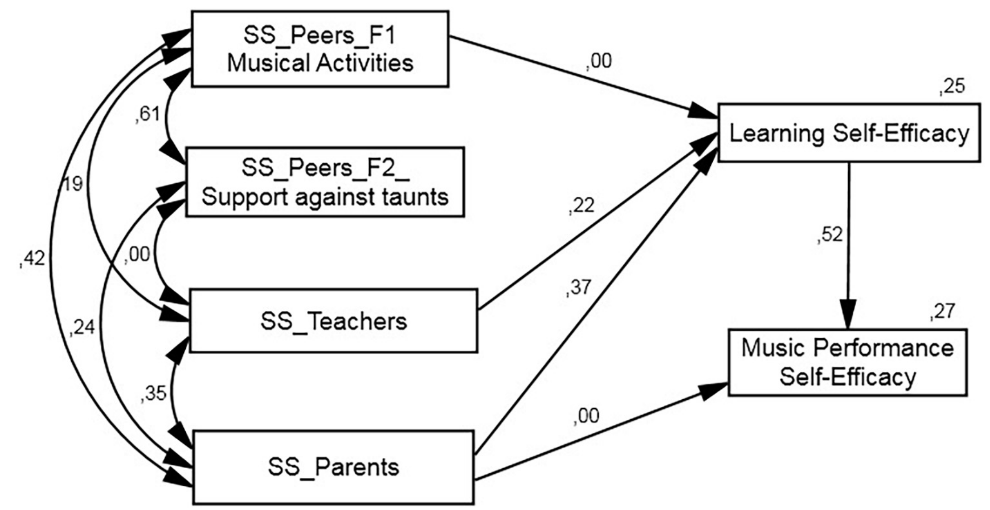



**FIGURE 2 |** Between 16 to 18 years old graph path.

## RESULTS

An initial result (**Table 1**) shows significant differences ($p < 0.05$) between the three age groups in all posited variables, except for F2, the factor of "peer social support in the face of taunts" ($F_{2-412} = 1.022$; $p = 0.361$), although the size effect reported by the $\eta^2$ is admittedly small. Comparing the group of youngest participants (<15 years old) with the aged 16–18, the *post hoc* tests reveal significant mean differences in the factors of self-efficacy for learning ($p = 0.001$) in parental social support ($p = 0.004$), and in teacher support ($p = 0.020$), with higher values in the younger groups (**Table 1**). In the factor of self-efficacy for public performance, the group of 16–18-year-olds scores lower than the two other groups ($p < 0.05$). Finally, in the F1 factor of "peer social support for musical learning," the oldest group of students (>19 years old) scores higher than the two other age groups, whereby the highest mean is in the oldest group ($p < 0.05$).

As shown in **Table 1**, a series of significant correlations can be observed in the three age groups among the factors posited for this study. The factors we attempt to explain by social support display significant relationships ($p < 0.003$) with social support coming from parents and teachers, particularly in the group of high-school-age students. Neither of the two older groups of students fulfills the established significance criterion. Differences appear in function of age groups. Parent support and self-efficacy for learning are more pronounced in the youngest group of students ($r = 0.455$) than in the 16–18-year-olds ($r = 0.297$) and in those over 19 ($r = 0.271$). In the case of peers, there is a strong correlation between the two types of support, but in neither case

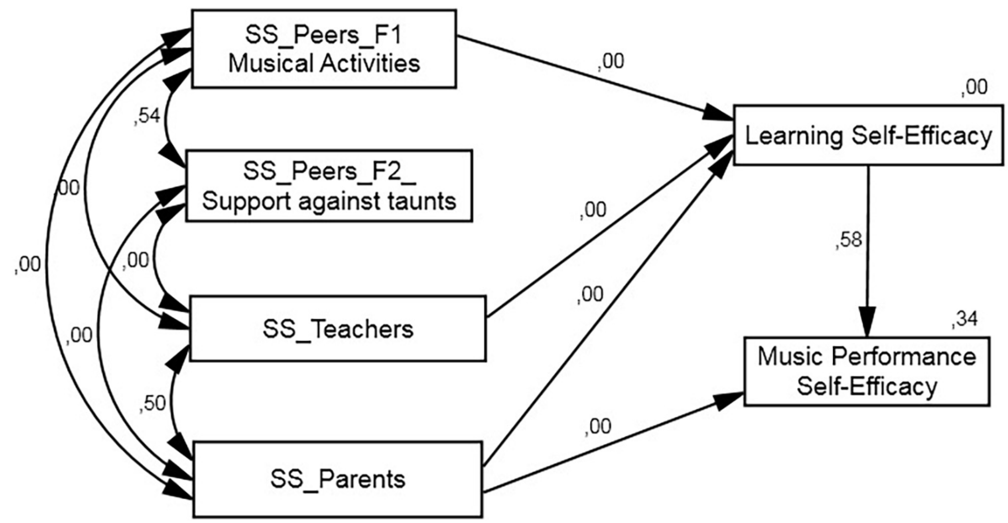



**FIGURE 3 |** From 19 years old graph path.

**TABLE 1 |** ANOVA self-efficacy and social support × age – level.

| Variable | Age–level | *N* | Mean | Standard deviation | *F* | Significant | $\eta^2$ |
|---|---|---|---|---|---|---|---|
| Learning self-efficacy | ≤15 years old secondary | 141 | 58.66 | 7.01 | 6.997 | 0.001 | 0.033 |
| | 16–18 years old secondary | 158 | 55.51 | 7.89 | | | |
| | ≥19 tertiary | 116 | 57.40 | 6.98 | | | |
| | Total | 415 | 57.11 | 7.46 | | | |
| Music performance self-efficacy | ≤15 secondary | 141 | 52.94 | 9.19 | 7.853 | 0.000 | 0.037 |
| | 16–18 secondary | 158 | 49.70 | 8.52 | | | |
| | ≥19 tertiary | 116 | 53.40 | 8.12 | | | |
| | Total | 415 | 51.84 | 8.79 | | | |
| Parents social support | ≤15 secondary | 141 | 56.10 | 5.90 | 4.685 | 0.010 | 0.022 |
| | 16–18 secondary | 158 | 53.63 | 7.75 | | | |
| | ≥19 tertiary | 116 | 55.42 | 7.83 | | | |
| | Total | 415 | 54.97 | 7.26 | | | |
| Teachers social support | ≤15 secondary | 141 | 50.49 | 7.87 | 5.489 | 0.012 | 0.021 |
| | 16–18 secondary | 158 | 47.59 | 10.04 | | | |
| | ≥19 tertiary | 116 | 50.21 | 9.40 | | | |
| | Total | 415 | 49.31 | 9.25 | | | |
| Peers social support F1 | ≤15 secondary | 141 | 23.83 | 6.62 | 5.797 | 0.003 | 0.027 |
| | 16–18 secondary | 158 | 24.41 | 5.65 | | | |
| | ≥19 tertiary | 116 | 26.37 | 6.31 | | | |
| | Total | 415 | 24.76 | 6.25 | | | |
| Peers social support F2 | ≤15 secondary | 141 | 15.91 | 5.73 | 1.022 | 0.361 | 0.005 |
| | 16–18 secondary | 158 | 15.35 | 5.03 | | | |
| | ≥19 tertiary | 116 | 14.96 | 5.32 | | | |
| | Total | 415 | 15.43 | 5.36 | | | |

*Peers social support F1: peers support musical activities. Peers social support F2: peers support face of taunts.*

do they exceed the limit established to consider the correlation as significant with self-efficacy, despite the values that appear in the intermediate age group ($r = 0.220$, $p = 0.007$). Correlations between self-efficacy for learning and self-efficacy for public performance are high in all age groups ($r = 0.553–0.624$), and there are no statistically significant differences among groups.

**Table 2** also displays the correlations we observed between age groups and sources of support. Most correlations, although

**TABLE 2 |** Correlations.

| | | Secondary level | | Secondary level | | Tertiary level | |
|---|---|---|---|---|---|---|---|
| | | ≤15 years old | | 16–18 years old | | ≥19 years old | |
| | | *r* | Significant | *r* | Significant | *r* | Significant |
| **Music learning self-efficacy** | | | | | | | |
| Parents SS | | 0.455 | <0.001 | 0.297 | <0.001 | 0.271 | 0.005 |
| Teachers SS | | 0.349 | <0.001 | 0.321 | <0.001 | 0.275 | 0.004 |
| Peers SS F1 | | 0.197 | 0.022 | 0.220 | 0.007 | 0.196 | 0.039 |
| Peers SS F2 | | 0.109 | 0.199 | 0.123 | 0.125 | 0.078 | 0.404 |
| **Music performance self-efficacy** | | | | | | | |
| Learning self-efficacy | | 0.553 | <0.001 | 0.587 | <0.001 | 0.622 | <0.001 |
| Parents SS | | 0.374 | <0.001 | 0.352 | <0.001 | 0.284 | 0.003 |
| Teachers SS | | 0.291 | <0.001 | 0.199 | 0.015 | 0.272 | 0.005 |
| Peers SS F1 | | 0.195 | 0.023 | 0.092 | 0.249 | 0.071 | 0.447 |
| Peers SS F2 | | 0.087 | 0.307 | 0.019 | 0.809 | –0.070 | 0.454 |
| **Sources of social support** | | | | | | | |
| Parents | Teachers | 0.360 | <0.001 | 0.308 | <0.001 | 0.505 | <0.001 |
| Peers F1 | Parents | 0.429 | <0.001 | 0.306 | <0.001 | 0.277 | 0.004 |
| Peers F1 | Teachers | 0.205 | 0.018 | 0.030 | 0.705 | 0.162 | 0.086 |
| Peers F1 | Peers F2 | 0.611 | <0.001 | 0.278 | <0.001 | 0.541 | <0.001 |
| Peers F2 | Parents | 0.253 | 0.004 | –0.006 | 0.939 | 0.198 | 0.037 |
| Peers F2 | Teachers | 0.032 | 0.709 | 0.053 | 0.511 | 0.123 | 0.19 |

*Peers social support F1: peer support for musical activities. Peers social support F2: peer support in the face of taunts.*

significant with $p \leq 0.05$, are no longer significant with the Bonferroni correction. Parental support correlates highly with teacher support, and peer support with musical activities. Magnitude of correlation is average, except, curiously, the correlation among parents and teachers in the oldest age group, which is greater ($r = 0.505$ vs. $r = 0.360$ and $r = 0.308$) than in the two younger groups. Conversely, parent support in the youngest group is associated with peer support for musical activities ($r = 0.429$ vs. $r = 0.306$ and $r = 0.207$). As was to be expected, the highest correlations can be found between peer support for musical activities and facing taunts, although the correlation in the 16–18-year-old group ($r = 0.278$) is much lower than in the youngest group ($r = 0.611$) and the oldest group ($r = 0.541$). Applying the Bonferroni correction, parent support and teacher support do not correlate with peer support.

The *SEM* model, constructed according to our postulated theoretical model (**Figure 1**), has optimal fit (**Table 3**), not only when the regression parameters are set equally for all three groups (structural weights model), but also when restrictions are introduced (unconstrained model). The unconstrained model nevertheless indicates that certain relationships established in the model have non-significant values. We therefore postulated a new model: structural weights without non-significant weights, in which those parameters are set at zero, whereby the remaining values are kept the same in all groups. On this model, which displays adequate fit values ($\chi^2 = 38.535$, $gl = 22$, $p = 0.016$, $\chi^2/gl = 1.752$, $CFI = 0.0966$, $RMSEA = 0.043$), we tested the hypothesis that regression weight could be different in some groups. To ascertain this, we freed up each parameter in order to ascertain whether the new model improved the former one's fit. **Table 3** displays the different options we tested. One of the models improves the previous model's fit ($\Delta\chi^2 = 5.371$, $gl = 1$, $p = 0.020$, $\Delta_{AIC} = 3.371$), and establishes that the regression weights of b7 (social support of parents for self-efficacy in learning) are different.

As **Figure 1** shows, teacher support (β = 0.22) and parent support (β = 0.37) in the group of youngest students was directly related with self-efficacy for learning, which, in turn, significantly mediates (β = 0.52) self-efficacy for music performance and explains 33.2% of its variance. In the case of self-efficacy for learning, self-efficacy for music performance explained 25% of its variance. In this mediation model, it is important to note that parent support (β = 0.193) and teacher support (β = 0.116) provided an indirect contribution to self-efficacy in public performance.

In the group of 16–18-year-olds (**Figure 2**), self-efficacy for learning was again related to teacher support (β = 0.26), but parents now had lower mediation than in the youngest group (β = 0.17). Although there are two differences compared to the model of the youngest group, peer support was related to self-efficacy for learning (β = 0.16), and explained 17% of that variable. Parents, for their part, had a direct influence on self-efficacy for public performance (β = 0.16), and this, along with self-efficacy for learning (β = 0.58), helped explain 42% of self-efficacy for public performance. Once again, one can note indirect effects of teacher support (β = 0.149), of parent support (β = 0.099), and of peer support (β = 0.094) on self-efficacy for interpretation. The total effect of parent support on public performance reached a total of β = 0.263.

None of the support sources provided a significant contribution to the model in the oldest group (**Figure 3**). The only relation that can be observe was that self-efficacy for

**TABLE 3** | Results of the *SEM* model.

| Model | $\chi^2$ | DF | Significant | CMIN/DF | *RMSEA* | *CFI* | TLI | Akaike |
|---|---|---|---|---|---|---|---|---|
| Unconstrained | 1.256 | 3 | 0.74 | 0.419 | 1.000 | 1.000 | 0.00 | 157.256 |
| Structural weights | 18.596 | 19 | 0.483 | 0.979 | 1.000 | 1.000 | 0.00 | 142.596 |
| Structural weights–without n.s. | 38.535 | 22 | 0.016 | 1.752 | 0.966 | 0.966 | 0.043 | 156.535 |
| Free_b8_1 | 37.824 | 21 | 0.014 | 1.801 | 0.965 | 0.965 | 0.044 | 157.824 |
| Free_b8_2 | 35.959 | 21 | 0.022 | 1.712 | 0.969 | 0.969 | 0.042 | 155.959 |
| Free_b8_3 | 37.679 | 21 | 0.014 | 1.794 | 0.965 | 0.965 | 0.044 | 157.679 |
| b6_1 and b6_2_free | 38.178 | 21 | 0.012 | 1.818 | 0.964 | 0.964 | 0.045 | 158.178 |
| b7_1 and b7_2_free | 33.164 | 21 | 0.044 | 1.579 | 0.975 | 0.975 | 0.037 | 153.164 |

*b8_1 = learning self-efficacy to music performance self-efficacy: group secondary level, ≤15 years old. b8_2 = learning self-efficacy to music performance self-efficacy: group secondary level, 16–18 years old. b8_3 = learning self-efficacy to music performance self-efficacy: group tertiary level, ≥19 years old. b6_1 and b6_2_free: teachers social support to learning self-efficacy groups 1 and 2. b7_1 and b7_2_free: parents social support to learning self-efficacy, groups 1 and 2.*

learning predicted self-efficacy for public performance ($\beta = 0.58$), explaining 34% of the variance.

## DISCUSSION AND CONCLUSION

This study was designed to examine the relationships between social support perceived by music students and their self-efficacy for learning, as well as for facing performance situations, using students of different ages and academic levels. We analyzed relationships within an *SEM* model in which support sources were the exogenous variables, and the two endogenous variables were self-efficacy for learning and self-efficacy for public performance, with a relationship of mediation between them. Results provide clear evidenced of an important relationship between self-efficacy for learning and self-efficacy for public performance in all three age groups. This result is especially relevant for the oldest group, where self-efficacy for learning is the sole predictor of self-efficacy for public performance. Our data have evidenced a relationship between social support and self-efficacy, but only for students in the two younger age groups. In other words, for older students, who have more experience, there is no evidence that social support effects self-efficacy.

These findings are important from the perspective of music education. Although previous studies have highlighted the importance of parents, teachers, and peers in students' musical training (Moore et al., 2003; McPherson, 2009; Lehmann and Kristensen, 2014), none have yet tested the relation between sources of support and one of the self-regulated learning model's main variables which has the closest relation with musical practice: self-efficacy (McPherson and Zimmerman, 2011; Varela et al., 2016), neither has the assumption been tested on different age groups and academic levels. At the same time, our study provides new evidence of the importance of social support in the development of a musical career and is in accord with previous studies that highlighted the importance of parents in the musical education of their children. McPherson (2009) postulated that such support needs to be integrated into a framework that equips the student with a wide array of strategies to help them meet the demands of a musical career and make progress therein: the framework of self-regulated learning. However, little evidence on these seemingly critical aspects has been gathered until now.

More concretely, we observed that parental support of secondary school students became the main predictor of self-efficacy (since in this study we did not gather responses from students over 18 years old enrolled in *conservatorios profesionales*). As mentioned above, little evidence had been previously gathered regarding the relationship between social support and self-efficacy, either as an influence on musical learning or on public performance. Other studies, however, have dealt with the importance of social support in early musical training stages (Howe and Sloboda, 1991; Davidson et al., 1996; McPherson and Davidson, 2002; Moore et al., 2003; Sichivitsa, 2007; Margiotta, 2011). More recently, Upitis et al. (2017b) demonstrated the importance of the involvement of families in their children's musical progress, from initial musical training to adolescence. They highlight day-to-day activities carried out by the parents to help their children's progress: study at home, setting weekly or yearly goals, providing instrumental support, contacting teachers, and teaching concrete strategies. Those authors likewise observed how parents gradually reduce the amount of support they provide as students grow older and become more autonomous. The same study showed that parents as well as teachers are sources of support (MacMillan, 2004; Creech, 2009). Given the importance generally attributed to the family as a source of support, this gives rise to a new debate over family variables that can affect support levels. Future studies could explore variables such as: family expectations or family beliefs about what consists in the necessary amount of practice, parental ability to help the child go on practicing, or the relationship of parents and other family members with music. McPherson and Davidson (2002) evidenced that mothers of 7–9-year-olds who were initiating musical training and who had more ability to determine the amount of support their children needed in order to practice on a regular basis could indeed increase their offspring's possibilities of pursuing training. In older students, Orejudo et al. (2020) found that the parents' relationship with music, either as professional musicians or as music teachers, is an important predictor of support as perceived by their children in the course of their musical training. However, it is important to point out that support provided by families can be quite different according to training level, as our study has revealed. Thus, McPherson and Davidson (2002) ascertained that parent support for 7–8-year-old children who are in initial training impinges decisively on whether those children will continue or not. This factor can be key, but not indispensable, for the pursual of a musical career, as evidenced by Howe and Sloboda (1991), who found that children of that age had a high degree of motivation that did not necessarily stem from a family environment that was closely associated with music.

Thus, the research presented here adds the importance of teacher social support to that of the family as a learning resource throughout musical training at secondary-school age. As noted above, few studies have been published on self-efficacy, but other papers do demonstrate its role in musical training (Moore et al., 2003). Upitis et al. (2017b) highlight the importance of the quality of teachers as the main factor that promotes student progress at that academic level. The study was carried out from the family perspective, but it is likewise corroborated from the vantage point of the students (Upitis et al., 2017c). As in the case of families, it is important to be able to analyze the tasks carried out by teachers that help students develop their self-efficacy. An analysis of the tasks involved in teacher support (Ryan et al., 2000) confirms that the involved factors are associated with the creation of a positive atmosphere in the classroom (questionnaire items: "makes music class interesting," "teaches music you like," "often gives you a chance to choose what musical activities you do," "wants you to try your best and not worry if you make mistakes"), but are likewise associated with the teachers' positive expectations of what their students can accomplish ("wants you to pass music exams," "thinks you could have a job in music when you get older"), with the way they value them globally as musicians ["thinks you are good at playing an instrument," "praises you (tells you 'well done') for the work you do in music class"] or with the way they deal with mistakes. In their daily encounters with teachers, students expect that the latter should become sources of support for them, not only in the area of learning, but also in terms of more global aspects of their wellbeing such as helping them learn to deal with the stressors involved in musical training (Perkins et al., 2017). At any rate, the behavior of teachers can vary in terms of the specific, individual characteristics of their students. For example, Waters (2020) points out how students with a lower degree of autonomy tend to mechanically reproduce the learning strategies suggested by their teachers without critically evaluating them.

A further contribution provided by this study concerns the evidence for the importance of peers in the reinforcement of self-efficacy. Until now, studies that dealt with peer support (Orejudo et al., 2020) have only described its sources and its relationships with other support sources but not information about the role it can exert in reinforcing self-efficacy. Our data shows that peers are indeed important, particularly in the 16–18-year-old stage, where adolescents become more independent from their families, and where also, the opinion of peers acquires a more significant weight. This is also the moment of choosing a profession: a point in time in which peers can play a fundamental role by reinforcing elements associated with self-efficacy, such as the selection of professional and educational goals related with music, such as daily practice, and one's own sense of self-worth. These latter aspects were taken into account in our self-efficacy tool, without ignoring the possible role they might play in certain elements of criticism of musical education, public performance, and the anxiety associated with the latter. Zarza-Alzugaray et al. (2020) have shown how peer support plays a key role in helping boys cope with performance anxiety. A similar idea can be found in the study by Hendricks (2014), who attributed an improvement in girls' perceived self-efficacy over the course of a 3-day music festival to the support they perceived from peers (among other factors, and only when the level of performance was more non-competitive). Among other concerns, Hewitt (2015) highlights peer context as one of the factors that can exert an influence on the development of self-efficacy – more concretely, self-evaluation. Thus, apart from the role of peers as a general support factor, they can also have an important part in the development of further self-efficacy elements: for example, by providing evaluative feedback.

One of our study's unexpected results was that we did not find a relationship between social support and self-efficacy in university-age students. There are several possible reasons for this result. Students who have already made progress in their musical career and have opted for a professional future in the field can already count on a considerable amount of social support from their family, from teachers, and from peers in earlier stages. In this context, there would hardly be any difference among these students in terms of the three support sources. Studies providing a contrast with this result are lacking. Hallam et al. (2012) and Perkins et al. (2017) nevertheless point out that music students at university level perceive a greater amount of support from teachers than students at lower academic levels. Such support, however, can be oriented toward other aspects, such as: general wellbeing, the handling of educational stressors, the cultivation of the students' professional identity, and the upkeep of their motivation to persevere in their musical education. Such support might thus not have a direct relation with the development of self-efficacy, which, most probably, would be affected by other sources such as the students' own performances, their comparison of themselves with peers, or their mastery of the curricular requirements of the institution of musical learning in which they are enrolled. It is also possible that the scale's lack of specificity regarding who provides the support – particularly teachers – can reduce the capacity of the tool we used in this study to identify university students' sources of support. At university level, these students have new teachers (professors) and different teaching/learning conditions; one-on-one classes, particularly with the professor specialized in their instrument, as well as group classes. Thus, the kind of support received from different professors can be thoroughly different. This aspect is identified by other Social Support Scales, such as, for example, Gluska (2011), which differentiates between social support on the part of the instrument teacher compared with that provided by other teachers.

Another complementary assumption that might explain our results could lie in the fact that when we specifically evaluate self-efficacy as an element pertaining to self-regulated learning, the only relationships that emerged were between self-efficacy for learning and self-efficacy for public performance. Students at university level only feel qualified to perform in public in cases where they are able to apply abilities, perceptions, study habits, and learning resources that they feel they have developed in previous stages: periods during which family, teacher, and peer support were indeed relevant, as our data show. This aspect is in accord with Upitis et al. (2017b), who showed that families gradually modified their type of support for their children as they continually improved their ability to self-regulate. It is to be expected that those students who have not sufficiently developed the necessary learning, motivation, and self-regulation abilities to devote themselves to music (McPherson and Zimmerman, 2011; Varela et al., 2016), will eventually abandon their training. This is confirmed by the revealed existence of feedback between self-efficacy and study habits from the onset. It is also well-known that a poor management of performance anxiety becomes a risk factor under which music students might either cease to make academic progress or abandon their studies and their career altogether (Orejudo et al., 2018). At the moment when the student embarks on university-level studies, a poor handling of stage fright will have new consequences (Casanova et al., 2018).

Thus, the main finding that aim to explain our results in university-level students leads us back to consider the central role played by self-efficacy in the self-regulated learning process. We postulate that self-efficacy acts as one of several essential factors in initial stages of behavior (a moment in which the student sets goals and lays out strategies), in the student's self-evaluation of achievements, in the internal attribution of results, and in planned behavior. Other studies have confirmed the central role of self-efficacy in this model. Waters (2020) shows how students possessing a greater sense of self-efficacy are able to select goals, lay out strategies, and work toward them. Conversely, the strategies deployed by students with low levels of self-efficacy are less adaptive; they perceive less control of the demands made upon them by their environment, as well as increased levels of discomfort and anxiety. Miksza and Tan (2015) found that students with a greater degree of self-efficacy tend to commit themselves more thoroughly to long-term music-related goals; they apply learning strategies that are more elaborate, and, most of all, they succeed in ensuring a greater degree of quality in their practice and learning, thus committing themselves more profoundly to the task. Hewitt (2015) as well as Miksza and Tan (2015) found a direct relationship between self-efficacy, public performance, and self-evaluation of goals: this is another of the key elements in the self-efficacy model, particularly associated with the phase of reviewing one's actions and reflecting upon them (Panadero and Alonso-Tapia, 2014). Finally, Bonneville-Roussy and Bouffard (2015) found that an individual's perception of their musical competency is one of the most significant predictors of "deliberate practice," a type of programmed, goal-oriented practice that is responsible for musical success.

It is thus possible that students with less self-perceived competency for music practice and for public performance will be less devoted to those activities, and that the probabilities of soon abandoning their musical career will increase. This could likewise explain the decrease in support perceived by 16–18-year-old students, who not only end up calling into question their personal value for musical professional activities (associated with self-efficacy), but who are likewise faced with a series of further decisions that have long-term implications for their professional outlook, thereby compelling them to make a choice between a career in music or to choose another field of pursuit. This finding could explain the great relevance acquired by social support in the group of 16–18-year-olds, an age phase in which it would be important for the social support variable to reinforce not only the student's self-worth as a musician (self-efficacy), but also the goal associated with it: namely, the choice to pursue a musical career. In other words, not only would self-efficacy be implied in this progression (a necessary condition to continue studying music), but also further dimensions of the self-regulated learning model (McPherson and Zimmerman, 2011; McPherson, 2022) which have not been addressed in our study. Indirect support for this is provided by the finding in this 16–18-year-old group: the higher correlation observed between parent and teacher support, was a necessary condition of great value for achieving progress in a musical career and for gaining access to university-level music studies (MacMillan, 2004; Creech, 2009; Upitis et al., 2017b). Upitis et al. (2017b) ascertain that adolescent music students gradually tend to abandon training to the same degree that they start getting involved in other activities which eventually become incompatible, or they start to find less pleasure in music-making.

Although our study reports a number of findings, it also has certain limitations. One lies in the age range, which only included music students who were pursuing a regular studying activity from a certain age on (11 years old). Other stages of musical training including elementary school, music schools, academies, and music education in normal school have not been analyzed herein. Since these are initial stages, it is likely that social support would serve as a highly significant source of motivation. The role of teachers and family is also essential (Upitis et al., 2017a,b). Furthermore, the scales we used in our study do not differentiate in terms of what kind of concrete support the students can perceive. The questionnaires did not differentiate among the kind of family relatives who can provide support: for instance, the gender of the person providing support can be relevant (McPherson, 2009). Our scales do not also differentiate among different types of teachers. Neither were we able to further explore the different types of support each of those sources can provide (Creech, 2009). A final limitation of our study lies in its very nature, self-report, which, as previously mentioned, cannot gather all the different music teacher/professor functions in different levels of music training: thus, it would be necessary or complimentary to apply other methodologies as well.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the CEIC Aragón (CEICA) en su reunión del día May 6, 2019, Acta No. 11/2019. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

SO: generation of ideas, literature review, conceptual framework, methodology, data collection, analysis of data, discussion and conclusion, first writing, second writing review, and funding. FZ-A: generation of ideas, literature review, conceptual framework, methodology, data collection, analysis of data, discussion and conclusion, first writing, and second writing review. OC: generation of ideas, literature review, conceptual framework, methodology, data collection, discussion and conclusion, first writing, second writing review, and funding. GM: generation of ideas, conceptual framework, discussion and conclusion, and second writing review. All authors contributed to the article and approved the submitted version.

## FUNDING

Educaviva Research Group, S57_20R-Government of Aragon. It offered financial support to cover the publication costs of this article. Chair United States Foreign Trade Institute of Advanced Intelligence at the University of Zaragoza supported the data collection process.

## REFERENCES

Baessler, J., and Schwarcer, R. (1996). Evaluación de la autoeficacia: adaptación española de la escala de Autoeficacia General. *Ansiedad y Estrés* 2, 1–8.

Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*. New York, NY: Freeman.

Bandura, A. (2006). Toward a psychology of human agency. *Perspect. Psychol. Sci.* 1, 164–180.

Barlow, D. (2000). Unraveling the mysteries of anxiety and its disorders from the perspective of emotion theory. *Am. Psychol.* 55, 1247–1263.

Bonneville-Roussy, A., and Bouffard, T. (2015). When quantity is not enough: disentangling the roles of practice time, self-regulation and deliberate practice in musical achievement. *Psychol. Music* 43, 686–704. doi: 10.1177/0305735614534910

Byrne, B. M. (2010). *Structural Equation Modeling with AMOS. Basic Concepts, Applications, and Programming*. New York, NY: Routledge.

Caplan, G. (1974). *Support Systems and Community Mental Health: Lectures on Concept Development*. New York, NY: Behavioral Publications.

Casanova, O., Zarza-Alzugaray, F. J., and Orejudo, S. (2018). Differences in performance anxiety levels among advanced conservatory students in Spain, according to type of instrument and academic year of enrolment. *Music Educ. Res.* 20, 377–389. doi: 10.1080/14613808.2018.1433145

Cobb, S. (1976). Social support as a moderator of life stress. *Psychosom. Med.* 38, 300–314.

Creech, A. (2009). Teacher-pupil-parent triads: a typology of interpersonal interaction in the context of learning a musical instrument. *Music. Sci.* 13, 387–413. doi: 10.1177/102986490901300208

Creech, A., and Hallam, S. (2003). Parent–teacher–pupil interactions in instrumental music tuition: a literature review. *Br. J. Music Educ.* 20, 29–44. doi: 10.1017/S0265051702005272

Davidson, J. W., Howe, M. J. A., Moore, D. G., and Sloboda, J. A. (1996). The role of parental influences in the development of musical performance. *Br. J. Dev. Psychol.* 14, 399–412. doi: 10.1111/j.2044-835X.1996.tb00714.x

Egilmez, H. O. (2015). Pre-service music teachers' piano performance self-efficacy belief inversely related to musical performance anxiety levels. *Educ. Res. Rev.* 10, 2558–2567. doi: 10.5897/ERR2015.2439

Girgin, D. (2020). Motivation, self-efficacy and attitude as predictors of burnout in musical instrument education in fine arts high schools. *Eurasian J. Educ. Res.* 85, 93–108. doi: 10.14689/ejer.2020.85.5

Gluska, A. A. (2011). "Skala wsparcia spoécznego uczniów szkó muzycznych [the scale of social support of students of music schools]," in *Badania Naukowe Nad Edukacj Artystyczn i Kulturow [Scientific Studies in Artistic and Cultural Education]*, eds W. A. Sacher and Y. A. Weiner (Bielsko-Biała: Y sza zko a Administra ji), 70–92.

Gruber, H., Lehtinen, E., Palonen, T., and Degner, S. (2008). Persons in the shadow: assessing the social context of high abilities. *Psychol. Sci.* 50, 237–258.

Hallam, S., Rinta, T., Varvarigou, M., Creech, A., Papageorgi, I., Gomes, T., et al. (2012). The development of practising strategies in young people. *Psychol. Music* 40, 652–680. doi: 10.1177/0305735612443868

Hendricks, K. S. (2014). Changes in self-efficacy beliefs over time: contextual influences of gender, rank-based placement, and social support in a competitive orchestra environment. *Psychol. Music* 42, 347–365. doi: 10.1177/0305735612471238

Hewitt, M. P. (2015). Self-efficacy, self-evaluation, and music performance of secondary-level band students. *J. Res. Music Educ.* 63, 298–313. doi: 10.1177/0022429415595611

Howe, M. J. A., and Sloboda, J. A. (1991). Young musicians' accounts of significant influences in their early lives. 1. The family and the musical background. *Br. J. Music Educ.* 8, 39–52.

Jeppsson, C., and Lindgren, M. (2018). Exploring equal opportunities: children's experiences of the swedish community school of music and arts. *Res. Stud. Music Educ.* 40, 191–210. doi: 10.1177/1321103X18773153

Kurtuldu, M. K., and Bulut, D. (2017). Development of a self-efficacy scale toward piano lessons. *Educ. Sci.* 17, 835–857. doi: 10.12738/estp.2017.3.0209

Lehmann, A. C., and Kristensen, F. (2014). Persons in the shadow" brought to light: parents, teachers, and mentors - how guidance works in the acquisition of musical skills. *Talent Dev. Excell.* 6, 54–70.

MacAfee, E., and Comeau, G. (2020). The impact of the four sources of efficacy on adolescent musicians within a self-modeling intervention. *Contributions Music Educ.* 45, 205–236.

MacMillan, J. (2004). Learning the piano: a study of attitudes to parental involvement. *Br. J. Music Educ.* 21, 295–311. doi: 10.1017/S0265051704005807

Margiotta, M. (2011). Parental support in the development of young musicians: a teacher's perspective from a small-scale study of piano students and their parents. *Aust. J. Music Educ.* 1, 16–30.

McCormick, J., and McPherson, G. (2003). The role of self-efficacy in a musical performance examination: an exploratory structural equation analysis. *Psychol. Music* 31, 37–51. doi: 10.1177/0305735603031001322

McPherson, G. E. (2009). The role of parents in children's musical development. *Psychol. Music* 37, 91–110. doi: 10.1177/0305735607086049

McPherson, G. E. (2022). "Self-regulated learning music microanalysis," in *The Oxford Handbook of Music Performance*, ed. G. E. McPherson (New York, NY: Oxford University Press).

McPherson, G. E., and Davidson, J. W. (2002). Musical practice: mother and child interactions during the first year of learning an instrument. *Music Educ. Res.* 4, 141–156. doi: 10.1080/14613800220119822

McPherson, G. E., and McCormick, J. (2000). The contribution of motivational factors to instrumental performance in a music examination. *Res. Stud. Music Educ.* 15, 31–39. doi: 10.1177/1321103X0001500105

McPherson, G. E., and McCormick, J. (2006). Self-efficacy and music performance. *Psychol. Music* 34, 322–336. doi: 10.1177/0305735606064841

McPherson, G. E., and Zimmerman, B. J. (2011). "Self-regulation of musical learning: a social cognitive perspective on developing performance skills," in *MENC Handbook of Research on Music Learning: Volume 2: Applications*, eds R. Colwell and P. R. Webster (Oxford: Oxford University Press), 130–175. doi: 10.1093/acprof:osobl/9780199754397.003.0004

McPherson, G. E., Davidson, J. W., and Faulkner, R. (2012). *Music in Our Lives: Redefining Musical Development, Ability and Identity*. Oxford: Oxford University Press.

McPherson, G. E., Osborne, M. S., Evans, P., and Miksza, P. (2019). Applying self-regulated learning microanalysis to study musicians' practice. *Psychol. Music* 47, 18–32. doi: 10.1177/0305735617731614

Mieder, K., and Bugos, J. A. (2017). Enhancing self-regulated practice behavior in high school instrumentalists. *Int. J. Music Educ.* 35, 578–587. doi: 10.1177/0255761417689921

Miksza, P., and Tan, L. (2015). Predicting collegiate wind players' practice efficiency, flow, and self-efficacy for self-regulation: an exploratory study of relationships between teachers' instruction and students' practicing. *J. Res. Music Educ.* 63, 162–179. doi: 10.1177/0022429415583474

Miksza, P., Blackwell, J., and Roseth, N. E. (2018). Self-regulated music practice: microanalysis as a data collection technique and inspiration for pedagogical intervention. *J. Res. Music Educ.* 66, 295–319. doi: 10.1177/0022429418788557

Moore, D. G., Burland, K., and Davidson, J. W. (2003). The social context of musical success: a developmental account. *Br. J. Psychol.* 94, 529–549. doi: 10.1348/000712603322503088

Nogaj, A. A., and Ossowski, R. (2015). Social support as a mediator for musical achievement. *Pol. Psychol. Bull.* 46, 300–308. doi: 10.1515/ppb-2015-0036

Orejudo, S., Candela, C., Casanova, O., and Cuartero, L. M. (2020). A social support scale for music students in music schools, academies, and conservatories: an adaptation into Spanish and a factorial invariance study. *Psychol. Music* doi: 10.1177/0305735620968626

Orejudo, S., Zarza, F. J., Casanova, O., Rodríguez, C., and Mazas, B. (2017). The relation of music performance anxiety (MPA) to optimism, self-efficacy and sensitivity to reward and punishment: testing the theory of personal vulnerability on a sample of Spanish music students. *Psychol. Music* 45, 570–583.

Orejudo, S., Zarza-Alzugaray, F. J., and Casanova, O. (2018). Music performance anxiety. Substance use and career abandonment in Spanish music students. *Int. J. Music Educ.* 36, 460–472. doi: 10.1177/0255761418763903

Osborne, M. S., McPherson, G. E., Miksza, P., and Evans, P. (2020). Using a microanalysis intervention to examine shifts in musicians' self-regulated learning. *Psychol. Music* 49, 972–988. doi: 10.1177/0305735620915265

Pajares, F., and Schunk, D. H. (2001). "Self-beliefs and school success: Self-efficacy, self- concept, and school achievement," in *Perception*, eds R. Ridingand and S. Rayner (London: Ablex Publishing), 239–266.

Panadero, E., and Alonso-Tapia, J. (2014). How do students self-regulate? Review of Zimmerman's cyclical model of self-regulated learning. *Ann. Psychol.* 30, 450–462. doi: 10.6018/analesps.30.2.167221

Papageorgi, I., Creech, A., and Welch, G. (2013). Perceived performance anxiety in advanced musicians specializing in different musical genres. *Psychol. Music* 41, 18–41. doi: 10.1177/0305735611408995

Papageorgi, I., Haddon, E., Creech, A., Morton, F., Bezenac, C., Himonides, E., et al. (2010). Institutional culture and learning II: inter-relationships between perceptions of the learning environment and undergraduate musicians' attitudes to performance. *Music Educ. Res.* 12, 427–446. doi: 10.1080/14613808.2010.520432

Papageorgi, I., Hallam, S., and Welch, G. (2007). A conceptual framework for understanding musical performance anxiety. *Res. Stud. Music Educ.* 28, 83–107. doi: 10.1177/1321103X070280010207

Perkins, R., Reid, H., Araújo, L. S., Clark, T., and Williamon, A. (2017). Perceived enablers and barriers to optimal health among music students: a qualitative study in the music conservatoire setting. *Front. Psychol.* 8:968. doi: 10.3389/fpsyg.2017.00968

Randles, C. (2011). "What is a good musician?" An analysis of student beliefs. *Arts Educ. Policy Rev.* 112, 1–8. doi: 10.1080/10632913.2010.490774

Randles, C., and Ballantyne, J. (2018). Measuring self-perceptions of creative identity: a cross-cultural comparison of the creative identities of pre-service music teachers in the US and Australia. *Music Educ. Res.* 20, 231–241. doi: 10.1080/14613808.2016.1249360

Randles, C., and Muhonen, S. (2015). Validation and further validation of a measure of creative identity among USA and Finland pre-service music teachers. *Br. J. Music Educ.* 32, 51–70. doi: 10.1017/S0265051714000151

Randles, C., and Smith, G. D. (2012). A first comparison of pre-service music teachers' identities as creative musicians in the United States and England. *Res. Stud. Music Educ.* 34, 173–187. doi: 10.1177/1321103X12464836

Randles, C., and Tan, L. (2019). Measuring pre-service music teachers' creative identities: a cross-cultural comparison of the United States and Singapore. *Br. J. Music Educ.* 36, 197–210. doi: 10.1017/S0265051719000172

Ritchie, L., and Williamon, A. (2007). "Measuring self-efficacy in music," in *Proceedings of the International Symposium on Performance Science 2007*, eds A. Williamon and D. Coimbra (Utrecht: Association Européenne des Conservatoires (AEC)), 307–312.

Ritchie, L., and Williamon, A. (2011). Measuring distinct types of musical self-efficacy. *Psychol. Music* 39, 328–344. doi: 10.1177/0305735610374895

Ritchie, L., and Williamon, A. (2013). Measuring musical self-regulation: linking processes, skills, and beliefs. *J. Educ. Train. Stud.* 1, 106–117. doi: 10.11114/jets.v1i1.81

Ryan, K. J., Boulton, M. J., O'Neill, S. A., and Sloboda, J. A. (2000). "Perceived social support and children's participation in music," in *Proceedings of the 6th International Conference on Music Perception and Cognition*, eds C. Woods, G. B. Luck, R. Brochard, F. Seddon, and J. A. Sloboda (Keele: Keele University).

Sarason, B. R., Pierce, G. R., and Sarason, I. G. (1990). "Social support: The sense of acceptance and the role of relationships," in *Social Support: An Interactional View*, eds B. R. Sarason I, G. Sarason, and G. R. Pierce (New York, NY: Wiley), 97–128.

Schermelleh-Engel, K., Moosbrugger, H., and Müller, H. (2003). Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods Psychol. Res. Online* 8, 23–74.

Schwarzer, R., and Jerusalem, M. (2010). The general self-efficacy scale (GSE). *Anxiety Stress Coping* 12, 329–345.

Sherer, M., and Adams, C. H. (1983). Construct validation of the self-efficacy scale. *Psychol. Rep.* 53, 143–148. doi: 10.2466/pr0.1983.53.3.899

Sichivitsa, V. O. (2007). The influences of parents, teachers, peers and other factors on students' motivation in music. *Res. Stud. Music Educ.* 29, 55–67. doi: 10.1177/1321103X07087568

Upitis, R., Abrami, P. C., Varela, W., King, M., and Brook, J. (2017c). Student experiences with studio instruction. *Music Educ. Res.* 19, 410–437. doi: 10.1080/14613808.2016.1202221

Upitis, R., Abrami, P. C., Brook, J., Boese, K., and King, M. (2017a). Characteristics of independent music teachers. *Music Educ. Res.* 19, 169–194. doi: 10.1080/14613808.2016.1204277

Upitis, R., Abrami, P. C., Brook, J., and King, M. (2017b). Parental involvement in children's independent music lessons. *Music Educ. Res.* 19, 74–98. doi: 10.1080/14613808.2016.1202220

Varela, W., Abrami, P. C., and Upitis, R. (2016). Self-regulation and music learning: a systematic review. *Psychol. Music* 44, 55–74. doi: 10.1177/0305735614554639

Waters, M. (2020). Perceptions of playing-related Discomfort/Pain among tertiary string students: a thematic analysis. *Music Educ. Res.* 22, 257–269. doi: 10.1080/14613808.2020.1765154

Watson, K. (2010). The effects of aural versus notated instructional materials on achievement and self-efficacy in jazz improvisation. *J. Res. Music Educ.* 58, 240–259. doi: 10.1177/0022429410377115

Zarza-Alzugaray, F. J., Casanova, O., McPherson, G. E., and Orejudo, S. (2020). Music self-efficacy for performance: an explanatory model based on social support. *Front. Psychol.* 11:1249. doi: 10.3389/fpsyg.2020.01249

Zelenak, M. S. (2015). Measuring the sources of self-efficacy among secondary school music students. *J. Res. Music Educ.* 62, 389–404. doi: 10.1177/0022429414555018

Zimmerman, B. J. (2000). Self-Efficacy: an essential motive to learn. *Contemp. Educ. Psychol.* 25, 82–91. doi: 10.1006/ceps.1999.1016

**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.

frontiers in Psychology

BRIEF RESEARCH REPORT
published: 03 November 2021
doi: 10.3389/fpsyg.2021.769831

Check for updates

# Effect of Bow Camber and Mass Distribution on Violinists' Preferences and Performance

*Aurélie Tomezzoli*[1]*, Benjamin Michaud*[2]*, Eric Gagné*[3]*, Mickaël Begon*[2,4] *and Sonia Duprey*[1]*

[1] Univ Lyon, Université Claude Bernard Lyon 1, Univ Gustave Eiffel, LBMC UMR_T9406, Lyon, France, [2] Laboratoire de Simulation et de Modélisation du Mouvement, École de Kinésiologie et des Sciences de l'Activité Physique, Faculté de Médecine, Université de Montréal, Montréal, QC, Canada, [3] Wilder & Davis, Montréal, QC, Canada, [4] Research Center, Sainte-Justine Hospital, Montréal, QC, Canada

Little is known about how bow mechanical characteristics objectively and quantitatively influence violinists' preferences and performance. Hypothesizing that the bow shape (i.e., camber) and mass distribution modifications would alter both violinists' appreciations of a bow and objective assessments of their performance, we recruited 10 professional violinists to play their own violin using 18 versions of a single bow, modified by combining three cambers and six mass distributions, in random order. A musical phrase, composed for this study, was played legato and spiccato at three octaves and two tempi. Each violinist scored all 18 bows. Then, experts assessed the recorded performances according to criteria inspired by basic musical analysis. Finally, 12 audio-descriptors were calculated on the same note from each trial, to objectivise potential acoustic differences. Statistical analysis (ANOVA) reveals that bow camber impacted the violinists' appreciations ($p < 0.05$), and that heavier bow tips gave lower scores for spiccato playing ($p < 0.05$). The expert evaluations reveal that playing with a lighter bow (tip or frog), or with a bow whose camber's maximum curvature is close to the frog, had a positive impact on some violinists' performance (NS to $p < 0.001$). The "camber-participant" interaction had significant effects on the violinists' appreciations ($p < 0.01$ to $p < 0.001$), on the expert's evaluation and on almost all the audio-descriptors (NS to $p < 0.001$). While trends were identified, multiple camber-participant interactions suggest that bow makers should provide a variety of cambers to satisfy different violinists.

**Keywords: violin, bow, preference, musical analysis, acoustic**

OPEN ACCESS

***Edited by:***
*Gary Edward McPherson,*
*The University of Melbourne, Australia*

***Reviewed by:***
*Nicola Di Stefano,*
*National Research Council (CNR), Italy*
*Caitlyn Trevor,*
*University of Zurich, Switzerland*

***Correspondence:**
*Sonia Duprey*
*sonia.duprey@univ-lyon1.fr*

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

***Received:*** *02 September 2021*
***Accepted:*** *12 October 2021*
***Published:*** *03 November 2021*

***Citation:***
*Tomezzoli A, Michaud B, Gagné E, Begon M and Duprey S (2021) Effect of Bow Camber and Mass Distribution on Violinists' Preferences and Performance. Front. Psychol. 12:769831. doi: 10.3389/fpsyg.2021.769831*

## INTRODUCTION

Violin bows' shape and materials have changed little since the end of the eighteenth century. However, trade restrictions have made the wood used in bow sticks, Pernambuco (*Paubrasilia echinata*, also termed brazilwood and derived mainly from Brazilian forests), less available and more expensive (Convention on International Trade in Endangered Species of Wild Fauna Flora, 2017). This increasing rarity raises the question of what makes a "good" pernambuco bow, if satisfactory alternatives are to be developed: both the point of view of violinists, through the assessment of their preferences, and of listeners, through the assessment of violinists' musical performance should be considered.

The scientific interest for string vibrations is ancient, as links between a monochord vibration and sound pitches are well known since the Antiquity (Abromont and de Montalembert, 2001), and is still relevant, as can be seen for instance with the studies of the non-steady parts of bowing, i.e., the attack and release of the sound also known as transients (Badeau, 2005; Castellengo et al., 2015). In further work, the relationships between the string vibration and the bow have been studied (Schoonderwaldt et al., 2007; Schoonderwaldt and Altenmüller, 2009). With the development of new technological possibilities appeared a growing interest for the upper limb biomechanics during the bowing (Schoonderwaldt and Altenmüller, 2011; Kelleher et al., 2013), as well as the description of the motion of the bow itself, which remains a technological challenge (Provenzale et al., 2021).

Most studies investigating violinists' preferences have focused on the violin itself (Saitis et al., 2012, 2015; Fritz and Dubois, 2015). Instrumentalists' rankings of violins have been shown to be reproducible (Saitis et al., 2012), and the more precise the musical task, the more the instrumentalists agree on their assessments of a violin. Violinists also discriminate better between violins, in terms of richness of sound, when playing rather than listening to them (Saitis et al., 2015). Such studies set the example in experimental design to assess instrumentalists' preferences. However, no study has addressed violin bow preferences. A single study (Caussé et al., 2001) explored potential connections between verbal descriptors used by instrumentalists and the static (measured) or dynamic (estimated by finite element modeling) mechanical properties of the bow. Their preliminary findings showed no evidence of a link. Other studies on the violin bow addressed its modal and acoustical responses (Ravina et al., 2008) and its spectral content (Schoonderwaldt and Altenmüller, 2009), but without considering how these might relate to violinists' preferences (Woodhouse, 2014). There is clearly a gap in our knowledge concerning the relationship between bow mechanical characteristics and preferences.

Determining whether a musical performance is satisfactory inevitably involves human perception. Although human evaluations are subjective by nature—i.e. dependent on personal musical culture, place and time—their reproducibility has been documented: music student grades have been reported to show an intra-rater correlation coefficient of 0.87–0.99 (Franzén, 1969; Beazley, 1981) and an inter-rater correlation of 0.95–0.95 (Beazley, 1981). Evaluations by experts were found to be most reproducible if the judges were experienced and judged from recordings rather than from real-time performances (Salvador, 2010). Furthermore, musicians were found to be more accurate than non-musicians in their perceptions (McAdams et al., 1999; Fritz et al., 2008).

A complementary approach to performance rating involves calculating audio-descriptors, which are numerical indicators that account for particular dimensions of a sound. The calculation is based either on sound energy envelopes or on frequency spectrums, using either the Fast Fourier Transform (FFT) or the harmonic model (**Table 1**). Some parameters are temporal, describing the variation of one parameter across time, while others are global, describing an entire sound (Peeters et al., 2011). Audio-descriptors are used in several ways in the literature. The approach taken by music information research, for example to automatically classify instruments, distinguishes between different sounds by including as many descriptors as possible, even those not relevant in terms of human perception. This approach has been found effective in this context, correctly classifying more than 90% of the instrumental sounds tested (Siedenburg et al., 2016). Another approach, used in musical psychology, assesses the dissimilarity between sounds using a small number of audio-descriptors whose relevance in terms of human perception has been shown (**Table 1**; Peeters et al., 2011). For the violin, research has focused on the correlation of audio-descriptors with verbal descriptors (Fritz et al., 2012; Saitis et al., 2012). As for bow assessment based on audio-descriptors, the impact of different bowing parameters, as speed and pressure, has been assessed, but without examining bow mechanical characteristics themselves. Schoonderwaldt and Altenmüller (2009) showed that the spectral centroid (**Table 1**) increases mainly with the pressure exerted by the bow, especially when played close to the bridge. Moreover, the fundamental frequency of the note decreases when bow pressure is high, especially when played far from the bridge. Conversely, fundamental frequency increases when a string is played rapidly near the bridge (Schoonderwaldt and Altenmüller, 2009). To date, audio-descriptors have not been used to assess the sound modifications generated by different bows.

The main objective of this study was to investigate the influence of bow mechanical characteristics, specifically its camber and mass distribution, on instrumentalists' preferences and on their musical performance. In keeping with the above literature, we chose to evaluate musical performance on the basis of both expert assessment and audio-descriptors.

# METHOD

## Data Collection

Ten professional musicians (9 women and 1 man, aged 27.7 ± 7.6 years), who had been playing the violin for 21.0 ± 8.5 years were recruited to participate in this study. The protocol was approved by the ethics committee of the University of Montreal (17-018-CERES-D). The violinists played their own personal violin, while a single bow (mass: 62 g; natural camber, i.e., bow shape: camber 1) was used, modified into 18 camber/mass distribution configurations.

A bow maker (coauthor EG) modified the bow behind a curtain, changing its camber and adding mass randomly in two steps [see the **Supplementary Material** and the video (S2M Lab, 2017) minute 0:42]. First, one of the three cambers was randomly selected. After being heated (see the **Supplementary Material**), the bow was bent to match one of three defined patterns. Then masses were randomly added at the tip (0, 1 or 2 g, i.e., 1.6 or 3.2% of the bow weight) and/or at the frog (0 or 2 g), using sticky gums. All camber and mass combinations were tested, i.e., 18 bow conditions for each violinist, except for one of the cambers (camber 1) for one of the violinists, due to lack of time. For 6 out of the 10 participants, one of the 18 conditions, randomly chosen for each camber, was repeated in immediate succession. The bow

**TABLE 1** | Main audio-descriptors as calculated in Peeters et al. (2011), variables transformed in our study, and camber-participant interaction effect on audio-descriptors.

| Sound representation | Audio-descriptor | Definition | Variable transformation | Camber—subject interactions |
|---|---|---|---|---|
| Temporal representation | RMS envelope | RMS of the amplitude of the temporal energy | ln (*x*) | $p < 0.001$ |
| | Attack | Attack duration | 1/x | $p < 0.01$ |
| | Release | Duration of the last phase of sound | $\sqrt{(1.55 + \ln(x))}$ | $p < 0.001$ |
| FFT representation | Spectral centroid | $\mu_1(t_m) = \sum_{k=1}^{K} f_k \cdot p_k(t_m)$ where $p_k(t_m) = [a_k(t_m)] \,/\, \sum_{k=1}^{K} a_k(t_m)$ | x | $p < 0.001$ |
| | Spectral variation (i.e., spectral flux) | $variation(t_m, t_{m-1}) = 1 - \frac{\sum_{k=1}^{K} a_k(t_{m-1}) a_k(t_m)}{\sqrt{\sum_{k=1}^{K} a_k(t_{m-1})^2}\sqrt{\sum_{k=1}^{K} a_k(t_m)^2}}$ | *x* | $p < 0.05$ |
| Harmonic representation | Fundamental frequency | $F_0$ | *x* | NS |
| | Inharmonicity | $inharm(t_m) = \frac{2}{f_0(t_m)} \frac{\sum_{h=1}^{H} (f_h(t_m) - hf_0(t_m))\, a_h^2(t_m)}{\sum_{h=1}^{H} a_h^2(t_m)}$ | ln (*x*) | $p < 0.05$ |
| | Noisiness | $noisiness(t_m) = \frac{E_N(t_m)}{E_T(t_m)}$ with $E_N(t_m) = E_T(t_m) - E_H(t_m)$; $E_T(t_m) = \sum_k a_k^2(t_m)$ | ln (1 – *x*) | $p < 0.001$ |
| | Odd-to-even harmonic ratio | $OER(t_m) = \frac{\sum_{h=1}^{H/2} a_{2h-1}^2(t_m)}{\sum_{h=1}^{H/2} a_{2h}^2(t_m)}$ | ln (*x*) | $p < 0.01$ |
| | Tristimuli (x3) | $T_1(t_m) = \frac{a_1(t_m)}{\sum_{h=1}^{H} a_h(t_m)}$ | ln (*x*) | NS |
| | | $T_2(t_m) = \frac{a_2(t_m) + a_3(t_m) + a_4(t_m)}{\sum_{h=1}^{H} a_h(t_m)}$ | *x* | NS |
| | | $T_3(t_m) = \frac{\sum_{h=5}^{H} a_h(t_m)}{\sum_{h=1}^{H} a_h(t_m)}$ | *x* | $p < 0.05$ |

*$f_k$, frequency of the partial k; $p_k$, normalized amplitude; $f_0$, fundamental frequency; h, multiple of the fundamental frequency; $a_h(t_n)$, $f_h(t_n)$, $\Phi_{h,0}(t_n)$, amplitude, frequency and initial phase; $E_T$ ($t_m$), total spectral energy; $E_N$ ($t_m$), noise energy.*

maker pretended to modify the bow as usual, ensuring that the violinist did not know the two conditions were identical.

The piece of music (**Figure 1**), especially composed to explore different facets of the bowing technique while keeping a certain degree of ecological validity, consisted of two sections, each involving a different type of articulation: first legato, then spiccato. Both were played over three octaves. The entire piece of music was played at two tempi, 60 beats per minute or 120 (an electronic metronome was used), in random order, for a total of 12 musical sections. The sound was recorded with a Zoom Q3HD recorder (condenser microphone, XY, 120° angle), set at a sampling rate of 44.1 kHz. The recorder was placed on the lectern, facing the violinist. The lectern position relative to the performer was kept relatively constant during recordings by floor markings.

## Evaluation

After each trial, the participants gave their overall appreciation of the bow, rating it from 0 to 10, as well as an appreciation specifically for playing legato and spiccato. No instructions were given regarding rating criteria.

From the recordings, two-step randomized by violinist and then by audio file, a first expert evaluated musical performance (coauthor TA, a composer). She searched for four types of defects, using a method inspired by musical analysis techniques: timbre errors (squeaking, wrong string played, unexpected bounced bowstroke), lack of pitch accuracy, lack of rhythm precision and articulation errors. She defined the defects in writing and via sound samples. Penalty scores (timbre score, pitch score, rhythm score and articulation score) were calculated for each bow configuration for each participant by incrementing them (+1) whenever a musical section contained at least one of the targeted defects. An overall penalty score was also calculated for the whole piece of music and for all the sections related to each musical parameter (types of articulation, octave and tempo). Penalty scores were then fitted to a scale ranging from 0 (flawless performance) to 10 (all musical sections were flawed). Then, 30 audio files (three per participant) were randomly selected. They were assessed again 1 month later by the same expert and a second expert (coauthor MB, a violinist and the composer of this piece of music) to assess grading repeatability, using coefficients of variation.

After averaging the two voices of the stereo recording, audio-descriptors that seemed the most relevant in terms of human perception were calculated using Matlab's Timbre Toolbox (Peeters et al., 2011). To avoid redundancy, we choose the descriptors which were the less linked to each other, i.e., that had a median rank correlation lower than 0.8 in Peeters et al. (2011) (see **Table 1**). As recommended by Peeters et al. (2011), the temporal audio-descriptors were characterized by their median value. Audio-descriptors were calculated on the C#3 (expected frequency of 277 Hz) spiccato tempo 60 (see **Figure 1**). This note was manually extracted using Audacity (Audacity Team, 2018).

## Statistics

To assess the repeatability of violinists' preferences, differences between their assessments of the two identical bow trials were calculated. Then, after checking the normality of these differences, the coefficient of variation between the first and


**FIGURE 1 |** The piece of music, consisting of two sections (spiccato, then legato), both played at three octaves. The C# analyzed using audio-descriptors is highlighted in yellow.

second trial was calculated. The coefficient of determination ($R^2$) of an analyse of variance (ANOVA) was used to assess the percentage of variance in preferences due to the difference between two repeated tests.

The influence of bow characteristics was assessed using the bow cambers, the mass added at the tip and the mass added at the frog as explanatory variables. Dependent variables were the violinists' preferences, the expert's penalty score and the audio-descriptors. The distribution of each variable was checked graphically to ensure its normality and perform any necessary transformations (**Table 1**). Homoscedasticity among subgroups was tested using a Levene test. After verification of these conditions of applicability, one-factor ANOVAs were performed. Due to high inter-participant variability, the factors were all re-tested with a two-factor ANOVA including a participant effect. More complex models were developed using this approach, i.e., by progressively adding the most significant variables (bottom-up approach).

Pearson's linear correlation coefficients were calculated among the violinists' different appreciations (overall, legato and spiccato playing) and among penalty scores and audio-descriptors, after normalizing each parameter by violinist (same mean and

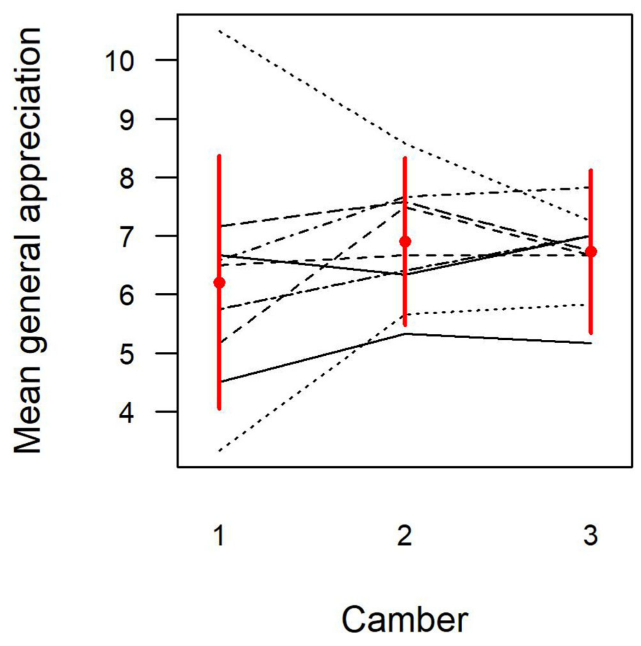



**FIGURE 2 |** Violinists' bow appreciations—average overall appreciation according to camber (1, 2 or 3) and violinist. Lines connect the values for each individual violinist.

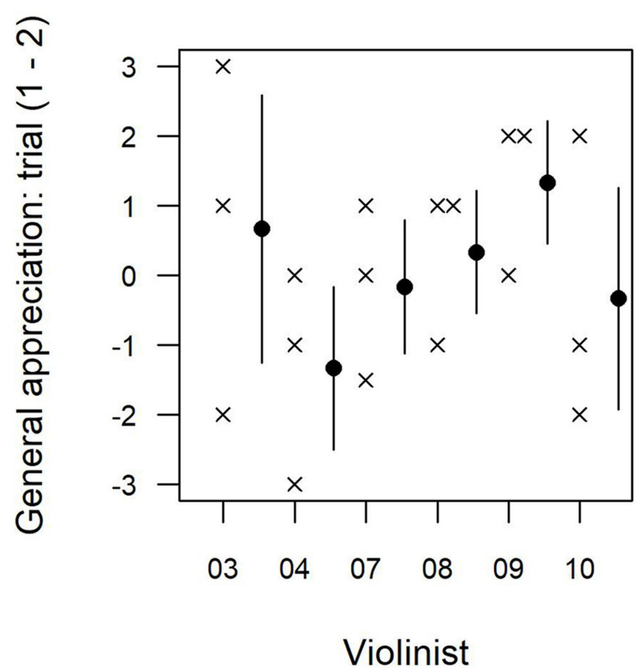



**FIGURE 3 |** Differences in violinists' general appreciations for the two tests performed under identical bow conditions (trial 1–trial 2) for six participants (3, 4 and 7–10).

standard deviation for each) and graphically verifying the shape of correlations on scatter plots.

# RESULTS

## Violinists' Preferences

The mass at the tip had a significant effect on appreciations of spiccato playing, and the camber had a significant effect on all appreciations. Since there was a major participant effect ($R^2 = 0.31–0.37$, $p < 0.001$, vs. $R^2 = 0.2$ at most, not significant, for the explanatory variables), it was taken into account simultaneously through two-factor ANOVA. Specifically, added mass at the tip of the bow reduced ratings for the spiccato on average (5.8, 6.5 and 6.6/10, for 2, 1 and 0 g addition, respectively, $p < 0.05$). Camber 1 was always the lowest rated (overall ratings of 6.2, 6.9 and 6.7/10, for cambers 1, 2 and 3, respectively), except by one violinist. However, there were differences in effect size and in preferences between cambers 2 and 3, generally explained more by a camber-participant interaction ($p < 0.01$ to $p < 0.001$) than by the camber itself ($p < 0.05$) (**Figure 2**). In the three-factor model, there was no interaction between violinists, added mass and camber.

In terms of repeatability of the violinists' appreciations, the mean of the absolute value of the difference between the two identical bow trials was 1.4/10 for the overall and legato assessments and 1.6/10 for the spiccato assessment. Half the participants raised their rating on average in the second trial, while the other half lowered their rating (**Figure 3**). The coefficient of variation between the values given for the first and second trials was 0.26 for overall assessment, 0.33 for spiccato, 0.27 for legato. The percentage of variance explained by difference between the two identical bow trials was 23% for the overall assessment, 30% for the legato assessment and 17% for the spiccato assessment.

After normalization by violinist, links between overall ranking and appreciations for spiccato and legato playing were respectively $r = 0.84$ and 0.87. Correlation between the latter two appreciations was $r = 0.62$.

## Experts' Penalty Ratings

The rating agreement between evaluations 1 month apart was very good (81.4%), indicating low intra-expert variability. The consistency of ratings between experts was good (70.1%), indicating reasonable inter-expert variability. Once again, all penalty ratings showed a major participant effect ($R^2 = 0.34–0.78$, $p < 0.001$, vs. at most $R^2 = 0.2$, NS, for the explanatory variables). After taking into account the participant effect, the main results were that a mass of 0 g at the tip had a positive effect on timbre ratings (5.6 vs. 6.3 and 6.0/10 for a 1 and 2 g mass, respectively, $p < 0.05$), and that a mass of 0 g at the frog had a positive effect on playing at high pitches (two last lines of the piece of music, **Figure 1**), depending on the violinist (mean: 4.1 vs. 4.2/10, $p < 0.05$).

Camber had an influence at low pitches and in fast passages ($p < 0.05$), in both of which cases camber 3 obtained the lowest penalty ratings (mean: 3.2 vs. 3.5 and 3.6/10 for camber 3, 1 and 2, respectively at low pitch and 3.4 vs. 3.8 and 3.7/10, respectively in fast passages). There were camber-participant interactions for rhythmic precision, for respect of articulation, at middle-range pitches ($p < 0.05$) and for overall ratings ($p < 0.001$).

After normalization by violinist, no links between timbre, pitch, rhythm and articulation penalty scores were found, nor between mutually exclusive musical parameter penalty scores (for example, between penalty scores for each of the three octaves).

### Audio-Descriptors

For the audio-descriptors that did not fit a normal distribution, variables were transformed to ensure normality (**Table 1**). For each audio-descriptor, there was a major participant effect ($R^2 = 0.18–0.78$, $p < 0.001$, vs. at most a non-significant $R^2 = 0.2$ for the explanatory variables). When the participant effect was taken into account, camber-participant interaction had a statistically significant effect on almost all audio-descriptors, whereas mass additions had no effect (**Table 1**).

After normalization by violinist, multiple links were found between audio-descriptors, computed two by two.

## DISCUSSION

The main objective of this study was to investigate the influence of bow mechanical characteristics on both instrumentalists' preferences and their musical performance. As hypothesized, bow modifications altered the violinists' appreciations, especially camber changes. Concerning musical performance, significant "camber-participant" interactions were found for each audio-descriptor. The experts' evaluations revealed that a camber with maximum curvature closest to the frog (Camber 3) had positive effects on playing, while added mass at the tip or frog had negative effects.

Acoustic response can differ across bows (Ravina et al., 2008). Specifically, we found that camber has a major impact, much more pronounced than bow mass distribution, on many sound parameters and on violinists' appreciations. The camber with maximum curvature closest to the tip (Camber 1) scored lowest, except with one violinist. Differences in violinists' preferences regarding the two other cambers, i.e., camber-participant interaction, could be explained by differences in the types of bow they habitually use, the ones with cambers most similar to theirs being possibly preferred. Documenting the characteristics of the bows generally used by the instrumentalists involved in experimental studies might therefore be worthwhile. On the other hand, these differing preferences could also reflect musicians' personal aesthetic aims. Actually, overall quality rating and preference rating are known to be poorly correlated (Wollman et al., 2014), likely suggesting different expectations regarding musical performance. In the end, since our violinists' appreciations and experts' evaluations often depended on the violinist (camber-participant interaction), a good general principle for bow makers would be to provide bows with different cambers (affecting overall bow stick design) to satisfy diverse violinists.

Throughout the history of bow making, a balance has been sought between bow length, mass and "nervosity," i.e., its ability to respond quickly to the violinist gesture (Bachmann et al., 2001), which seems likely to refer to its dynamic properties. In the present study, the mass added at the tip and frog affected both the violinists' appreciations of the bow and their performance. At the tip, added mass had a negative impact both on instrumentalists' ratings for spiccato playing and on the frequency of timbre flaws as evaluated by the expert. A deleterious effect on the spiccato was expected: the bow "bounces" less well with more weight at the tip, since the center of mass may have shifted and/or the bow may be more difficult to control. Adding 2 g to the frog also had a negative impact, i.e., higher expert penalty ratings, with some players producing more defects in the treble. Bow mass and nervosity are linked to the nature of the materials used and to the geometry of the bow (Bachmann et al., 2001). Our data, together with the new prospect of using carbon fibers to make lighter bows, encourage reconsideration of the total weight of bows, as well as their mass balance and their dynamic behavior. This field of research is complex, as the bow material and geometry also impact the risk for bow fatigue failure (Bachmann et al., 2001). The expectations of different musicians have to be accounted, as highlighted above, and the musical patterns used for assessing the bow quality can also be adjusted to specific musical repertoires (Koechlin, 1956; Penesco, 1986).

While the present protocol was designed in a comprehensive way, taking into account instrumentalists' preferences, experts' evaluations and effects on sound, some limitations remain. In terms of expert evaluation, the reproducibility, even though it is reasonable (consistency of 81.4 and 70.1% for intra- and inter-expert assessments), could be improved. As for audio-descriptors, calculations could be extended to include more musical notes (Tomezzoli et al., 2019). Furthermore, given the multiple links found between the audio-descriptors used here, their number could be reduced. Finally, the highly significant inter-participant effect found here reflects not only the effect of inter-individual variations between participants but also differences in the violin and sound recording. Studies with larger samples of instrumentalists will need to be performed before these findings can be generalized to the entire population of violinists.

## AUTHOR'S NOTE

The results reported in the present article were partly presented at the 44th congress of the French Biomechanics Society (see Tomezzoli et al., 2019). With respect to the conference paper, we have detailed the Methods section to allow replication studies and have included experts' evaluations of violinists' performance in the dataset.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

The study involving human participants was reviewed and approved by the Ethics Committee of the University of Montreal

(17-018-CERES-D). The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

AT, BM, MB, and SD contributed to conception and design of the study. BM, EG, MB, and SD contributed to the acquisition and analysis or interpretation of data for the work. AT performed the statistical analysis and wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

## ACKNOWLEDGMENTS

This study was carried out within the framework of the Associated International Laboratory EVASYM.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2021.769831/full#supplementary-material

## REFERENCES

Abromont, C., and de Montalembert, E. (2001). *Guide de la théorie de la musique*. Paris: Fayard H. Lemoine.

Audacity Team (2018). *Audacity (2.2.2) (Computer Software)*. Boston: Free Software Foundation. Availability online at: http://audacity.sourcefOrge.net/

Bachmann, W., Seletsky, R., Boyden, D., Liivoja-Lorius, J., Walls, P., and Cooke, P. (2001). "Bow," in *Grove Music Online*, eds. S. Sadie and G. Grove (Oxford: Oxford University Press). doi: 10.1093/gmo/9781561592630.article.03753

Badeau, R. (2005). *Méthodes à haute résolution pour l'estimation et le suivi de sinusoïdes modulées. Application aux signaux de musique*. Dissertation Thesis. Paris: Télécom ParisTech. Available online at: https://pastel.archives-ouvertes.fr/tel-00009321 (accessed October 6, 2021).

Beazley, H. V. (1981). Development and validation of a music education competency test. *J. Res. Music Educ.* 29, 5–10. doi: 10.2307/3344674

Castellengo, M., Liénard, J.-S., and Bloch, G. (2015). *Écoute musicale et acoustique: avec 420 sons et leurs sonagrammes décryptés*. Paris: Eyrolles.

Caussé, R., Maigret, J. P., Dichtel, C., and Bensoam, J. (2001). "Study of violin bow quality," in *International Symposium on Musical Acoustics*.

Convention on International Trade in Endangered Species of Wild Fauna and Flora (2017). Available online at: https://cites.org/eng/disc/text.php#II (accessed September 23, 2018).

Franzén, B. (1969). Entrance test validity. *J. Res. Music Educ.* 17, 62–68. doi: 10.2307/3344187

Fritz, C., Blackwell, A. F., Cross, I., Woodhouse, J., and Moore, B. C. J. (2012). Exploring violin sound quality: investigating English timbre descriptors and correlating resynthesized acoustical modifications with perceptual properties. *J. Acoust. Soc. Am.* 131, 783–794. doi: 10.1121/1.3651790

Fritz, C., Cross, I., Moore, B., and Woodhouse, J. (2008). Perceptual thresholds for detecting modifications applied to the acoustical properties of a violin. *J. Acoust. Soc. Am.* 122, 3640–3650. doi: 10.1121/1.2799475

Fritz, C., and Dubois, D. (2015). Perceptual evaluation of musical instruments: state of the art and methodology. *Acta Acust. United Acust.* 101, 369–381. doi: 10.3813/AAA.918833

Kelleher, L. K., Campbell, K. R., and Dickey, J. P. (2013). Biomechanical research on bowed string musicians: a scoping study. *Med. Probl. Perform Art* 28, 212–218.

Koechlin, C. (1956). *Traité de l'orchestration*. Paris: M. Eschig.

McAdams, S., Beauchamp, J. W., and Meneguzzi, S. (1999). Discrimination of musical instrument sounds resynthesized with simplified spectrotemporal parameters. *J. Acoust. Soc. Am.* 105(2 *Pt* 1), 882–897. doi: 10.1121/1.426277

Peeters, G., Giordano, B. L., Susini, P., Misdariis, N., and McAdams, S. (2011). The Timbre Toolbox: extracting audio descriptors from musical signals. *J. Acoust. Soc. Am.* 130, 2902–2916. doi: 10.1121/1.3642604

Penesco, A. (1986). *Les Instruments du quatuor: technique et interprétation*. Paris: la Flûte de Pan.

Provenzale, C., Di Stefano, N., Noccaro, A., and Taffoni, F. (2021). Assessing the bowing technique in violin beginners using MIMU and optical proximity sensors: a feasibility study. *Sensors* 21:5817. doi: 10.3390/s21175817

Ravina, E., Silvestri, P., and Airenti, A. (2008). Experimental modal analysis of bow. *J. Acoust. Soc. Am.* 123:3659

S2M Lab (2017). *L'Archet et son effet sur la biomécanique*. Available online at: https://www.youtube.com/watch?v=0-NHGJz6vTw (accessed October 08, 2021).

Saitis, C., Giordano, B., Fritz, C., and Scavone, G. (2012). Perceptual evaluation of violins: a quantitative analysis of preference judgments by experienced players. *J. Acoust. Soc. Am.* 132, 4002–4012. doi: 10.1121/1.4765081

Saitis, C., Scavone, G., Fritz, C., and Giordano, B. (2015). Effect of task constraints on the perceptual evaluation of violins. *Acta Acust. Unit. Acust.* 101, 382–393. doi: 10.3813/AAA.918834

Salvador, K. (2010). How can elementary teachers measure singing voice achievement? a critical review of assessments, 1994–2009. *Appl. Res. Music Educ.* 29, 40–47. doi: 10.1177/8755123310378454

Schoonderwaldt, E., and Altenmüller, E. (2009). The violinist's sound palette: spectral centroid, pitch flattening and anomalous low frequencies. *Acta Acust. United Acust.* 95, 901–914. doi: 10.3813/AAA.918221

Schoonderwaldt, E., and Altenmüller, E. (2011). "Mastering the violin: motor learning in complex bowing skills," in *Proceedings of the International Symposium on Performance Science*, 649–654.

Schoonderwaldt, E., Guettler, K., and Askenfelt, A. (2007). "Schelleng in retrospect—a systematic study of bow force limits for bowed violin strings," in *International Symposium on Musical Acoustics*, Barcelona, Spain

Siedenburg, K., Fujinaga, I., and McAdams, S. (2016). A comparison of approaches to timbre descriptors in music information retrieval and music psychology. *J. New Music Res.* 45, 27–41. doi: 10.1080/09298215.2015.1132737

Tomezzoli, A., Michaud, B., Begon, M., and Duprey, S. (2019). Bow camber and mass distribution: violinist's preferences and sound descriptors. *Comput. Methods Biomech. Biomed. Eng.* 22, S240–S242. doi: 10.1080/10255842.2020.1714255

Wollman, I., Fritz, C., Poitevineau, J., and McAdams, S. (2014). Investigating the role of auditory and tactile modalities in violin quality evaluation. *PloS ONE* 9:e112552. doi: 10.1371/journal.pone.0112552

Woodhouse, J. (2014). The acoustics of the violin: a review. *Rep. Prog. Phys.* 77:115901. doi: 10.1088/0034-4885/77/11/115901

**Conflict of Interest:** EG was employed by Wilder & Davis.

The remaining 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.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Psychology

HYPOTHESIS AND THEORY
published: 21 December 2021
doi: 10.3389/fpsyg.2021.732729

# Constraint-Based Sound-Motion Objects in Music Performance

*Rolf Inge Godøy*[1,2*]

[1]*Department of Musicology, University of Oslo, Oslo, Norway,* [2]*RITMO Centre for Interdisciplinary Studies in Rhythm, Time and Motion, University of Oslo, Oslo, Norway*

The aim of this paper is to present principles of constraint-based sound-motion objects in music performance. *Sound-motion objects* are multimodal fragments of combined sound and sound-producing body motion, usually in the duration range of just a few seconds, and conceived, produced, and perceived as intrinsically coherent units. Sound-motion objects have a privileged role as building blocks in music because of their duration, coherence, and salient features and emerge from combined instrumental, biomechanical, and motor control constraints at work in performance. Exploring these constraints and the crucial role of the sound-motion objects can enhance our understanding of generative processes in music and have practical applications in performance, improvisation, and composition.

**Keywords: sound objects, performance, intermittency, motor gestalt, constraints**

OPEN ACCESS

***Edited by:***
*Oscar Casanova,*
*University of Zaragoza, Spain*

***Reviewed by:***
*Clément Canonne,*
*Centre National de la Recherche Scientifique (CNRS), France*
*Frederic Bevilacqua,*
*Institut de Recherche et Coordination Acoustique Musique (IRCAM), France*

****Correspondence:***
*Rolf Inge Godøy*
*r.i.godoy@imv.uio.no*

***Specialty section:***
*This article was submitted to Performance Science, a section of the journal Frontiers in Psychology*

***Received:*** *29 June 2021*
***Accepted:*** *23 November 2021*
***Published:*** *21 December 2021*

***Citation:***
*Godøy RI (2021) Constraint-Based Sound-Motion Objects in Music Performance. Front. Psychol. 12:732729. doi: 10.3389/fpsyg.2021.732729*

## INTRODUCTION

### Object Perspectives

A rapid harp glissando, a fast drum fill, a harpsicord ornament, a trumpet signal, or a tremolo crescendo on a cymbal are all examples of sound-motion objects in the sense that they combine salient sound energy envelopes with salient sound-producing body motion shapes into strongly coherent entities, typically in the duration range of 0.3–5 s. The main features of sound-motion objects are that they are multimodal, meaning that they include sensations of both sound and sound-producing body motion, and that they are conceived and perceived holistically as strongly coherent and stable units and hence may be called *objects* in musical performance.

Besides enhancing our understanding of the multimodal nature of music, such sound-motion objects are privileged both in the sound domain and in the human motion domain, by making present holistic features not found at shorter timescales, nor easy to focus on at longer timescales, for instance: An ornament object does not exist at timescales smaller than the duration of the entire ornament, and at larger timescales, the ornament will just be one element in a larger context, not enabling a focus on the features of that particular ornament. And in view of body motion constraints, there is, as we shall see, converging evidence that the typical duration of sound-motion objects may be optimal for both control and effort in sound-producing body motion.

The idea of sound-motion objects stems from our research on music-related body motion (see Godøy et al., 2016 for an overview), but is derived from Pierre Schaeffer's theory of *sound objects* (Schaeffer, 1998, 2012, 2017; Chion, 2009; Godøy, 2021). The main idea of

Schaeffer's sound objects is that they have a limited duration and convey salient features linked with their overall dynamic, timbral, and pitch-related shapes. In Schaeffer's sound object theory, these features may be further differentiated by a top-down and increasingly detailed classification scheme, however, always based on the subjective perceptions of these features as shapes (Godøy, 2021). The sound object focus of Schaeffer, although originating in connection with electroacoustic music of the 1940s and 1950s (Schaeffer, 2012), is very general and applicable to most kinds of music, regardless origin and genre, provided the point of departure for any inquiry is the subjectively perceived dynamic, timbral, textural, and pitch-related shapes (cf. a number of audio examples from different kinds of music in Schaeffer, 1998).

In addition, our own and other work of the last couple of decades on music-related body motion suggest that there are strong links between perceptual images of musical sound and sensations of body motion, be that as detailed images of *sound-producing motion* of performers, or be that as so-called *sound-accompanying motion* by listeners with more generic energy envelope resemblances between sound and other kinds of body motion, such as in walking, dancing, or gesticulating (Godøy and Leman, 2010; Maes et al., 2014). Actually, it may often be difficult to tease apart what are sound sensations and what are body motion sensations in music perception and/or musical imagery. For instance, in the case of ferocious drumming, what in our perception is due to the energy of the drum sound, and what is due to the sensation of energetic mallets, hands, arms, shoulders, torso, etc., motion of the performer that we see or just imagine? Or is our perception of the drumming made up of some combination of all these sensations?

From the perspective of the so-called *motor theory of perception* (Godøy, 2003; Galantucci et al., 2006), the answer is that we tend to perceive sound, albeit variably so, in terms of some mental images of how we believe the sound has been produced. To find out more about this, we previously made studies of so-called *air instrument performance* (Godøy et al., 2006) and of so-called *sound-tracing* (Nymoen et al., 2013). It seemed that listeners, with different levels of musical expertise, all tended to render some feature of musical sound by body motion; however, expert musicians tended to render more details of assumed sound-producing body motion.

In line with the motor theory claims of close connections between sound and assumed sound-producing body motion, the concept of *constraints* should be understood here not only in a limiting sense of possible vs. impossible, or easy vs. difficult, in performance, but should also be understood in a positive sense of providing the basic conditions for music performance. This means that various constraints may become an integral part of musical creativity and expression, even to the extent that we as listeners may actually expect the music to be in line with familiar constraints.

One outcome of this approach will be constraint-based shifts between muscle contractions and muscle relaxation and associated parsing of performers' sound-producing body motion into chunks, i.e., into what we may call motion objects. In a striking parallel to sound objects, some human movement researchers have suggested a similar object perspective in the form of so-called *motor gestalts* (Klapp and Jagacinski, 2011). Motor gestalts are preprogrammed chunks of body motion that are triggered and carried out as holistic entities, thus optimizing motor control. In a similar vein, other research on human motor control has suggested that there is an intermittent, or discontinuous, element in body motion, meaning that there is a point-by-point triggering of motion based on a piecewise anticipatory planning of the motion to come (Loram et al., 2011). The combination of sound objects with intermittently triggered motor gestalts in body motion is then the basis for coining the term *sound-motion objects* in our research (Godøy et al., 2016).

The use of the word *object* in our present context is based on the observation that although both sound and motion are ephemeral and temporally distributed phenomena, they seem to leave some more solid traces in our minds, as well as to be conceived and perceived holistically at discontinuous points in time. This is the core meaning of "object" here, namely *a more or less solid mental image of a delimited fragment of unfolding sound and motion*. Crucially, this use of the word "object" does not signify something that is opposed to "subjective" sensations, but on the contrary takes individual-subjective perceptual images of sound and motion as integral to musical experience; however, there may in many cases be inter-subjective agreements about the experience of various features, as was indeed the point of departure for Schaeffer's sound object theory. The term "object" here is then a collective term for sound and motion features as *shapes*, and the corresponding research method here is that of exploring shapes associated with sound-motion objects as well as the possible *correlations* between subjective object images and recorded sound and motion data (such correlation mapping is also a core element in Schaeffer's theory, see Godøy (2021) on this). The epistemological approach here may be regarded as "object-focused" (the term "object-oriented" already taken and used in programming languages), in the sense that object images of sound and motion are the core components of our theory.

This object-focused approach touches on long-standing enigmatic relationships between notions of continuity and discontinuity in both philosophy and psychology, and which the concept of sound-motion objects seeks to address by what can be called the *intermittency hypothesis*. This hypothesis suggests that at the timescale of sound-motion objects (as mentioned, in the 0.3–5 s range), we may find compressed overview images of both sound and motion. This hypothesis also suggests that such "all-at-once" or "instantaneous" images of any segment of sound and motion can work both retrospectively (as recollections) and prospectively (as anticipatory control) and, hence, be temporally bidirectional.

The main aim of this paper is then to present the theoretical basis for constraint-based sound-motion objects in music performance, including the intermittency hypothesis as an attempt to work around some of these constraints. Given the explorative nature of this main aim, the present paper will be a so-called *hypothesis and theory paper*, but with some illustrative examples from our work on sound-producing motion.

For a start, it could be useful in the next two subsections of the present introductory main section to consider issues of *timescales* as well as of *constraints* in music performance. These subsections will be followed by a main section focused on *sound-producing motion*, containing four subsections concerned, respectively, with *motion features* and phenomena essential for the fusion of sound and motion into coherent entities, so-called *phase transition* and *coarticulation,* as well as the phenomenon of *idioms* of sound-producing motion, i.e., what may be considered particularly successful sound-motion objects. Then follows a main section on *control theories*, containing three subsections concerned with *motor control*, *motor gestalts* and the so-called *posture-based theory*, a theory suggesting that motion is centered on salient postures at selected moments in time. All this will be seen to converge in the main section on *musical intermittency*, containing subsections on *the intermittency hypothesis*, the *triggering* of motion chunks, ending with the concept of *sound-motion objects*. Finally, there will be a discussion section on the various advantages and challenges of the sound-motion objects perspective on music performance.

## Timescales

The concept of sound-motion objects depends crucially on duration criteria. To see why this is the case, it could be useful to have an overview of the different timescales for sound and body motion features in music performance. These timescales extend from those in the sub-millisecond range to those in the range of several seconds, minutes, and beyond, here grouped into three main categories (Godøy, 2006):

- *Micro*, denoting the below ≈ 300 ms event timescale and encompassing stationary sound features such as pitch, timbre, dynamics, as well as some fast transients and fluctuations, including salient timbral-textural features in the duration region of around 250 ms (Gjerdingen and Perrott, 2008).
- *Meso*, typically the timescale of sound-motion objects in the mentioned duration range of 0.3–5 s. This is the timescale of the most salient sound object features in Schaeffer's theory as well as in more traditional Western music theory, such as motives, ornaments, contours, and modality conveying features of style and affect, as well as sense of corresponding body motion. There may very well be a general predisposition for the meso timescale as suggested by Ernst Pöppel in that the approximately 3-s time window is optimal for human perception and cognition in several domains (Pöppel, 1997). The meso timescale also corresponds to so-called *short-term memory* where basic processing and feature extraction are assumed to go on (Snyder, 2000).
- *Macro*, denoting the longer than meso timescale, with durations in the range of up to minutes and even hours. However, the efficacy of these large-scale forms as proclaimed by mainstream Western music theory, as well as by Schenkerian analysis, could be questioned (Eitan and Granot, 2008). There have been debates between proponents of more meso timescale views and more macro timescale views, for instance with Jerrold Levinson's "concatenationism" vs. Peter Kivy's "architectonicism" (Levinson, 2015). In our present context of constraint-based sound-motion objects, the main point is to see the convergence of several salient sound and motion features in the meso timescale of sound-motion objects, hence that there is an affinity here with Levinson's ideas, yet also recognizing that there may always be some larger context for these objects, and that the focus in our perception may vary between different timescales.

For sound-motion objects, we may also have different nested timescales, extending from that of the entire sound-motion object down to that of its internal spectral content, but where the meso timescale is privileged in manifesting the convergence of several different perceptual-cognitive features. Interestingly, the initial reason for Schaeffer's focus on meso timescale sound objects was the following two experiences (Chion, 2009; Schaeffer, 2017):

- The experience of countless repeated listening to looped fragments of sound in the early days of electroacoustic music, the so-called *closed groove* experience, emphasizing the crucial importance of the overall dynamic, timbral, and pitch-related shapes of the sound objects.
- The experience of the so-called *cut bell*, of how the different parts of any unfolding sound contribute to the total perceptual image of the sound, in particular how the dynamic shapes of the attack segments color the perceptual image of the subsequent stationary segments of sounds.

These two experiences demonstrated that sound features are temporally distributed, and hence, that a sound object should take the entire sound fragment into account.

For these features to become perceptually manifest, Schaeffer added the constraint of the so-called *suitable object* which included some duration and content criteria. Durations should be sufficient to encompass salient sequentially unfolding events, e.g., attack and sustain segments, yet not too long, nor too diverse, or too static, and the examples given in (Schaeffer, 1998) are mostly in the mentioned 0.3. to 5 s range, sometimes shorter and in exceptional cases longer, i.e., up to 30 s. The too short ones can become acceptable with pauses between the objects, so that a pause actually becomes part of an object. Furthermore, an important feature of sound objects is what Schaeffer called *facture*, essentially denoting the energy envelope of the sound object and making a link with body motion, and hence, also making a link also with the timescales of body motion (Godøy, 2006).

Furthermore, timescales concern not only duration issues, but also issues of continuity vs. discontinuity in perception and cognition. This was much discussed in philosophy and psychology toward the end of the nineteenth and beginning of the twentieth century. Edmund Husserl argued that perception and cognition proceed in a discontinuous manner by a series of so-called now-points, points in time for a cumulative and prospective overview of a segment of sensory unfolding, and that without such intermittent stepping out of the sensations stream, we would not be able to extract any meaning from our experiences (Husserl, 1991; Godøy, 2010).

Husserl's model of a point-by-point overview image of past and future events is relevant for a number of present issues in perception and cognition, *cf.* Schaeffer's view of sound object perception needing to take the entire fragment into account. And in the case of motor control, Husserl's model resembles a discontinuous, point-by-point scheme of anticipatory motor control as manifest in the ideas of motor gestalts and intermittent control. This means that timescale issues are closely linked with core issues of motor control, and as we shall see, in particular with the intermittency hypothesis, for instance, the idea of "segmented control" (Sakaguchi et al., 2015) seems to resemble this "now-point" view of cognition.

At the object timescale, motion control and motion effort have been referred in some of the intermittency literature as "serial ballistic," "open loop," or "feedforward," terms essentially expressing the view that at this timescale, there is a discontinuity of control and effort, yet seen from the outside, the resultant sound and motion may appear to be continuous.

## Constraints

Taking instrument and body motion constraints of music performance into account means taking a concrete, that is, a non-abstract, approach to music cognition. Furthermore, this means regarding symbols of Western music notation as a sparse script for concrete sound-producing motion to be manifested on concrete instruments, necessitating a transformation and an adaptation to a set of combined instrument-motion constraints.

As for instrument constraints, they include both acoustic and ergonomic features. The various modes of excitation, such as hitting, plucking, stroking, and blowing and corresponding modes of resonance, form the basis for sound-motion objects. Singular instrument sound objects may in turn have multiple internal features, e.g., the grainy quality of bowed deep double bass tone, or the buzzing sound of a snare drum.

On top of the instrument (and room) constraints, body motion constraints are shaping the sound-motion objects, first of all by the obvious fact that all body motion takes time, i.e., that there is no instantaneous displacement of the sound-producing effectors (fingers, hands, arms, feet, vocal apparatus). This means that sound events are embedded in motion events, and that there may be fusion of motion events by the phenomena of so-called *phase transition* and *coarticulation*:

- Phase transition in motion contexts denotes a categorical change in motion mode based on incremental changes in speed or rate of motion (Haken et al., 1985), for instance, between protracted singular strokes and fast tremolo motion in bowing, with emergent constraint-based changes in the motion amplitude of the bow.
- Since all body motion takes time, there will always be a contextual smearing in the form of coarticulation, meaning the contextual fusion of otherwise separate sound and motion events into new and longer sound-motion objects (Godøy, 2014).

There will be more details on these fusion phenomena later, and for now, we should not forget some other constraints of sound-producing motion, such as limitations of speed and endurance, need for rests, and need for motion strategies to avoid strain injury (Altenmüller et al., 2006), all contributing to the shaping of sound-motion objects. In terms of optimization of sound-producing body motion, we may also note:

- There seems to be a tendency toward minimization of energy expenditure in expert musicians (Winges et al., 2013).
- Fluency is a hallmark of experts' minimization of effort as opposed to non-experts' clumsiness (Gonzalez-Sanchez et al., 2019).
- Concrete, non-abstract, and logistic-ergonomic motion is often needed, e.g., in drum set performance (Godøy et al., 2017).
- Patterns of sound-producing motion that are particularly successful in generating good-sounding results with minimal effort, known as *idioms*, are important in the context of sound-motion objects because they testify to the extensive fusion of sound and motion optimization.
- On the other hand, implementing motion constraints may drastically change features of the output sound (Rozé et al., 2020).

It may sometimes be difficult to distinguish what are basically cognitive control constraints from what are more biomechanical constraints; however, there seems to be agreement that there is a control constraint with the so-called *psychological refractory period* (PRP). The PRP is believed to impose a limitation of around 0.5 s for initiating new motion in the course of any currently ongoing motion (Klapp and Jagacinski, 2011) and has the following consequences:

- A workaround solution to the PRP constraint may be anticipatory motor control in the form of the mentioned motor gestalts, meaning that an entire motion chunk may be carried out without the need for attention to details. This means that with PRP making continuous feedback control impossible, PRP leads to intermittent control, and hence, intermittency is constraint-based (Loram et al., 2014).
- Recognizing that all control processes in human motion take time because of inherent speed limitations of the neurocognitive apparatus, there have been long-lasting discussions of so-called *open loop* (no continuous feedback) vs. *closed loop* (continuous feedback) in human motor control; however, there may now be some kind of half-way agreement, cf. (Hanneton et al., 1997; Desmurget and Grafton, 2003), as also suggested by the intermittency hypothesis.
- Differentiating timescales could be a solution to *feedback* vs. *feedforward* disagreements in the sense that they can coexist as interleaved phenomena, i.e., assuming feedforward control being the case for rapid and continuous motion; however, it can alternate with feedback control at intermittent intervals (and with subsequent error correction), as suggested by the principle of "observe continuously, act intermittently" in intermittent control (Loram et al., 2011, p. 317).

In sum, constraints concern both motion and the control of motion; hence, the next sections will be about basic motion features, first from the more biomechanical and bottom-up constraints point of view, followed by the more top-down control point of view of constraints, and after that, there will be a focus on what are basically optimization elements, including the intermittency hypothesis.

## SOUND-PRODUCING MOTION

### Motion Features

Music-related body motion is often differentiated into the main categories of sound-producing motion, e.g., such as hitting, stroking, kicking, bowing, blowing, and sound-accompanying motion, e.g., dancing, walking, gesticulating (Godøy and Leman, 2010). Although the boundaries between these two main categories may sometimes be blurred, such as in cases of theatrical motion by musicians, the main feature of sound-producing motion is that of contributing to the generation of musical sound. This will first of all include what we call *excitatory motion*, i.e., motion that transfers energy from the body to an instrument, but it will also include so-called *modulatory motion*, such as for changing pitch or timbre, as well as *ancillary motion*, for instance, for optimizing postures and help avoid strain injury as well as help in shaping musical expression (Cadoz and Wanderley, 2000).

Furthermore, there are a number of attributes of sound-producing motion such as trajectory shape, amplitude, velocity, acceleration, and periodicity that may all variably contribute to the features of the output sound, so much so that sound-producing motion and output sound may fuse into our sound-motion objects. With available technologies and analytic tools, it is possible to zoom in on details of motion, for instance, into the finger acceleration rates that pianists use for different types of articulation (Palmer, 2013).

In rather broad terms, we can differentiate sound-producing motion into what we call *typological categories* as was suggested by Schaeffer for sound objects (Schaeffer, 2017). By way of the mentioned *facture*, this typology reflects distinct biomechanical and motor control features with its three main categories:

- *Sustained*, denoting a continuous, protracted sound corresponding to a continuous transfer of energy from the body to the instrument such as in bowing or blowing.
- *Impulsive*, meaning a short and abrupt sound, such as produced by hitting or plucking.
- *Iterative*, denoting a rapidly repeated sound, such as in a tremolo or trill, produced by a corresponding rapid shaking or rotating motion.

There are categorical thresholds between these main types, and we may move from one to another by the earlier mentioned phase transition. For instance, a sustained sound and motion may turn into an impulsive sound and motion if shortened, and a series of impulsive sounds and motion may turn into an iterative sound and motion with increasing rate. Examples of such phase transitions can be seen in **Figures 1** and **2**.

The main purpose of this motion typology is to point out the close links between musical sound features and motion features. This is in particular relevant in view of the emergence of coherent sound-motion objects due to the sound-producing motion with intermittent control and energy input.

### Phase Transition

The expression *phase transition* has been used to signify qualitative changes in physical substances on the basis of continuous or incremental change in some underlying parameter, e.g., the change from ice to water to steam on the basis of changes in temperature, but has also come to signify qualitative change in other phenomena such as mode of motion, for instance, with changes from walk to trot to gallop on the basis of the animal's speed of motion. Furthermore, phase transition has also been used to signify emergence of various cognitive phenomena (Haken et al., 1985; Spivey et al., 2009).

Taking a wide view of phase transition, we could include several perceptual phenomena ranging from low-level features such as the flicker-fusion threshold in vision and the similar threshold in auditory perception between impulses and pitch, to higher level categorical phenomena such as the identification of timbre, interval categories, various rhythmical patterns, and melodic contours with the common element of categorical threshold crossings resulting from continuous or incremental change of some underlying parameter.

Given the constraints of human body motion, we may assume that phase transition can be found in music

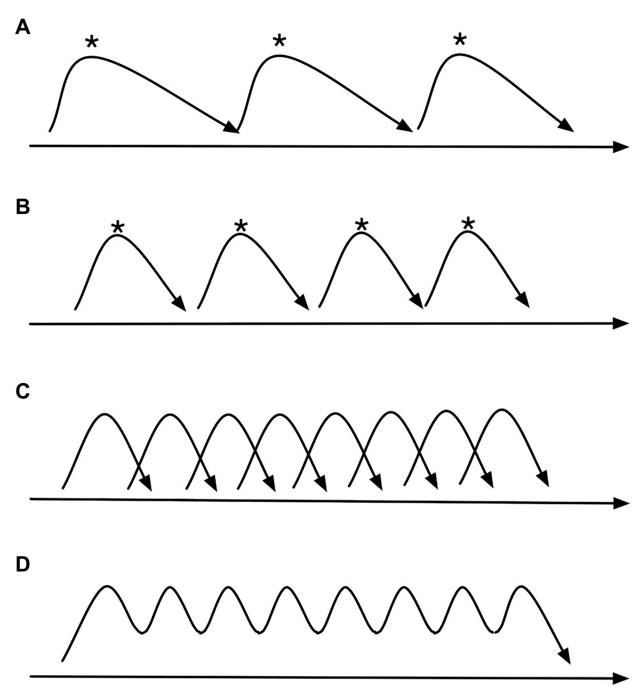



**FIGURE 1 |** A schematic image of emergent phase transition. If a series of separate motion events, each with a goal posture marked with an asterisk, lined up along a timeline from left to right such as in **(A)**, are moved closer together in time like in **(B)**, and even closer still, the events will fuse together like in **(C)**, becoming more like an iterative pattern where the originally separate events have become transformed into a rapid wavy motion like in **(D)**.

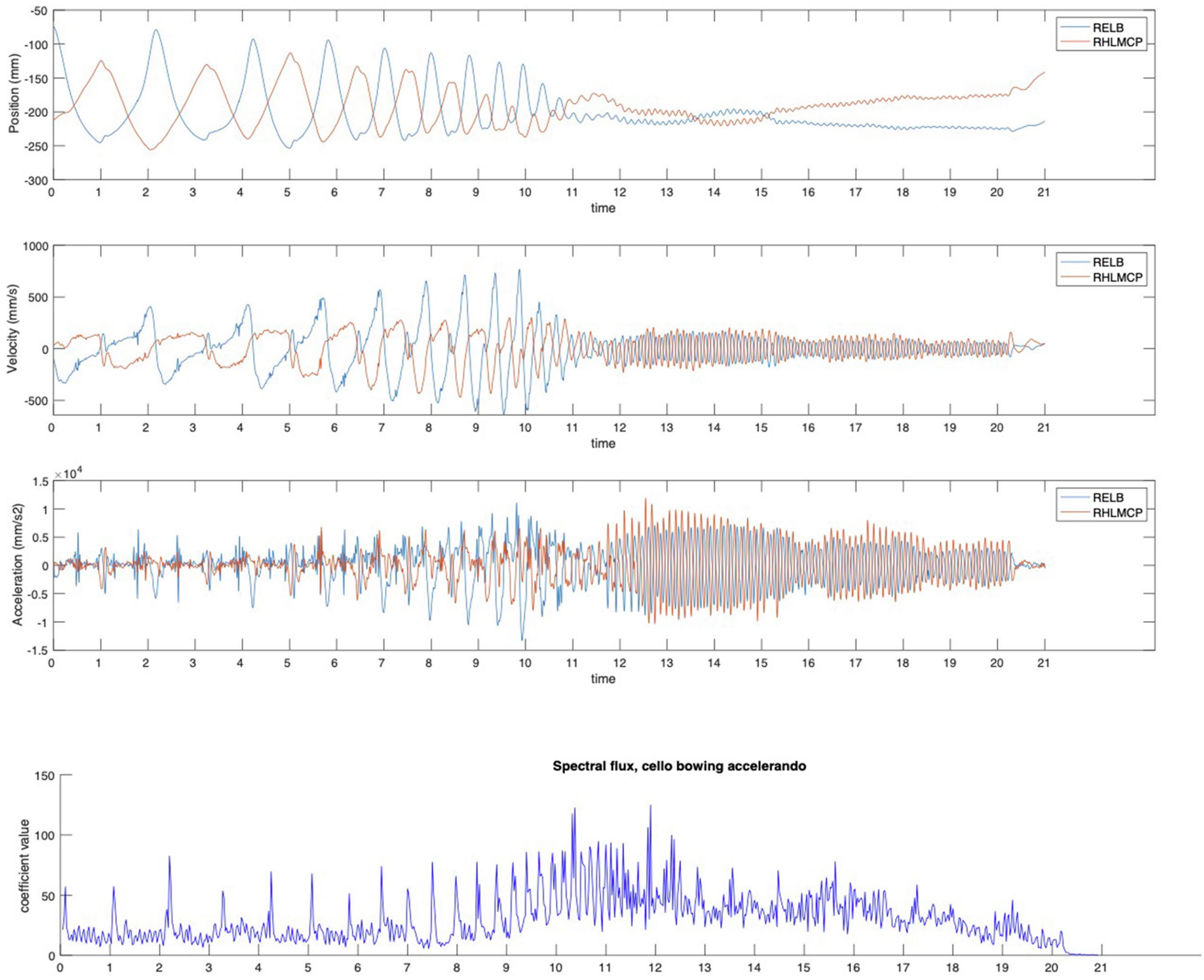



**FIGURE 2 |** Phase transition in bowing. The figure displays the markers placed on the elbow and wrist (the right hand metacarpophalangeal articulation, next to the left hand little finger) tracing the back-and-forth motion of cello bowing, accelerating from slow to very fast (as fast as in a tremolo), is displaying phase transition of the motion patterns due to the constraints of the high speed, i.e., a change to smaller amplitude motion because large amplitude motion would be impossible at high speeds, eventually ending up in just a shaking of the wrist at the highest speed. Also, note the changes in the spectral flux (rate of change of spectral features) due to an increase in the noise components in turn due to the increase in transients because of the increase in the rate of bow direction changes. The motion trajectories of the two markers were recorded using a Qualisys motion capture system at a frame rate of 300 fps and analyzed with the MoCap Toolbox software (Burger and Toiviainen, 2013), and the spectral flux of the sound of the F#4 was calculated using the MIRtoolbox software (Lartillot and Toiviainen, 2007).

performance, in particular with repetitive motion such as in bowing (Rasamimanana et al., 2007). We can see two illustrations of this in the present paper. In **Figure 1**, there is a schematic depiction of phase transition starting from a series of separate events, and if these events are shortened and moved closer, e.g., if there is a tempo increase in the music, we see that the events will overlap and the series of distinct events will be transformed into a stream of undulating motion. In **Figure 2**, we see a similar case of phase transition, but here with accelerated bowing on a cello. In the initial slow tempo, we see the long bowing motion of both the elbow and the wrist, but with the acceleration, the bowing motions become shorter until the bowing reaches the speed of a fast tremolo and where bowing is reduced to just a small-amplitude wrist shaking motion. Interestingly, this phase transition of bowing motion has also resulted in significant timbral changes, with the sound becoming more noise-dominated as can be seen from the graph of spectral flux, i.e., an indication of degree of spectral change along the temporal axis.

Phase transition is about emergent (and forced) fusion of initially separate events (both sonic and sound-producing), but may just as well be about the fission of coherent events into separate events. For instance, we may also come across phase transition from fast to slow, sometimes also showing the difficulties of moving slowly (Park et al., 2017).

## Coarticulation

The term *coarticulation* signifies the fusion of otherwise distinct motion events into larger and more coherent motion events. Coarticulation is found in several domains of human motion, first of all in linguistics (Hardcastle and Hewlett, 1999), but also in other areas such as typing, hand writing, sign language, and various everyday human activities, see (Rosenbaum, 2009) and (Grafton and Hamilton, 2007) for general presentations, as well as (Sosnik et al., 2004) for emergent coarticulation by practice and optimization. But we also have some publications on music-related coarticulation such as the following:

- In piano playing, with fingers moving to optimal positions before hitting keys (Engel et al., 1997).
- Coarticulation and so-called *muscle synergy*, i.e., cooperation between muscles groups, in piano performance (Winges et al., 2013).
- In string playing, with left hand fingers in place in position well before playing of tones (Wiesendanger et al., 2006) and in the contextual smearing of bowing movements (Rasamimanana and Bevilacqua, 2008).
- In drumming, where a drummer in some cases may start to prepare an accented stoke well in advance (Dahl, 2006).
- For some examples of our own work with motion capture data of piano and marimba performance (see Godøy et al., 2010; Godøy, 2014).

Coarticulation may involve different elements, and all of them may contribute to the fusion of otherwise separate elements in music performance as in the following (Godøy, 2014):

- *Temporal coarticulation*: otherwise singular events become embedded in a context.
- *Spatial coarticulation*: motion in one effector (e.g., hand) recruits motion in other effectors (e.g., arm, shoulder, torso).
- *Spillover effects*: past events influence present events, i.e., position and shape of effectors are determined by recent motion.
- *Anticipatory effects*: future events influence present events, i.e., position and shape of effectors are determined by preparation for future motion.

Coarticulation seems to be quite common in music performance, and here just two illustrations of this. In **Figure 3**, there is a schematic depiction of how we may think of coarticulation as emerging from a context of sound-producing events, initially consisting of a series of separate events, but when these events are moved closer together, there will be a contextual spillover effect that is the hallmark of coarticulation, successively blurring the boundaries between the events until the original events fuse to become more like a continuous motion trajectory. In **Figure 4**, we can see how a series of trills result in a similar smearing of finger motion because of the need for rapid motion in playing the trills.

As for the more detailed workings of coarticulation, the score-like depictions of muscle synergies of (d'Avella and Lacquaniti, 2013), showing how different muscles participate at different moments in time, are highly relevant for how we believe coarticulation for more complex sound-producing body motion may be implemented and is something to be explored further, cf. attempts in this direction in (Winges et al., 2013).

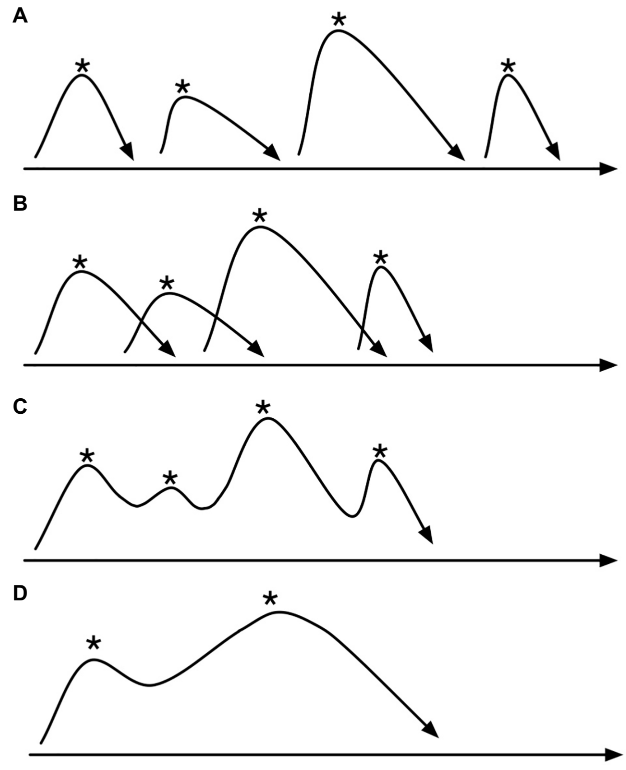



**FIGURE 3** | A schematic image of emergent coarticulation, starting with four separate events along a time line from left to right such as in **(A)**, each event centered on a goal posture marked with an asterisk, then moved closer together in time in **(B)**, with the consequence that the partially overlapping trajectories become more like a continuous trajectory in **(C)**, and resulting in an even more smoothed trajectory in **(D)**, thus illustrating the event fusion phenomenon of coarticulation.

As an ubiquitous phenomenon in music, coarticulation provides a testimony of constraint-based contextual smearing of events. Coarticulation means that performers are constantly preparing for the upcoming sound-motion events, unless there is a pause (with no motion) ahead. This contextual smearing of events due to coarticulation is in a sense what makes music (and spoken language) sound "natural," and recuperating the "original" events behind the smeared manifestations of coarticulation can be very challenging. We have seen some interesting work in this direction (Bevilacqua et al., 2016), and furthermore, this recuperation could also resemble recovering underlying intermittency components from emergent continuous sound and motion.

## Idioms

The term *idiom* in connection with music performance denotes particularly successful sound-motion objects in the sense of

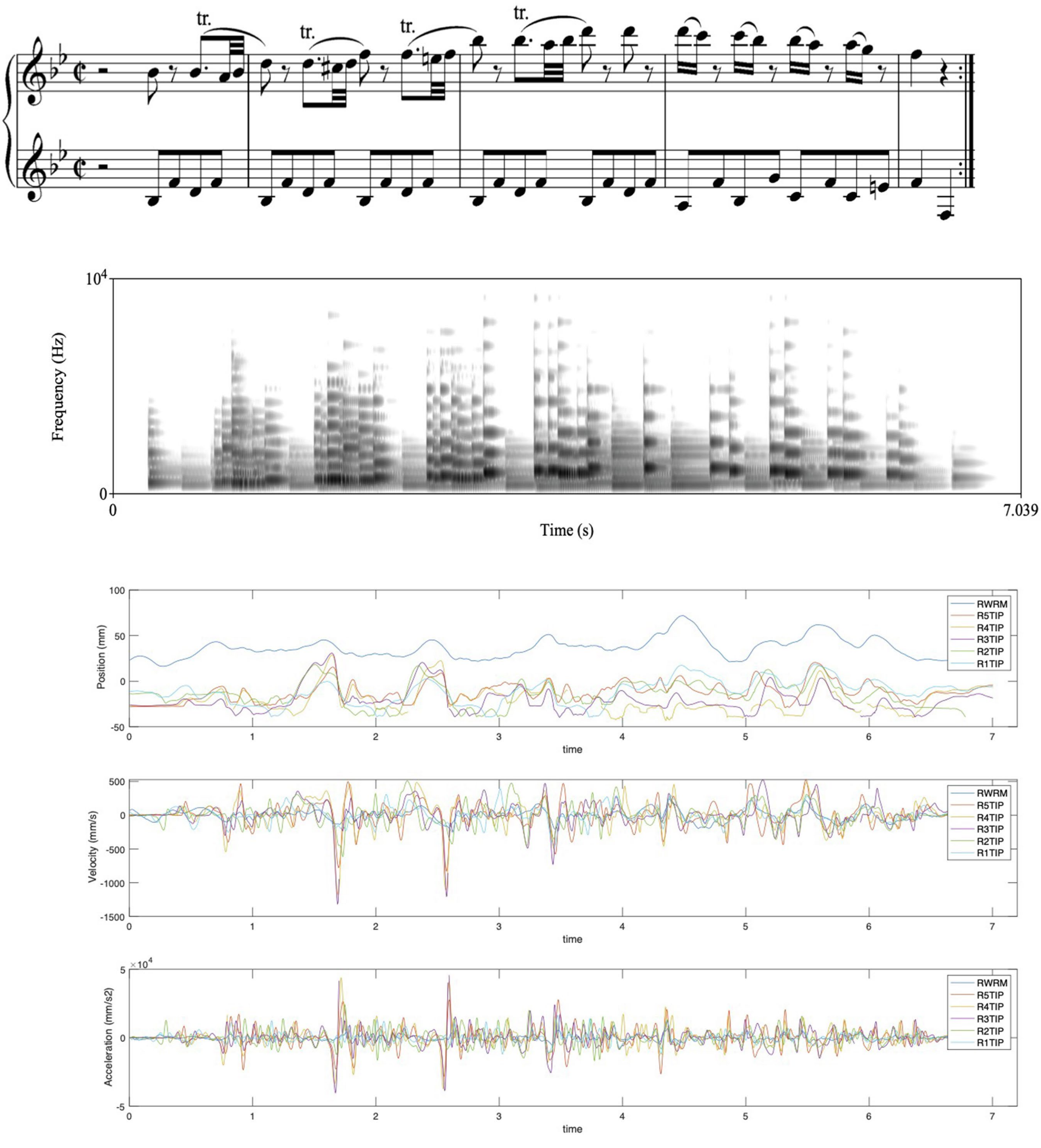



**FIGURE 4 |** Coarticulation in a piano performance, here of the first four measures of Mozart's Variations KV500, showing the notation, a spectrogram of a performance of these measures, and the vertical motion data (the "z" dimension) of the five fingers and the wrist of the right hand from this performance, including the positions, velocity, and acceleration of the fingers/wrist. The burst of finger motion for each of the trills is most clearly seen in the velocity and acceleration plotting. The motion trajectories of the five fingertips and the wrist of the right hand markers were recorded using a Qualisys motion capture system at a frame rate of 250 fps and analyzed with the MoCap Toolbox software (Burger and Toiviainen, 2013), and the spectrogram of the sound was made with the Praat software (Boersma, 2001). The notation (score) is added for reference purposes and aligns only approximately with the recorded data.

having maximally aesthetically pleasing audible results with minimal effort, in view of both biomechanics and control, of sound-producing motion. As any quality textbook on orchestration can tell us, idioms in music is a huge topic, yet it should be mentioned in our context of sound-motion objects because it is crucial for how musical expression can emerge from constraints, as well as for how composers' and arrangers' considerations of such constraints may influence the music they produce.

Consider, e.g., the case of tremolo by back-and-forth bow motion for repeated tones as easy in string instrument performance (cf. the cello bowing phase transition example above), but difficult in a keyboard performance, whereas a tremolo by an octave rolling motion on keyboards would be easy. Conversely, a rapid octave back and forth motion on a string instrument would be difficult and hence is something that would better be rendered as single tone repetitions on a bowed string instrument (cf. Godøy, 2018 for examples of this). Also, same tone repetitions can be done with relative ease on some other instruments, e.g., brass instruments, but not so easily on instruments that have the mouthpiece inside the mouth, e.g., woodwind instruments. Furthermore, so-called *flatterzunge* is easy on some and difficult on other wind instruments. Related, vocally, making an alternating sound such as *ti-ku-ti-ku* is easier than a non-alternating sound such as *ti-ti-ti-ti*.

These examples concern biomechanical issues, but there are also types of sound-producing motion that are challenging more on the control cognitive side, e.g., some tones, typically in the extreme registers, may be difficult or risky to produce, whereas others are failsafe, such as open strings tones and lower natural harmonic tones on string instruments. Also, there are cognitive control issues, i.e., mental load because of too many and/or rapidly changing patterns, e.g., unconventional bowing patterns exemplified in (Rosenbaum, 2009, p. 124), or rhythmically extremely complicated patterns that place a heavy cognitive load on the musicians. Additionally, there are cases of active cultivation of idioms for the sake of maximal output with minimal effort (both biomechanical and cognitive), such as in the mature works of Rimsky-Korsakov where practically every single part in the orchestral texture has been designed in view of optimal idiom-compliance.

In Western music culture, the idea of going from score to performance by way of so-called *interpretation* could here be understood as an optimalization of sound-producing motion by way of the mentioned principles of phase transition and coarticulation, i.e., minimizing effort by constraint-compliance. Interestingly (Furuya and Yokota, 2018) demonstrated optimization of effort by rhythmical variation, denoted "neuromuscular adaptation," to find the most favorable mode of motion, and the removal of unnecessary motion components by expert musicians can increase the sense of fluency in performance (Gonzalez-Sanchez et al., 2019).

## CONTROL THEORIES

### Motor Control

One core idea of sound-motion objects is that of a within-object coherence, both of the sound and of the corresponding body motion. To broaden our understanding of this coherence, it could be useful to have a look at coherence-enhancing elements of motor control in general, followed by the more specific theory of *motor gestalts*, as well as a theory of so-called *goal postures*, which are postures at particularly salient moments in human motion. Motor gestalts and goal postures are both elements of human motion that lead to the *intermittency hypothesis*, i.e., the hypothesis of the discontinuous, point-by-point emergence of sound-motion objects.

From an overview of human motor control (Rosenbaum, 2009), and more specifically of music-related motor control (Altenmüller et al., 2006), there seems to be a basic agreement that skilled behavior necessitates extensive practice as well as using perceptual corrective feedback during performance; however, there seem to be divergent opinions about the rate and speed of such corrective feedback. Also, there is variable concern about the biomechanical constraints involved in skilled behavior, i.e., there seems to be not always a focus on the physical effort required, cf. the motion constraints mentioned above.

What there seems to be agreement on though is that there are inherent limitations on the speed of the human motor control system, particularly concerning reaction times, hence the need for some kind of anticipatory element to work around such limitations. Yet there seems to be disagreement about the duration of such reaction times as well as about the extent of the corresponding anticipatory pre-programing in human motion. Some researchers have seen this as a debate lasting more than a century (Elliott et al., 2001), dating back to Woodworth's seminal 1899 paper on the so-called *initial impulse* in motion, designating a first and typically coarse motion trajectory that needs to be corrected when homing in on a target.

A related issue is the need for simplification of human motor control by way of hierarchies with the idea of a top-down chain of command starting with the overall executive goal being passed on to various lower-level motor control components, resulting in coordinated sequences of muscle contractions and relaxations that finally make the wished for motion to happen (Grafton and Hamilton, 2007). Following the pioneering writings on goal-directed motion by Nikolai Bernstein, it is suggested that motion is controlled in a hierarchical manner which also includes the integration of different motor units into a single functional unit, manifest as a coarticulation cooperation between the motor units.

Another recurring and relevant concept is that of *motor programs* (Summers and Anson, 2009), denoting the principle of some kind of mental script for a motion sequence, which also concerns the degree of pre-planning and degree and rate of corrective feedback in the course of implementing the motion sequence (Desmurget and Grafton, 2003). The topic of corrective feedback is related to general *control theory*, a topical area situated at the intersection of engineering and human movement science. Control theory spans from rather simple regulatory systems such as thermostats to highly complex and human-like control systems and includes key concepts such as *open loop*, *closed loop*, *feedforward*, and *feedback*, now also elements of predictive and intermittent human motor control. Most of these concepts are relevant for motor control in music performance; however, we need as always to consider the timescale in question: Some features of music happen so fast that closed loop feedback will be impossible, making musicians rely on preprogramming and open loop, feedforward control, and hence, on motor gestalts in combination with intermittent corrective feedback.

## Motor Gestalts

In (Klapp and Jagacinski, 2011), we find a strong claim for pre-programing in their notion of *motor gestalts*. The gist of their reasoning is that motion control shares several features with gestalt perception, first of all in working by holistic entities, something that is regarded as the quintessential feature of perceptual gestalts.

The extensive and long-lasting relationship between gestalt theory and music attests to the affinity of basic gestalt principles and music perception. It should be remembered that gestalt theory actually had its beginnings in music perception with the work of Ehrenfels, Stumpf, and others (Smith, 1988), and ensuing gestalt-related publications attest to a continued relationship (e.g., Tenny and Polansky, 1980; Bregman, 1990; Leman, 1997). Gestalt theory was also influential in the development of Schaeffer's sound object theory (Schaeffer, 2017), and although there have been a number of divergent opinions on details of gestalt theory, the main idea is clear, namely the "all-at-once" presence of fragments of sensory experience as objects, underlining our organisms penchant for organizing also sound and motion as objects. The idea of motor gestalts is based both on the general features of gestalt and on the need for an efficient and reliable way to control body motion given the limitations of the human motor control system. The most serious challenge for the human motor control system is the mentioned PRP; hence, a reliable possibility of pre-programming would be advantageous.

The core of gestalt theory is the following: "Perhaps the most fundamental aspect of the gestalt tradition is the notion of a gestalt as a *holistic pattern* that is 'more than the sum of its parts.' Thus, it is a unit rather than isolated components." (Klapp and Jagacinski, 2011 p. 444). A crucial feature of gestalt is then that of something solid, and in the case of motor gestalt, something that is ready to be instantiated very quickly. Measuring so-called reaction time, or RT for short, has been used for probing the degree of preprogramming. So-called *Simple RT* designated the triggering of a ready response, whereas *choice RT* designated the need to assess and choose the most suitable response. In our context, simple RT seems to be the most relevant as an indicator of ready-made motor gestalts. If the RT increases, it is assumed to be due to having to choose between different responses.

Motor gestalts are seen to have the following four principles in action, as summarized in (Klapp and Jagacinski, 2011):

1. "A brief chunk or motor gestalt is a holistic pattern." (*Ibid*, p. 458). It has previously been suggested that alternatively, the word *chunk* could be used to designate a motor gestalt provided that "it (a) is processed as a unit and not as separable components and (b) functions to enable a coordinated action." (*Ibid*, p. 444).
2. "Abstract coding permits perceptual constancy; an abstract action code is not tied to specific effectors, thereby permitting constancy in motor control." (*Ibid*, p. 458). Such a flexibility of effectors seems similar to what is sometimes called *motor equivalence*, i.e., that different effectors can make similar motion patterns, as well as do so at different scales, e.g., signature on a small piece of paper or on a large blackboard.
3. "Motor gestalts are mutually exclusive; only one gestalt can be programmed at any moment." (*Ibid*, p. 458). This is crucial for rhythmical patterns in that a series of sound onset motions cannot at the same time belong to two different metrical organizations, e.g., 6/8 and 3/4, by the principle of so-called *exclusive allocation*. In another publication, there is the following passage concerning polyrhythm: "The limitation to only one motor Gestalt may be analogous to limits that arise with visual patterns such as the Necker cube. That figure can be perceived in only one of its configurations at any given instant. In either configuration, however, all of the lines of the cube are perceived simultaneously as one pattern. Thus, the Gestalt is not restricted in terms of the number of lines that can be perceived. Instead, the limit is that only one organization can be activated. Similarly, the limit in concurrent motor actions is assumed not to lie in the number of muscles that can be controlled, but, instead, the limit is that only one action pattern can be active." (Klapp et al., 1998, p. 318). In other words, even rather complex motion patterns may be conceived and perceived as single gestalts in motor control.
4. "The organization of an action sequence can be either integrated or streamed." (Klapp and Jagacinski, 2011, p. 458). This is related to so-called *streaming* in auditory scene analysis (Bregman, 1990), meaning that a sequence of sound events may either be seen as belonging to several parallel streams, or as belonging to one single stream, as suggested for a polyrhythmic pattern transformed from two streams to one single integrated entity, e.g., the 3 against 4 streams transformed to a single punctuated rhythmical pattern. Interestingly, the suggestion that "Speed of action can influence organization" (*Ibid*) seems similar to categorical shifts due to the above-mentioned phase transition.

The fundamental meaning of motor gestalts in music is then that it is conceived and perceived as a holistic entity, and as having a coherent motion trajectory from start to end, hence, also contributing to the coherence of sound-motion objects in music performance.

## Posture-Based Theory

One way to look at sound-motion objects is that they are hierarchically organized around goals. This goes back to the pioneering work of Bernstein with his idea of the complexities of human motion and the multiple degrees of freedom and associated redundancy (i.e., multiple solutions possible to any task) needing to be regulated by some kind of hierarchical control scheme. There are different models of hierarchies in motor control, but usually with some scheme for simplification and automatization of lower level control tasks. Such top-down control schemes have the advantage of optimizing the motion components that go into any sound-motion object, hence, also linked with coarticulation [see (Grafton and Hamilton, 2007) on this as well as an interesting discussion of Bernstein goal-directed behavior]. A similar focus on goals has been a leading idea in David Rosenbaum's work on motor control, such as in the case of dance:

> "Dance, for its appearance of being a *continuous* activity, is actually controlled, or is supposed to be controlled, by aiming for one target position after another. Insofar as this method is endorsed by dance coaches and proves useful for dancers, it probably reflects a deeper principle about the control of physical action. That deeper principle, according to the posture-based motion planning theory developed by my colleagues and me, is that a reference condition for goal postures is established for positioning movements before movements to those goal postures are planned." (Rosenbaum, 2010, p. 44)

Furthermore, according to Rosenbaum: "Dance, for its appearance of being a *continuous* activity, is actually controlled, or is supposed to be controlled, by aiming for one target position after another." (*Ibid*, p. 44). Designated by the term *goal postures*, this is depicted as a general element of motor control in the work of Rosenbaum, also related to the phenomenon of *keyframes* in animation (Rosenbaum et al., 2007). The role of keyframes in animation is to establish salient moments in the narrative, with the aim of then making continuous motion between these keyframes, which given the keyframes becomes a much simpler task. Actually, some choreographers and dancers have been using a similar scheme in rehearsing new scripts called *marking,* denoting a sparse running through of dance sequences with basically just moving from posture to posture, not doing much motion between the postures, both in order to focus on the overall structure of any dance sequence, and to save energy during rehearsals (Kirsh, 2011; Warburton et al., 2013).

More recently, Rosenbaum has proposed this focus on postures as a general theory of motor control called *posture-based theory* (PBT), in particular associated with manual postures in body motion (Rosenbaum, 2017). Posture-based theory is then a theory of hierarchies in motor control, where motion is planned by way of the goal postures. It is interesting to see the similarity here with coarticulation in that there is a transition from one posture to another, i.e., that there is a temporal smearing effect in the motion trajectories. Furthermore, there is a link with intermittency with the temporally distinct goal postures, hence, there is a convergence here of intermittency, postures, motor gestalts, and coarticulation.

In sum, PBT is about the primordial importance of shape and position of effectors, and we can adapt this to music performance, primarily as postures of hands, but possibly also of other effectors such as arms, shoulders, torso, and feet, and not to forget the vocal apparatus. The position and posture shapes of effectors in relation to pitch space, as well as multidimensional timbre space, at salient moments in time, are that which (following the idea of PBT) could make the control of subsequent motion between postures easier. With both motor gestalts and PBT converging in intermittency, this may be included in what we call *shape cognition* (Godøy, 2019), in turn related to the general shape cognition of so-called *morphodynamical theory*, an extensive theory on the primordial role of shapes in perception and cognition (Thom, 1983; Petitot, 1990; Godøy, 1997).

## MUSICAL INTERMITTENCY

### The Intermittency Hypothesis

The word *intermittent* has been defined as "not happening regularly or continuously; stopping and starting repeatedly or with periods in between" (Cambridge English Dictionary, 2021), hence as basically signifying discontinuity. Using the expression *musical intermittency* in this paper is then motivated by the need to have a general concept for discontinuity in musical experience.

In our context of music performance, such discontinuity may seem paradoxical if we think of music as a continuous stream of sound and motion sensations, cf. the mentioned similar paradox in dance (Rosenbaum, 2010). In our present perspective, it is actually this *coexistence of continuity and discontinuity* that is the crucial attribute of sound-motion objects, in that we may have discontinuity between the objects combined with subjective experiences of musical performances as a continuous stream.

Continuity vs. discontinuity may in our context be regarded as relative to timescale of observation, with the idea that intermittency, and hence discontinuity, is valid when we consider music as a series of sound-motion objects, and that we also have a within-object continuity, recall that this is typically in the 0.3–5 s duration range. Also, recall that the holistic perception of sound objects was one of the crucial features of Schaeffer's theory, because the features of sound objects are spread throughout the sound object, and that the perception of such sequentially unfolding sound requires a cumulative, holistic perception, as demonstrated by the cut bell and closed groove experiences.

The relationships between discontinuity and continuity in music can been seen in relation to a more extensive debate in human motor control. Since the mentioned essay by Woodward from 1899 and ensuing debates (Elliott et al., 2001), a view of motion control as discontinuous emerged with the concept of intermittent control, initially proposed by Kenneth Craik in his posthumous publication (Craik, 1947) and in publications by his associate Margaret Vince (Vince, 1948). The term "intermittent ballistic" was used to denote control actions as being intermittent and having a character of ballistic motion, i.e., of an impulse followed by energy dissipation, as for instance when tapping a joystick. Similar ideas of intermittency in human motor control were presented by Navas and Stark (1968) and later on by Ronco (1999), and by several others in the ensuing decades (Gawthrop et al., 2011; Loram et al., 2011, 2014; Karniel, 2013; Sakaguchi, 2013; Sakaguchi et al., 2015). There are also intermittency-related ideas in other research, such as in research on hierarchical control in human movement (Grafton and Hamilton, 2007) and in research on muscle synergies (d'Avella and Lacquaniti, 2013). *Muscle synergies* here basically denote scripts of time-dependent contraction and relaxation of muscles ensembles, a muscular cooperation needed to produce the desired motion events, hence also related to coarticulation.

These different ideas can be seen to converge in what could be called the *intermittency hypothesis*, implying intermittency in control as well as in effort (or energy input), when applied to sound-motion chunks. The basic model for this hypothesis is that of an intermittent burst of effort followed by a prolonged phase of continuous motion, as suggested by the "serial ballistic" expression of Craik.

The reasons given for intermittency in the literature referred to above are first of all that an open loop control scheme with only intermittent feedback may be an efficient workaround for the slow and noise-prone motor control system, so that a series of intermittent motor control input points may be better able to handle the demands of performance than any attempt to have a continuous feedback, or a closed loop, control scheme. This is partly supported by behavioral evidence, but there are still substantial challenges of method in detecting these intermittent control points in time. For instance, when intermittent control and energy inputs are quite close in time, there will be an emergent sensation of continuity. In the words of (Gawthrop et al., 2011, p. 31): "It is shown that when event thresholds are small and sampling is regular, the intermittent controller can masquerade as the underlying continuous-time controller and thus, under these conditions, the continuous-time and intermittent controller cannot be distinguished. This explains why the intermittent control hypothesis is consistent with the continuous control hypothesis for certain experimental conditions."

A major challenge in research on intermittency is then to detect and qualify intermittency in human body motion data (Loram et al., 2014). One solution is to look for discontinuities in the motion trajectories, something that was done already in the pioneering work of Craik and Vince several decades ago (Craik, 1947; Vince, 1948), and which has been done again recently with more developed technologies and analysis methods (Sakaguchi, 2013). There is additionally the phenomenon of the so-called *pre-motion silent period*, meaning that before the onset of a ballistic muscle contraction, there is a relaxation of the muscles that can be detected in the EMG signals (Aoki et al., 1989).

Applied to our context of sound-motion objects, we may understand music performance as concatenations of sound-motion objects, where the output may be perceived as a continuous stream, but where the control and energy input at the object timescale may be intermittent. However, this raises an important question:

If we think of intermittency as associated with salient points in time, what are then these salient point in time in music? We could mention downbeats and other salient points in time such as melodic peaks. However, it seems that downbeat is a strangely little researched topic, and one possibility could then be that it is associated with impacts, i.e., what can be seen as *velocity reversals* in a motion trajectory, something that seems to make good sense in some cases (see Godøy, 2013 for how downbeats in a waltz pattern reflects this velocity reversal). More cases of velocity reversals and other discontinuities in the motion signals could be interesting to examine, e.g., as suggested in (Sakaguchi, 2013).

In our own work, we are presently looking at discontinuities in the motion capture trajectories (e.g., the mentioned velocity reversals), mostly focused on short, rapid, and highly pre-programmed sound-motion objects in the form of ornaments and other figures. However needless to say, there are several challenges in getting good motion capture data here because we need to use a fairly large number of reflective markers and relatively high framerates in order to capture small-scale and rapid sound-producing effector motion (Song and Godøy, 2016).

## Triggers

Accepting the existence of preprogrammed motion chunks and intermittency, the next question is how these motion chunks are triggered. Unfortunately, this seems to be a not well researched topic in human motor control, yet a crucial topic for understanding the workings of skilled behavior in music performance. Assuming there is a volitional initiation of body motion in music performance, the question becomes: *What is actually the triggering mechanism in such time-sensitive tasks and how does it work?*

We have in recent years seen research on so-called entrainment in music, demonstrating that our organism picks up salient motion-inducing patterns in musical sound which in turn may result in body motion, e.g., in dancing, walking, or gesticulating (Clayton et al., 2013), and highly synchronized triggering is also documented in group performance contexts (Palmer, 2013). Yet it seems that the question still remains of what is the initial impulse to start the assumed ready-planned motion chunks.

An impulse in our context can be understood as a short burst of energy, what in physiological terms may be called a *ballistic muscle contraction*. An EMG signal with a steep attack slope could then be an indicator of a trigger, resembling a salient rhythmic articulation by the shape of the beat (Elliott et al., 2009). More generally, it could be interesting to look at the so-called *startle reactions*, as these reactions are hypothesized to work by high degrees of pre-programming activated by some loud noise or other sudden sensory impulse. It has been suggested that these triggers and the corresponding extremely fast reaction times are an integral part of our motor control system and could also be at work in more ordinary, non-startle induced motion (Valls-Solé et al., 1999, 2008).

Fast triggering of motion can be found in various sports, and an important feature here is that the triggering impulse may be rather simple, yet activate quite complex patterns of musculoskeletal motion, i.e., be part of an eminently hierarchical control scheme as suggested in (Karniel, 2013). This seems also to be the case in writing (Plamondon et al., 2013) and graffiti motion (Berio et al., 2017), research that is summarized as follows: "In our work we rely on a family of models known as the Kinematic Theory of Rapid Human Movements, mainly developed by R. Plamondon et al. in an extensive body of work since the 90's."… "They show that if we consider that a movement is the result of the parallel and hierarchical interaction of a large number of coupled linear systems, the impulse response of such a system to a centrally generated command asymptotically converges to a lognormal function. This assumption is attractive from a modelling perspective because it abstracts the high complexity of the neuromuscular system in charge of generating movements with a relatively simple mathematical model, which further provides state of the art reconstruction of human velocity data." (*Ibid*, p. 2).

Lastly, some intermittency research has made a distinction between clock-based (internal) and event-based (external, adaptive) intermittent control (Sakaguchi et al., 2015), suggesting that the latter mode of triggering is more flexible and well suited to real-world demands.

The nature and workings of triggering are still a major outstanding issue, and hopefully, it will be possible to design experimental paradigms in music performance for exploring this further. Triggers are about feedforward and open loop kinds of motor control. Triggering is also about recognizing the typical sound-motion object timescale where intermittent control and energy input are optimal, as well as recognizing the possibly negative interference of input in the course of an ongoing sound-motion chunk. It seems better to leave an ongoing sound-motion object alone and let it run its course, as has been suggested in connection with chunking and the basal ganglia: "Chunks take their advantage from being manipulable as entities, and the intervention of consciousness or attention might actually disrupt their smooth implementation." (Graybiel, 1998, p. 131).

## Sound-Motion Objects

From the various research and ideas presented so far in this paper, we can now summarize the main ideas of sound-motion objects as follows:

1. Musical features are intrinsically multimodal, comprising sensations of both sound and corresponding body motion.
2. The most salient perceptual features of both sound and motion are found at the meso timescale of approximately 0.3–5 s.
3. The intermittency hypothesis suggests that motion chunks, and the resultant sound chunks, are optimally conceived and triggered discontinuously.

Needless to say, there are several substantial challenges here:

Firstly, in understanding in more detail the interaction of sound and motion in our minds and bodies, and to have more well-informed notions of *what is what* in musical features such as rhythm and textures: What is the sound sensation and what is the motion sensation, e.g., in listening to drumming or to a string ensemble?

Secondly, there are substantial challenges in understanding the workings of timescales of both sound and body motion. The basic hypothesis here is that the meso timescale of sound-motion objects is privileged in that it encompasses the most salient sound features and the most salient motion features, i.e., that smaller timescales just do not have these features and that larger timescales do not allow the same focus on these features as was one of the main arguments for Schaeffer in favor of the sound object (Godøy, 2021).

Thirdly, there is the challenge of better understanding the initiation, or triggering, of sound producing motion. Clearly, there is preplanning going on in musical performance, but what does such a preplan look like? Is it a kind of compressed mental image of motion trajectories to be triggered? And is there some kind of ultra-rapid triggering for such preprogrammed motion chunks as suggested by the mentioned startle research?

In spite of these and similar outstanding questions, what seems to be reasonably well supported is that in music performance, there is a shaping of motion control and motion effort going on, manifest in the mentioned elements of phase transition, coarticulation, muscle synergies, idiom use, and optimization of energy use. In Western music, this shaping means transforming otherwise discrete tone events of common practice notation into coherent coarticulated sound-motion objects. The advantage of the sound-motion object idea in music performance is that of focusing on the motion optimization at the object timescale, guided by the inherent constraints of both instruments and body motion, as well as the perceptual constraints favoring holistic objects at the meso timescale. Notably so, this includes most sound features unfolding within the confines of a sound-motion object, including timbral, harmonic, and modal flavors (Persichetti, 1962), determined by the cumulative impressions of characteristic intervals over a certain time stretch.

The idea of sound-motion objects as presented in this paper is primarily concerned with motion and sound sensations, but there is no denial that such sound-motion object may have multiple significations, for instance, as is the claim of ecological acoustics (Gaver, 1993) and largely in line with ideas of sonic event perception (Rocchesso and Fontana, 2003) as well as research on auditory object perception (Bizley and Cohen, 2013). In addition, links have been made between an object focus and semiotics, as found in the UST project (Delalande et al., 1996), with the idea of "temporal semiotic units" (*unités sémiotiques temporelles*). The semiotic aspect here could also be seen as related to more semantic and/or hermeneutic aspects of sound-motion objects, needless to say important and extensive areas of music research that hopefully could be the focus of future research.

Also, there are of course features at larger timescales, i.e., the macro timescale, which could likewise be overviewed in an "all-at-once" manner, as suggested by Paul Hindemith of an entire composition: "If we cannot, in the flash of a single moment, see a composition in its absolute entirety, with every pertinent detail in its proper place, we are not genuine creators." (Hindemith, 2000, p. 61). Zooming back and forth between such different levels of resolution is clearly a possibility, an important topic that deserves more systematic research.

## DISCUSSION

From various available evidence, it seems reasonable to infer that sound-motion objects can play an important role in music performance, yet that the corollary notions of a fundamental discontinuity and the associated intermittency hypothesis may go against established ways of thinking. Hence, we can list here some elements that are in favor of the sound-motion object approach, followed by some elements that go against this approach, lastly followed by some ideas on how to continue exploring sound-motion objects:

- Various motion and motor control constraints seem to favor the meso timescale approach to sound-producing motion, including the intermittency hypothesis, because there is the need to work around the constraints of PRP by pre-programming.
- Sound-motion objects seem well-suited to help us understand rhythmical, textural, modal, contoural, and timbral patterns, because they provide a local context that fuse individual events into strongly coarticulated and coherent entities.
- The basically physical-physiological and motor control approach of sound-motion objects to musical features could open up for more cross-cultural assessments of musical expression.

In current music research, we have seen a focus on features of music performance such as nuances in timing and dynamics, research often using advanced measurements and processing methods and leading to interesting findings about timing-related issues. But what seems to be less focused on is the meso timescale shapes of sound-motion objects, meaning features of the object considered as a whole, as was one of the main aims of Schaeffer's sound object theory.

But surely, there are significant unsolved problems with this sound-motion objects approach, such as the following:

- There seem to be divergent opinions on feedforward and intermittency in motor control research, and it may be that some of the ideas presented above will be contradicted by other research.
- Body motion and motor control constraints have not been much focused on in mainstream music research; hence, the idea of including various constraints in the analysis of music may require new ways of thinking in music-related research.
- The view of music as a series of sound-motion objects may go against the idea of music as a continuous flow, yet the concept of sound-motion objects could be well reconciled with the ideas of music as continuous sensory streams, provided we have an awareness of the different timescales simultaneously at work here.
- There are substantial challenges in collecting more precise motion capture data and more extensive EMG data on effort distribution in support of the intermittency hypothesis.

As for further work, there are three main areas that we are hoping to work in:

- *Conceptual-analytic work*: Clearly, there are inherited issues in music theory and also in performance research that may need critical assessment, in particular in view of how meso timescale salient sound and body motion features are handled. Said differently, there seems to be a lack of object focus in mainstream music analytic thinking, hence the need to exercise shifts of timescale perspective, i.e., ask questions about what we are focusing on, as was the strategy of Schaeffer, and then going on to represent salient features as holistic shapes. Also in terms of motor control, there seems to be a need for more critical assessment of inherited control theory, in particular in view of how meso timescale object-focused features are taken care of in the control schemes.
- *Exploratory modeling*: Given the difficulties in detecting intermittency in motor control, it could be useful with a reverse engineering approach of simulating impulse-driven motion chunks by combining pre-programmed shapes with intermittent, point-by-point energy input. This could be a heuristic strategy to discover what would be the workings and requirements in real-life intermittent control situations for sound-producing motion. We are presently making some simplified toy models of intermittency, inspired by so-called impulse-response modeling, i.e., basically by having an energizing impulse that dissipates its energy through a stationary shape by convolution.
- *Testable hypotheses*: This is about formulating various sound-producing motion tasks for performers that may be carried out either intermittently or continuously, and assessing the results. We envisage using a combination of motion capture and EMG to record both the output motion and the muscular effort that go into the motion. Participants will concretely be asked to move as abruptly as possible (i.e., ballistic motion) followed by relaxation phases, or as smoothly as possible (i.e., fluency) with continuous effort, hopefully providing insights into intermittent vs. more continuous energy input in sound-producing motion.

As for possible practical applications of ideas of constraint-based sound-motion objects, the following could be suitable exploratory schemes in cooperation with musicians:

- In instrument practice experiments, the phenomena of phase transition and coarticulation could be explored by alternating between very different tempi, enabling more direct observations of the effects of tempi on sound-motion object cohesion and other emergent features.
- In improvisation, and by extension also in composition, work specifically with sound-motion objects as combined motor

gestalts and sound objects, focusing on the overall typological shapes and sense of motion, reminiscent of David Sudnow's book on motion-based improvisation (Sudnow, 1978).

In summary, the core idea of this paper is due to an object perspective in music, originally proposed by Pierre Schaeffer and applied to the *musique concrète*, then successively to other kinds of musical sound, and now also to body motion with the concept of sound-motion objects. The motivation for all this is the belief that sound-motion objects are optimal for both the generation and the perception of music and could also be the source of novel analytic and creative tools.

## DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: The dataset is available upon application to RITMO for research purposes. Requests to access these datasets should be directed to RG, r.i.godoy@imv.uio.no.

## AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

## FUNDING

The work on this paper has been partially supported by the Research Council of Norway through its Centres of Excellence scheme, project number 262762 and by the University of Oslo.

## ACKNOWLEDGMENTS

Many thanks to colleagues at the RITMO Centre of Excellence for interesting discussions on topics of sound and motion in musical experience, and many thanks to the two reviewers for lucid and very useful comments.

## REFERENCES

Altenmüller, E., Wiesendanger, M., and Kesselring, J. (2006). *Music, Motor Control and the Brain*. Oxford: Oxford University Press.

Aoki, H., Tsukahara, R., and Yabe, K. (1989). Effects of pre-motion electromyographic silent period on dynamic force exertion during a rapid ballistic movement in man. *Eur. J. Appl. Physiol.* 58, 426–432. doi: 10.1007/BF00643520

Berio, D., Akten, M., Leymarie, F. F., Grierson, M., and Plamondon, R. (2017). "Calligraphic Stylisation Learning with a Physiologically Plausible Model of Movement and Recurrent Neural Networks." in *Proceedings of MOCO'17, London, United Kingdom*. June 28–30, 2017.

Bevilacqua, F., Caramiaux, B., and Françoise, J. (2016). "Perspectives on real-time computation of movement Coarticulation." in *3rd International Symposium on Movement and Computing, Jul 2016, Thessaloniki, Greece*. pp. 1–5.

Bizley, J. K., and Cohen, Y. E. (2013). The what, where and how of auditory-object perception. *Nat. Rev. Neurosci.* 14, 693–707. doi: 10.1038/nrn3565

Boersma, P. (2001). PRAAT, a system for doing phonetics by computer. *Glot Inter.* 5, 341–345.

Bregman, A. (1990). *Auditory Scene Analysis*. Cambridge, Mass: The MIT Press.

Burger, B., and Toiviainen, P. (2013). "MoCap toolbox - A Matlab toolbox for computational analysis of movement data." in *Proceedings of the Sound and Music Computing Conference 2013, SMC 2013*. ed. R. Bresin (Berlin: Logos Verlag Berlin), 172–178.

Cadoz, C., and Wanderley, M. M. (2000). "Gesture – music," in *Trends in Gestural Control of Music*. eds. M. Wanderley and M. Battier (Paris: IRCAM, CD-ROM).

Cambridge English Dictionary (2021). https://dictionary.cambridge.org/dictionary/english/intermittent (Accessed November 22, 2021).

Chion, M. (2009). Guide to sound objects (English translation by John Dack and Christine North). ElectroAcoustic Resource Site (EARS). Available at: http://ears.huma-num.fr

Clayton, M., Dueck, B., and Leante, L. (Eds). (2013). *Experience and Meaning in Music Performance*. New York: Oxford University Press.

Craik, K. J. W. (1947). Theory of the human operator in control systems; the operator as an engineering system. *Br. J. Psychol. Gen. Sect.* 38, 56–61.

Dahl, S. (2006). "Movements and analysis of drumming," in *Music, Motor Control and the Brain*. eds. E. Altenmüller, M. Wiesendanger and J. Kesselring (Oxford: Oxford University Press), 125–138.

d'Avella, A., and Lacquaniti, F. (2013). Control of reaching movements by muscle synergy combinations. *Front. Comput. Neurosci.* 7:42. doi: 10.3389/fncom.2013.00042

Delalande, F., Formosa, M., Frémiot, M., Gobin, P., Malbosc, P., Mandelbrojt, J., et al. (1996). *Les unités sémiotiques temporelles: Éléments nouveaux d'analyse musicale*. Marseille: Éditions MIM-Documents Musurgia.

Desmurget, M., and Grafton, S. (2003). "Feedback or feedforward control: end of a dichotomy," in *Taking Action: Cognitive Neuroscience Perspectives on Intentional Acts*. ed. S. H. Johnson-Frey (Cambridge, Mass: The MIT Press), 289–338.

Eitan, Z., and Granot, R. Y. (2008). Growing oranges on Mozart's apple tree: 'inner form' and aesthetic judgment. *Music. Percept.* 25, 397–418. doi: 10.1525/mp.2008.25.5.397

Elliott, D., Helsen, W. F., and Chua, R. (2001). A century later: Woodworth's (1899) two-component model of goal-directed aiming. *Psychol. Bull.* 127, 342–357. doi: 10.1037/0033-2909.127.3.342

Elliott, M. T., Welchman, A. E., and Wing, A. M. (2009). Being discrete helps keep to the beat. *Exp. Brain Res.* 192, 731–737. doi: 10.1007/s00221-008-1646-8

Engel, K. C., Flanders, M., and Soechting, J. F. (1997). Anticipatory and sequential motor control in piano playing. *Exp. Brain Res.* 113, 189–199. doi: 10.1007/BF02450317

Furuya, S., and Yokota, S. (2018). Temporal exploration in sequential movements shapes efficient neuromuscular control. *J. Neurophysiol.* 120, 196–210. doi: 10.1152/jn.00922.2017

Galantucci, B., Fowler, C. A., and Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychon. Bull. Rev.* 13, 361–377. doi: 10.3758/BF03193857

Gaver, W. W. (1993). What in the world do we hear? An ecological approach to auditory event perception. *Ecol. Psychol.* 5, 1–29. doi: 10.1207/s15326969eco0501_1

Gawthrop, P., Loram, I., Lakie, M., and Gollee, H. (2011). Intermittent control: a computational theory of human control. *Biol. Cybern.* 104, 31–51. doi: 10.1007/s00422-010-0416-4

Gjerdingen, R. O., and Perrott, D. (2008). Scanning the dial: The rapid recognition of music genres. *J. New Music Res.* 37, 93–100. doi: 10.1080/09298210802479268

Godøy, R. I. (1997). *Formalization and Epistemology*. Oslo: Scandinavian University Press.

Godøy, R. I. (2003). Motor-mimetic music cognition. *Leonardo* 36(4), 317–319. doi: 10.1162/002409403322258781

Godøy, R. I. (2006). Gestural-sonorous objects: embodied extensions of Schaeffer's conceptual apparatus. *Organ. Sound* 11(2), 149–157. doi: 10.1017/S1355771806001439

Godøy, R. I. (2010). "Thinking now-points in music- related movement," in *Concepts, Experiments, and Fieldwork. Studies in Systematic Musicology and Ethnomusicology*. eds. R. Bader, C. Neuhaus and U. Morgenstern (Frankfurt am Main: Peter Lang), 245–260.

Godøy, R. I. (2013). "Quantal Elements in Musical Experience," in *Sound - Perception - Performance*. ed. R. Bader (Berlin, Heidelberg: Springer), 113–128.

Godøy, R. I. (2014). "Understanding Coarticulation in musical experience," in *Lecture Notes in Computer Science*. eds. M. Aramaki, M. Derrien, R. Kronland-Martinet and S. Ystad (Berlin: Springer), 535–547.

Godøy, R. I. (2018). Motor constraints shaping musical experience. Music Theory Online, 24.

Godøy, R. I. (2019). "Musical shape cognition," in *The Oxford Handbook of Sound and Imagination*. *Vol. 2*. eds. M. Grimshaw, M. Walther-Hansen and M. Knakkergaard (New York: Oxford University Press), 237–258.

Godøy, R. I. (2021). Perceiving sound objects in the *Musique Concrète*. *Front. Psychol.* 12:672949. doi: 10.3389/fpsyg.2021.672949

Godøy, R. I., Haga, E., and Jensenius, A. R. (2006). "Playing 'air instruments': mimicry of sound-producing gestures by novices and experts," in *GW2005, LNAI 3881*. eds. S. Gibet, N. Courty and J.-F. Kamp (Berlin, Heidelberg: Springer-Verlag), 256–267.

Godøy, R. I., Jensenius, A. R., and Nymoen, K. (2010). Chunking in music by coarticulation. *Acta Acust. united Ac.* 96, 690–700. doi: 10.3813/AAA.918323

Godøy, R. I., and Leman, M. (Eds.) (2010). *Musical Gestures: Sound, Movement, and Meaning*. New York: Routledge.

Godøy, R.I., Song, M., and Dahl, S. (2017). "Exploring sound-motion textures in drum set performance." in *Proceedings of the SMC Conferences*. 145–152.

Godøy, R. I., Song, M.-H., Nymoen, K., Haugen, M. R., and Jensenius, A. R. (2016). Exploring sound-motion similarity in musical experience. *J. New Music Res.* 45, 210–222. doi: 10.1080/09298215.2016.1184689

Gonzalez-Sanchez, V., Dahl, S., Hatfield, J. L., and Godøy, R. I. (2019). Characterizing movement fluency in musical performance: toward a generic measure for technology enhanced learning. *Front. Psychol.* 10:84. doi: 10.3389/fpsyg.2019.00084

Grafton, S. T., and Hamilton, A. F. (2007). Evidence for a distributed hierarchy of action representation in the brain. *Hum. Mov. Sci.* 26, 590–616. doi: 10.1016/j.humov.2007.05.009

Graybiel, A. M. (1998). The basal ganglia and chunking of action repertoires. *Neurobiol. Learn. Mem.* 70, 119–136. doi: 10.1006/nlme.1998.3843

Haken, H., Kelso, J., and Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biol. Cybern.* 51, 347–356. doi: 10.1007/BF00336922

Hanneton, S., Berthoz, A., Droulez, J., and Slotine, J. J. (1997). Does the brain use sliding variables for the control of movements? *Biol. Cybern.* 77, 381–393. doi: 10.1007/s004220050398

Hardcastle, W. J., and Hewlett, N., (Eds.) (1999). *Coarticulation: Theory, Data and Techniques* Cambridge: Cambridge University Press.

Hindemith, P., (2000). *A composer's World: Horizons and Limitations*. Mainz: Schott.

Husserl, E. (1991). *On the Phenomenology of the Consciousness of Internal Time, 1893-1917*. English Translation by BroughJohn Barnett. Doredrecht/Boston/London: Kluwer Academic Publishers.

Karniel, A. (2013). The minimum transition hypothesis for intermittent hierarchical motor control. *Front. Comput. Neurosci.* 7:12. doi: 10.3389/fncom.2013.00012

Kirsh, D. (2011). How marking in dance constitutes thinking with the body. *Versus* 112-113, 183–214.

Klapp, S. T., and Jagacinski, R. J. (2011). Gestalt principles in the control of motor action. *Psychol. Bull.* 137(3), 443–462. doi: 10.1037/a0022361

Klapp, S. T., Nelson, J. M., and Jagacinski, R. J. (1998). Can people tap concurrent bimanual rhythms independently? *J. Mot. Behav.* 30, 301–322. doi: 10.1080/00222899809601346

Lartillot, O., and Toiviainen, P. (2007). "A Matlab toolbox for musical feature extraction from audio." in *International Conference on Digital Audio Effects, Bordeaux, 2007*.

Leman, M. (Ed.) (1997). *Music, Gestalt, and Computing*. Berlin, Heidelberg: Springer.

Levinson, J. (ed.) (2015). "Concatenationism, Architectonicism, and the appreciation of music," in *Musical Concerns: Essays in Philosophy of Music* (Oxford: Oxford University Press).

Loram, I. D., Golle, H., Lakie, M., and Gawthrop, P. J. (2011). Human control of an inverted pendulum: is continuous control necessary? Is intermittent control effective? Is intermittent control physiological? *J. Physiol.* 589, 307–324. doi: 10.1113/jphysiol.2010.194712

Loram, I. D., van De Kamp, C., Lakie, M., Gollee, H., and Gawthrop, P. J. (2014). Does the motor system need intermittent control? *Exerc. Sport Sci. Rev.* 42, 117–125. doi: 10.1249/JES.0000000000000018

Maes, P. J., Leman, M., Palmer, C., and Wanderley, M. M. (2014). Action-based effects on music perception. *Front. Psychol.* 4:1008. doi: 10.3389/fpsyg.2013.01008

Navas, F., and Stark, L. (1968). Sampling or intermittency in hand control system dynamics. *J. Biophys.* 8, 252–302. doi: 10.1016/S0006-3495(68)86488-4

Nymoen, K., Godøy, R. I., Jensenius, A. R., and Torresen, J. (2013). Analyzing correspondence between sound objects and body motion. *ACM Trans. Appl. Percept.* 10, 1–22. doi: 10.1145/2465780.2465783

Palmer, C. (2013). "Music performance: movement and coordination," in *The Psychology of Music*. ed. D. Deutsch (Amsterdam: Elsevier Academic Press), 405–422.

Park, S. W., Marino, H., Charles, S. K., Sternad, D., and Hogan, N. (2017). Moving slowly is hard for humans: limitations of dynamic primitives. *J. Neurophysiol.* 118, 69–83. doi: 10.1152/jn.00643.2016

Persichetti, V. (1962). *Twentieth Century Harmony*. London: Faber and Faber.

Petitot, J. (1990). *'Forme' in Encyclopædia Universalis*. Paris: Encyclopædia Universalis.

Plamondon, R., O'Reilly, C., Rémi, C., and Duval, T. (2013). The lognormal handwriter: learning, performing, and declining. *Front. Psychol.* 4:945. doi: 10.3389/fpsyg.2013.00945

Pöppel, E. (1997). A hierarchical model of time perception. *Trends Cogn. Sci.* 1, 56–61. doi: 10.1016/S1364-6613(97)01008-5

Rasamimanana, N., Bernardin, D., Wanderley, M., and Bevilacqua, F. (2007). "String bowing gestures at varying bow stroke frequencies: a case study," in *Gesture-Based Human-Computer Interaction and Simulation. GW 2007. Lecture Notes in Computer Science. Vol. 5085*. eds. M. Sales Dias, S. Gibet, M. M. Wanderley and R. Bastos, (Berlin, Heidelberg: Springer).

Rasamimanana, N., and Bevilacqua, F. (2008). Effort-based analysis of bowing movements: evidence of anticipation effects. *J. New Music Res.* 37, 339–351. doi: 10.1080/09298210902731386

Rocchesso, D., and Fontana, F. (2003). The Sounding object. *IEEE Multimedia* 10, 42–52, doi: 10.1109/MMUL.2003.1195160

Ronco, E. (1999). Open-Loop Intermittent Feedback Optimal Control: A Probable Human Motor Control Strategy. Technical Report: EE-98005. Systems and Control Laboratory, University of Sydney.

Rosenbaum, D. A. (2009). *Human Motor Control* (second edition). Burlington, MA: Elsevier.

Rosenbaum, D. A. (2010). "Shall we dance? Action researchers and dancers can move together," in *The Neurocognition of Dance: Mind, Movement and Motor Skills*. eds. B. Bläsing, M. Puttke and T. Schack (New York: Psychology Press), 41–52.

Rosenbaum, D. A. (2017). *Knowing Hands: The Cognitive Psychology of Manual Control*. Cambridge: Cambridge University Press.

Rosenbaum, D. A., Cohen, R. G., Jax, S. A., Weiss, D. J., and van der Wel, R. (2007). The problem of serial order in behavior: Lashley's legacy. *Hum. Mov. Sci.* 26, 525–554. doi: 10.1016/j.humov.2007.04.001

Rozé, J., Aramaki, M., Kronland-Martinet, R., and Ystad, S. (2020). Cellists' sound quality is shaped by their primary postural behavior. *Sci. Rep.* 10, 13882. doi: 10.1038/s41598-020-70705-8

Sakaguchi, Y. (2013). "Intermittent brain motor control observed in continuous tracking task," in *Advances in Cognitive Neurodynamics (III)*. ed. Y. Yamaguchi (Dordrecht: Springer).

Sakaguchi, Y., Tanaka, M., and Inoue, Y. (2015). Adaptive intermittent control: A computational model explaining motor intermittency observed in human behavior. *Neural Netw.* 67, 92–109. doi: 10.1016/j.neunet.2015.03.012

Schaeffer, P. (1998). *Solfège de l'objet sonore*. Paris: INA/GRM.

Schaeffer, P. (2012). *In Search of a Concrete Music (English Translation by Christine North and John Dack)*. Oakland, CA.: University of California Press.

Schaeffer, P. (2017). *Treatise on Musical Objects (English Translation by Christine North and John Dack)*. Oakland, CA.: University of California Press.

Smith, B. (Ed.) (1988). *Foundations of Gestalt Theory*. München/Vienna: Philosophia Verlag.

Snyder, B. (2000). *Music and Memory: An Introduction*. Cambridge, MA: MIT Press.

Song, M.-H., and Godøy, R. I. (2016). How fast is your body motion? Determining a sufficient frame rate for an optical motion tracking system using passive markers. *PLoS One* 11:e0150993. doi: 10.1371/journal.pone.0150993

Sosnik, R., Hauptmann, B., Karni, A., and Flash, T. (2004). When practice leads to co-articulation: the evolution of geometrically defined movement primitives. *Exp. Brain Res.* 156, 422–438. doi: 10.1007/s00221-003-1799-4

Spivey, M. J., Anderson, S. E., and Dale, R. (2009). The phase transition in human cognition. *New Math. Nat. Comput.* 5, 197–220. doi: 10.1142/S1793005709001234

Sudnow, D. (1978). *Ways of the Hand*. Cambridge, MA: Harvard University Press.

Summers, J. J., and Anson, J. G. (2009). Current status of the motor program: revisited. *Hum. Mov. Sci.* 28, 566–577. doi: 10.1016/j.humov.2009.01.002

Tenny, J., and Polansky, L. (1980). Temporal gestalt perception in music. *J. Music Theory* 24, 205–241. doi: 10.2307/843503

Thom, R. (1983). *Paraboles et catastrophes*. Paris: Flammarion.

Valls-Solé, J., Kumru, H., and Kofler, M. (2008). Interaction between startle and voluntary reactions in humans. *Exp. Brain Res.* 187, 497–507. doi: 10.1007/s00221-008-1402-0

Valls-Solé, J., Rothwell, J. C., Goulart, F., Cossu, G., and Muñoz, E. (1999). Patterned ballistic movements triggered by a startle in healthy humans. *J. Physiol.* 516, 931–938. doi: 10.1111/j.1469-7793.1999.0931u.x

Vince, M. A. (1948). The intermittency of control movements and the psychological refractory period. *Br. J. Psychol.* 38, 149–157. doi: 10.1111/j.2044-8295.1948.tb01150.x

Warburton, E. C., Wilson, M., Lynch, M., and Cuykendall, S. (2013). The cognitive benefits of movement reduction: evidence from dance marking. *Psychol. Sci.* 24, 1732–1739. doi: 10.1177/0956797613478824

Wiesendanger, M., Baader, A., and Kazennikov, O. (2006). "Fingering and bowing in violinists: a motor control approach," in *Music, Motor Control and the Brain*. eds. E. Altenmüller, M. Wiesendanger and J. Kesselring (Oxford: Oxford University Press), 109–123.

Winges, S. A., Furuya, S., Faber, N. J., and Flanders, M. (2013). Patterns of muscle activity for digital coarticulation. *J. Neurophysiol.* 110, 230–242. doi: 10.1152/jn.00973.2012

**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.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Psychology

ORIGINAL RESEARCH
published: 29 April 2022
doi: 10.3389/fpsyg.2022.733615

Check for updates

# Oral Tradition as Context for Learning Music From 4E Cognition Compared With Literacy Cultures. Case Studies of Flamenco Guitar Apprenticeship

*Amalia Casas-Mas[1]*, Juan Ignacio Pozo[2] and Ignacio Montero[2,3]*

*[1] Faculty of Education, Department of Didactics of Languages, Arts and Physical Education, Complutense University of Madrid, Madrid, Spain, [2] Faculty of Psychology, Department of Basic Psychology, Autonomous University of Madrid, Madrid, Spain, [3] Faculty of Psychology, Department of Social Psychology and Methodology, Autonomous University of Madrid, Madrid, Spain*

OPEN ACCESS

**Edited by:**
*Laura Herrero,*
*Camilo José Cela University, Spain*

**Reviewed by:**
*Julian Cespedes Guevara,*
*ICESI University, Colombia*
*Veronika Diaz Abrahan,*
*Interdisciplinary Laboratory of Cognitive Neuroscience, Argentina*

***Correspondence:**
*Amalia Casas-Mas*
*amacasas@ucm.es*

**Specialty section:**
*This article was submitted to Performance Science, a section of the journal Frontiers in Psychology*

**Received:** *30 June 2021*
**Accepted:** *17 March 2022*
**Published:** *29 April 2022*

**Citation:**
*Casas-Mas A, Pozo JI and Montero I (2022) Oral Tradition as Context for Learning Music From 4E Cognition Compared With Literacy Cultures. Case Studies of Flamenco Guitar Apprenticeship. Front. Psychol. 13:733615. doi: 10.3389/fpsyg.2022.733615*

The awareness of the last 20 years about embodied cognition is directing multidisciplinary attention to the musical domain and impacting psychological research approaches from the 4E (embodied, embedded, enactive, and extended) cognition. Based on previous research regarding musical teaching and learning conceptions of 30 young guitar apprentices of advanced level in three learning cultures: Western classical, jazz, and flamenco of oral tradition, two participants of flamenco with polarised profiles of learning (reproductive and transformative) were selected as instrumental cases for a prospective *ex post facto* design. Discourse and practice of the two flamenco guitarists were analysed in-depth to describe bodily issues and verbal discourse on the learning practice in their natural contexts. Qualitative analysis is performed on the posture, gestures, verbal discourse, and musical practice of the participants through the System for the Analysis of Music Teaching and Learning Practices (SAPIL). The results are organised attending: (a) the *Embodied* mind through differential postures and gestures of flamenco participants that showed a fusion among verbal, body language, and musical discourse with respect to the musical literacy cultures; (b) the *Embedded* mind and a detailed description of circumstances and relationships of the two flamenco participants, and how music is embedded in their way of life, family and social context, and therefore transcends musical activity itself; (c) the *Enactive* mind, regarding the active processes that make differences between the reproductive and the transformative flamenco apprentices, then tentative relationship are observed in the discourse of each apprentice and the way in which they practice; finally, (d) the *Extended* mind through the bodily, technical and symbolic tools they use during learning. Flamenco culture of oral tradition made use of listening, and temporary external representations instead of notational, but also the body played a central role in a holistic rhythm processing through multimodality, such as singing, playing, and dancing. Conclusions point out the embodied mind as a result of the culture of learning reflected through the body and the gesture in instrumental learning.

**Keywords: oral learning, instrumental learning, embodiment, discourse about practice, musical practice, flamenco culture of learning**

## INTRODUCTION

Traditionally, analyses on musical apprenticeship processes began from academia and were mostly focused on the acquisition and transfer of Western European classical music. As soon as apprenticeship learning processes began to be studied in popular and traditional music cultures multifactorial nuances and differences were observed. These factors had repercussions on the activation of different mechanisms by the apprentices which at times involved a different focus or approach to the music. Often the new apprenticeship cultures studied, which did not have to be actually new (Leung, 2018), were described as "otherness" or cultural alterity, i.e., in reference to the previously defined standard of the academic area. Otherness is an epistemological stance that discursively explores the image of cultures that built up their space in the periphery or other intermediate cultural spaces (Sosa, 2009). It is a type of analysis that works as a discourse in specific historic contexts and is deployed in the social, legislative, institutional, and material order through processes that gain meaning through language and in this study through gesture and practice as well.

We cannot simplify or reduce the idea that formal music learning is synonymous with Western classical music using scores, and that informal learning is restricted to popular music transmitted by ear (Middleton, 1981; Folkestad, 2006). What is learned and how the elements are interconnected are not shaped by the type of music itself, but rather by a given approach to music that uses certain tools (representations, processes, contents, and conditions) as mediators for people developing within a *milieu*. These approaches are what we call learning cultures (Olson and Bruner, 1996), defined as the set of shared beliefs and practices on how to organise and promote learning, in this case, musical learning. We approach them without placing anyone above the other, simply directing our attention to elements that may not have been previously taken into consideration. Thanks to previous research on popular music (Green, 2002, 2008; Shah, 2006) educational spheres shifted the focus from teaching to learning, and thus from teacher to apprentice (Casas-Mas, 2022). In addition, today they provide us with a series of analysis categories in which differences are established with respect to formal realms, such as goals, planning, participants, motivation, place, learning style, leadership, and intentionality. But we still need to develop what that sociocultural construction affects in the embodied minds of the members of these communities or environments.

This research aims to provide a novel and therefore exploratory qualitative approach due to being an investigation that delves into the body, the gesture, the posture, the musical production, and the verbal discourse without antecedents, and also comparing these aspects for the first time to cultures with and without musical literacy. Firstly, a comparative disclosure of bodily and gestural issues is performed in three cultures of music learning and then to justify and make an in-depth description of the elements that characterise flamenco guitar learners in the context of the oral tradition. Our approach uses immersion and follow-up processes of students and their environments, in order to describe the culture from the words and gestures of the participants. This is achieved through their personal communicative processes, being treated with the greatest respect, and always monitoring the type of observation and interview.

For this reason, we begin by profiling the framework of 4E (embodied, embedded, enactive, and extended) cognition in relation to music, and then the relations between discourse and practice, or the differences of verbal and action productions. We will then present some nuances through reproductive and transformative learning conceptions regarding different cultures of learning music and the external representations or those "signs" each culture uses to either communicate or self-regulate in apprenticeship and that provide us with vital clues on the construction that could occur on a cognitive, emotional and corporal level. Finally, we will explain several common ideas about the culture of flamenco learning in communities of oral tradition; a non-uniform culture, but with some common music learning strategies because explicit forms of music notation, such as sheet music, are not used. We are interested in the connection with the body, sounds, and gestures that help us understand the experience and externalisation of music from a perspective not yet investigated in-depth within flamenco culture. Cognition 4E help to understand the social aspects of music, particularly those found on the oral transmission, since the body plays its part through action where verbal discourse does not predominate, and the body is an extension of the mind (or vice versa), and even the musical instrument is also an added extension (Nijs, 2017; Schiavio and van der Schyff, 2018).

### Embodied, Embedded, Enactive, and Extended Cognition and Musical Demeanour

The dualism of the classic computational models that depicted both psychological research and education were characterised by splitting the mind and the body, prioritising the first and leaving as a second category those learnings coming from psychomotor action, movement, emotion, and ignoring the specific context in which they are inserted (Claxton, 2015; Pozo, 2017). The philosophical roots on which this approach is based have been highly questioned in the last three decades until the statement that not only the mind is in the body but that the body is also in the mind (Varela et al., 1991; Damasio, 1994). The processes that the musical domain activates in human beings are already an embodied form of action that we adopted to make sense of certain properties of the environment that are discovered in a deep continuity between cognitive and biological processes (Maturana and Varela, 1980; Varela et al., 1991; Schiavio et al., 2017). If we assume some principles of a deep continuity between cognitive and biological processes, through feedback processes of imitation (Donald, 1991; Mithen, 2005), then a dualistic division closes the possibilities of understanding a multimodal and holistic phenomenon as music.

In the analytical effort of science there are authors who divide the notion of the body in two; on the one hand the physical, measurable, and "objective" body, and on the other the body from the phenomenological point of view, which occupies the metabolic, emotional, and other aspects

(Schiavio and van der Schyff, 2018). To say that cognition is *Embodied* means for them that our mental life depends directly on these two (objective and phenomenological) descriptions of the body. For example, the student who is learning to play the violin can become aware of his bodily representations if he records himself while playing (his body tension, the relationship between the movement of the bow and the sound obtained, etc.) (Pozo et al., 2020b). This in turn feeds back into the neurological activity, because it plays an important role in training the musician's physical abilities and thus developing his sense of himself as a skilled and embodied agent (Høffding and Schiavio, 2021).

That new 4E mind (Rowlands, 2010; Nagy, 2017) would not only be *Embodied*, but also *Enactive*, or based on action rather than on the contemplation of the world (e.g., Bowman, 2004; Pozo, 2017; Schiavio et al., 2020). In the same example above, the actual action performed, even mimetic, and only emulated without the instrument (Pozo et al., 2020b), provides the essential content of the musical representation that links sound and action. Students learn better when they play music instead of seeing someone else playing it (Schiavio and van der Schyff, 2018). An *Embedded* cognition is due to it being always situated in an environment from which it cannot be separated without losing its meaning. The learning and production of music cannot be separated from the cultural context in which it is produced. As already mentioned, various studies have shown that music is conceived, lived, and felt differently in the contexts of formal and informal education (Green, 2002; Casas-Mas et al., 2015a).

Finally, Newen et al. (2018) defend that a cognitive process is strongly embodied by bodily processes if it is essentially based on processes in the body that are not in the brain, or is strongly embodied by extra-bodily processes if it is partially constituted by extra-bodily processes. In the example mentioned before, when the student records himself to analyse his bodily actions and their relationship with the sound achieved, the video constitutes an external representation from which to redescribe his own action (Pozo et al., 2020b). At a different level, the scores constitute a resource that allows the performer to reconstruct their own bodily action; such as the sequence of actions to perform and fingering (Casas and Pozo, 2008). Thus, the mind is *Extended* since the mental action would be expanded beyond the brain, not just to the rest of the body, but to the extracorporeal material and symbolic resources in which it is supported. The mind emerges in relation to devices and environments that co-constitute music-like behaviours, and not only "afford" them (Clark and Chalmers, 1998; Schiavio et al., 2017).

## Coherence Between Discourse and Practice

The social and communicative aspects of music, especially through the oral transmission learning and extralinguistic elements could be a priority research objective in music due to the repercussion they have on the body and the musical instrument as extensions and feedback of the mind (Nijs, 2017; Schiavio and van der Schyff, 2018). The music fits into movement, dialectic, emotion, and shared experiences with other beings because it is an inherently social phenomenon and this is precisely what lends it meaning (Cross, 2003). Nevertheless, a new division between the terms of theory and practice appears, between what beings report and what they do in reality, and this begins with the ways in which we approach the phenomenon from research.

Traditionally, the study of conceptions and approaches to learning music were based on verbal information uptake (e.g., López-Íñiguez and Pozo, 2014; Marín et al., 2012), either showing interest in finding out what musicians focus their attention on during the different stages of studying and practicing sheet music (Williamon et al., 2002; Chaffin et al., 2003) or without a score (Green, 2002, 2008). Recently practices are gaining greater emphasis (López-Íñiguez and Pozo, 2016; Boucher et al., 2020), because they inform us about elements that the *Embodied* mind may be hiding behind verbal speech, in addition to appearing to be going at a different speed compared to what a person expresses verbally. Studies of analysis regarding consistency between discourse (as verbal productions) and practice (as sequence and structure of actions) reveal the existence of a major gap in the area of instruction (Torrado, 2006; Clarà and Mauri, 2010; Buehl and Beck, 2015; Clarà, 2019; Úcar, 2021) where verbal expression is higher in complexity than teaching practices. Pozo et al. (2006) observe a frequently produced dissociation between implicit embodied representations of teaching and learning (the somatosensory states associated with physical actions or emotions) and our explicit knowledge of them (the attributions or explanations that we give ourselves over the reasons for these often biased emotional states or physical actions). Examples of research and intervention proposals are focused on developing metaphoric experiences (Zbikowski, 2015; Gibbs, 2017) and reflection on practice (Bain et al., 2002). In relation to musical learning, it could be possible that the coherence between verbal discourse and body language may depend more on the conceptions of learning, for example, reproductive versus transformative, as different types of *Enactive* mind, rather than on the culture of musical learning (Casas-Mas, 2020).

## Reproductive and Transformative Learning in Music

The premise of cultural preservation could easily be related to the context of oral tradition as a necessity for preserving the passing on of information. However, this also occurs in traditions with highly developed literacy, such as the Western European classical music learning culture, precisely because it shares a mode of production with flamenco (Pozo et al., 2022a) in which the piece reaches the apprentice as closed, and the apprentice is expected to "correctly" reproduce it. In the former, the piece reaches the apprentice as closed through the musical score and mediation of the teacher and in the flamenco of oral tradition mediating by the teacher through the oral-procedural tradition. This differs from learning cultures where the pieces reach the apprentice and they have to complete them either individually or collectively through improvisation, such as in jazz (Murphy, 1994; Gioia, 1997).

Different modes of music production refer to the internal structure that every group learning situation possesses and that organises social interaction during practice (Tagg, 1982; Kagan, 1989). The musical production is the process that transforms a hypothetical silence into a complete cultural product (Baño, 2018). The Western European classical and flamenco cultures of oral tradition are opposite in the formal-informal continuum, but have similar modes of production, because the piece is closed to the learner, either because of the score or because of the teacher (Baño, 2018; Torrado and Mengual, 2020). Therefore, the purpose of instruction tends to different reproduction degrees of the piece. On the other hand, in the jazz culture of learning the mode of production is more open and ends through the person or group that improvises (Murphy, 1994; Gioia, 1997; Hickey, 2015), in our case the apprentices. This is an example where reproduction of the theme is not a priority but the use of characteristic elements of the language or style.

That said, the open or closed production model has many nuances, depending on the restrictions that the individual or group context imposes, through the teaching and learning conceptions. Put in another way: this is why within the same culture we may find more reproductive or more transformative apprenticeship processes. In the reproductive approach the apprentice is limited to faithfully copying the notational material, whilst in the transformative approach – also called exploratory – the apprentice needs to integrate the information. This difference was established by Scardamalia and Bereiter (1987) in relation to written language, which was later applied to the field of classical music and musical scores (Cantwell and Millard, 1994; Hallam, 1995; Hultberg, 2002). More superficial approaches to the learning of scores are shown here, in contrast to approaches that required greater autonomy and reflection by the apprentice to integrate the information from different sources.

In previous studies these apprenticeship polarities have been classified as direct and constructive (e.g., Bautista et al., 2010; López-Íñiguez and Pozo, 2014). The direct conception starts from a realistic epistemology in which knowledge reflects the object with fidelity, although with varying degrees of completeness or exhaustiveness. There is partial and complete knowledge. A linear and direct relationship is established between certain conditions (age, motivation, amount of practice, etc.) and the learning outcomes. Instead, the constructive conception implies a qualitative change because it is based on a relativistic epistemology in which knowledge is a construction elaborated in a social and cultural context in relation to certain goals. That construction provides tentative and alternative models for interpreting the object. Learning is understood from complex relationships between components that are part of a system. Nevertheless, in this study, it makes more sense to speak of either more reproductive or more transformative apprenticeships respectively because within the flamenco learning culture of oral tradition no comparable constructive profile has been found to that of other learning cultures. Furthermore, no previous studies have described in detail the possible nuances that could arise within the oral tradition of musical knowledge transmission.

Although classical and flamenco guitar traditions have common antecedents, such as the first guitar treatise played by Amat (1758) or the baroque and classical-romance guitar treatise (Torres, 2012), one singular element of flamenco guitar compared with the classical guitar is the fact that it continues requiring song and dance accompaniment training (Calahorro, 2019). This separates it from most academic musical practice, fostering the creativity and ability to adapt to the situation, that requires the development of intuition and spontaneity in musical production. Another element we believe could differentiate the academic tradition from the oral tradition is that the latter necessarily requires support in practice for its transmission. The former is based on musical notation and has therefore created verbal discourse through a process of development of elements and knowledge. This activates a type of music apprenticeship like a second language instead of like a mother tongue emanating from the oral-procedural tradition (Casas-Mas, 2018). As a result, the apprentice's autonomy, their decision-making and their more or less reproductive use of material are key elements to take into consideration. These can be observed not only from their verbal discourse but also from their gestures, private singing, and how the learner uses the external representations (Casas-Mas et al., 2019). Therefore, the analysis needs to be made of these representations that conform to the *Extended* mind in the oral learning culture.

## External Representations in Music

Vygotsky (1933/1978), conceives of the human as a social and cultural being, and from here develops his theory on social training of the mind through the use of signs (or external representations). These are culturally created tools that are modified by human nature itself. External representations are those that express an international relationship between a series of signs with visual-spatial qualities but refer to another reality present or not, e.g., sound, thought, etc. (Martí and Pozo, 2000). Some are also deployed in a spatial dimension, e.g., notations based on (a) flexible combinations of rules typical of graphics, visual images, or designs on some surface, such as illustrations, maps, musicograms; and (b) the stricter notational systems, like alphabetical, metric or traditional musical scores, which remain outside of the content of this article. In the flamenco learning culture of oral tradition, we concentrate on those signs or external representations which are deployed temporarily and have a little structural definition, like the observable body and gesture representations and oral language, as classified by Pérez-Echeverría and Scheuer (2009). The *Extended* mind has been highlighted through gestures and mimesis (Goldin-Meadow, 2003) because they imply a different level of abstraction from notational systems since the gesture is always linked to the context of production. It is difficult to separate the gesture from the communicative or self-regulating situation that elicits it.

It is true that in the case of non-notational external representations, cognitive psychology has not yet dedicated much attention to describing and organising phenomena and we will base our theories on external representation studies and on research that analyses the self-regulating function of the self-referring gestures and private speech from which we will attempt to make our contribution. Private speech is an original Vygotskian concept defined as speaking to oneself for communication, self-guidance, and self-regulation of behaviour. It is neither intended for nor directed at others. Private speech is a crucial window providing insight into how language mediates and regulates thought processes (Lantolf and Thorne, 2006).

However, we will make use of music for the wealth of signs or marks it can provide thanks to its historical role as a socio-cultural symbolic system, converting it into a tool for regulating social interactions and group behaviour (Livingstone and Thompson, 2009; Tarr et al., 2014). We take this concept to the musical field and see that the use of private singing, and vocal, guttural, whistling, etc., have also been observed, and we propose it has two main functions: (a) self-guidance and self-regulation, e.g., cultural tool in the regulation of a child's behaviour in situations when they are falling asleep (Mead and Winsler, 2019), or (b) when learning a musical instrument (Casas-Mas et al., 2019), or managing waiting times (Winsler et al., 2011) with regard to the previous musical experience.

Thibodeaux et al. (2019) found a relationship albeit incipient, between the production of private singing and private speech. During tasks of selective attention when one appears, so does the other, with several similar self-regulating uses. There were no children who, when producing the first, in the form of song or humming did not produce the second, as speech in its overt form, as a self-guiding tool whilst carrying out a task. Even into adulthood, overt speech can be observed during challenging cognitive tasks (Duncan and Cheyne, 2001; Alarcón-Rubio et al., 2013). In general, we would expect children to move toward increased inner speech with age (Winsler et al., 2006), but this may depend on the type of task or area of executive function skill they are currently developing. It remains to be explored whether *private singing* (defined as a qualitatively self-regulatory and self-guidance tool involving some form of singing, guttural, or vocal sound) has a similar pattern. For the moment, it has been observed that in cultures that make use of explicit musical notation, such as scores (either traditional or in the form of jazz charts), greater use is made of private singing elements, some more explicitly, sonorous, and intelligible, while others more internalised, such as whispered or silent, guttural sounds, and verbal lip movements (Casas-Mas et al., 2019). The same authors indicate a more internalised use in popular musical cultures, like the one treated in this article, rather than in the academic ones.

## Flamenco Music, the Culture of Learning in Communities of Oral Tradition in Spain

Flamenco music is the product of cross-hybridisation when the traditional agrarian culture became transformed into today's urban folklore, and popular religious festivals were transformed at the beginning of the 20th century into show business (Steingress, 2004). The initial reference to the existence of "new flamenco music" was made in 1853 (Sneeuw, 1991) in late nineteenth-century Paris. This was where the "dance schools" of the 18th century came to be called "Bolera School dances" (Junta de Andalucía, 2003) and involved an artistic re-interpretation of folk dances and theatrical repertoire of dances "on stage," also at the grassroots level, with a greater professionalisation of performers. Here we are talking of flamenco music which like other popular Mediterranean types of music was defined as a piece of urban music at the beginning of the XX century. It has recently been acknowledged as the *Intangible Cultural Heritage of Humanity* by United Nations Educational, Scientific and Cultural Organization [UNESCO] (2010), but there are many aspects remaining, particularly urban flamenco, of its forms of production and transmission, which have still not been researched in Spain. This article analyses and attaches value to the recovery of the forms of flamenco apprenticeship within the communities in which it is orally transmitted.

The Flamenco culture of learning for instrumental learning had no formal academic accreditation in the country until very recently; it belonged to the oral tradition, and here we describe it to situate the *Embedded* mind in which its members develop. The incorporation of instrumental and vocal flamenco to academic studies in the Conservatory is very incipient in Spain. The Act of the regulation of the general educational system (Jefatura del Estado, 1990) was the first to mention, in Jefatura del Estado (1995), the specialty of flamenco and instrumental flamenco guitar and the Organic Law on Education (Jefatura del Estado, 2006) when a greater systematisation was carried out. Currently, the flamenco guitar at a tertiary educational level is poorly integrated into a few conservatories throughout the country, in cities, such as Córdoba, Murcia, Barcelona, Malaga, and recently, Jaen. Despite institutional neglect sometimes flamenco is now considered as a mark of identity in Spanish citizens of Roma ethnicity, despite the fact that there are many non-Roma professional flamenco musicians (Payos[1]). This article will spotlight the Roma ethnicity community in Madrid to reflect on apprenticeship issues which could also be present in non-Roma ethnicity communities in certain regions of Spain where flamenco music is acquired as a mother tongue.

The European Commission *Country Report Spain 2020* recognises that the Roma ethnicity community of today is inclined toward considerable social vulnerability. Only 17% of teenagers finish the first cycle of secondary education and 63% of Roma ethnic young people do not work, study or receive any training [European Unión Agency of Fundamental Rights [FRA], 2018; Roma Secretariat Foundation (FSG), 2019]. Early dropout from secondary education in Spain is high, at 16%, compared with the European average, at 14.9% (Ministry of Education and Professional Training, 2019), and more common than we would like to think in Roma communities, at 63.4% [Roma Secretariat Foundation (FSG), 2013]. García Fernández et al. (2017) performed an exhaustive review of primary and secondary school textbooks in Spain, concluding that the Roma population has been systematically omitted, and if it is mentioned, the tone is paternalistic toward or prejudiced against it. This has some improvements in the new 2022 Act of the regulation, that the Spanish government has established for Primary and Secondary Education. This type of omission has already been described in research studies with other communities (Lave, 2011; Karlsen, 2012; Sexton, 2012), and perhaps it is one of the factors to take into account in the high dropout rate.

Some authors have resonated this history of discrimination up to the present day [e.g., Casas-Mas, 2018; European Unión Agency of Fundamental Rights [FRA], 2018; Roma Secretariat Foundation (FSG), 2019], highly internalised in the culture. Even the Dictionary of the Royal Academy of the Spanish Language insists on maintaining the definition of Roma people as

[1] *Payo* is the name that members of the Roma community assign to people who do not belong to it.

"trapacero" (sb. who with craftiness, falsehoods and lies attempts to trick someone in a matter), although at the request of the Roma community it incorporated in 2015 a usage note that warns against the meaning "as offensive or discriminatory"[2]. Despite the conceptions and conditions surrounding this population, which is not monolithic, the Roma ethnicity community continues to attempt to strengthen its cultural cohesion ties that preserve some of the traits found in its origin. Its media are the preservation and replication of several values which are possibly regarded with bewilderment or disconcertion from "outside." However, this reinforces and protects them from languages, ideas, policies, or behavioural patterns prevailing in their immediate environment.

This research study portrays two case studies of two participants belonging to the oral tradition of guitar learning in Roma ethnicity communities and whose speech and practice was remarkable from a previous sample of a multiple case study with guitarists from different learning cultures. The previous studies allowed us to hypothesise within this flamenco tradition and the main aim of this article is to probe into the contrast of this flamenco of oral tradition with other cultures of learning music regarding the body and gestures, and the learning differences within the same tradition. More specifically, we have formulated the following goals for this article:

1. First, to observe the peculiarities of guitar learners related to posture, gesture, and body, belonging to a flamenco learning culture of oral tradition, based on a previous comparative sample of Western classical, jazz, and flamenco guitar learners of oral tradition (Embodied mind).
2. Second, to describe the environment where these apprentices develop, taking into account the relations they establish with teachers, family, and peers (Embedded mind).
3. Third, to focus on observing if there are differences in the speech and learning practice between two flamenco guitarists selected as instrumental cases from the initial sample, with polarised approaches to learning, reproductive (R) and transformative (T), as differential active processes. The hypothesis is that intuitive conception about what it is to learn constitutes the starting point for the development of practice and action. We will observe whether or not there is coherence between the discourse and the practice of these learners (Enactive mind).
4. Finally, the observation of the forms and uses of external representations by the learners of oral tradition constitutes another of the main objectives of this study, because it provides a lot of information about the body, sound, and gesture, which has not been the objective of traditional research, more focused on verbal language. We examine such issues here which, to our knowledge, is the first article on this topic (Extended mind).

[2]Dictionary of the Royal Academy of the Spanish Language, the fifth meaning of Gitano (Roma):http://lema.rae.es/drae/?val = gitano (January 2022).

# METHODS

In previous research, through the analysis of variance and the lexicometric analysis, three learning cultures were outlined as significantly different in their conceptions and verbal discourse about teaching and learning music (Casas-Mas et al., 2014, 2015a). It was carried out through a prospective *ex post facto* design, where the musical learning cultures of Western classical, flamenco, and jazz were representative of formal, informal, and non-formal educational environments, respectively (Trilla, 1997; Casas-Mas, 2013). Their teaching structure and intentionality made up the non-manipulable independent variable. Then, we set out to observe the discourse and practice of learners belonging to these cultures, which would be our dependent variables in the research. To highlight the nuances between forms of learning within each culture and not present them in a uniform way and as a single prototype, we selected the two participants with the most polarised conceptions and discourse (reproductive or transformative) from each of the three cultures. Thus, we implemented a multiple case study in which we observed and analysed their music learning practices in depth through the preparation of a piece and the situated discourse on practice. There were then six instrumental case studies (Stake, 1998) selected by the variable conception of learning within each culture of musical learning. This allowed us to observe the coherence between verbal discourse, discourse situated on musical practice, and practice (Clarà and Mauri, 2010). These flamenco case studies are drawn from this context to be illustrated in this article.

## Participants

The original sample was comprised of 30 guitar apprentices coming from Western classical, jazz, and flamenco of oral tradition as formal, non-formal, and informal cultures of learning, respectively, following Trilla (1997) and Green (2002). The categorisation of learning cultures covered: the degree and institutionalisation of teaching; the structure of the studies established in the national curriculum or not; whether they had teacher selection through competition or only merits; the regularity and intentionality of teaching, and the evaluation leading to accreditation. The Western classical culture of learning met all requirements of the formal realm. In the jazz culture of learning in Spain, there are learning structures, such as jam sessions, that break one or some of the formal realms which are essential to learner training, and they use chart-type musical scores that give a general notion of the piece beforehand. In the flamenco culture of oral tradition children from a very early age practice music (singing, dancing, playing) many hours in a family setting and with peers, during which they exchange exercises, *falsetas,* and pieces inherited from teachers or composed by themselves. Nowadays they often use video/audio recordings as tools, particularly on mobile phones, but they do not use any kind of music notation. There are no exams or specific accreditations.

The participants were guitarists at a semi-professional stage of learning who had educational and semi-professional experience in only one of the cultures of musical learning considered in this study. Based on this criterion, the sample is configured as follows: ten participants are from the classical culture of learning (5 men, 5 women), aged 19 to 29 years ($M = 24.9$; $SD = 3.48$), have spent more than 10 years studying music in formal realms, such as the conservatory, and are studying for a tertiary degree (equivalent to a Bachelor of Music degree) or master of music studies. They also have pre-university or university studies, and most of their families have some relationship with amateur music. Among the participants, ten are from the jazz culture of learning (9 men, 1 woman), aged 26 to 42 years ($M = 29.6$; $SD = 4.93$), with a college degree in non-musical studies, and most have studied for a professional degree in music (prior to tertiary studies). Their families have no relation to learning music. Furthermore, ten participants are from the flamenco culture of learning (10 men), aged 15 to 25 years ($M = 16.82$; $SD = 2.96$), all of Roma ethnicity, and most have not completed compulsory secondary education. They have been studying the guitar for 1 to 5 years with a specific teacher. There are professional musicians in all their families.

Considering the reproductive and transformative conceptions of learning, the two most opposite learners from each of the three cultures were selected to make an in-depth multiple case study with constant comparison analysis of their learning practice and their situated discourse about practice. After contrasting the hypothesis, some peculiar characteristics about teaching and learning in the oral tradition flamenco culture came to light, which showed it to be far removed from literate music cultures (Casas-Mas et al., 2014, 2015a). In this article, we will describe the two participants of the flamenco apprenticeship culture selected as instrumental case studies. The first is the apprentice with a more reproductive learning profile (hereinafter "R") and the second with a more transformative profile (hereinafter "T") in the Roma ethnicity communities observed.

## Materials and Procedure

### Musical Piece

The pieces prepared by the two students were a *falseta*[3] each by bulerías. Bulería is a fast flamenco rhythm (around 240 bpm) in 12 beats that may also be broken down into a measure of 6/8 followed by a measure of 3/4. It is most commonly in an A-Phrygian mode, with a sharpened third. It is traditional from Jerez de la Frontera, in Western Andalusia. Although on occasions it is played with *air* (speed) not excessively fast, its spirit is burlesque and playful. The twelve beats are usually those most indicative of flamenco, compared with other types of music; for this reason, we believe their inclusion to be most relevant, in addition to being those which the participants chose to work with. The *falseta* of the first student was from the maestro Víctor Monge, "Serranito," based on a text by Federico García Lorca. The *falseta* of the second student was a composition by his teacher "El Viejín."

[3]The *falseta* is a fragment of music in the complete sense that it has a beginning, middle and end. Duration is variable, and it is used as a combinable module with other musical sections which adapt well in rhythm, speed and tone. The *falsetas* are pieces taught by teachers to students, or between equals. Many are the creation of the emblematic maestros themselves and others already form part of the tradition. In some cases, they are composed by the actual apprentices.

### Structured Interviews With Open-Ended Questions

An initial interview with each participant was conducted. This took a generic and global look at their learning expectations during the three learning phases of their musical piece: initial, intermediate, and final. After each of the three practice sessions (PS) a post-session interview was conducted, three in total, which was different from the initial one and was more detailed, where their opinion regarding learning from the previous session was sought and their expectations regarding the following PS (the interview protocol is included in the **Supplementary Material**).

### Practice Sessions

We proceeded to record the PS of learning a piece of music for each apprentice. We recorded 3 PS for them: the first at the beginning of learning, the second after 2 weeks of preparation (in the middle), and the third when they were ready to play in public, around 3 weeks later, based on previous research (Nielsen, 1999; Williamon et al., 2002; Chaffin et al., 2003). We analysed the recordings of the PS in which the researcher also took notes *in situ*, to facilitate the appearance of private speech, and we fitted both pieces of information collected together.

### Reflection on Practice Interview

We also combined verbal information collected and learning practice observation with reflection on practice interviews (RPI). In these, we watched the video of the PS with the apprentice, and we asked questions regarding points of interest from the research, and the apprentice also commented upon what was brought to his attention when watching his PS (interview protocol included in **Supplementary Material**). These interviews proved to be very rich in the information given by the apprentices and served to clarify thoughts, emotions, and decision-making with regards to what was observed by the researcher. The participants provided their written informed consent to participate in this study in accordance with the Regulations of the Research Ethics Committee of the Spanish university. Participants had the right to discontinue participation at any moment and were not compensated for their time.

This article is based on the System for the Analysis of Music Teaching and Learning Practice (SAPIL), which is widely developed in Pozo et al. (2022b). Categories are divided into three main sections: First, it deals with categories of learning conditions; that refers to the nature of the sociocultural environment that conditions the practice of learning from the human mediators that make it possible. Second, the processes of learning that include the explicitly or implicitly cognitive processes managed by the teachers and students, the cognitive demand of tasks, the level of the meta-cognitive management the students require of them, the memory or recovery strategies used. Finally, the results of learning or what is learned in a certain learning culture, what it values and encourages both explicitly and implicitly. The system contemplates the distinction between symbolic (where we include external representations), procedural, and attitudinal results of learning. The inter-reliability agreement is Fleiss' Kappa $> 0.8$ for the system of analysis, SAPIL. In addition, we have carried out a triangulation to test validity through the convergence of information from different collecting information techniques: Structured Open-ended Interviews, Practice Sessions, Reflection on Practice Interviews, and triangulation with the apprentices at the same time.

## RESULTS

The structure of the results section will be presented according to the 4E cognition; embodied, embedded, enacted, and extended. We begin by comparing the posture and gestures of the participants in the three cultures of learning as the *Embodied* mind, then we follow with the social environment and circumstances of the participants and a deep description of two flamencos, from the *Embedded* mind framework. In addition, a detailed description through microanalysis of practice and discourse about the practise of musical processes of learning, as *Enactive* mind. Finally, we analyse the use of external representations, but instead of a spatial notation system, we focus on the analysis of temporal ones, such as body and gesture and oral language for learning music as an *Extended* mind. These bodily, postural, and gestural categories are mentioned in this analysis system but are not specified because they are emerging from the results of this research. That is why we will begin and finish by explaining the results regarding them.

### Posture and Gestures in Three Cultures of Learning (Embodied Mind)

We will begin by describing the learning position with respect to the teacher, the posture, and the gestures that accompany the verbal discourse of these advanced guitar learners observed during the investigation. Teacher and student were positioned in an arc in the jazz and classical cultures of learning so they could both see the notational material, the score, located between them. Nevertheless, in the absence of notation communication between teacher and learner through body language was very important in flamenco, so the teacher was usually positioned opposite the student, who was accustomed to decoding the teacher's movements from a contralateral position. The teacher's eyes, eyebrows, shoulders, and head were involved in imbuing the musical content with significance, used very expressively for emphasising while accompanying musical performance.

The first issue to mention is that posture while playing an instrument is a relevant cultural acquisition that configures the embodied mind. We found different postures had been adopted in the three cultures during PS: the posture of classical participants could be defined as ergonomic, aimed at preventing pathologies due to lengthy exposure to practice (Carreras et al., 1996). The sitting posture of jazz participants looked relaxed and introspective, with legs completely crossed and the body leaning over the guitar as if drawing knowledge from the inside out. Flamenco participants' posture was somewhere in-between: legs crossed, like jazz players, though in a more open position, with right ankle almost resting on left knee, and although they leaned over the guitar, their posture was more active than in jazz players, more similar to classical players. These three postures can be seen in **Figure 1**.

Second, going into details regarding the distinctive position of apprentice and teacher in this flamenco of oral tradition culture, teacher and student sit opposite one another, so that the teacher's guitar fretboard can be seen by the student sat opposite. This is not a mirror, it is so the student gets used to deciphering the teacher's movements and performing them inversely from their own (**Figure 2**). The student is learning through what we could call "osmosis"[4], since they reproduce practically in real-time and simultaneously, the material they have not previously seen. We believe the importance of gestural communication and the teacher's facial expressions is of great interest. The student usually watches the teacher's guitar fretboard whilst reproducing what he plays, but the teacher mainly looks at the student's eyes, as well as checking the student's guitar fretboard from time to time.

The third and relevant question is the frequency of appearance of gestures that accompany verbal speech in these three learning cultures. Classical participants did not use the guitar while speaking but often gestured as if playing an *air guitar*. Jazz participants gave many examples of what they told us about, by playing them on the guitar, which they usually had nearby, but taking the guitar and leaving it in the case after each example. Flamenco participants displayed a fusion among verbal, body language, and musical discourse: their discourse was completed in the three forms while remaining incomplete in any of the three forms separately. They always held the guitar, making it sound constantly, without any exemplifying purpose. Sometimes they would replace the beginning of the end of a verbal phrase with a musical phrase on the guitar, in which the prosody usually matched the verbal language it was substituting (**Supplementary Video 1** in **Supplementary Material**). Another example was the way in which flamenco participants stamped a foot to support what they were explaining in verbal discourse, as an important repeated gesture that was not used by apprentices in other cultures. A comparative frequency of gestures in these three cultures of learning can be observed in **Table 1**.

### The Socio-Educational Context in Which Learning Cultures Are Inserted (Embedded Mind)

From the *ex post facto* study we have been able to observe some variables in common among the participants that outline the three learning cultures as embedded minds. The classical participants have completed high school studies and are in the process of university studies. Their specific musical studies are either undergraduate or in some cases postgraduate. Their families often include amateur musicians or music students. Up to seven of the jazz participants have university-level academic studies not necessarily musical. Their musical studies span twelve years or are equivalent to a professional degree in music, in the majority of formal classical studies that they have left. The study highlights that in their families there are usually no members who have had specific musical training. The apprentices of this culture are, on average, a decade older than the apprentices of

[4] Osmosis is a term taken from biology that defines a reciprocal influence between two individuals or elements that are in contact. In this case the contact is metaphorical.


**FIGURE 1 |** Learner posture during practice in three learning cultures.

the flamenco culture. Flamenco participants were educated to a maximum level of secondary school, most of them incomplete; with specific non-formal guitar musical studies of up to five years. They are very young apprentices who can already perform in public with a few years of specific training in the guitar. In all their families there are professional musicians. In flamenco culture, the perception that apprentices have of music cannot be separated from their family and social context, music is embedded in their way of life and therefore transcends musical activity itself, unlike the other two cultures where musical activity is more restricted to the context of musical production itself.

We will now describe in-depth the flamenco guitar school and the two flamenco participants selected for this study, to gain a better understanding of their environment. It is important to point out that in the case of the school of master "Entri," PS are daily and collective. All the guitarists go from Monday to Friday to group PS lasting approximately 2 h, in a large room and all the apprentice guitarists of all ages sit in a circle with the guitar in their hands (**Figure 3**). The room has a dance stage in the centre of the floor and is presided over by the teacher expertly playing the guitar. He also forms part of the circle, presiding over it and proposing routines and exercises which they then play either one by one until the circle is completed or all at the same time. Occasionally the teacher asks an advanced student to guide part of the session whilst he gives individual attention to a student in the adjacent room. Unlike what happens in other musical cultures (Casas-Mas, 2018), flamenco apprentices are embedded

**FIGURE 2 |** One-to–one class between the flamenco teacher Entri with student R, sitting opposite, with the teacher looking at the apprentice's fretboard and the apprentice looking at the teacher's fretboard.

**TABLE 1 |** Frequency of appearance of gestures that accompany verbal speech in three learning cultures.

| Codes | CL-R | CL-T | JZ-R | JZ-T | FL-R | FL-T |
|---|---|---|---|---|---|---|
| Play guitar without exemplifying purpose | 2 | 0 | 0 | 3 | 35 | 13 |
| Play guitar with exemplifying purpose | 0 | 5 | 32 | 36 | 24 | 25 |
| Gesture of playing guitar to set an example (without guitar) | 14 | 12 | 0 | 5 | 0 | 0 |
| Play guitar replacing verbal discourse | 0 | 0 | 0 | 0 | 10 | 1 |
| Foot kicks to support the verbal discourse | 0 | 0 | 0 | 0 | 4 | 7 |

*CL, Classical apprentice; JZ, Jazz apprentice; FL, Flamenco apprentice; R, Reproductive apprentice; T, Transformative apprentice.*

in a community of practice that has, however, a clear hierarchical organisation, from the teacher and advanced students to novices.

The first apprentice is 17 years old and a flamenco guitar student at the Academia de Caño Roto[5] directed by the teacher, "maestro," Aquilino Jiménez, "El Entri," and both are of Roma ethnicity. The level of this student, described by his teacher, is advanced. He has studied with the teacher for four years and has already performed in public and professional events although he continues instrument training. He sometimes replaces the teacher in the group guitar classes in the school. He has finished his third year of statutory secondary education and the middle grade of a vocational training course in computing. He comes from a Roma ethnic family of music professionals, among whom is the teacher. His father plays the guitar on a regular basis in Evangelist church gatherings. We would classify his discourse as closest to traditional, repetitive learning.

The second apprentice is 15 years old and is a flamenco guitar student of the maestro José Jiménez, "el Viejín," and both are of Roma ethnicity. For almost a year he was previously a student at the Academia de Caño Roto directed by the maestro Aquilino Jiménez, "El Entri." He has performed at several public events as a guitarist, despite being very young at the time of this research. Since he was very young he has played the flamenco drum box. He comes from a Roma ethnic family with professional musicians, among them the maestro "Entri." His grandfather sang fandangos in a professional show and his younger cousins also play instruments and sing. His father is pastor of the Evangelist church. He is studying the second year of statutory secondary education and intends to complete his education. He also likes blues and jazz music. His discourse is qualitatively different from the previous participant and is closer to constructive, comprehensive, and transformative learning.

[5]Caño Roto is an area belonging to the Latina district, in the south of the Madrid capital of Spain. It is part of a social housing project carried out from the 1950s to bring order to spontaneous shantytown settlements. An extremely high percentage of citizens residing in this neighbourhood did not finish primary education and it occupies the fifteenth place in the income level of the capital (Government Area of Economy and Employment, 2012).

**FIGURE 3 |** Daily group flamenco guitar class in the *Escuela del Maestro Entri* in Cañorroto, Carabanchel, Madrid, Spain.

During interviews apprentice, R never alludes at any time to learning issues in relation to other apprentices, but apprentice T mentions the possibility of practicing the piece with his peers as something that is easily within his reach and real. Therefore, this situation demands that all participating parties adjust to one another, those who are clapping, singing, or dancing, and those who are playing an instrument.

> Practise the beat, if I say that to any girl [his sisters] come and get into the beat [plays], or to my cousins who are usually more often with me, they do not get tired because they probably like it as well. (T: Initial I.)

This is an indication of an element that is highly relevant for the understanding of the musical apprenticeship phenomenon in this culture but is expressed at a very low frequency in the individual discourse. No importance is attached to it because for them it does not fit into the "learning" category which we are asking questions about, it is understood as fun.

## Differences in the Activity of Learning Processes Within Flamenco Guitar Apprentices in the Oral Tradition (Enactive Mind)

Learning processes are neither passive nor reactive. They are the consequence of several neuronal connections that imply some kind of activity and these connections configure perception, emotion, attribution, movement, or the disposition to act (Noë, 2004). We will now describe specific elements of the individual PS of the two flamenco apprentices, that refer to this activity of the mind. The complete duration of the individual PS ranged between 6 and 19 min. They are brief sessions compared with academic cultures (that are around 60 min), since a great part of the practice takes place in a group setting, not individually. The main idea of the session was literal retrieval of the *falseta* that they had learned with their teachers. While the session of apprentice R was essentially reproductive with the creation of material just when the session was finished, the PS of apprentice T involved up to 12 interruptions of that literal retrieval in which the material played was totally new. The time distribution percentages of the PS of both apprentices are contained in **Table 2**.

The modification in literal retrieval in apprentice T was due to three types of reasons and each one, therefore, fulfils a different function (**Figure 4**). Firstly, the apprentice's mobile where he had recorded the class with the teacher broke, so that he was unable to use it at home and could not retrieve the information properly. Under these circumstances, where memory is not entirely faithful, rather than stop the musical discourse the apprentice completed it with new information of his own. Secondly, the *falseta* started in an unconventional beat, the eighth of a total of twelve, which is difficult. Instead of always starting from silence the apprentice linked this *falseta* with rhythmic pattern sequences and a rhythmic *rasgueado* technique, as a "run-up" or by way of a filler. Thirdly, due to so many repetitions of the same material, the apprentice used his creations as a form of digression, i.e., without direct reference to the *falseta* in question, as an intrinsic motivation strategy to continue persisting in the task later.

We will now explain the arguments of this apprentice in the RPI in these three circumstances. In the first instance he was not concerned about having learned it in a different way to the original, or "with mistakes," and this was not a worry to him at all. The second function was a strategy for playing a *falseta* starting on the eighth beat. Here self-regulation was detected from whispering to counting the beat, like the self-guidance described in detail in the categories which Casas-Mas et al. (2019) termed as *private singing* (**Supplementary Video 2** in **Supplementary Material**). He used the numbering of beats to adjust the beginning of the *falseta*. This process differed from that of apprentice R, where we found the private signing was very internalised, emerging as a highly guttural whispering when confronted with difficulty, e.g., a negative assessment of what he was playing and the need to visualise the fragment again to try to reproduce it exactly as it was in the video. The third strategy was the use of digression as a source of intrinsic motivation strategy to continue persisting in the task later.

We observed a tentative relationship between the discourse of each apprentice and the way in which they practice, but it was truly surprising to find such differences in action, having selected the participants from their discursive production. There were, therefore, concordances between discourse and observation of posterior practice between participants. This represents the connection of activity between the verbal mind and the mind in procedural action (Casas-Mas, 2020). However, within each participant, it was striking that in the case of apprentice R we found there was incongruence between his discourse and his practice which we did not find in participant T. When we asked questions in the RPI apprentice R sometimes did not realise what had previously been practiced. For example, situations in which the researcher had observed that were possibly difficult for him, he denied knowledge of. The researcher then showed him one of these difficulties in the video recording asking for his opinion. This was when he reacted and played the fragment "properly," but he did not acknowledge that it had been difficult in the past. On the contrary, for apprentice T there were no contradictions between discourse and practice. He easily recognised situations and actions he had not done "correctly." His practice was ahead of or on a par with his discourse. He had automated strategies in his practice which he easily was aware of in the RPI. This difference between the reproductive and the transformative apprentice was also found in apprentices of other learning cultures (Casas-Mas, 2013).

**TABLE 2 |** Dedication of time in percentages of the practice session of flamenco apprentices with Reproductive (R) approach, and with Transformative (T) approach.

| Distribution of the PS | R (in minutes) | T (in minutes) |
|---|---|---|
| Literal retrieval and linking between fragments | 97.4 | 80.9 |
| No literal retrieval, adaptations instead | 0 | 16.9 |
| Creations with new material | 2.6 | 2.2 |

## External Representations in Flamenco Learning of Oral Tradition (Extended Mind)

With regard to external representations, there are some outstanding features distinguishing the participants of flamenco culture from those of the classical culture, closer to the jazz culture. Here, the flamenco culture made use of temporary external representations, such as a drum machine to define the tempo and speed of typical *palos flamencos*, and videos, that were closely related to beat and rhythm, common in popular cultures. In jazz culture, for example, audio editing software (Band in a Box) and metronome are frequently used.

The apprentice R uses his mobile to record a video of the teacher (**Figure 5**) and then reproduces it to learn from it. The mobile phone functions as an extended mental resource that enables new cognitive functions, in this case as a mnemonic tool. In his own words, he describes the action of recording his teacher and reproducing it at home on this computer where the screen size is bigger for viewing with greater precision. The teacher makes a special version for recording characterized by lowering the usual speed of the *falseta* (**Supplementary Video 3**).

However, apprentice T describes getting knowledge mainly from listening. In his discourse, he stresses more what he hears than what he sees. Listening is his main source of knowledge and he associates it with repetition for generating a mental representation. He describes this representation as "having it in his head" and "getting the concept." It is also an activity that intrinsically motivates him. This apprentice also says he uses the mobile for recording the teacher on video and then reproduces it to finish learning it. However, he specifies that the recording is support for long-term memory but the learning is done with the teacher on site. None of the apprentices, mentions listening to cover versions or any kind of notation since the latter is not used in oral tradition flamenco.

> You have to have this in your head, you have to listen to it a lot [. . .] I like getting out and inventing things [. . .] (T: Initial I.)

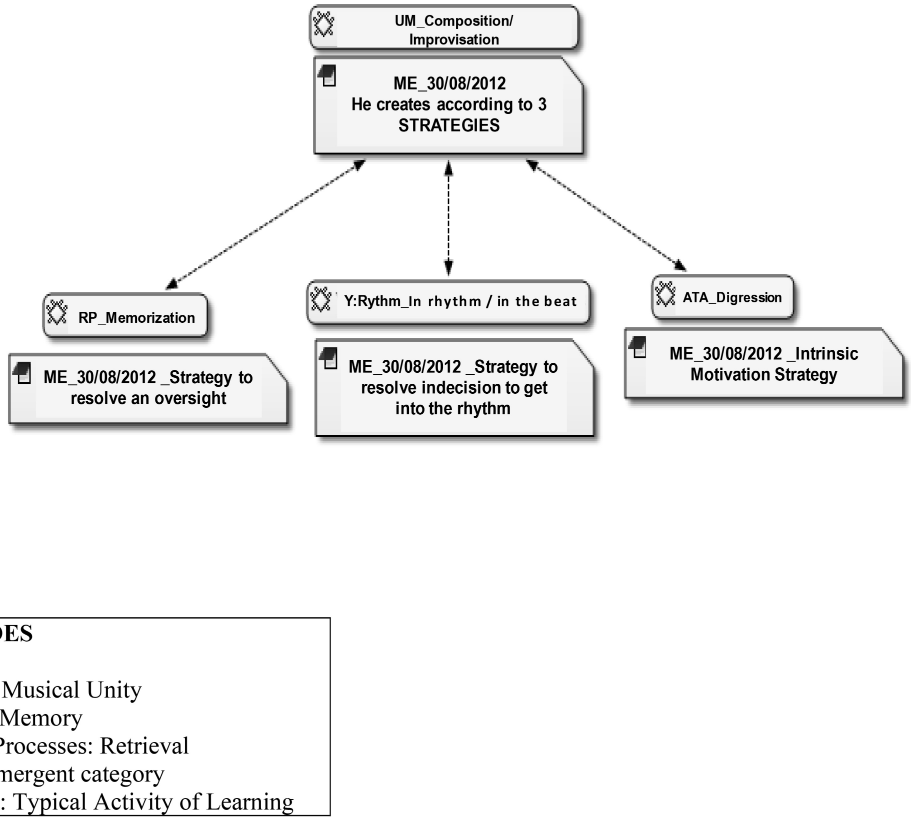



**FIGURE 4 |** Sequence function where apprentice with Transformative approach establishes three strategies of musical creation.

**FIGURE 5 |** Apprentice Flamenco R records a video on his mobile of his teacher's performance.

The crucial difference between both apprentices is undoubtedly what they do with the auditory material. Whilst apprentice R only mentions the analytical processing, referring to technical guitar issues and pointing to certain sections of the piece, apprentice T also mentions the expressive and holistic processing. What we have called expressive processing refers to those procedural contents of an interpretative-intuitive nature that help us to embellish musical discourse. This is widely mentioned by this participant and is not linked to notation but to a large extent depends on the auditory appreciation. It is highly connected with considerations of rhythm, such as accents and entries or closures in certain specific beats.

We also found that it was very much linked to holistic processing, which is the one most characterised by this participant. In other words, with contents of an interpretative-intuitive nature, alluding to the holistic type comprehension or referring to the composer, style, and overall nature of the piece.

> Tomate[6] is pure rhythm. For example, Tomatito in all of his *falseta*s, you perform a Tomatito *falseta* and it won't sound the same, because what he does with all of his notes is he tries to get in the rhythm, so it's as if he is dancing all the time [plays an example]. The rhythm [beats with his foot] and the accents [beats with his foot], [. . .] notice where it goes faster [plays], where the rhythm is held [plays], so that it comes out as if it were dancing [plays], you see? (T: Post-Practice I. 2)

[6] *Tomate* refers to the guitar master Juan Manuel Muñoz who was born in Córdoba in 1970 and was the son of the mythical guitarist Juan Muñoz, who was also nicknamed "Tomate" and brother of the three flamenco-pop singers "las Ketchup," who popularised the song *Aserejé* in 2002.

The idea of creating applies both to the guitar performance and to other more global psychomotor procedures in the form of body extensions; like the dance that is another characteristic of their discourse and that he describes as something common to all people of the community:

> [. . .] but I think that all [those of Roma ethnicity] create small choreographies, I mean any one of them you meet on the street [plays], anyone of Roma ethnicity you see and you say give us a *patada*[7] and they will do one [plays], come what may. (T: Post-practice I.2)

Apprentice R gives us explicitly sung examples of what he is explaining in the verbal discourse on two occasions. The rest of the time we only observe guttural sounds accompanying verbal discourse and musical discourse, especially when he makes a mistake. However, apprentice T uses singing instead of using his guitar up to four times. He is an apprentice who integrates different modes of expression to communicate his musical idea. The mode of communication, such as dance, singing, or instrument interpretation, is only a means of this transmission, not an end. He explains that one of the difficulties he has found has been to focus on this expressive content of the *falseta*, and the solution he offers encompasses the previous ideas holistically.

> [To resolve that difficulty] I play it by ear, like the teacher did, and I sing it to myself as if it was being danced. (T: Post-practice I.2)

## DISCUSSION AND CONCLUSION

The discussion will be presented by embracing the four objectives of this research and attending to the four factors that constitute the 4E Cognition in musical learning: Embodied, Embedded, Enactive, and Extended mind.

The first objective of this work was to observe the peculiarities of the body, posture, and gestures of advanced guitar learners comparing three cultures of learning, from formal to informal and with differences in musical literacy. The main conclusion is the importance of the body, verbal, and extralinguistic aspects in music learning, some of the main issues that the 4E highlights. This article allows us to conclude about the conformation of the embodied mind as a result of the culture of learning reflected through the body and the gesture in instrumental learning. In general terms, flamenco learners differed significantly from other cultures with a more sparing verbal speech (Baño, 2018; Casas-Mas, 2020), and compensated for this by greater use of facial and body gestures, as well as by playing the instrument constantly during their verbal communication. What is representative in them is a fusion between the body, music, and verbal discourse. We want to draw attention to emerging elements in relation to facial, eye, and eyebrow expression, as elements to be investigated in the future. The body eases the meaning-making process of learning music from early childhood (Nijs and Bremmer, 2019), and that process is the result of the limits that the body imposes on it, as well as the body being configured through its use. The posture of the flamenco and classical apprentices during practice was similar, no doubt due to their common origin. The crossed legs and closed-chest in the practice of these urban jazz participants might be hinted at as a reflection of greater introspection for creation. It could be the difference between drawing knowledge from within the individual or receiving it from without (Casas-Mas et al., 2015b). Although flamenco culture of learning is based on the literal transmission of musical knowledge between instrument teacher and student, we have been able to observe this fusion of communicative modalities, such as gestures, the body, the music, and verbal discourse, also in creative processes in other cultures (Besada et al., 2021). This multimodality has repercussions on their sense of identity or self, a state of mind of pre-reflective or embodied subjectivity (Høffding and Schiavio, 2021), and on the capacities or competencies that they integrate into their discourse, such as dance.

The second objective was to describe the environment where these apprentices develop, taking into account the relations they establish with teachers, family, and peers, or their *Embedded* minds. This is a brain-body system in continuous interplay with its niche where the organism enacts rich patterns of action and perception in a complex musical domain, as in some Roma communities, that dynamically unfold and shape each other, revealing new properties of a fertile environment that can be acted upon (Høffding and Schiavio, 2021). The study of "otherness," as a culture of oral tradition, is relevant because of the learning richness it provides, and resources are needed for in-depth analysis of these communities from the inside. Previous research studies have already focused on socio-economically disadvantaged groups to prove that they have limited means of buying musical instruments and the repercussion this has had on their identity and enrolment in music conservatoires (Benedict et al., 2015; Burland, 2020). Furthermore, greater attention is increasingly given over to how indigenous musicians are overlooked by urban citizens, mass media, and even the national curriculum (e.g., O'Neill, 2009; Cain, 2015; Holguín and Shifres, 2015), coinciding with the analysis of the school texts by García Fernández et al. (2017) in Spain. This points out the barriers that Roma girls and boys face in their educational trajectory. It is equally the case that there are some descriptions of successful student models belonging to the Roma community (Abajo and Carrasco, 2004). Meanwhile, just as the "other" may be the flamenco apprentices compared with musical literacy apprentices, so may they also, and why not, be instrumentalist women in the area of this same learning culture. Incipient attention has been paid during the last decade to women flamenco guitar players (Cifredo, 2014; Recio Lineros, 2020), and the corporeity and roles of flamenco women in dance (Cruces Roldán, 2002, 2005, 2015) but much research remains to be completed in this field.

The third objective was to observe the differences in the speech and learning practice between flamenco guitarists selected as instrumental cases from the initial sample, with polarised approaches to learning. It was expected that observation would

[7] *Patada de baile* is a small choreography that involves body movement and footwork, which each person adapts, composes or improvises in their own way, so they are often valued for their creativity and flow as much as for their technical prowess. It usually has a duration of musical phrase according to the different rhythms or *palos*.

provide information on the activity or on the *Enactive* minds of the two apprentices, which could be differentiated according to their conception of learning and having an impact of this on their practice. This was another major factor for establishing such an in-depth qualitative study. Something in common is that both apprentices are highly motivated toward collective practice and classes with the teacher. However, solitary individual practice makes less sense and they are not very persistent in it. They show essentially hetero-regulated learning when they are working with the teacher, but very co-regulated learning when they are practising with peers, which is a huge part of their daily life, although they do not usually identify the latter as for learning but just for fun. Maybe because of that the discourse of both flamenco participants was usually focused on positive emotions, as found in learners of other cultures with a more constructive, or in this case transformative, profile (Casas-Mas et al., 2015b; López-Íñiguez and Pozo, 2016; López-Íñiguez and McPherson, 2020, 2021). We found there was a fair bit of similarity between these apprentices when observing the relationship between discourse and practice, in what they said they did and what they really did. In other words, the gap between explicit knowledge and more implicit or corporal knowledge may be lower than usual compared with learners of other cultures. Nevertheless, some differences between the two flamenco participants emerged during repetitive practice, when apprentice T was totally unfazed by having learned in a different way, "with mistakes" with regard to what the teacher had proposed, whilst R denied any difficulties or did not recognise any mistakes. We also have exemplified the perception of self-efficacy that apprentice T expresses when evaluating his own performance in the video recordings, coinciding with Zarza-Alzugaray et al. (2020).

The fourth objective refers to the use of symbolic or specific tools as *Extended* mind. With the flamenco community of oral tradition, it would not make sense to consider the mind extended through scores, or the theoretical mind (Donald, 1991), but rather the mind extended through gestures, or the mimetic mind (Goldin-Meadow, 2003) and the different levels of explicitness or abstraction that it implies. In the discourse on both flamenco apprentices, they emphasised playing by ear, although both appreciated the image and recordings in the video, mentioning how they greatly focused on the sound. Nevertheless, the crucial difference is that whilst apprentice R only mentioned analytical type elements or technical issues of the guitar in certain sections of the piece, T also mentioned contents of an interpretative-intuitive nature which alluded to a holistic type of comprehension, like dance or movements, or those referring to the emotion or the style and global nature of the piece, related to a multimodal explicit conception. The possibility of learning from video recording has an impact on the activation of learning processes has been observed in previous research (Boucher et al., 2020; Volioti and Williamon, 2021) and can contribute to important implications for music education in both formal and informal learning.

The gesture is always linked to the context of production; that is why these apprentices when interviewed sing music more than talk about it. This affects self-regulation processes using private speech or singing during learning, coinciding with that defended by McPherson and Gabrielsson (2002) regarding the primacy of auditory over the symbolic-notational image, these apprentices construct only the auditory image. However, we observed that they do not specify the melody, neither by singing, humming, or whistling that we saw in other cultures (Casas-Mas et al., 2019). We only detected very guttural sounds in apprentice R and whispering of counting beats as self-guidance in apprentice T. This suggests a profound internalisation of private speech and song, coinciding with that observed in other popular cultures, with less use of notational material. The body provides us with a lot of information about these learners of oral tradition and paves the way for future exploration in a field that continues generating interest in the area of educational psychology (Mulvihill et al., 2020).

Our main conclusion is, as Schiavio et al. (2021) were emphasising, that there is a need to examine within ecologically valid settings, how musical ability, e.g., to correctly synchronise, emerges and flourishes in musically rich contexts like the flamenco community, where the rhythm and the beat may take years to develop. The flamenco players who participated on this occasion were learning guitar with a teacher but prior to this, they had already learned music as a mother tongue since infancy, as described by Lave (2011), Rogoff (2012) in other domains. It is as if these flamenco apprentices needed the perception of a participating audience in order to practice music in a process of participatory sense-making (De Jaegher and Di Paolo, 2007; Krueger, 2014). In this realm singing and dance were not contradictory with playing an instrument but were even taught and lauded from the social environment, as an actual lived experience (Reybrouck, 2021). This is different from learning music as a second language, i.e., symbolising with musical signs, decoding the meaning of the musical score, and coding the material again into musical productions, just as the formation of musical language continues prioritising in the Western classical tradition and in contexts, such as the conservatory.

The limitations of this work are that the findings come from a small number of participants and therefore cannot be immediately generalised to other apprentices or cultures of learning. However, this method provides information on situated learning processes that are explicitly verbal first hand but are also implicitly informed due to its focus on embodied knowledge. Some regularities between learning cultures, and a detailed description that had not been previously made are the main contributions, that may open possibilities for future lines of research in other cultures or focusing on external representations from the 4E Cognition approach.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Autonomous University of Madrid. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin. Written informed consent was obtained from the individual(s), and minor(s)' legal guardian/next of kin, for the publication of any potentially identifiable images or data included in this article.

## AUTHOR CONTRIBUTIONS

AC-M organized the database and wrote the first draft of the manuscript. All authors performed the analysis, and contributed to the conception, design of the study, and manuscript revision, read, and approved the submitted version.

## FUNDING

This study was supported by the Universidad Complutense de Madrid and Universidad Autónoma de Madrid. Project Learning and instruction in specific domains: the role of the body and external representations, funded by the Spanish Ministry of Science and Innovation (PID2020-114177RB-I00).

## ACKNOWLEDGMENTS

We acknowledge the selfless participation of all the students who cooperated by giving the interviews and the teachers who allowed us into their classrooms, especially the teachers "El Entri" and "El Viejín" with the flamenco participants, and Escuela Superior de Canto de Madrid and Real Conservatorio Superior de Música with the jazz and classical participants.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.733615/full#supplementary-material

## REFERENCES


Abajo, J. E., and Carrasco, S. (2004). *Experiencias y Trayectorias de Éxito Escolar de Gitanas y Gitanos en España[Experiences and Trajectories of School Success of Roma Women and Men in Spain].Colección Mujeres en la Educación, 4*. Madrid: CIDE Instituto de la Mujer.

Alarcón-Rubio, D., Sánchez-Medina, J. A., and Winsler, A. (2013). Private speech in illiterate adults: Cognitive functions, task difficulty, and literacy. *J. Adult Dev.* 20, 100–111. doi: 10.1007/s10804-013-9161-y

Amat, J. C. (1758). *Guitarra Española, y Vandola, en dos Maneras, Castellana, y Valenciana, de cinco Ordenes, la qual enseña a templar, y tañer rasgado, todos los puntos naturales, y b, mollados, con estilo maravilloso [Spanish guitar, and Vandola]*. València: Agustín Laborda.

Bain, J. D., Mills, C., Ballantyne, R., and Packer, J. (2002). Developing reflection on practice through journal writing: impacts of variations in the focus and level of feedback. *Teach. Teach.* 8, 171–196. doi: 10.1080/13540600220127368

Baño, L. F. (2018). *Culturas Musicales. Análisis de la influencia de los modos de producciónmusical en prácticas grupales. Hacia la Música Ecológica [Musical cultures. Analysis of the influence of musical production modes in group practices. Towards Ecological Music]. Doctoral dissertation*. Madrid: Universidad Autónoma de Madrid.

Bautista, A., Pérez Echeverría, M. D. P., and Ignacio Pozo, J. (2010). Music performance teachers' conceptions about learning and instruction: a descriptive study of Spanish piano teachers. *Psychol. Music* 38, 85–106. doi: 10.1177/0305735609336059

Benedict, C., Schmidt, P., Spruce, G., and Woodford, P. (eds). (2015). *The Oxford Handbook of Social Justice in Music Education*. Oxford: Oxford University Press.

Besada, J. L., Barthel-Calvet, A. S., and Pagán Cánovas, C. (2021). Gearing time toward musical creativity: conceptual integration and material anchoring in Xenakis' Psappha. *Front. Psychol.* 11:611316. doi: 10.3389/fpsyg.2020.611316

Boucher, M., Creech, A., and Dubé, F. (2020). Video feedback and the choice of strategies of college-level guitarists during individual practice. *Musicae Sci.* 24, 430–448.

Bowman, W. (2004). "Cognition and the body: perspectives from music education," in *Knowing Bodies, Moving Minds: Towards Embodied Teaching and Learning*, ed. L. Bresler (Dordrecht: Kluwer Academic Press), 29–50.

Buehl, M. M., and Beck, J. S. (2015). "The relationship between teachers' beliefs and teachers' practices," in *International Handbook of Research on Teachers' Beliefs*, eds H. Fives and M. G. Gill (New York, NY: Routledge), 66–84.

Burland, K. (2020). Music for all: identifying, challenging and overcoming barriers. *Music Sci.* 3:205920432094695. doi: 10.1177/2059204320946950

Casas-Mas, A. (2013). *Musical Learning Cultures: Learning Conceptions, Processes and Practices in Classical, Flamenco and Jazz*. Doctoral dissertation. Madrid: Universidad Autónoma de Madrid.

Casas-Mas, A. (2018). "Developing an approach to the flamenco learning-teaching culture: an innovative (Traditional) learning," in *Traditional Musics in the Modern World: Transmission, Evolution, and Challenges*, ed. B. W. Leung (Cham: Springer), 25–39.

Casas-Mas, A. (2020). "The connection of the verbal and bodily mind," in *Proceedings of the 23rd International Seminar of the ISME Music in Schools and Teacher Education Commission (MISTEC)*, eds S. A. Ruthmann and M.-L. Bowe (Nedlands, WA: International Society for Music Education), 175–188.

Casas-Mas, A. (2022). "Learning outside the music classroom: from informal to formal learning as musical learning cultures," in *Learning and Teaching in the Music Studio. A Student-Centred Approach. Landscapes: Arts, Aesthetics*, eds J. I. Pozo, M. P. Pérez-Echeverría, G. López-Íñiguez, and J. A. Torrado (Singapore: Springer), 327–343.

Casas-Mas, A., López-Íñiguez, G., Pozo, J. I., and Montero, I. (2019). The function of private singing in instrumental music learning: a multiple case study of self-regulation and embodiment. *Music. Sci.* 23, 442–464. doi: 10.1177/1029864918759593

Casas-Mas, A., Montero, I., and Pozo, J. I. (2015b). Discourse on practice of a semi-professional jazz guitarist: a case study of constructive musical learning. *IASPM J.* 5, 54–80. doi: 10.5429/ij.v5i1.722

Casas, A., and Pozo, J. I. (2008). ¿Cómo se utilizan las partituras en la enseñanza y el aprendizaje de la música? [How are musical scores used in teaching and learning music?]. *Cult. Educ.* 20, 49–62. doi: 10.1174/113564008783781503

Casas-Mas, A., Pozo, J. I., and Montero, I. (2014). The influence of music learning cultures on the construction of teaching-learning conceptions. *Br. J. Music Educ.* 31, 319–342. doi: 10.1017/S0265051714000096

Casas-Mas, A., Pozo, J. I., and Scheuer, N. (2015a). Musical learning conceptions as sociocultural productions. Lexicometrical analysis of discourse in classical, flamenco and jazz cultures. *J. Cross Cult. Psychol.* 46, 1191–1225. doi: 10.1177/0022022115603124

Cain, M. (2015). Musics of 'The Other': creating musical identities and overcoming cultural boundaries in Australian music education. *Br. J. Music Educ.* 32, 71–86. doi: 10.1017/S0265051714000394

Calahorro, M. Á (2019). La legislación en torno a la guitarra flamenca: recorrido diacrónico y propuesta de mejora con base en su naturaleza popular. *Rev. Centro Invest. Flamenco Telethusa* 12, 5–17.

Cantwell, R. H., and Millard, Y. (1994). The relationship between approach to learning and learning strategies in learning music. *Br. J. Educ. Psychol.* 64, 45–63. doi: 10.1111/j.2044-8279.1994.tb01084.x

Carreras, J., Orozco, D. L., and Solé, E. J. (1996). *Tecnopatías del músico: Prevención y Tratamiento de las Lesiones y Enfermedades Profesionales de Instrumentistas y Cantantes; Introducción a la Medicina de la Danza [Technopathies of the musician: Prevention and Treatment of Injuries and Occupational Diseases of Instrumentalists and Singers; Introduction to Dance Medicine]*. Barcelona: Aritza.

Chaffin, R., Imreh, G., Lemieux, A. F., and Chen, C. (2003). "Seeing the big picture": piano practice as expert problem solving. *Music Percept.* 20, 465–490. doi: 10.1525/mp.2003.20.4.465

Cifredo, A. (2014). *Tocaoras [Female Guitarists Performers]*. Available online at: https://www.filmin.es/pelicula/tocaoras (accesed March 2, 2022).

Clarà, M. (2019). The theory-practice problem in teacher education models: a psychological perspective. *Estud. Pedagogicos* 45, 179–195. doi: 10.4067/S0718-07052019000200179

Clarà, M., and Mauri, T. (2010). A discussion about practical knowledge and its relationship to theoretical knowledge and practice. *Infanc. Aprendizaje* 33, 199–207. doi: 10.1174/021037010791114643

Clark, A., and Chalmers, D. (1998). The extended mind. *Analyses* 58, 7–19. doi: 10.1093/analys/58.1.7

Claxton, G. (2015). *Intelligence in the Flesh*. New Haven, CT: Yale University Press.

Cross, I. (2003). "Music and biocultural evolution," in *The Cultural Study of Music: A Critical Introduction*, eds M. Clayton, T. Herbert, and R. Middleton (New York, NY: Routledge), 19–30.

Cruces Roldán, C. (2002). "De cintura para arriba. Hipercorporeidad y sexuación en el flamenco [From waist to top. Hyper-corporeality and sexuation in flamenco]," in *Paper Presented at the Entretejiendo saberes: Actas del IV Seminario de la Asociación Universitaria de Estudios de Mujeres, 1-2 [Paper presentation]*, Sevilla.

Cruces Roldán, C. (2005). *Mujeres Flamencas, Etnicidad, Educación y Empleo ante los Nuevos Retos Profesionales [Roma Women, Ethnicity, Education and Employment in the Face of New Professional Challenges].Ministerio de Trabajo y Asuntos Sociales*. Sevilla: Instituto de la Mujer.

Cruces Roldán, C. (2015). Género y bailes. Geografías corporales en los orígenes del flamenco [Gender and dances. Body geographies in the origins of flamenco]. *Cult. Cienc. Deporte* 10, 75–86.

Damasio, A. (1994). *Descartes's Error. Emotion, Reason and the Human Brain*. New York, NY: Avon Books.

De Jaegher, H., and Di Paolo, E. (2007). Participatory sense-making: an enactive approach to social cognition. *Phenomenol. Cogn. Sci.* 6, 485–507. doi: 10.1007/s11097-007-9076-9

Donald, M. (1991). *Origins of the Modern Mind. Three Stages in the Evolution of Culture and Cognition*. Cambridge, MA: Harvard University Press.

Duncan, R. M., and Cheyne, J. A. (2001). Private speech in young adults: task difficulty, self-regulation, and psychological predication. *Cogn. Dev.* 16, 889–906. doi: 10.1016/S0885-2014(01)00069-7

European Unión Agency of Fundamental Rights [FRA] (2018). *Segunda Encuesta de la Unión Europea Sobre las Minorías y la Discriminación (EU-MIDIS II 2016). La Población Romaní: Resultados Principales*. Vienna: European Unión Agency of Fundamental Rights.

Folkestad, G. (2006). Formal and informal learning situations or practices vs. formal and informal ways of learning. *Br. J. Music Educ.* 23, 135–145.

García Fernández, J. A., Moreno Herrero, I., García Medina, R., Salguero Juan y Seva, J. M., Casas-Mas, A., García Serrano, M., et al. (2017). *La Cultura del Pueblo Gitano en el Currículo de la Educación Obligatoria, a Través de su Presencia, Ausencia y Percepción en los libros de texto [The Culture of the Gypsy People in the Compulsory Education Curriculum, Through its Presence, Absence and Perception in Textbook]*. Madrid: Spanish Ministry of Education, Culture and Sport. Available online at: https://sede.educacion.gob.es/publiventa/la-cultura-del-pueblo-gitano-en-el-curriculo-de-la-educacion-obligatoria-a-traves-de-su-presencia-ausencia-y-percepcion-en-los-libros-de-texto/cultura-gitanos/22079

Gibbs, R. W. (2017). "The embodied and discourse views of metaphor: why these are not so different and how they can be brought closer together," in *Metaphor: Embodied Cognition and Discourse*, ed. B. Hampe (Cambridge: Cambridge University Press), 319–334. doi: 10.1017/9781108182324.018

Gioia, T. (1997). *The History of Jazz*. Oxford: Oxford University Press.

Goldin-Meadow, S. (2003). *Hearing Gesture: How Our Hands Help Us Think*. Cambridge, MA: Harvard University Press.

Government Area of Economy and Employment (2012). *Caño Roto*. Madrid: Madrid city council.

Green, L. (2002). *How Popular Musicians Learn: A Way Ahead for Music Education*. London: Ashgate Press.

Green, L. (2008). *Music, Informal Learning and the School: A New Classroom Pedagogy*. London: Ashgate.

Hallam, S. (1995). Professional musicians' orientations to practice: implications for tea-ching. *Br. J. Music Educ.* 12, 3–19. doi: 10.1017/S02650517Year2357

Hickey, M. (2015). Learning from the experts: a study of free-improvisation pedagogues in university settings. *J. Res. Music Educ.* 62, 425–445.

Høffding, S., and Schiavio, A. (2021). Exploratory expertise and the dual intentionality of music-making. *Phenomenol. Cogn. Sci.* 20, 811–829. doi: 10.1007/s11097-019-09626-5

Holguín, P., and Shifres, F. (2015). Escuchar música al sur del Río Bravo: desarrollo y formación del oído musical desde una perspectiva latinoamericana [Listening to music in the south of Rio Grande: development and training of the musical ear from a Latin American perspective]. *Calle 14 Rev. Invest. Campo Arte* 10, 40–53. doi: 10.14483/udistrital.jour.c14.2015.1.a04

Hultberg, C. (2002). Approaches to music notation: the printed score as a mediator of meaning in Western tonal tradition. *Music Educ. Res.* 4, 185–197. doi: 10.1080/1461380022Year11902

Jefatura del Estado (1990). *Ley Orgánica General del Sistema Educativo 1/1990 of October 3rd [General Organic Law of the Educational System]*. Madrid: Spanish Government.

Jefatura del Estado (1995). *Royal Decree 617/1995 of April 21st, por el que se Establecen los Aspectos Básicos del Currfculo del Grado Superior de las Enseñanzas de Música y se Regula la Prueba de Acceso a estos estudios [The Basic Aspects of the Curriculum of the Tertiary Degree of Music Education and the Regulation of the Access to These Studies]*. Madrid: Spanish Government.

Jefatura del Estado (2006). *Ley Orgánica de Educación 2/2006, of May 3rd. [Organic Law of Education]*. Madrid: Spanish Government.

Junta de Andalucía (2003). *La Escuela Bolera [Bolera School]*. Available online at: https://www.juntadeandalucia.es/cultura/flamenco/content/expediente-t%C3%A9cnico-sobre-la-escuela-bolera (accessed March 30, 2022).

Kagan, S. (1989). The structural approach to cooperative learning. *Educ. Leadersh.* 47, 12–15.

Karlsen, S. (2012). Multiple repertoires of ways of being and acting in music: immigrant students' musical agency as an impetus for democracy. *Music Educ. Res.* 14, 131–148. doi: 10.1080/14613808.2012.685460

Krueger, J. (2014). Affordances and the musically extended mind. *Front. Psychol.* 4:1003. doi: 10.3389/fpsyg.2013.01003

Lantolf, J. P., and Thorne, S. L. (2006). "Sociocultural theory and second language acquisition," in *Theories in Second Language Acquisition*, eds B. van Patten and J. Williams (Mahwah, NJ: Erlbaum), 201–224.

Lave, J. (2011). *Apprenticeship in Critical Ethnographic Practice*. Chicago, IL: The University of Chicago Press.

Leung, B.-W. (Ed.) (2018). *Traditional Musics in the Modern World: Transmission, Evolution, and Challenges, Landscapes: The Arts, Aesthetics, and Education 24*. Cham: Springer International Publishing AG. doi: 10.1007/978-3-319-91599-9_3

Livingstone, R. S., and Thompson, W. F. (2009). The emergence of music from the theory of mind. *Musicae Sci.* 13, 83–115. doi: 10.1177/1029864909013002061

López-Íñiguez, G., and McPherson, G. E. (2020). Applying self-regulated learning and self-determination theory to optimize the performance of a concert cellist. *Front. Psychol.* 11:385. doi: 10.3389/fpsyg.2020.00385

López-Íñiguez, G., and McPherson, G. E. (2021). Regulation of emotions to optimize classical music performance: a quasi-experimental study of a cellist-researcher. *Front. psychol.* 12:789. doi: 10.3389/fpsyg.2021.627601

López-Íñiguez, G., and Pozo, J. I. (2014). Like teacher, like student? Conceptions of children from traditional and constructive teachers regarding the teaching and learning of string instruments. *Cogn. Instr.* 32, 219–252. doi: 10.1080/07370008.2014.918132

López-Íñiguez, G., and Pozo, J. I. (2016). Analysis of constructive practice in instrumental music education: case study with an expert cello teacher. *Teach. Teach. Educ.* 60, 97–107. doi: 10.1016/j.tate.2016.08.002

Marín, C., Pérez Echeverría, M. P., and Hallam, S. (2012). Using the musical score to per-form: a study with Spanish flute students. *Br. J. Music Educ.* 29, 193–212. doi: 10.1017/S0265051712Year46

Martí, E., and Pozo, J. I. (2000). Beyond mental representations: the acquisition of external systems of representation. *Infanc. Aprendizaje* 90, 11–30. doi: 10.1174/021037000760087946

Maturana, H., and Varela, F. (1980). *Autopoiesis and Cognition: The Realization of the Living*. Boston, MA: Reidel.

McPherson, G. E., and Gabrielsson, A. (2002). "From sound to sign," in *The Science and Psychology of Musical Performance: Creative Strategies for Music Teaching and Learning*, eds R. Parncutt and G. E. McPherson (Oxford: Oxford University Press), 99–115.

Mead, D. L., and Winsler, A. (2019). Change over time in the type and functions of crib speech around the fourth birthday. *Lang. Commun.* 67, 29–44. doi: 10.1016/j.langcom.2018.12.001

Middleton, R. (1981). "Popular music, I, 4: Europe & North America: genre, form, style," in *Popular Music. Folk or Popular? Distinctions, Influences, Continuities*, eds R. Middleton and D. Horn (Cambridge: Cambridge University Press).

Ministry of Education and Professional Training (2019). *Encuesta de Población Activa y Abandono Escolar [Survey of the active population and school dropout]*. Available online at: http://www.educacionyfp.gob.es/prensa/actualidad/2019/01/20190129-abandonoescolar.html (accessed May 05, 2021).

Mithen, S. (2005). *The Singing Neanderthals: The Origins of Music. Language, Mind and Body*. London: Weidenfeld and Nicolson.

Mulvihill, A., Carroll, A., Dux, P. E., and Matthews, N. (2020). Self-directed speech and self-regulation in childhood neurodevelopmental disorders: current findings and future directions. *Dev. Psychopathol.* 32, 205–217. doi: 10.1017/S0954579418001670

Murphy, D. (1994). Jazz studies in American schools and colleges: a brief history. *Jazz Educ. J.* 26, 34–38.

Nagy, Z. (2017). *Embodiment of Musical Creativity: The Cognitive and Performative Causality of Musical Composition*. Abingdon: Routledge.

Newen, A., Gallagher, S., and De Bruin, L. (2018). "4E cognition," in *The Oxford Handbook of 4E Cognition*, eds A. Newen, L. de Bruin, and S. Gallagher (Oxford: Oxford University Press). doi: 10.1093/oxfordhb/9780198735410.013.1

Nielsen, S. (1999). Regulation of learning strategies during practice: a case study of a single church organ student preparing a particular work for a concert performance. *Psychol. Music* 27, 218–229. doi: 10.1177/0305735699272015

Nijs, L. (2017). "The merging of musician and musical instrument: incorporation, presence and the levels of embodiment," in *The Routledge Companion to Embodied Music Interaction*, eds M. Lesaffre, P. J. Maes, and M. Leman (New York, NY: Routledge), 49–57.

Nijs, L., and Bremmer, M. (2019). "Embodiment in early childhood music education," in *Music in Early Childhood: Multi-disciplinary Perspectives and Interdisciplinary Exchanges, International Perspectives on Early Childhood Education and Development*, Vol. 27, eds S. Young and B. Ilari (Cham: Springer Nature), 87–102. doi: 10.1007/978-3-030-17791-1

Noë, A. (2004). *Action in Perception*. Cambridge, MA: MIT Press.

O'Neill, S. A. (2009). Revisioning musical understandings through a cultural diversity theory of difference. *Explor. Soc. Justice* 4, 70–89.

Olson, D. R., and Bruner, J. S. (1996). "Folk psychology and folk pedagogy," in *Handbook of Education and Human Development: New Models of Learning, Teaching and Schooling*, eds D. R. Olson and N. Torrance (Malden, MA: Blackwell), 9–27.

Pérez-Echeverría, M. P., and Scheuer, N. (2009). "Introduction: externally supported symbolic practices as learning tools," in *Representational Systems and Practices as Learning Tools in Different Fields of Knowledge*, eds C. Andersen, N. Scheuer, M. P. Pérez Echeverría, and E. Teubal (Rotterdam: Sense), 1–18.

Pozo, J. I., Pérez-Echeverría, M. P., López-Íñiguez, G., and Casas-Mas, A. (2022b). "A system for the analysis of music teaching and learning practices," in *Learning and Teaching in the Music Studio. A Student-Centred Approach. Landscapes: Arts, Aesthetics*, eds J. I. Pozo, M. P. Pérez-Echeverría, G. López-Íñiguez, and J. A. Torrado (Singapore: Springer), 141–179. doi: 10.1111/bjep.12026

Pozo, J. I, Scheuer, N., Pérez Echeverría, M. P., Mateos, M., Martín, E., and de la Cruz, M. (Eds). (2006). *Nuevas Formas de Pensar la Enseñanza y el Aprendizaje: las Concepciones de Profesores y Alumnos [New Ways of Thinking about Teaching and Learning: The Conceptions of Teachers and Students]*. Barcelona: Graó.

Pozo, J. I., Torrado, J. A., and Baño, L. (2022a). "La enseñanza de la música instrumental: viejas tradiciones, nuevos enfoques [Teaching instrumental music: old traditions, new approaches]," in *Learning and Teaching in the Music Studio. A Student-Centred Approach. Landscapes: Arts, Aesthetics*, eds J. I. Pozo, M. P. Pérez-Echeverría, G. López-Íñiguez, and J. A. Torrado (Singapore: Springer), 21–46.

Pozo, J. I., Torrado, J. A., and Pérez-Echeverría, M. P. (2020b). Aprendiendo a interpretar música por medio del smartphone: la explicitación y reconstrucción de las representaciones encarnadas [Learning to play music through the smartphone: the explanation and reconstruction of embodied representations]. *Cuad. Cent. Estud. Diseño Comun. Ensayos* 89, 236–260.

Pozo, J. I. (2017). Learning beyond the body: from embodied representations to explicitation mediated by external representations. *Infanc. Aprendizaje* 40, 219–276. doi: 10.1080/02103702.2017.1306942

Recio Lineros, F. (2020). *La Dona en la Guitarra Flamenca: Paper i Visibilitat [The Woman in the Flamenco Guitar: Paper and Visibility]*. Final Degree Project. Girona: Universidad de Girona.

Reybrouck, M. (2021). *Musical Sense-Making: Enaction, Experience and Computation*. London: Routledge.

Rogoff, B. (2012). Learning without lessons: Opportunities to expand knowledge. *Infanc. Aprendizaje* 35, 233–252. doi: 10.1174/021037012800217970

Roma Secretariat Foundation (FSG) (2013). *El Alumnado Gitano en Secundaria: un Estudio Comparado [Roma Students in Secondary School: A Comparative Study]*. Madrid: Ministerio de Educación, Cultura y Deporte.

Roma Secretariat Foundation (FSG) (2019). *Estudio Comparado Sobre la Situación de la Población Gitana en España en Relación al Empleo y la Pobreza 2018 [Comparative Study on the Situation of the Roma Population in Spain in Relation to Employment and Poverty 2018]*. Madrid: Fundación Secretariado Gitano.

Rowlands, M. (2010). *The New Science of the Mind*. Cambridge, MA: The MIT Press.

Scardamalia, C., and Bereiter, M. (1987). *The Psychology of Written Composition*. Hillsdale, NJ: Lawrence Erlbaum.

Schiavio, A., and van der Schyff, D. (2018). 4E music pedagogy and the principles of self-organization. *Behav. Sci.* 8, 1–15. doi: 10.3390/bs8080072

Schiavio, A., Moran, N., van der Schyff, D., Biasutti, M., and Parncutt, R. (2020). Processes and experiences of creative cognition in seven western classical composers. *Musicae Sci.* 1–23. doi: 10.1177/1029864920943931

Schiavio, A., Stupacher, J., Xypolitaki, E., Parncutt, R., and Timmers, R. (2021). Musical novices perform with equal accuracy when learning to drum alone or with a peer. *Sci. Rep.* 11:12422. doi: 10.1038/s41598-021-91820-0

Schiavio, A., van der Schyff, D., Kruse-Weber, S., and Timmers, R. (2017). When the sound becomes the goal. 4E cognition and teleomusicality in early infancy. *Front. Psychol.* 8:1585. doi: 10.3389/fpsyg.2017.01585

Sexton, F. (2012). Practitioner challenges working with informal learning pedagogies. *Br. J. Music Educ.* 29, 7–11. doi: 10.1017/S0265051711000507

Shah, S. M. (2006). Popular music in Malaysia: education from the outsider. *Int. J. Music Educ.* 24, 132–139.

Sneeuw, A. C. (1991). *El Flamenco Descrito en 1850 por François Gevaert [Flamenco described in1850 by François Gevaert]. Candil, Revista de Flamenco, 74*. Jaén: Peña flamenca de Jaén.

Sosa, E. (2009). La otredad: una visión del pensamiento latinoamericano contemporáneo [Otherness: a vision of contemporary latin american thought]. *Letras* 51, 349–372.

Stake, R. E. (1998). *Investigación con Estudio de Casos [Case Study Research]*. Madrid: Ediciones Morata.

Steingress, G. (2004). "El flamenco como arte popular y popularizado," in *Introducción al Flamenco en el Currículum Escolar*, ed. M. López (Madrid: Akal), 181–199.

Tagg, P. (1982). Analysing popular music: theory, method and practice. *Popular Music* 2, 37–65.

Tarr, B., Launay, J., and Dunbar, R. I. (2014). Music and social bonding: "self-other" merging and neurohormonal mechanisms. *Front. Psychol.* 5:1096. doi: 10.3389/fpsyg.2014.01096

Thibodeaux, J., Bock, A., Hutchison, L. A., and Winsler, A. (2019). Singing to the self: children's private speech, private song, and executive functioning. *Cogn. Dev.* 50, 130–141. doi: 10.1016/j.cogdev.2019.04.005

Torrado, J. A. (2006). "Del dicho al hecho: de las concepciones sobre el aprendizaje a la práctica de la enseñanza de la música [From saying to doing: from the conceptions about learning to the practice of teaching music]," in *Nuevas Formas de Pensar la Enseñanza y el Aprendizaje: Las Concepciones de Profesores y Alumnos. Crítica y Fundamentos*, Vol. 12, eds J. I. Pozo, N. Scheuer, M. P.

Pérez-Echeverría, M. Mateos, E. Martín, and M. de La Cruz (Barcelona: Graò), 205–230.

Torrado, J. A., and Mengual, L. F. B. (2020). "La enseñanza de la música instrumental: viejas tradiciones, nuevos enfoques [Teaching instrumental music: old traditions, new approaches]," in *Aprender y Enseñar Música. Un Enfoque Centrado en los Alumnos*, eds J. I. Pozo, M. P. Pérez-Echeverría, J. A. Torrado, and G. López-Íñiguez (Madrid: Ediciones Morata), 43–68.

Torres, N. (2012). El estilo "rasgueado" de la guitarra barroca y su influencia en la guitarra flamenca: fuentes escritas y musicales (siglo XVII). *Rev. Invest. sobre Flamenco "La Madrugá"* 6, 1–46.

Trilla, J. (1997). *Relaciones Entre la Educación Formal, la no Formal y la Informal. En La Educación Fuera de la Escuela [Relations Between Formal, Non-Formal and Informal Education]*. México: Ariel, 187–196.

Úcar, X. (2021). Theoretical and practical knowledge in social pedagogy: levels and agents of the pedagogical knowledge production. *Oxf. Rev. Educ.* 1–19. doi: 10.1080/03054985.2021.2013794

United Nations Educational, Scientific and Cultural Organization [UNESCO] (2010). *Convention for the Safeguarding of the Intangible Cultural Heritage. Nomination file no. 00363*. Paris: United Nations Educational, Scientific and Cultural Organization (UNESCO).

Varela, F., Thompson, E., and Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge: MIT Press.

Volioti, G., and Williamon, A. (2021). Performers' discourses on listening to recordings. *Res. Stud. Music Educ.* 43, 481–497. doi: 10.1177/1321103X19899168

Vygotsky, L. S. (1933/1978). *Mind in Society. The Development of Higher Psychological Processes*, eds M. Cole, V. John-Steiner, S. Scribner, and E. Souberman (Cambridge, MA: Harvard University Press).

Williamon, A., Valentine, E., and Valentine, J. (2002). Shifting the focus of attention between levels of musical structure. *Eur. J. Cogn. Psychol.* 14, 493–520. doi: 10.1080/09541440143000221

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

Winsler, A., Naglieri, J., and Manfra, L. (2006). Children's search strategies and accompanying verbal and motor strategic behavior: developmental trends and relations with task performance among children age 5 to 17. *Cogn. Dev.* 21, 232–248. doi: 10.1016/j.cogdev.2006.03.003

Zarza-Alzugaray, F. J., Casanova, O., McPherson, G. E., and Orejudo, S. (2020). Musicself-efficacyfor performance: anexplanatorymodelbasedon social support. *Front. Psychol.* 11:1249. doi: 10.3389/fpsyg.2020.01249

Zbikowski, L. M. (2015). Des herzraums abschied: Mark Johnson's theory of embodied knowledge and music theory. *Theory Pract.* 22-23, 1165–1170.

**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.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

# 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:** frontiersin.org/about/contact


**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