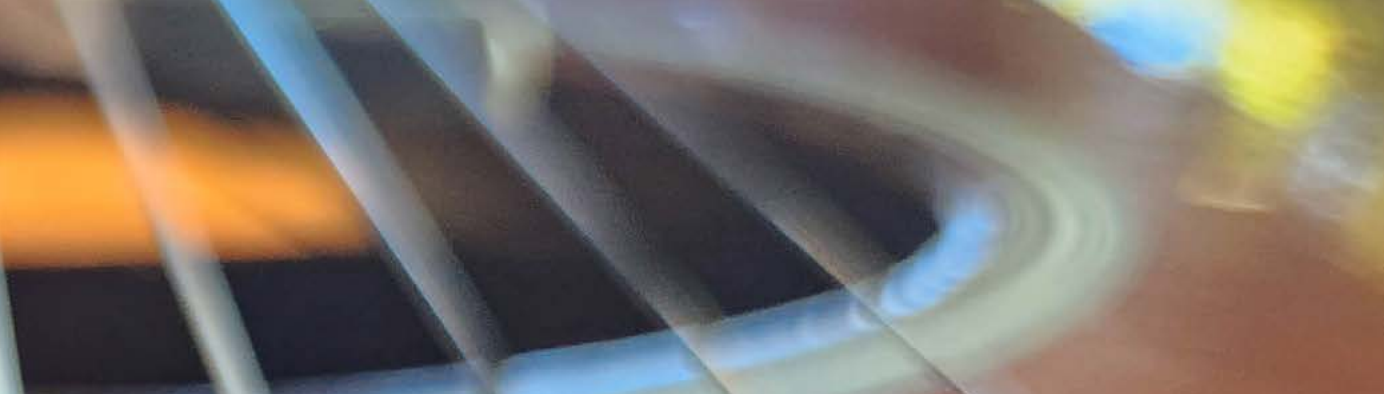




IntechOpen

Systematic Musicology

Current Topics and Interdisciplinary
Approaches

Edited by Christiane Neuhaus



Systematic Musicology
– Current Topics and
Interdisciplinary
Approaches

Edited by Christiane Neuhaus

Published in London, United Kingdom

Systematic Musicology – Current Topics and Interdisciplinary Approaches

<http://dx.doi.org/10.5772/intechopen.1005932>

Edited by Christiane Neuhaus

Contributors

Abirlal Gangopadhyay, Adrián García Riber, Christiane Neuhaus, Clemens Maidhof, Eliana de Toledo, Filippo Pasqualitto, Ildar Khannanov, Jörg Fachner, Keith Oatley, Mateus Henrique de Oliveira, Prithvi Ravi Kantan, Tim Ziemer

© The Editor(s) and the Author(s) 2026

The rights of the editor(s) and the author(s) have been asserted in accordance with the Copyright, Designs and Patents Act 1988. All rights to the book as a whole are reserved by INTECHOPEN LIMITED. The book as a whole (compilation) cannot be reproduced, distributed or used for commercial or non-commercial purposes without INTECHOPEN LIMITED's written permission. Enquiries concerning the use of the book should be directed to INTECHOPEN LIMITED rights and permissions department (permissions@intechopen.com).

Violations are liable to prosecution under the governing Copyright Law.



Individual chapters of this publication are distributed under the terms of the Creative Commons Attribution 4.0 License which permits commercial use, distribution and reproduction of the individual chapters, provided the original author(s) and source publication are appropriately acknowledged. If so indicated, certain images may not be included under the Creative Commons license. In such cases users will need to obtain permission from the license holder to reproduce the material. More details and guidelines concerning content reuse and adaptation can be found at <http://www.intechopen.com/copyright-policy.html>.

Notice

Statements and opinions expressed in the chapters are these of the individual contributors and not necessarily those of the editors or publisher. No responsibility is accepted for the accuracy of information contained in the published chapters. The publisher assumes no responsibility for any damage or injury to persons or property arising out of the use of any materials, instructions, methods or ideas contained in the book.

First published in London, United Kingdom, 2026 by IntechOpen

IntechOpen is the global imprint of INTECHOPEN LIMITED, registered in England and Wales, registration number: 11086078, 167-169 Great Portland Street, London, W1W 5PF, United Kingdom

For EU product safety concerns: IN TECH d.o.o., Prolaz Marije Krucifikse Kozulić 3, 51000 Rijeka, Croatia, info@intechopen.com or visit our website at intechopen.com.

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

Systematic Musicology – Current Topics and Interdisciplinary Approaches

Edited by Christiane Neuhaus

p. cm.

Print ISBN 978-1-83635-683-7

Online ISBN 978-1-83635-682-0

eBook (PDF) ISBN 978-1-83635-684-4

If disposing of this product, please recycle the paper responsibly.

Meet the editor



Christiane Neuhaus is a Privatdozentin at the Institute of Systematic Musicology at the University of Hamburg, Germany. Her research focuses on the cognitive neuroscience of music. She gives interdisciplinary lectures on music-related issues, including aesthetics, memory, evolutionary, medical, and neuroscience aspects. She previously worked as a Post-doctoral Researcher at the Max Planck Institute for Human

Cognitive and Brain Sciences in Leipzig for 7 years. She has published research in international peer-reviewed journals and many book chapters. She holds a Habilitation degree. The idea of her Ph.D. thesis (about processing musical scales from a cross-cultural perspective) was awarded at the European ESCOM conference. She also has a strong background in classical music, having studied piano at the Hochschule für Musik Detmold. In 2014, she was invited as a guest professor at Tübingen University.

Contents

Preface	IX
Section 1	
Audio Technology	1
Chapter 1	3
Musical Data Sonification: Expanding the Boundaries of Data Representation <i>by Tim Ziemer, Adrián García Riber and Prithvi Ravi Kantan</i>	
Section 2	
Music Theory	31
Chapter 2	33
Cognitive Music Theory: The Current State of Research <i>by Ildar Khannanov</i>	
Section 3	
Music Psychology	65
Chapter 3	67
Music and Emotion: Theory, Evidence, and Comparison with Literary and Visual Art <i>by Keith Oatley</i>	
Chapter 4	79
Music and Emotions: Spectrum of Possibilities in Esthetics Sports <i>by Eliana de Toledo and Mateus Henrique de Oliveira</i>	
Section 4	
Music and Neuroscience	99
Chapter 5	101
Hyperscanning and Mirror Neurons: Musical Ensemble Playing from a Neuroscientific Perspective <i>by Christiane Neuhaus</i>	

Section 5	
Music Therapy	125
Chapter 6	127
Music Therapy in Classical Ayurveda: A Reconstruction from Ancient Sanskrit Sources	
<i>by Abirlal Gangopadhyay</i>	
Chapter 7	153
Music and Psychoactive Substances: Cultural Practices, Adolescent Development and Music Therapy in Addiction Treatment	
<i>by Jörg Fachner, Filippo Pasqualitto and Clemens Maidhof</i>	

Preface

This is a book about current topics in Systematic Musicology. The term is not very well known in everyday English. So, what does it mean – Systematic Musicology? It is research on music, about its acoustic foundations and its effects on psyche and body, about music's nature, its functions, and its healing potential. Unlike Historical Musicology, it is less concerned with unique musical masterworks, composers' biographies, and the history of music. Instead, it is interdisciplinary by its very nature.

Searching for answers to questions about how music is processed, experienced, and performed is also a (mainly) empirical endeavor for which a plethora of methods exist, borrowed from the natural, social, and life sciences. Today, research work is extremely fast-paced, and new methods and software applications are developed almost every year.

The present book comprises seven chapters. Each aims at presenting the current state of research. Despite all the progress, one has to keep in mind that many issues of the present time have their roots in antiquity, in ancient Greek and ancient Indian concepts. A prime example is Pythagoras' doctrine that all things are based on numbers (around 500 BCE). Indeed, numerical proportions are fundamental to music. They are its essence, the inherent characteristic of intervals, rhythms, and scales.

Furthermore, since Pythagoras was convinced that this doctrine could also be applied to the universe, the distances between celestial bodies, as well as their planetary speeds, were also calculated based on analogical reasoning. Thus, the beginnings of cross-disciplinary thinking reach back into antiquity. Today, dealing with the concept of number is a matter of course, a precondition for successful work in disciplines such as acoustics, music theory, and astronomy (and this 'arguing by analogy' is also still in use).

In parallel with such developments in ancient Greece, insights into expressed and induced emotions emerged in ancient India. An elaborated description of those *bhāvas* and *rasas* can be found in a well-known ancient Indian treatise called *Natyashastra* (around 300 BCE to 300 CE), thus suggesting that the performing arts, including music and dance, have an impact on listeners' souls. Starting from this concept, a second set of domains emerged, this time comprising disciplines such as music psychology, aesthetics, and performance studies.

It took centuries, however, before any systematization of all music-related groups was suggested. It was Guido Adler, an Austrian musicologist, who first distinguished between the historical and the systematic branches of musicology, as laid out in his famous journal article "The Scope, Methods, and Aims of Musicology" ("Umfang, Methode und Ziel der Musikwissenschaft"), published in 1885. In Adler's view,

Systematic Musicology should comprise three main disciplines: Music Theory, Music Aesthetics, and Music Pedagogy. After World War II, a different classification was suggested by the Austro-German (music-)psychologist Albert Wellek with Acoustics in the first place, followed by Music Psychology, Music Aesthetics/Music Philosophy, and the Sociology of Music.

The current book aims to present the most recent state of research in Systematic Musicology. It consists of seven chapters assigned to five sections: *Audio Technology*, *Music Theory*, *Music Psychology*, *Music and Neuroscience*, and *Music Therapy*.

While, in principle, following Wellek's classification, new section titles are aligned with current issues and trends. Of course, the limited number of chapters cannot give more than a short impression of what Systematic Musicology is about. However, every chapter has been carefully prepared, thanks to the authors who delved deeply into the respective subject matter.

Section 1 – *Audio Technology*, is represented by a chapter about Musical Data Sonification. In principle, sonification is a transformation process by which data, in this case taken from astrophysics and from human movement science, are turned into sound. The chapter shows how music (with its parameters) can be integrated into sonification designs to use sonification data more effectively in artistic and therapeutic contexts. The chapter is written by Tim Ziemer, Interim Professor for Systematic Musicology at the University of Hamburg (Germany), with research experience in spatial audio, sonification, and psychoacoustics (gained, e.g., at the National Institute of Informatics, Tokyo). The chapter is co-authored by Adrián García Riber, Associate Professor of Audio Equipment and Systems at Pompeu Fabra University, Barcelona (Spain). He is also an Associate Researcher in the Spanish Virtual Observatory Project. Another co-author is Prithvi Ravi Kantan, a PostDoc at Aalborg University, Copenhagen (Denmark), who specializes in the development of novel sonification technology for motor rehabilitation.

In Section 2, *Music Theory*, two fields, music theory and cognitive musicology, are brought together to form a new hybrid called Cognitive Music Theory. A method of trace observation is developed by which one can deduce from the score how listeners may perceive and experience a piece of music. This is exemplified by analyzing the modulating path of an Allemande by Johann Sebastian Bach. The chapter is written by Ildar Khannanov, Associate Professor of Music Theory at Johns Hopkins University, Baltimore (USA). He has authored numerous publications, including books, that reflect his scholarly interests in music theory, philosophy, and Russian music. He was educated at the Moscow State Tchaikovsky Conservatory (music theory, piano, composition) and at the University of California, Irvine (philosophy).

Section 3, *Music Psychology*, consists of two chapters about the relationship between Music and Emotion. Both are, in a way, complementary: Chapter 3 focuses on induced emotions (in composers, poets, and recipients), whereas Chapter 4 focuses on expressed emotions (as shown by performing artists during choreographic work).

In Chapter 3, the emotional experiences with music are compared with those related to the sister arts. It emphasizes that artwork is often founded on a strong feeling that

is transformed into music, a poem, or a painting to better understand it. The chapter is written by Keith Oatley, Professor Emeritus of Cognitive Psychology at the University of Toronto (Canada). He is the originator of the principal textbook, *Understanding Emotions*, and author of seven other books on psychology. Keith Oatley is also a prize-winning novelist.

Chapter 4 is about the role of music and emotions in artistic sports, such as figure skating, artistic swimming, and artistic gymnastics. The relationship between art and sport is discussed from the viewpoints of athletes, coaches, and judges, each time focusing on the expressive and aesthetic components. The chapter is written by Eliana de Toledo, Professor of Sports Science and Leader of the Laboratory of Research and Experiences in Gymnastics at the University of Campinas, Limeira, State of São Paulo (Brazil). She has worked as a coach (gymnastics), judge (gymnastics), and as a Gymnastics Federation member (until today). The chapter is co-authored by Professor Mateus Henrique de Oliveira, a member of the Laboratory for Research and Experiences in Gymnastics and the Gymnastics Research Group at the University of Campinas (Brazil).

Section 4, *Music and Neuroscience*, consists of a chapter on ensemble playing from the micro- and macro-level perspectives of neuroscience. Part one takes a closer look at the functioning of mirror neurons, which play a central role in quickly understanding others' movements, ensuring smooth interactions between musicians on stage. Part two discusses a new method called EEG hyperscanning, which allows simultaneous recording of brain activity from two or more players. The chapter is written by Christiane Neuhaus, the editor of this book. She is a Privatdozentin at the University of Hamburg (Germany) with profound expertise in music and cognitive neuroscience, gained, e.g., at the Max Planck Institute for Human Cognitive and Brain Sciences in Leipzig.

Section 5, *Music Therapy*, consists of two chapters that are unrelated to each other but that together show the range of topics in this field.

Chapter 6 is about music therapy in classical Ayurveda, reconstructed from ancient Sanskrit sources. This topic is, for the most part, unknown to the Western readership. The time period covered is 600 BCE to 1,500 CE, yet an outline of modern Indian music therapy is also given. The chapter is written by Abirlal Gangopadhyay, an Independent Researcher in Sanskrit Studies, West Bengal (India), with scholarly interests in Sanskrit Poetics, Music Therapy, and Ancient Indian Musicology. He is dedicated to research that reconstructs and validates ancient Indian textual methods. He holds a doctoral degree (highest honour) from the University of Hyderabad.

Chapter 7 is about music and psychoactive substances, including music therapy as a treatment for drug addiction. Its central part discusses the pros and cons of drug consumption as seen from anthropology, culture, developmental psychology, and neuroscience, thus contributing to a better understanding of substance use and abuse. In part two, the effectiveness of music therapy in addiction treatment is described in detail. Chapter 7 is written by Jörg Fachner, Professor of Music, Health and the Brain and Co-Director of the Cambridge Institute of Music Therapy Research at Anglia Ruskin University, Cambridge (United Kingdom). With over 25 years of

experience in music therapy, psychology, and medical research across Germany, Finland, and the UK, he specializes in research that bridges the social, medical, and musical domains. He is the author of two influential books: *Music and Altered States* (2006) and *Music Therapy and Addictions* (2010).

The chapter is co-authored by Filippo Pasqualitto, now a PostDoc at the University of Ferrara (Department of Neuroscience and Rehabilitation) and the Italian Institute of Technology (IIT). The second co-author is Clemens Maidhof, who is a senior research fellow at the Cambridge Institute of Music Therapy Research. His research focuses on music therapy and neuroscience, particularly the social, cognitive, and affective processes that underlie therapeutic relationships and musical interaction.

A special thanks to Mirna Papuga, the Publishing Process Manager of this book, and to Marijana Josipovic for their kind and immediate support during all phases of the publishing process.

Christiane Neuhaus
Institute of Systematic Musicology,
University of Hamburg,
Hamburg, Germany

Section 1

Audio Technology

Chapter 1

Musical Data Sonification: Expanding the Boundaries of Data Representation

Tim Ziemer, Adrián García Riber and Prithvi Ravi Kantan

Abstract

Music and sonification have significant overlaps in terms of tools, parameters, and concepts, but they are, in essence, different. The most striking distinction between the two is their objective: while sonification functions similarly to natural language (in terms of communication and knowledge representation), music tends to aim at emotional expression, social engagement and integration, entertainment, or enjoyment. However, these objectives are not mutually exclusive. In this chapter, we provide an overview of the relationship between music and sonification. We offer examples of how music can be used for data sonification and explore the extent to which sonification can serve as a musical element. Specific topics covered include how music-making can contribute to sonification design, the use of deep learning in generating accurate and engaging musical sonifications, and how music can serve as a useful medium to sonify human movement in artistic and rehabilitative applications. All in all, understanding the affordances and limitations of musical structure and sound is crucial to using them effectively within the context of sonification design across application domains.

Keywords: sonification, auditory display, musification, psychoacoustics, audio engineering, sound design

1. Introduction

Although music and sonification often share the same technologies, parameters, and techniques, they differ in fundamental intent. Sonification is primarily concerned with conveying information – transforming data into sound to support understanding or action [1]. Its goal is communicative, akin to language. Music, in contrast, is generally aimed at emotional expression, aesthetic experience, and social connection [2]. These aims are not mutually exclusive, but they shape the expectations we bring to each, and they impose different possibilities and limitations on design choices.

Increasingly, researchers and practitioners are exploring how musical structure can enhance the intelligibility, memorability, or appeal of sonified data, particularly in contexts where affective or embodied responses are desirable. For musical

sonification, musical aspects – such as rhythm, harmony, and timbre – are not primarily shaped to please the listener but to convey information. Conversely, for sonification-inspired music, sonified data or processes are being used as compositional material. Advances in machine learning have further enabled mappings between complex datasets and musical parameters in ways that are responsive, adaptive, and often aesthetically refined.

This duality raises conceptual and methodological challenges. If sonification borrows the tools of music, does it also carry its cultural and emotional baggage [3–5]? When music is generated from data, does it remain music, or does it become something else – an aestheticized form of data analysis [6, 7]? These are not merely academic questions; they shape how we evaluate meaning, usefulness, and impact in applications ranging from science communication to artistic performance [8, 9].

To address these issues, we begin by distinguishing between three domains in which sound operates: the physical (acoustical), the perceptual (auditory), and the musical. This framework allows us to more clearly identify when sound functions as information, when it becomes experience, and how it may move between these roles.

The subsequent sections will explore how deep learning techniques contribute to the design of musical sonification, how sonification and musification differ in functions and possibilities, how the logic of audio mixing parallels that of sonification design, and how movement sonification utilizing musical structure can support rehabilitative and artistic practices. All these topics are supported by *ad hoc* multimedia examples¹ that can be found at: [10.5281/zenodo.16581819](https://doi.org/10.5281/zenodo.16581819).

2. Boundaries of sonification and music

In this section, we introduce the background of sonification, adhering strictly to the *sound terminology for sonification* [10]: We distinguish the physical domain (acoustical) from the perceptual domain (auditory) and the musical domain, and name the respective discipline whenever there is a danger of confusion (e.g., musical pitch vs. perceived pitch).

2.1 Boundaries of music

Music is a cultural phenomenon that has physical, psychological, and social aspects. What music is and what music can be is illustrated in **Figure 1**. Abstractly speaking, music takes place in a personal bubble. Physically speaking, heard music requires acoustic waves that propagate from a sound source to the listener. Both heard and imagined music have a physical component in terms of the electricity of neural activity. These physical requirements of music are self-evident. Physiological boundaries already determine if a sound is audible at all and to what aspects of the auditory experience it can contribute. Then, psychological boundaries may hinder even more to what extent sound can be understood as music. Lastly, the experience of the individual determines what he or she actually understands as music. In **Figure 1**, the green-circled region is his or her music bubble. Naturally, the individual experience grows over a lifetime. Every individual's bubble size and region are rather unique. With the right experience, some psychological boundaries

⁴ Complementary materials (permanent): [10.5281/zenodo.16581819](https://doi.org/10.5281/zenodo.16581819).

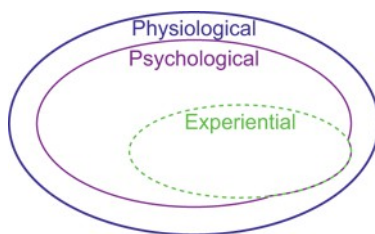


Figure 1.
Boundaries of music. There are hard physiological boundaries (blue), extendible experiential boundaries (green), and softer psychological boundaries (purple) that can expand through experience. The principle is valid for sonification too, except that the green region is comparably tiny.

can also be widened. Many musicians aim at widening their own and their listeners' bubbles. But this widening has to happen in small steps. When new music lacks overlap with our music bubble, we are not eager to learn to understand and love it. This is because we can hardly imagine that the reward is worth the time and effort.

The following details help to understand the concept of the three boundaries:

2.1.1 Physiology

Firstly, there are physiological boundaries. These have been investigated from the viewpoint of biology and psychoacoustics [11, Chapters 3 and 4]. Pressure fluctuations are potentially heard when they deflect the hair cell stereocilia in the organ of Corti in the inner ear by a certain amount with the right timing, which only happens with the right constellation of amplitude, frequency, and time. In psychoacoustic terms, what we perceive is sound with certain properties, such as aspects of pitch (like height, pitch strength, pitch changes, multiple pitches), timbre (like brightness, roughness, fullness), and audible space (like auditory event location and size, or imagined size and surface properties of its enclosing space).

2.1.2 Psychology

One example for psychological boundaries is the capacity of the working memory [11]. For a duration of 10 to 15 seconds, listeners can put sounds in relation to each other to form rhythms, melodies, phrases, voices, etc. Large-scale properties, such as the musical form or musical architecture, are perceived comparably vague. If the density of distinct acoustic events exceeds what working memory can handle, the resulting sound perception may be chaotic. Many more psychological constraints have been described phenomenologically, referred to as auditory scene analysis [12, Chapter 4]. For example, several frequencies with synced onsets may be grouped into one auditory stream and produce the perception of one musical tone, that is, an auditory event from a certain perceived location with a certain spatial extent, pitch, loudness, duration, and timbre. In other words, the frequencies are integrated vertically into one auditory stream. Milliseconds later, another set of frequencies may be integrated vertically, too. And if both vertically integrated streams do not deviate too much, for example, in terms of pitch, or timbre, the two may be integrated horizontally into a new auditory stream. The combination of vertical and horizontal integration is the main requirement of forming a melody

inside the listener's mind. But whether such a melody is musical for the individual or not is a matter of experience.

Demonstrations of how physiological and psychological boundaries affect your recognition and perception of music are provided in our [YouTube playlist] (<https://www.youtube.com/playlist?list=PLVv3BMS8IIXG8ahjhd7NUztfPeFLN2nBYouTu> be playlist²).

2.1.3 Experience

The experiential boundary can be considered our *music bubble*. Naturally, experience with music grows over the course of one's life, so our bubble grows. There is certainly some overlap between every personal music bubble – music that everybody understands as music. Other music may lie outside one's own music bubble. For example, listeners may initially perceive electroacoustic music [12, Chapter 2] as random noise. With some more experience, and maybe some specific education, the same listener may understand it as organized sound. And with even more experience, this organized sound may finally be understood as music, capable of evoking expectations, emotions, identification, and enjoyment.

Moreover, experience can enhance perceptual abilities. For example, violin players may develop a finer pitch resolution than others, simply as a matter of training. Another example is listening to scales from foreign cultures; recognizing melodies is easier if you are familiar with the tone system. It also helps you feel the intended emotion. New musical experiences can widen your psychological capabilities. You can learn to understand something as music that you had not recognized as music before. Musical experience can improve your sound discrimination skills, and it can change the way music affects you emotionally.

2.2 Boundaries of sonification

The boundaries of music mostly apply to sonification, too. However, there are two striking differences: the experience bubble and the relevance of detail level. These two differences are explained in more detail below.

The most striking difference is that the experience bubble tends to be much smaller for sonification than it is in the case of music. Even sonification experts have certainly listened to music far more in their lives than to sonification.

The other difference is that even though musical education may change the way we perceive music [2], we do not necessarily need to understand the musical structure in detail to be able to enjoy it. In contrast, if we do not understand the information represented by a sonification, the sound is music or organized sound for us at best, or a cacophony or noise at worst, but not a sonification. To be understandable, the information has to lie inside all three boundaries: Physiologically, the information has to be audible, that is, relevant information must not lie in inaudible sound magnitudes or changes; psychologically, the information has to be interpretable, for example, in terms of working memory capacity and the principles of auditory scene analysis. This is already a challenging task that has been discussed a lot [13, 14]. Experientially, a sonification has to work with principles that are familiar enough to the user. And here, the absence of sonification in everyday life and the lack of analytic listening

⁶ Available under <https://www.youtube.com/playlist?list=PLVv3BMS8IIXG8ahjhd7NUztfPeFLN2nB>

experience are a real hurdle. The bubble is extremely small. As with new music, sonification has to have overlap with previous experiences so that people are able to understand and eager to learn it. But this typically means that the sonification has to be information-poor, since nobody has much experience with the routine use of information-rich sonification.

This issue has been addressed by [15], who conclude that simple sonifications are ideal because the majority can use them. And indeed, this offers the chance to widen the sonification bubble for many people to an extent that will one day allow the utilization of more informative, more complex sonifications. *Sonification for the masses* was the theme of the International Conference on Auditory Display in 2023 [16]. In addition, the *Tiltification* app [17] aimed at reaching the masses through a certain spirit level app that combined an unfamiliar, complex sonification with a simple, familiar task⁷, which has led to various spirit level sonifications [18].

The downside of simple sonifications is that they are intrinsically information-poor. They offer little benefit over visualizations. While aural beginners cannot handle an information-rich sonification that lies clearly outside their sonification bubble, some aural adepts might become annoyed when exposed to primitive sounds that do not allow them to fully utilize their aural capacities, which degrades the user experience instead of enriching it. A trade-off between musicality and informativeness has been discussed in various publications [7, 19, 20]. Concerning sonification, there are four clear needs:

1. Aural adepts need highly informative sonification to unlock their capabilities and make extensive data accessible to them.
2. Latent aural adepts need tools that allow them to discover and improve their aural capabilities.
3. Aural beginners need tools that help them gain the aural skills required for the usage of highly informative sonification.
4. Developers and users of computer interfaces need to become aware of sonification so that it becomes a standard rather than an exception in human-computer interaction.

And music is an ideal catalyst: A sonification that lies within everyone's music bubble is a pleasant starting point for aural beginners. By adding information to sonify, the sonification can grow outside the listener's music bubble. As with new music, this can be exciting and motivating, increasing the music bubble and the sonification bubble together. The inverse principle is meaningful for aural adepts: Starting with a highly informative sonification and moving it closer to the listener's music bubble maximizes the user experience. It has been stated repeatedly that long-term sonification studies with comparable concepts are needed to investigate the progress using this method [21–23]. One example of a sonification that starts very musical and eventually adds more and more information is presented in Sec. 4.1

⁷ For Android and iOS smartphones, many spirit level apps exist, utilizing accelerometers and sometimes gyroscopes. What is special about *Tiltification* is that the two tilt angles of the smartphones are not displayed via graphics, but also via sonification.

3. Random preludes

One of the earliest known works in Western music composition that expanded the music and sonification bubbles described in Section 2.2 was the *Würfelwalzer*, or *Dice Waltz*. This type of composition introduced elements of chance by rolling a die to determine which musical phrases of a prepared composition were concatenated for the playing [24]. From 1757 to 1812, at least 20 dice games were published in Europe, one of the most popular being Wolfgang Amadeus Mozart’s *Musikalisches Würfelspiel*, or *Musical Dice Game*, published posthumously in 1792 [25].

Dice games can be considered a simplified form of “musification,” in which the numerical outcomes of the dice rolls are mapped onto musical fragments to obtain an auditory output. In other words, the introduction of random variations contributes to the expansion of the musical bubble of the original pieces, while, on the other hand, the underlying musical structure expands the sonification bubble through the inclusion of esthetic contexts.

The process behind a dice game begins with the composition of a musical piece or fragment. In the following simulation, the piece is called *Lissi* (Riber, 2016). It is a loopable harmonic *Allegro* of forty-four 4/4 bars in C minor, with a rhythmic pattern based on sixteenth-notes. The musical notes are chosen around a central value (B4) to smooth transitions between the fragments, as an attempt to hide the random generative process behind the resulting pieces. **Table 1** shows *Lissi*’s assignment matrix based on the one used by Mozart in his piece *K.516f* [26]. The 11 rows correspond to the results obtained when rolling two dice. The eight columns, represented with letters, are the bars incorporated in the final score. The numbers are assigned to each bar of the original piece; notice that each bar is repeated twice to complete the matrix.

	A	B	C	D	E	F	G	H
2	1	2	3	4	5	6	7	8
3	9	10	11	12	13	14	15	16
4	17	18	19	20	21	22	23	24
5	25	26	27	28	29	30	31	32
6	33	34	35	36	37	38	39	40
7	41	42	43	44	1	2	3	4
8	5	6	7	8	9	10	11	12
9	13	14	15	16	17	18	19	20
10	21	22	23	24	25	26	27	28
11	29	30	31	32	33	34	35	36
12	37	38	39	40	41	42	43	44

Table 1.
Lissi’s assignment matrix based on Mozart’s k.516f [26]. The rows represent the random results of rolling two dice. The columns represent, with, a reference for the bars to be included in the final score (to generate longer pieces, looping over the 8 bars with new dice rolls is necessary). The numbers are assigned to each bar of the original piece (44 bars in this example).

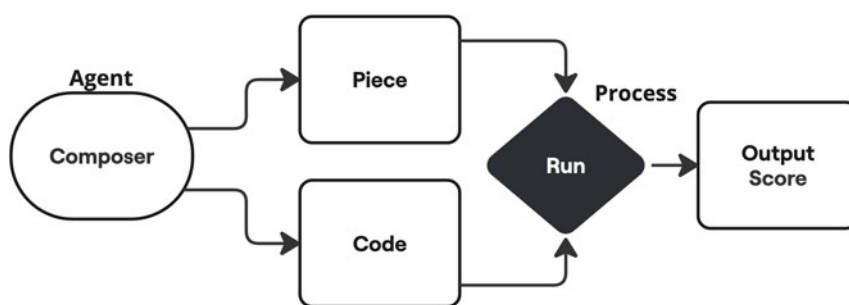


Figure 2. Agent analysis and conceptual modeling for *Lissi* (Riber, 2016). The composer prepares musical phrases and runs a stochastic process to play them in random order.

The following videos show *Lissi*'s initial piece and two different output scores of 20 bars generated with a virtual dice roller notebook⁴, and rendered using MuseScore, the free software for composition and notation.

<https://vimeo.com/1016354640> [Riber, 2016]

<https://vimeo.com/1016354654> – <https://vimeo.com/1016354648>

Figure 2 shows the conceptual model behind the generative process of *Lissi* to illustrate the dual role of the composer (musician and programmer), who provides a mechanism by which the piece is expanded into a collective experience. In other words, with one simple action (rolling dice), participants without significant musical knowledge become involved in the composition process to generate their own pieces.

In this approach, randomness is more structural than performative, but it provides a clear example of how chance can be used to represent the evolution of a musical composition over time. This idea of indeterminacy can be considered as the seed for aleatory and stochastic music, deeply explored in the twentieth century by composers such as Henry Cowell, John Cage, and Iannis Xenakis.

4. Deep learning sonification and musification

Another interesting form of algorithmic music that has occupied the minds of mathematicians, musicians, and astronomers from the time of Ancient Greece to the present can be found around the philosophical topic of *Sphärenharmonie*, or the *Harmony of the Spheres*. This idea, based on the belief that the movements of the celestial bodies can generate a form of music, transcended from Pythagoras to the late Renaissance, when Johannes Kepler materialized the concept by relating musical intervals to the distances and orbital velocities of the planets [27].

Updating this concept to the currently available technology based on Virtual Observatory (VO) interoperable technology [28–30], the connection between astronomical data and the generation and control of sound variables constitutes a promising research line in Astronomy and Astrophysics. On one hand, the approach has the potential to provide a better understanding of a phenomenon [31]

⁹ *Lissi*'s initial score and virtual dice roller are available at: <https://github.com/AdrianGRiber/Musical-data-sonification-Lissi.git>.

by allowing holistic interpretations of the data [32] and by complementing graphical representations in analytical tasks [22]. On the other hand, the use of sound for the representation of astronomical data opens a wide range of possibilities for Sound Design and Music Composition, which can be used in scientific outreach or as a musical outcome in itself.

4.1 Astronomical data-driven compositions

Framed in this context, *Galaxy Spectra Suites* is a collection of six short suites for electronics, which illustrate the possibilities of astronomical data-driven musical composition based on deep learning. Deep learning is a subset of machine learning that uses artificial neural networks with many layers (often called deep networks) to automatically learn hierarchical representations of data.

The proposed piece uses the underlying musical chord structure of Bach’s Cello Suites (BWV 1007–1012) to provide a frame for the sequential representation of galaxy spectra. The network, trained with hundreds of hours of MIDI piano scores from the MAESTRO⁵ dataset [33], creates a musical structure with notes predicted from the input notes of the Cello Suites. Finally, it generates a completely original, astronomical, data-driven soundtrack. Vice versa, it also provides a musical exploration of the actual galaxy spectra from the CALIFA⁶ survey. The CALIFA project [34–36] aims to enable the study of fundamental questions related to galaxy evolution, providing spatially resolved spectroscopic observations (around 1,000 spectra per galaxy) from 895 galaxies in the nearby universe.

To illustrate the dataset used as generative sound sources for the *Galaxy Spectra Suites*, **Figure 3** shows a spatial pixel matrix (left) and the spectrum corresponding to the center of the galaxy (right) for NGC 768. A spectrum is the representation of the object’s light flux as a function of wavelength, which is used to study the physical properties and chemical composition of stars and galaxies.

The autonomous composition system⁷ [37, 38] that allowed the generation of this work uses deep learning algorithms to convert the galaxy spectra into musical chords, cross-matches them via *Pitch Class Set Theory*⁸ with a score generated by a recurrent neural network (triggered by Bach’s Baroque masterpiece), and finally generates the proposed composition with the matching “galaxy chords.” The visuals, launched in real time, show the image of each object and its central spectrum, providing a sequential multimodal representation of the galaxy survey. The instrumentation and final sound design of each piece are performed by the authors, avoiding any changes in the generated score. A video of the complete work is available at: <https://vimeo.com/902748176>.

⁵ <https://magenta.withgoogle.com/datasets/maestro>.

⁶ The CALIFA (Calar Alto Legacy Integral Field Area) survey (<https://califa.caha.es/>) is a project based on the observations from the Centro Astronómico Hispano Alemón (CAHA) at Calar Alto, operated by the Max-Planck-Institut für Astronomie and the Instituto de Astrofísica de Andalucía (CSIC).

⁷ <https://github.com/AuditoryVO/AI-rmonizer>.

⁸ The *Pitch Class Set Theory* identifies and classifies pitches assigning them a number related to their intervallic distance to C (from C = 0 to B = 11), and without distinguishing octaves (such as A3 and A4), and enharmonic equivalents (such as A3 sharp and B3 flat). This reduction allows the analysis and comparison of atonal compositions, making it useful to establish a correlation between the chords generated from the galaxy spectra and those created by the underlying musical structure.

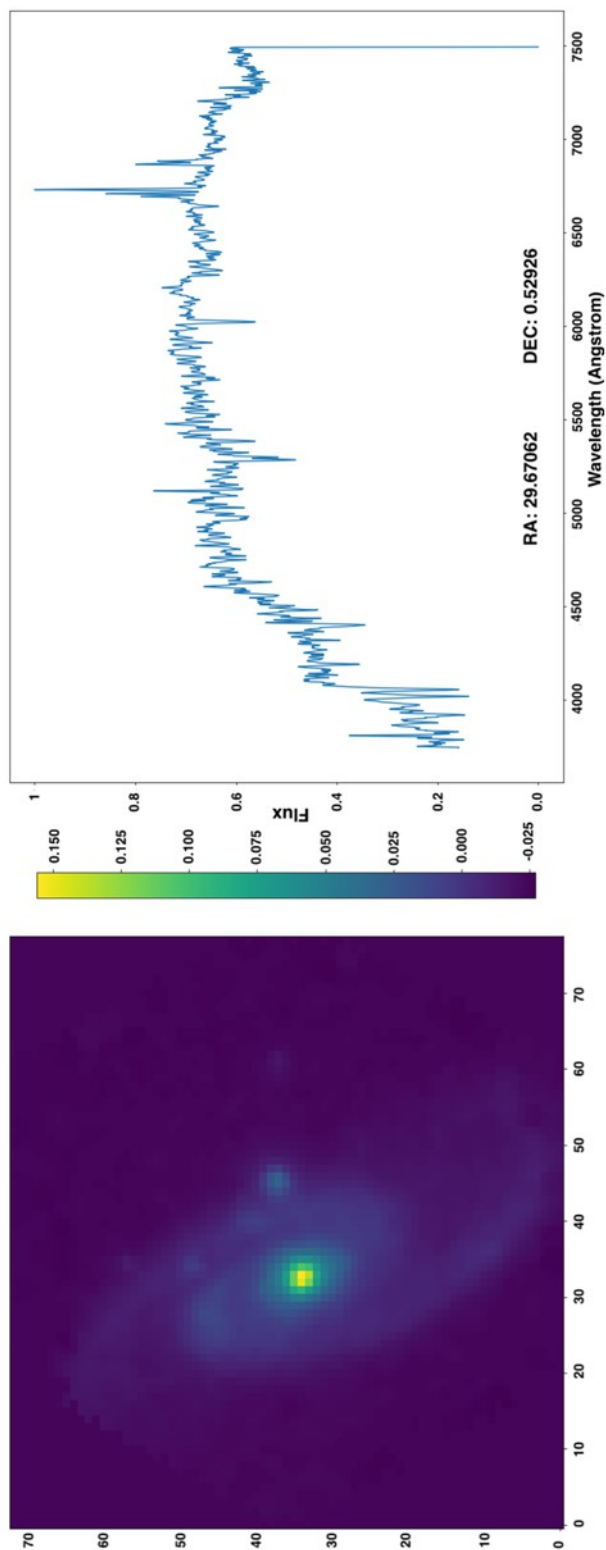


Figure 3. Galaxy image spaxels, or spatial pixels (left), and a spectrum representing the normalized flux vs. wavelength (right) for the central spaxel of NGC 768 (RA: 29.67062, DEC: 0.52926, ICRSD).

As can be seen in **Figure 4**, each galaxy is represented by a chord of six notes, which is generated from the latent values calculated by an autoencoder. This generative architecture allows the reduction of the 1877 points of each spectrum to a lower-dimensional representation called a latent vector. Each latent vector represents, with only six latent values, the spectral features of each galaxy, such as flux distribution or absorption lines. Each latent value of these vectors is interpreted as the fundamental frequency of each note that generates the final chord, whose duration is obtained from the sum of the six values. The name of each source galaxy is included in the lyrics line of the score to provide a direct reference to the generated chords.

Figure 5 shows the conceptual model behind the generative process of *Galaxy Spectra Suites* as an example of musification, that is, converting data into musical representations. An autoencoder reduces the central spectrum of each galaxy to six-dimensional vectors, thus providing the auditory representation of the data. In a parallel process, a neural network with an attention mechanism holds the composition of the underlying musical structure triggered by MIDI transcriptions⁹ of Bach’s Cello Suites [39]. Finally, a cross-matching process based on the Pitch Class Set Theory generates, “in real time,” the sequence of auditory and graphical representations with the matched galaxy spectrum chords. The researcher/composer then acts on the final sound design, choosing the instrumentation and textures to give the pieces the desired aesthetics, thus closing the interaction loop between artificial intelligence and human creativity to generate original soundtracks.

4.2 Sonification–musification dichotomy

Although sonification and musification share numerous concepts and parameters, most of them related to human sound perception, they mainly differ in their final function [1], which significantly influences the design decisions and practical approaches to both sound-based methodologies. While the purpose of sonification is functional and aids in the understanding of complex datasets through audio representations, the goals of musification are more aesthetic, focusing on entertainment and translating data into structured compositions that adhere to musical principles such as melody, rhythm, tension, and surprise. **Table 2** illustrates the applications and fundamental functions of both approaches.

To illustrate this sonification–musification dichotomy, the following video provides a demonstration of *ViewCube* [41], a multimodal Integral Field Spectroscopy (IFS) tool developed in collaboration with Rubén García Benito that incorporates the same deep learning-based encoding process described in Section 4.1: <https://vimeo.com/1005208084>.

While in *ViewCube*, the autoencoder is used to provide “objective” sonifications of the datacubes from the CALIFA survey to be used in interactive scientific analysis; in *Galaxy Spectra Suites*, the same spectra are musically ordered to generate a sense of emotion that has the potential to connect with the audience.

As shown in **Figure 6**, the sonification approach implemented in *Viewcube* uses OSC¹⁰ to send six non-quantified fundamental frequencies to an additive synthesizer, while the musification approach yielding the *Galaxy Spectra Suites* uses

¹² http://www.jsbach.net/midi/midi_solo_cello.html.

¹⁰ OSC (Open Sound Control) is a network communication protocol designed for transmitting control data between multimedia devices (<https://ccrma.stanford.edu/groups/osc/index.html>).

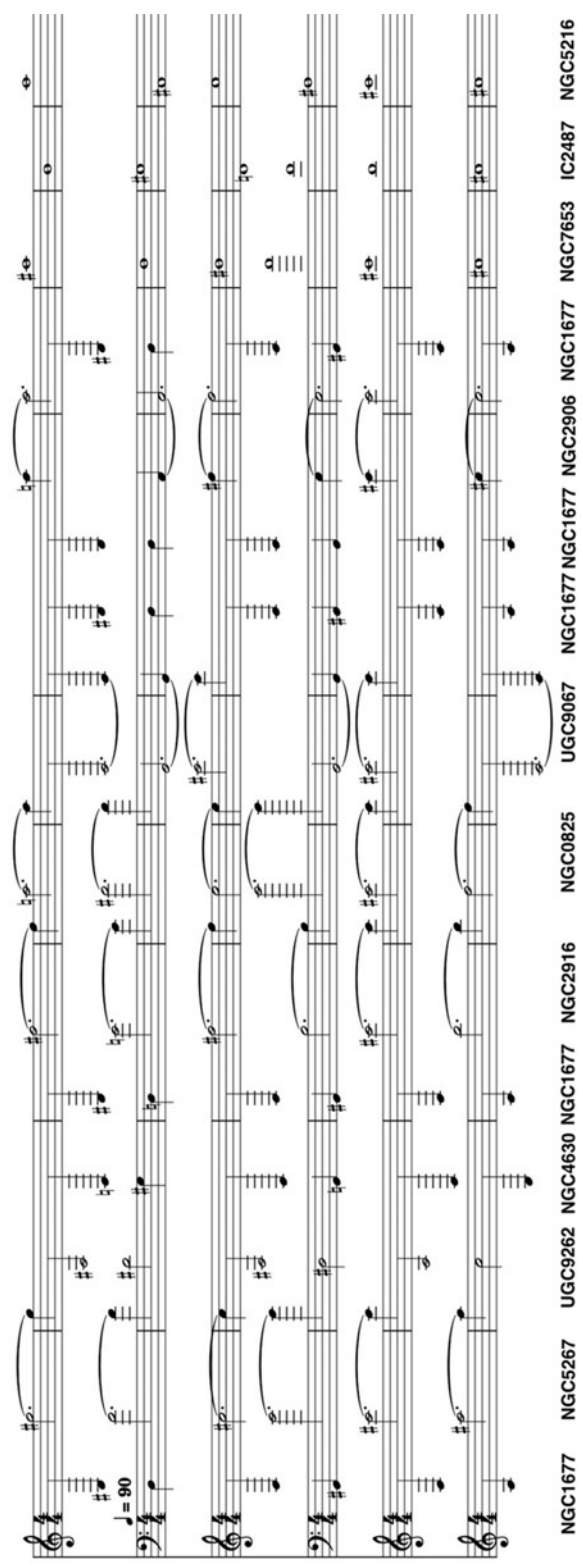


Figure 4.
First bars of Galaxy Spectra Suite No. 1, triggered by the chords of Johann Sebastian Bach's Cello Suite No. 1 in G Major, BWV 1007. Each chord is generated from the central spectrum of each referenced galaxy.

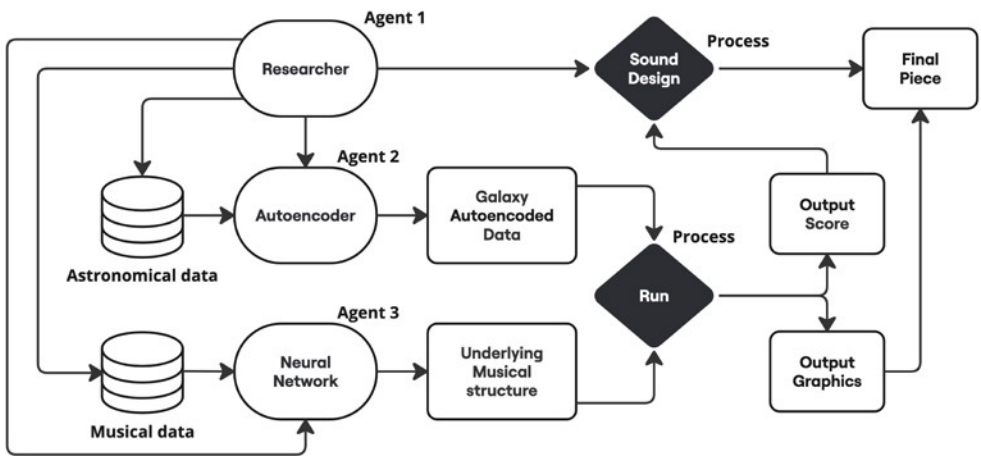


Figure 5.
Agent analysis and conceptual modeling for Galaxy Spectra Suites (Riber and Serradilla), premiered at Audio Mostly & ICAD Joint Conference, AM.ICAD 2025, Coimbra.

	Musification	Sonification
Advantages	General representation	Objective representation
	Public engagement	Accuracy
Disadvantages	Lack of accuracy	Abstraction
	Information loss	Learning curve
Applications	Scientific outreach	Scientific analysis
	Original soundtracks	Accessibility
		Inclusion
		User Interfaces
Functions	*Communication	*Communication
	Symbolic representation	Symbolic representation
	Emotional	*Knowledge representation
	Enjoyment	
	Entertainment	

Asterisks correspond to the main functions of natural language [40].

Table 2.
Advantages, disadvantages, main applications, and fundamental functions of musification and sonification for the CALIFA case study.

MIDI to generate six-note chords from the latent vectors of the autoencoder. This seemingly small difference in the encoding process greatly determines the final functionalities of the approaches. The use of MIDI introduces a loss of information associated with its intrinsic pitch quantization, which reduces the accuracy of the approach but allows a direct mapping to standard digital instruments. The incorporation of musical structures in the representation introduces emotional

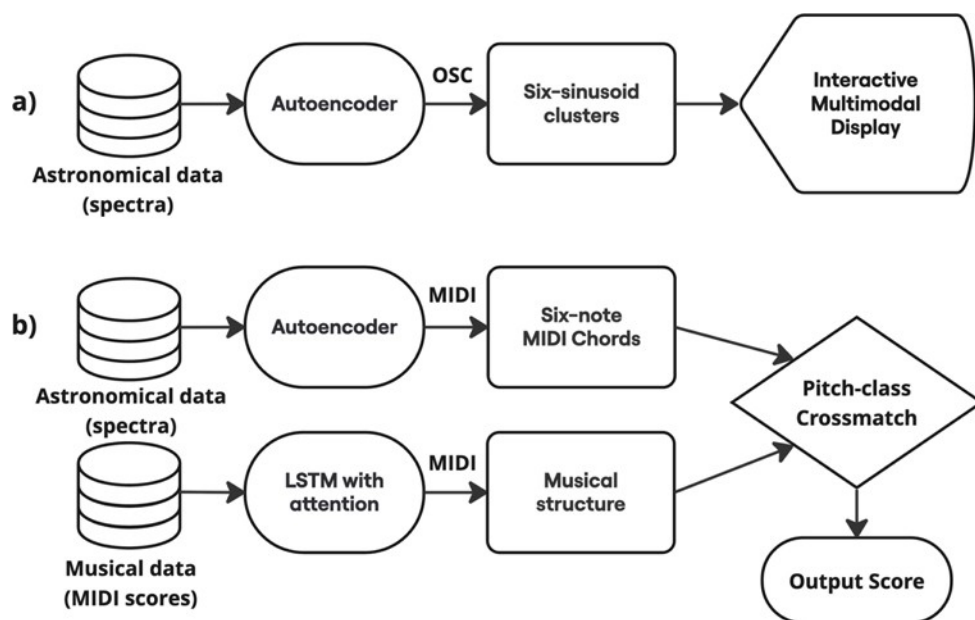


Figure 6. Similarities in the sonification–musification dichotomy. A) The galaxy spectra of the CALIFA survey are converted by an autoencoder into six-sinusoids to allow interactive multimodal analysis. B) The same dataset is converted by the same autoencoder into six-note MIDI chords. Additionally, the MAESTRO dataset is used to train the model that generates the underlying musical structure of the final piece, which is triggered by Bach’s cello suites.

factors, useful for public engagement and entertainment but dispensable in scientific interactive analysis.

5. Music mixing vs. sonification design

5.1 Parallels between design processes

There are clear parallels between the design processes of sonification and music making – particularly with regard to the art of mixing. Mixing is the creative act of defining an algorithm, or a combination of algorithms, that converts multitrack audio into a stereo or multichannel product for eventual consumption [42, Chapter 1]. The mixing engineer devises audio processing operations to apply to raw tracks prior to playback, with decisions shaped by both technical and artistic requirements and guided by critical listening [42–44]. The sum of these *DSP* (*digital signal processing*) operations forms a complex transfer function – that is, the mix – that links the raw multitrack input to the final playback medium. In sonification design, the objective is to define a mapping function that links data variables to audio signal parameters [14, 45]. This can be viewed as an analogous process to mixing: combining and shaping data signals to convey their informational content through audible changes in sound. In this sense, the “mix” in music mixing corresponds to the mapping function in sonification. While the two practices differ fundamentally in purpose and context (see [46] for a detailed comparison), the mature discipline of

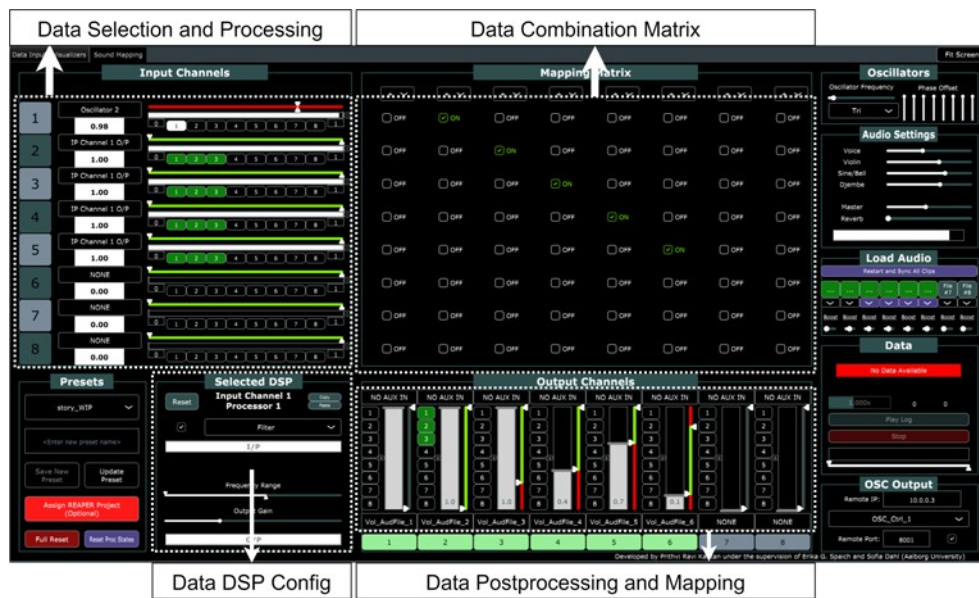


Figure 7.

The mix-n-map interface with labeled key blocks (data selection/processing, data combination matrix, data postprocessing/mapping, data DSP configuration).

music mixing offers valuable insights for sonification – particularly through its central technological instrument: the audio mixer.

5.2 Tool inspiration

Software-based music mixing platforms have been widely used for decades and exist in many forms [47, Chapter 6]. Despite differences in functionality, layout, and workflow, most share two core features: a track view for visual editing and playback control, and a mixer view for applying DSP operations and configuring signal routing [48, Chapter 11]. Some sonification platforms have adapted these layouts to accommodate non-audio data [47–49]. While many platforms since the 1990s have drawn inspiration from mixer architectures, few have explicitly incorporated core elements for sonification purposes. Given the mixer’s pivotal role in audio production, we argue that its explicit adaptation could similarly benefit sonification design. A notable example is Mix-N-Map [46], a platform that directly implements key mixer components for data-driven sound design (Figure 7). It supports real-time signal processing and mapping manipulation, features normalized internal signals for combining diverse data streams, and includes discrete input/output channels with multiple DSP slots for operations like filtering or quantization. A flexible mapping matrix allows complex parameter routing, and users can choose from built-in synthesizers, load custom audio files, or transmit control signals to external software via OSC. Mapping configurations can be stored and recalled, enabling rapid prototyping and iterative refinement. A demonstration of Mix-N-Map is provided in the supplementary video A1 (see video repository)¹¹.

¹⁶ <https://drive.google.com/drive/folders/11FZLgLEuLGC8azSghBI2ev55lom41tcy?usp=sharing>.

6. Music in movement sonification

We now turn to a new application domain of musical sonification. Generally speaking, movement sonification is the conversion of human motion data into sound. It is most commonly applied in real time to increase self-awareness of movement in various contexts, thereby enhancing motor learning and execution [50–52]. As it typically involves the mapping of movement-relevant kinematic or dynamic parameters onto audio synthesis or processing parameters, it is a case of *interactive parameter mapping sonification* [45, 53]. Movement sonification has been explored within several application contexts, such as rehabilitation, sport movement learning, video games, surgical training, and dance [50, 52, 54–56], by exploiting the strong existing neural coupling between action and auditory perception [57]. Music readily lends itself to such applications due to its universal ability to motivate movement, reduce perceived exertion/pain, and improve mood during exercise and motor training [58]. Moreover, studies have shown that interactive music (which changes in response to movement) is more effective than passive (non-interactive) music in this regard [59, 60]. These benefits can contribute to improving user adherence to training, which is directly linked to better functional outcomes in several contexts [61].

6.1 How can music be integrated into movement sonification designs?

Although music has clear advantages for movement-specific applications, the arbitrary or indiscriminate use of musical sonification can lead to suboptimal results for sonification users. In principle, the conveyance of movement variables through richly layered musical stimuli may be tempting for aesthetic reasons, since such sounds typically contain a combination of non-stationary melodic and percussive entities that overlap in time and frequency. Moreover, users vary greatly in terms of musical literacy, hearing ability, and cognitive function, meaning that they are not equally adept at perceiving and comprehending such mappings [20]. Furthermore, the music bubble – and thus, tastes, expectations, and behaviors of users – is culturally influenced and extremely diverse [62], which is highly relevant given the considerable subjectivity present in the interpretation of sonifications [63]. Finally, it can be difficult to distinguish between sound changes resulting from one's movements and those resulting from the natural evolution of the musical stimuli. The end result is often a reduction in clarity combined with an increase in acquisition time [64, 65].

Successful past applications of musical sonification for movement training have primarily employed musical interaction as a means of modulating affect, mood, and perceived exertion by promoting an external focus of attention and a sense of musical agency [59, 66–68]. For example, Fritz *et al.* [59] found that providing musical feedback during machine-based gym workouts had an enhancing effect on mood compared to workouts with passive music listening. However, it is important that users can clearly relate sonic changes to the actions that precipitate them [51]. Two general design approaches can be taken in order to ensure this:

- Sonifications in which every single aspect of musical evolution (melody, harmony, rhythm, timbral changes, etc.) is controlled by movement variables based on intuitive crossmodal correspondences (e.g., vertical position mapped

onto pitch, force/speed mapped onto sound intensity), meaning that there is no element that evolves independently of the movement. Videos E1–E5 exemplify this approach through knee extension and sit-to-stand movements (refer to [69] for details), with mappings that effectively represent each movement in the form of multitrack musical pieces with melodic and percussive elements.

- Sonifications, in which movement variables are represented through sonic changes to an evolving music piece, are clearly distinguishable from the changes resulting from the basic musical structure itself. Examples where weather data are mapped to the depth of various audio effects on a musical piece are provided on [YouTube] (<https://www.youtube.com/playlist?list=PLVv3BMS8IIXGLGrSCX1ANAcKv1OKStAV-YouTube>). The sonic changes must typically have a high degree of perceptual salience in order to be reliably discerned. For instance, undesirable movement patterns are ideally represented through undesirable changes to the music (e.g., interruption, unpleasant distortion). The use of nonlinear or stepwise mapping functions can help ensure that the auditory changes corresponding to the movement sonification are perceptually salient. Videos B1–B4 and C1–C2 show how such a mapping approach can be applied to full songs, clips, or stems (which are audio files containing groups of related tracks within a song – drums, guitars, voices) to reinforce a variety of training goals, such as achieving or maintaining a particular walking speed, step tempo, total distance, or knee flexion pattern (see [70] for details) through clear changes in the music playback parameters. Comparable approaches are seen in [68, 71].

An increasingly adopted sonification design ethos is based on embodied cognition [51, 72]. It exploits embodied associations that humans acquire through lifelong sensorimotor experience [73], which are, thus, intuitive and naturalistic (e.g., a ‘heart rate data sonification design’ where the data is mapped onto the tempo of a recorded heartbeat sound). Roddy and Bridges’ *Embodied Sonification Listening Model* advocates for mappings based on *conceptual metaphors*, that is, essentially using familiar conceptual domains (like musical melody or real-world sound behaviors) as filters/extractors [72] to represent digital data, which can enable faster and more accurate auditory display comprehension [73]. These principles allow motor learners to readily apply familiar embodied image schemata to decode movement information from musical sound. Music as a sonification medium provides both an aesthetic baseline as well as a sonic grammar that people can decode even without training [7, 58, 66, 74]. Melodic sonification is known to help structure timed actions and recover complex movement patterns [75]. Based on Maes et al.’s 3MO model, musical biofeedback, that is, the interactive musical sonification of movement, can *motivate*, *monitor*, and *modify* movement [58].

In a non-musical sonification context (e.g., when using simple sine waves or noise), movement data dimensions are usually mapped onto audio processing parameters pertaining to signal attributes such as fundamental frequency, overall amplitude, envelope characteristics, or repetition rate [54, 76]. Note that similar signal manipulations can also be applied to musical stimuli (synthesized or recorded) in musical sonifications. If so, this affords an additional set of musical parameters for mapping purposes, namely note articulation, consonance/dissonance, instrument layering, rhythmic/melodic patterning, harmonic tension/resolution, and rhythmic

synchronicity. Thus, it is possible to achieve more musical results by mapping data onto musical parameters through mathematical transfer functions that are time-discrete and staircase-shaped, hence conforming to rules of musical organization in time and frequency. For instance, a simple mapping to the fundamental frequency can be modified so that it is quantized to yield frequencies corresponding to pitches of a given musical scale, with changes occurring at specific rhythmic intervals (e.g., beat instants or 16th notes). Such a modification naturally results in a degree of timing inaccuracy and information loss but can bring considerably more appealing sonic results. In D4, the knee angle is mapped to the musical pitch of scale tones in such a way that straightening the knee raises the pitch within a predefined scale, whereby note triggering is always locked in phase with the rhythmic evolution of the background ensemble, thus always yielding a musically “correct” sonic result. In general, the available mappable parameters directly depend on the sound generation method used, specifically when the musical stimulus is a pre-recorded piece (single stereo audio file), a collection of pre-recorded audio stems, a mono/polyphonic output of a sample sequencer/real-time synthesizer, or a combination of these.

- Pre-recorded music unfolds in a fixed temporal manner, and the primary parameters available for sonification mapping are audio-level DSP controls such as volume, panning (spatial positioning in stereo or surround “image”), equalization, compression, distortion, and reverberation. Playback parameters – such as playback rate (which affects pitch and tempo), play/pause, and scrubbing (manual navigation through an audio file, typically by dragging a cursor or playhead across the waveform) – can also be mapped. For example, F1 (see 10.5281/zenodo.16581819) is a VR + music system for stationary biking that slows music down when the user falls behind a target position (rabbit) and speeds it up when overtaking [77]; C1 (see examples) applies a similar approach to walking. Modifying pitch independently of tempo is challenging due to the polyphonic textures in music and the latency introduced by frequency-domain pitch shifters like Melodyne [80, Chapter 1]. However, consonance and dissonance can be coarsely altered by mapping onto the strength of ring modulators applied to the audio [78]. While pre-recorded audio files are limited in how they can be transformed, they offer high flexibility for user tailoring, as diverse genres are easily accessible and mappings can be applied in real time to any clip. A key caveat is that music clips vary in spectral and temporal characteristics [79], so applying the same mappings (e.g., equalization) to different clips may yield perceptually inconsistent results depending on the underlying musical material.
- Pre-recorded stems (audio files that have to be played together to become the final song, like a drum section file, a bassline file, and some brass section files) offer similar mapping possibilities in comparison to full song clips, along with a set of additional affordances related to the layering and mixing of the stems. For instance, different data variables can be linked to processing parameters of different stems, or stems can be muted or unmuted depending on the data variable value over time. *Mixing Levels* [19] is a spirit level app where a music mix is only balanced correctly when the smartphone is made perfectly horizontal. Inter-stem synchronicity can also be manipulated, and musical scale (or tonality) modulation can be achieved by transposing the stems with melodic or

harmonic content. Although stems can be obtained for a wide range of genres, these are less freely available than ordinary mixed tracks. Examples B1–B4 illustrate how stem balance can be altered to support a variety of walking goals related to speed, step tempo, and total distance.

- Sample sequencers that generate mono- or polyphonic output from symbolic musical sequences (e.g., MIDI notation) using pre-recorded samples enable data to be mapped at the symbolic rather than audio level, opening up a distinct domain of mapping possibilities. In other words, such sequencers generate a music mix based on rules, which are easier to access and manipulate before rather than after mixing the music. Moreover, data can directly control musical features such as melodic contour (e.g., D4), harmonic progression (e.g., E3), rhythmic structure (e.g., D3, D6), event density, and their long-term evolution – without introducing audio artifacts (e.g., clicks, pops, stuttering). Musical structures can be sourced from MIDI files or generated in real time using rule-based or machine learning algorithms [80]. A variety of sample packs, MIDI files, and generative systems also allow for diverse and evolving outputs, though not all genres are equally compatible with this approach. Genres outside the conventions of Western notation (e.g., Hindustani classical music) may be difficult to represent in MIDI-based sequencers [81], since the realism of sequencer output is limited by the inherent quantization applied to timing and articulation parameters [82].
- Real-time synthesizers that generate audio in accordance with symbolically represented sequences (e.g., MIDI notation) afford maximum flexibility in terms of manipulating both musical content and audio generation specifics. Depending on the synthesis methods used (e.g., additive, subtractive, FM, granular, physics-based) [83, 84], there can be a variety of mappable synthesis parameters. As these parameters control the very source of musical audio signals, their manipulation can affect the audible output far more dramatically than the mere modification of existing audio signals (see examples D1, D2, D5). For instance, mapping a data variable onto the modulation depth or modulator/carrier frequency ratio of an FM synthesizer can result in sound changes not achievable through traditional DSP on pre-recorded waveforms [85]. The main potential disadvantages of real-time synthesis are related to increased computational load and limited sound variety (both depending on the synthesizer in question) [83]. Additionally, it is rarely possible to exactly reproduce user-preferred music tracks through real-time synthesis, which undershoots the positive potential of preferred music during exercise and movement [86].

6.2 Perspectives

As low-cost motion-sensing and music-generation tools proliferate in conjunction with rising computational power in mobile and desktop devices, there will be growing opportunities to musically sonify human movement for artistic, scientific, and therapeutic purposes. Prompt-based music generation platforms (e.g., Stable Audio, Udio, Suno) can rapidly produce high-quality music in various styles, though their underlying AI models currently lack real-time control. Converting movement or physiological data – such as heart rate or EEG signals – into

generative music that adapts to a user's physical and psychological state is a promising sonification application, given AI's rapid advancement.

The digital domain unlocks sonic interaction possibilities that are currently difficult to fully express or comprehend. However, this power must be exercised judiciously to ensure that such technologies serve users meaningfully. With regard to musical sonification, the role of music within a given application must be clearly defined, and designs should be user-centered, that is, addressing the needs, preferences, and abilities of the target group [87]. Mapping strategies grounded in embodied conceptual metaphors (e.g., using an actual heartbeat sound with tempo mappings to represent information related to heart rate) [72] support ubiquitously meaningful movement-music associations, while iterative development and formative evaluation help meet the evolving design requirements. Ultimately, musical sonification enables interaction beyond the limits of physics, offering new pathways for therapy and artistic expression.

7. Conclusion

As digital technologies expand the boundaries of what sound can represent and how it can be shaped, the convergence of music and sonification offers both expressive power and communicative precision. Yet, harnessing this potential demands clarity of intent and attention to the user's context. This chapter has shown that music is more than an aesthetic layer – it is a rich structural and experiential resource that can deepen engagement, enhance comprehension, and support embodiment when thoughtfully integrated into sonification designs. On the other hand, sonification is more than just a methodology to convey information through sound – it is a quantitative ground that provides meaning to data-driven music and sound art. By distinguishing between acoustical, auditory, and musical domains, we have clarified where music and sonification intersect and how they diverge. We explored how deep learning enables musically coherent data representations, how audio mixing principles offer design insights, and how movement-based sonification can benefit from musical structure in therapeutic and artistic settings. Across all these domains, music remains a uniquely versatile and culturally resonant medium. When applied with care, it transforms sonification from a tool of representation into a medium of interaction – capable of connecting data, body, and emotion in ways that extend both human understanding and creative expression.

Author details

Tim Ziemer^{1*}, Adrián García Riber² and Prithvi Ravi Kantan³

1 Institute of Systematic Musicology, University of Hamburg, Hamburg, Germany

2 Pompeu Fabra University, Barcelona, Spain

3 Department of Architecture, Design and Media Technology, Aalborg University, Copenhagen, Denmark

*Address all correspondence to: tim.ziemer@uni-hamburg.de

IntechOpen

© 2026 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Scaletti C. Sonification ≠ music. In: Dean RT, McLean A, editors. *The Oxford Handbook of Algorithmic Music*. New York, NY: Oxford University Press; 2018. p. 363–386. DOI: 10.1093/oxfordhb/9780190226992.013.9
- [2] Bader R. *How Music Works. A Physical Culture Theory*. Cham: Springer; 2021. DOI: 10.1007/978-3-030-67155-6
- [3] Barrass S. The aesthetic turn in sonification towards a social and cultural medium. *AI and Society*. 2012;27:177–181. DOI: 10.1007/s00146-011-0335-5
- [4] Vickers. *Ways of listening and modes of being: Electroacoustic auditory display*. arXiv. 2013. DOI: 10.48550/arXiv.1311.5880
- [5] Worrall D. Sonification: A prehistory. In: 24th International Conference on Auditory Display, 177–182, 2018
- [6] Worrall D. *Sonification Design*. Switzerland: Springer Nature; 2019. DOI: 10.1007/978-3-030-01497-1
- [7] Vickers P, Hogg B. Sonification Abstraite/Sonification Concrète: An ‘Aesthetic perspective space’ for classifying auditory displays in the ARS Musica domain. In: 12th International Conference on Auditory Display, 211–216, 2006
- [8] Filella NB. *Data Sonification in Creative Practice*. PhD thesis, University of Plymouth, 2019
- [9] Middleton J, Hakulinen J, Tiitinen K, Hella J, Keskinen T, Huuskonen P, Linna J, Turunen M, Ziat M, Raisamo R. Sonification with musical characteristics: A path guided by user engagement. In: *International Conference on Auditory Display*, 35–41, 2018. DOI: 10.21785/icad2018.006
- [10] Ziemer T. Sound terminology in sonification. *Journal of the Audio Engineering Society*. 2024;72(5): 274–289. DOI: 10.17743/jaes.2022.0133
- [11] Neuhaus C. Processing musical form: Behavioural and neurocognitive approaches. *Musicae Scientiae*. 2013;17(1):109–127. DOI: 10.1177/1029864912468998
- [12] Ziemer T. *Psychoacoustic Music Sound Field Synthesis*. Cham: Springer; 2020. DOI: 10.1007/978-3-030-23033-3
- [13] Ziemer T, Schultheis H. PAMPAS: A PsychoAcoustical Method for the Perceptual Analysis of multidimensional Sonification. *Frontiers in Neuroscience*. 2022;16. DOI: 10.3389/fnins.2022.930944
- [14] Hermann T. *Sonification for Exploratory Data Analysis*. PhD thesis, Bielefeld University, 2002
- [15] Groß-Vogt K, Kajetan Enge, and IOhannes m. zmölnig. Reflecting on qualitative and quantitative data to frame criteria for effective sonification design. In: 18th International Audio Mostly Conference. Edinburgh, UK, October 2023. p. 93–100. DOI: 10.1145/3616195.3616233
- [16] Weger M, Ziemer T, Rönneberg N, editors. *Sonification for the Masses: Proceedings of the 28th Annual International Conference on Auditory*

Display. Norrköping, Sweden. June 2023. DOI: 10.21785/icad2023.000

[17] Asendorf M, Kienzle M, Ringe R, Ahmadi F, Bhowmik D, Chen J, Huynh K, Kleinert S, Kruesilp J, Lee YY, Wang X, Luo W, Jadid NM, Awadin A, Raval V, Schade EES, Jaman H, Sharma K, Weber C, Winkler H, Ziemer T. Tiltification — An accessible app to popularize sonification. In: 26th International Conference on Auditory Display. Virtual Conference, 2021. p. 184–191. DOI: 10.21785/icad2021.025

[18] Ziemer T. The sonic tilt competition. In: Proceedings of the 28th International Conference on Auditory Display (ICAD 2023), Norrköping, Sweden: June 2023. p. 203–206. DOI: 10.5281/zenodo.14513443

[19] Ziemer T. Mixing levels – A rock music spirit level app. In: International Conference on Auditory Display, Norrköping, Sweden: 06 2023. DOI: 10.48550/arXiv.2302.08249

[20] John GN. Is Sonification Doomed to Fail? In: 25th International Conference on Auditory Display, Newcastle upon Tyne: June 2019. p. 327–330. DOI: 10.21785/icad2019.069

[21] Ziemer T. Results from the CURAT sonification game. In: Proceedings of the 29th International Conference on Auditory Display (ICAD 2024). Troy, NY, USA: June 2024. p. 42–49. DOI: 10.21785/icad2024.007

[22] Walker BN, Nees MA. Theory of sonification. In: Hermann T, Hunter A, Neuhoﬀ JG, editors. *The Sonification Handbook*. Berlin: Logos Verlag: 2011. p. 9–39. Available from: <https://sonification.de/handbook/>

[23] John GN. Perception, cognition and action in auditory displays. In: Hermann T, Hunt A, Neuhoﬀ JG, editors. *The Sonification Handbook*, Chapter 4. Berlin: Logos Verlag: 2011. p. 63–85. Available from: <https://sonification.de/handbook/>

[24] Albini G. Combinatorics, probability and choice in music composition: Towards an aesthetics of composing systems for non-musicians. In: *Proceedings of Bridges 2018: Mathematics, Art, Music, Architecture, Education, Culture*, 395–398, 2018

[25] Hedges SA. Dice music in the eighteenth century. *Music and Letters*. 1978;**59**(2):180–187. DOI: 10.1093/ml/59.2.180

[26] Noguchi H. Musical game in C K. 516f. *Mozart Studies Online*. 1997. Available from: <https://www.asahi-net.or.jp/~RB5H-NGC/e/k516f.htm>

[27] Kepler J. *Harmonices Mundi Libri V*. Bologna: Forni; 1969. p. 207

[28] Quinn PJ, Górski KM, editors. In: *Toward an International Virtual Observatory: Proceedings of the ESO/ESA/NASA/NSF Conference Held at Garching, Germany, 10–14 June 2002*. Springer, Berlin, Heidelberg, 2004. DOI: 10.1007/b82981

[29] Hanisch RJ, Berriman GB, Lazio TJW, Bunn SE, Evans J, McGlynn TA, Plante R. The virtual astronomical observatory: Re-engineering access to astronomical data. *Astronomy and Computing*. 2015;**11**:190–209. DOI: 10.1016/j.ascom.2015.03.007

[30] Dowler P, Evans J, Arviset C, Gaudet S, Group TC, et al. Ivoa architecture version 2.0. 2021;**36**:1101

- [31] Enge K, Elmquist E, Caiola V, Rönnerberg N, Rind A, Iber M, Lenzi S, Lan F, Höldrich R, Aigner W. Open your ears and take a look: A state-of-the-art report on the integration of sonification and visualization. *Computer Graphics Forum*. 2024;**43**(3):e15114. DOI: 10.1111/cgf.15114
- [32] Jeff Cooke WD-M, Foran G, Hannam J, Garcia B. Exploring data sonification to enable, enhance, and accelerate the analysis of big, noisy, and multi-dimensional data: Workshop 9. *International Astronomical Union*. 2017;**14**(S339):251–256. DOI: 10.1017/S1743921318002703
- [33] Hawthorne C, Stasyuk A, Roberts A, Simon I, Huang C-ZA, Dieleman S, Elsen E, Engel J, Eck D. Enabling factorized piano music modeling and generation with the maestro dataset. *arXiv*. 2018. DOI: 10.48550/arXiv.1810.12247
- [34] Sánchez SF, Kennicutt RC, Gil De Paz A, Van de Ven G, Vílchez JM, Wisotzki L, Walcher CJ, Mast D, Alfonso J, Aguerri L, Albiol-Pérez S, et al. CALIFA, the calar alto legacy integral field area survey-I. Survey presentation. *Astronomy and Astrophysics*. 2012;**538**(A8):A8. DOI: 10.1051/0004-6361/201117353
- [35] Walcher CJ, Wisotzki L, Bekeraite S, Husemann B, Iglesias-Páramo J, Backsmann N, Ballesteros JB, Cristina Catalán-Torrecilla CC, Olmo AD, et al. CALIFA: A diameter-selected sample for an integral field spectroscopy galaxy survey. *Astronomy and Astrophysics*. 2014;**569**(A1). DOI: 10.1051/0004-6361/201424198
- [36] Sánchez SF, García-Benito R, Stefano Zibetti CJW, Husemann B, Mendoza MA, Galbany L, Falcón-Barroso J, Mast D, Aceituno J, et al. CALIFA, the calar alto legacy integral field area survey-iv. third public data release. *Astronomy and Astrophysics*. 2016;**594**(A36). DOI: 10.1051/0004-6361/201628661
- [37] García Riber A, Serradilla F. Al-monies of the spheres. In *International Conference on Computational Intelligence in Music, Sound, Art and Design*, 132–147. Springer, 2023. DOI: 10.1007/978-3-031-29956-8_9
- [38] García Riber A, Serradilla F. Toward an auditory virtual observatory. *Journal of the Audio Engineering Society*. 2024;**72**(5):341–351. Available from: <https://secure.aes.org/forum/pubs/journal/?elib=22425>
- [39] David JG. Bach's cello suites midi transcriptions, 1997. Available from: https://www.jsbach.net/midi/midi_solo_cello.html
- [40] Schubert E. The fundamental function of music. *Musicae Scientiae*. 2009;**13**(2_suppl):63–81. DOI: 10.1177/1029864909013002051
- [41] García Riber A, García-Benito R, Serradilla F. Interactive multimodal integral field spectroscopy. *RAS Techniques and Instruments*. 2024;**3**(1):748–758. DOI: 10.1093/rasti/rzae049
- [42] Lefford MN, Bromham G, Moffat D. Mixing with intelligent mixing systems: Evolving practices and lessons from computer assisted design. *Audio Engineering Society Convention* 148. 2020
- [43] Ziemer T. Source width in music production. Methods in stereo, ambisonics, and wave field synthesis. In Schneider A, editor. *Studies in Musical Acoustics and Psychoacoustics*. Cham:

Springer International Publishing; 2017. p. 299–340. DOI: 10.1007/978-3-319-47292-8_10

[44] Ziemer T, Kiattipadungkul P, Karuchit T. Acoustic features from the recording studio for music information retrieval tasks. *Proceedings of Meetings on Acoustics*. 2020;**42**(1):035004. DOI: 10.1121/2.0001363

[45] Grond F, Berger J. Parameter mapping sonification. In: Hermann T, Hunt A, Neuhoff JG, editors. *The Sonification Handbook*. Berlin: Logos Verlag; 2011. p. 363–399. Available from: <https://sonification.de/handbook/>

[46] Kantan PR, Dahl S, Spaich EG. Adapting audio mixing principles and tools to parameter mapping sonification design. In *International Conference for Auditory Display*, 148–155, June 2024. DOI: 10.21785/icad2024.034

[47] Phillips S, Cabrera A. Sonification workstation. In *25th International Conference on Auditory Display*, 184–190, Newcastle Upon Tyne, UK, 2019. DOI: 10.21785/icad2019.056

[48] Peng T, Choi H, Berger J. SIREN: Creative and extensible sonification on the web. In *International Conference on Auditory Display*, 78–84, Norrköping, Sweden, 2023. DOI: 10.21785/icad2023.6167

[49] Lindetorp H, Falkenberg K. Sonification for everyone everywhere. evaluating the WebAudioXML Sonification Toolkit for Browsers. In: *26th International Conference on Auditory Display*, 15–21, 2021. DOI: 10.21785/icad2021.009

[50] Sigrist R, Rauter G, Riener R, Wolf P. Augmented Visual, Auditory, Haptic, and Multimodal Feedback in

Motor Learning: A Review. *Psychonomic Bulletin and Review*. 2012;**20**:20–53. DOI: 10.3758/s13423-012-0333-8

[51] Dyer JF, Stapleton P, Rodger MWM. Sonification as Concurrent augmented feedback for motor skill learning and the importance of mapping Design. *The Open Psychology Journal*. 2015;**8**(1):192–202. DOI: 10.2174/1874350101508010192

[52] Schaffert N, Janzen TB, Mattes K, Thaut MH. A review on the relationship between sound and movement in sports and rehabilitation. *Frontiers in Psychology*. 2019;**10**(244). DOI: 10.3389/fpsyg.2019.00244

[53] Degara N, Hunt A, Hermann T. Interactive Sonification [Guest editors' introduction]. *IEEE Multimedia*. 2015;**22**(1):20–23

[54] Ziemer T, Schultheis H. The CURAT sonification game: Gamification for remote sonification evaluation. In: *26th International Conference on Auditory Display*, 2021. p. 233–240. DOI: 10.21785/icad2021.026

[55] Vajsbaher T, Ziemer T, Schultheis H. A multi-modal approach to cognitive training and assistance in minimally invasive surgery. *Cognitive Systems Research*. 2020;**64**:57–72. DOI: 10.1016/j.cogsys.2020.07.005

[56] Ziemer T. Three-dimensional sonification as a surgical guidance tool. *Journal on Multimodal User Interfaces*. 2023;**17**(4):253–262. DOI: 10.1007/s12193-023-00422-9

[57] Lahav A, Saltzman E, Schlaug G. Action representation of sound: Audiomotor recognition network while listening to newly acquired actions.

Journal of Neuroscience. 2007;27(2):308–314. DOI: 10.1523/jneurosci.4822-06.2007

[58] Maes P-J, Buhmann J, Leman M. 3MO: A model for music-based biofeedback. *Frontiers in Neuroscience*. 2016;10(548). DOI: 10.3389/fnins.2016.00548

[59] Fritz TH, Hardikar S, Demoucron M, Niessen M, Demey M, Giot O, Li Y, Haynes J-D, Villringer A, Leman M. Musical agency reduces perceived exertion during strenuous physical performance. *Proceedings of the National Academy of Sciences*. 2013;110(44):17784–17789. DOI: 10.1073/pnas.1217252110

[60] Fritz TH, Bowling DL, Contier O, Grant J, Schneider L, Lederer A, Höer F, Busch E, Villringer A. Musical agency during physical exercise decreases pain. *Frontiers in Psychology*. 2018;8(2312). DOI: 10.3389/fpsyg.2017.02312

[61] Dimitriadis T, Porta DD, Perschl J, Evers AWM, Magee WL, Schaefer RS. Motivation and music interventions in adults: A systematic review. *Neuropsychological Rehabilitation*. 2024;34(5):649–678

[62] North AC. Individual differences in musical taste. *The American Journal of Psychology*. 2010;123(2):199–208

[63] Wolfe K. Sonification and the mysticism of negation. *Organised Sound*. 2014;19(3):304–309

[64] Kantan PR. Comparing sonification strategies applied to musical and non-musical signals for auditory guidance purposes. In 19th Sound and Music Computing

Conference, SMC 2022, 279–286. Sound and Music Computing Network, 2022. DOI: 10.5281/zenodo.6573218

[65] Rönnerberg N. Sonification for conveying data and emotion. In 16th International Audio Mostly Conference, 56–63, 2021. DOI: 10.1145/3478384.3478387

[66] Bergstrom I, Seinfeld S, Arroyo-Palacios J, Slater M, Sanchez-Vives MV. Using music as a signal for biofeedback. *International Journal of Psychophysiology*. 2014;93(1):140–149

[67] Van den Berghe P, Derie R, Bauwens P, Gerlo J, Segers V, Leman M, De Clercq D. Reducing the peak tibial acceleration of running by music-based biofeedback: A quasi-randomized controlled trial. *Scandinavian Journal of Medicine and Science in Sports*. 2022;32(4):698–709

[68] Lorenzoni V, Van den Berghe P, Maes P-J, De Bie T, De Clercq D, Leman M. Design and validation of an auditory biofeedback system for modification of running parameters. *Journal on Multimodal User Interfaces*. 2019;13:167–180. DOI: 10.1007/s12193-018-0283-1

[69] Kantan PR, Spaich EG, Dahl S. A metaphor-based technical framework for musical sonification in movement rehabilitation. In: 26th International Conference on Auditory Display. 2021. p. 176–183. DOI: 10.21785/icad2021.008

[70] Kantan PR, Dahl S, Røvsing Møller Jørgensen H, Khadye C, Spaich EG. Designing sonified feedback on knee kinematics in hemiparetic gait based on inertial sensor data. In: SoniHED 2022: Conference on the Sonification of

Health and Environmental Data, 2022. p. 12–19. DOI: 10.5281/zenodo.7243766

[71] Schedel M, Weymouth D, Pinkhasov T, Loomis J, Morris IB, Vasudevan E, Muratori L. Interactive sonification of gait: Real-time biofeedback for people with Parkinson's disease. 5th Interactive Sonification Workshop, 2016. p. 94–97

[72] Roddy S, Bridges B. Mapping for meaning: The embodied sonification listening model and its implications for the mapping problem in sonic information design. *Journal on Multimodal User Interfaces*. 2020;**14**(2):143–151. DOI: 10.1007/s12193-020-00318-y

[73] Antle AN, Corness G, Droumeva M. What the body knows: Exploring the benefits of Embodied Metaphors in hybrid physical digital environments. *Interacting with Computers*. 2009;**21**(1-2):66–75

[74] Maes P-J, Leman M, Palmer C, Wanderley MM. Action-based Effects on Music Perception. *Frontiers in Psychology*. 2014;**4**(1008). DOI: 10.3389/fpsyg.2013.01008

[75] Dyer JF, Stapleton P, Rodger MWM. Advantages of melodic over rhythmic movement sonification in bimanual motor skill learning. *Experimental Brain Research*. 2017;**235**:3129–3140. DOI: 10.1007/s00221-017-5047-8

[76] Schwarz S, Ziemer T. A psychoacoustic sound design for pulse oximetry. In: *Proceedings of the 25th International Conference on Auditory Display (ICAD 2019)*, Newcastle Upon Tyne; 2019. DOI: 10.21785/icad2019.024

[77] Kantan PR, Rojo A, Olmo-Fajardo T, Matas LM, Spaich EG, Dahl S, Moreno JC. Developing an integrated vr + musical feedback system for stationary biking in endurance training. In *9th International Conference on Movement and Computing*, 1–6, 2024. DOI: 10.1145/3658852.3659086

[78] Kantan PR, Dahl S. An interactive music synthesizer for gait training in neurorehabilitation. In: *16th Sound and music computing conference*, 159–166. *Sound and Music Computing Network*, 2019. DOI: 10.5281/zenodo.3249297

[79] Lee C-H, Shih J-L, Yu K-M, Lin H-S. Automatic music genre classification based on modulation spectral analysis of spectral and cepstral features. *IEEE Transactions on Multimedia*. 2009;**11**(4):670–682

[80] Dash A, Agres K. AI-based affective music generation systems: A review of methods and challenges. *ACM Computing Surveys*. 2024;**56**(11):1–34

[81] Rothstein J. *MIDI: A Comprehensive Introduction*. Vol. 7. Middleton, Wisconsin: AR Editions, Inc.; 1995

[82] Morrison L. Rhythm quantization: Notes on the history of a technocultural practice. In: *The Oxford Handbook of Time in Music* Mark Doffman, Emily Payne, Toby Young. New York, NY: Oxford University Press; 2021. p. 341–365

[83] Cook P. Sound synthesis for auditory display. In: Hermann T, Hunt A, Neuhoff JG, editors. *The Sonification Handbook*, Chapter 9. Logos, Berlin: 2011. p. 197–235. Available from: <https://sonification.de/handbook/> [Accessed: 2025-March-1]

- [84] Ziemer T, Schultheis H. Three orthogonal dimensions for psychoacoustic sonification. arXiv. 2019. DOI: 10.48550/arXiv.1912.00766
- [85] Chowning JM. The synthesis of complex audio spectra by means of frequency modulation. *Journal of the Audio Engineering Society*. 1973;21(7):526–534
- [86] Ballmann CG. The influence of music preference on exercise responses and performance: A review. *Journal of Functional Morphology and Kinesiology*. 2021;6(2):33. DOI: 10.3390/jfmk6020033
- [87] Lesaffre M. Investigating Embodied Music Cognition for Health and Well-Being. In: Bader R, editor. *Springer Handbook of Systematic Musicology*. Berlin, Heidelberg: Springer; 2018. p. 779–791. DOI: 10.1007/978-3-662-55004-5_36

Section 2

Music Theory

Chapter 2

Cognitive Music Theory: The Current State of Research

Ildar Khannanov

Abstract

Cognitive music theory, in contrast with cognitive musicology, is a new direction by following which music theorists may obtain the perceptual aspect of music that is currently missing in their analyses. Indeed, music theory since Pythagoras denied any participation of an active subject, and this attitude has become mainstream, while such important contributions like that of Aristoxenus were downplayed and remained on a recessive track. Cognitive musicologists, on the other hand, used the general scientific language of description, but had formal training only in music. They (Carol Krumhansl, Fred Lerdahl, Lawrence Zbikowski, and David Huron) relied on the discoveries of cognitive psychology (see, for example, Ulric Neisser's book *Cognitive Psychology*), and on experimental research, while, at the same time, in-depth evaluation of music theoretical categories, such as function, motive and harmony, is missing. This chapter intends to theorize these two positions on the example of *modulation* in tonal music. The author will discuss the most representative accounts of modulation, such as A.B. Marx's and Hugo Riemann's on the music-theoretical side and Carol Krumhansl's on the cognitive one. The author will then offer a hybrid of the two, exemplified by an analysis of J. S. Bach's modulatory path in his *Allemande* from the *French Suite no. 5* based upon the method of trace observation to deduce from the score how listeners may experience this piece, by gathering what the author calls *the circumstantial evidence*. By establishing a long path to a closely related key, Bach seems to be preoccupied not with writing out a modulation but with guiding the listener's cognition.

Keywords: cognitive music theory, cognitive musicology, music theory, cognitive aspects of modulation, cognitive excess

1. Introduction

From the outset, I must state that there are two well-established fields of music scholarship: music theory and cognitive musicology. The former is an old enterprise; it spans 2600 years and covers vast regions globally, Pythagoras and his school being one of its earliest representatives. From the outset, music theory engaged with

mathematics and philosophy, thus all three born in the same place at the same time. In contrast to that, cognitive musicology is a relatively new domain, the result of a collaboration of music scholars with the representatives of a new stage in experimental psychology, namely, *cognitive psychology*, which is commonly defined as a part of the so-called “cognitive revolution of the 1950s.” It coincided with the emergence of computer science that opened a new perspective for comparisons between human information processing and that of a computer. The term “cognitive science” was coined in 1973 by Christopher Longuet-Higgins and was adopted by musicians in the form of cognitive musicology.

Noteworthy is the detail: they use the term *musicology* and not *theory* in the name of their field. In North American tradition, the divide between music theory and musicology is significant. It is a traditional division of labor: music theory in the United States has always been narrowly defined and rigorously relied on positivist knowledge; musicology, in contrast, embraced all the adjacent fields and has been interdisciplinary by default. That explains why the cognitive studies historically joined the corpus of musicological studies while music theory remained isolated from this process. Due to this vast difference in history and context, the theory of music remains predominantly foreign to cognitive musicology. Theorists, on the other hand, traditionally reject what they called “psychologization of theory.”¹ Such an attitude was established by the followers of the Pythagorean doctrine that has always been exclusively objective and cosmological.² The fashion for numerology was established by Pythagoras already around 500 BC, and this doctrine of interval proportions has become especially pronounced in the eighteenth-century theory of harmony and in the nineteenth-century German music theory and its derivatives. A widely known “Hanslick against Wagner controversy” was the most pronounced manifestation of this divide [2]. All things considered, the intentional disregard for perception and cognition has always been a deficiency of traditional music theory. As already mentioned, the example of Aristoxenus is an exception: in contrast with the overwhelming position of most music theorists he maintained that listening (perception by ear) is the starting point for judging sound and for forming aesthetic opinions.³

¹ Ellon Carpenter in her Ph. D. Dissertation divides theory into hard and soft. The former is theory *per se*; the latter is the product of psychologization [1]. Yuri Kholopov mentioned this in his lectures at Moscow Conservatory as a major problem; Pieter C. van den Toorn maintained the same attitude toward psychologization of music theory. One of the most fervent proponents of an anti-psychological image of music theory was Milton Babbitt. Noteworthy is the approach of Hugo Riemann: although he was the one who introduced the interdisciplinary fields into music scholarship, when it came to discussions of harmony and form, he avoided psychological dimension.

² Aristoxenus, as it is known, was not accepted by the academia and his views have been rejected as being based on “poor knowledge of mathematics.” In general, two tracks of thought started in the pre-Socratic era, dominant from Pythagoras and recessive from Heraclitus. Aristoxenus can be grouped together with Heraclitus: both viewed harmony as a dynamic force related more to human perception than to mere relationships of numbers.

³ Such adherence of music theorists to a non-psychological sphere is dictated by profound reasons. In comparison with their colleagues from precise sciences, music theorists deal with highly subjective matters. That fact carries the dangers of slipping into subjectivity and losing the strong grasp on reality. Therefore, theorists have always been cautious when dealing with music and imitated the objective style of mathematics and geometry.

Cognitive musicology has its own shortcomings as well. As seen in several recent publications, the cognitivists seem to start from ground zero in their search for the cognitive foundations of “sound relationships” (Krumhansl) and “tonal hierarchies” (Lerdahl), by establishing their terminology from *tabula rasa*. Although the field of cognitive psychology exists and its critique is not within the author’s expertise, the path to its concepts proves to be complicated for those not trained properly. To be a psychologist, one must have a solid educational background, preferably in medicine and psychiatry. It is in this sense the terms *tabula rasa* and *ground zero* are used here: they represent the large field that professionally trained musicians must cross before reaching the level of expertise in cognitive psychology.

Cognitive musicologists rely heavily on the metalanguage of sciences and approach music as an object of scientific research, often following the tradition of the natural sciences of the nineteenth century. Thus, the reliance on experiments with fixed objects, the criterion of repeatability, and the quantification of results, all the postulates of science that seemed undeniable before the revolution of Karl Popper [3] mark the practical and empirical achievements of cognitive musicologists.

Thus, the conclusion with regard to the current output of both fields is not optimistic: both approach the same subject matter—the creative design of music—but the non-convertible character of terminology of each field remains an obstacle in communication and mutual enrichment. This chapter will be an attempt to bridge this gap.

Well-established concepts by Carol Krumhansl, Fred Lerdahl, Eugene Narmour, and David Huron, enclosed in rigorous scientific terms, comprise the core of cognitive music studies. Their basic assumptions rely upon a seemingly unequivocal phenomenon of what is commonly perceived as *musical structure*. However, in practice, cognitive musicology deals with “single tones, and chords and keys in notated form,” as suggested by Krumhansl in her seminal treatise *Cognitive Foundations of Musical Pitch* [4]. She searches for systematic relationships between these basic elements based on cognitive principles; indeed, she creates a new system in the place that has already been occupied by the corpus of knowledge of music theory. These elements are viewed in the same way precise sciences approach, typical of natural phenomena. However, due to the specificity of music, none of these elements ought to be analyzed as an objective matter. Tones, chords and keys are not natural phenomena: they are theoretical concepts. Their analysis inevitably involves reflection on their origins: Music theory since its inception 2600 years ago (if one considers pre-Socratic Greek thought as a point of its departure) has been preoccupied not with “single tones, and chords and keys in notated form” but with larger categories, such as harmony (Giuseffo Zarlino), divine numeric proportion (Pythagoras and his followers), harmonic function (Rameau, Riemann), formal function (A. B. Marx, Igor Sposobin, William Caplin) melody, rhythm (Riemann, Kholopova, Lerdahl and Jackendoff), genre (Heinrich Bessler, Arnold Sokhor), syntax (Riemann) and form (in reference to Plato’s theory of ideas, morphe, eidos, idea). Also, the “notated form of music” was developed by music theorists, monks who were preoccupied with fixing in time the elusive substance of Gregorian chant, from which music notation evolved throughout the millennia. Notation is not a natural object either: it is a trace of the acts of analysis and theorizing, also serving as a kind of

mnemotechnique – since Notker Balbulus, notation was used as a tool for memorization and storage of musical information.

Single tones of the scale are not a gift of nature either, although they are related to natural phenomena in a mediated way (say, to the frequencies of the natural overtone series, the phenomenon that has been described only by the end of the seventeenth century by Joseph Sauveur). All the attempts at relating them to natural phenomena failed. Rather, the tones of a melody are the product of the human psyche, subjected to theoretical reflection. Likewise, music theorists formulated the concept of *sound*, fixed in height and time, already 2600 years ago (*tonos* of early Greek theory), the systematic relationship of two such sounds (*systema*, Pythagorean tradition), a line-up of such intervals (tetrachord, *systema teleion* in a number of Greek approaches of Classical and Hellenistic periods), and, centuries later, the simultaneous relationship of more than two sounds (*armonia*, *cordo* in Zarlino *Istitutioni harmoniche*, 1558). Before cognitivists were able to set their experiments with the “seven scale-steps” music theorists, composers and performers also battled for centuries over the ideal tuning system. All in all, theorists made massive references to philosophy, mathematics, geometry, and politics. In consequence, music theory, backed up by philosophy, has reached a level of understanding by which the notational signs (notes, scale steps, scales, intervals) are not treated as objects; rather, they are signs that refer to another plane. According to Ferdinand de Saussure, the nature of a sign is arbitrary; the nature of the signified is not [5]. In this regard, the fundamental difference between a note and what it signifies does not allow to deal with the systematic relationships of tones in isolation from the plane of musical reality. That is, even if there is the adjacency in a printed score, it may not exist in musical space. For example, the leading tone and scale step V are not located on one and the same plane. They may look comparable on paper in the score, but that is deceiving. Scale step five is part of a quartal-quintal structure from which the Western concept of *mode* was developed. It is the invention of the fifth century B.C.E that occurred in ancient Greece. The leading tone, by contrast, is an adjacency that emphasizes the tonic. It appeared in music history in the eighteenth century, first in France as a *tone sensible*. These two participate in different processes and belong to different phenomenologies, although, at a certain point in history of music, they were united, despite their heterogeneity, into one entity, the dominant seventh chord. That makes graphic representation of them both in the same place misleading. This also means that the dominant seventh chord is internally contradictory. In general, music theory has reached the point of realization of the fundamental binary opposition of *res extensa* and *res cogitans*. Discussed by René Descartes in his *Principia philosophiae* [6], this distinction allows to separate the plane of reality from the transcendental one, or the plane of immanence according to Gilles Deleuze [7].

In its well-developed stage, music theory has switched its attention to musical work and musical composition as a whole, hence the title of the discipline in the nineteenth century: *Die Kompositionslehre*. In the course of its evolution, music theory has reached a high level of abstraction and is also engaged with the questions of musical content, such as expression, semantics, rhetoric, narrative (Eero Tarasti, Robert Hatten) and dramaturgy (the latter is developed by Russian theorists, such as Boris Asafiev, Victor Bobrovsky, Tatiana Tchernova etc.). These are by default large-scale categories, placed on top of a highly systematic and hierarchical body of knowledge. In this regard, what cognitive musicologists work with is related more closely to the subfield of early-stage pedagogy, also known as elementary music theory or

rudiments.⁴ As a result, neither the chosen subject matter, nor the conclusions of cognitive musicology rise up to the level of discussion of music common in traditional music theory.

Yet, there is a paradox: music theorists while teaching large-scale complex categories find themselves in need of cognitive explanations. For example, why does the dominant seventh chord exhibit the dominant function and which neural networks are engaged in resolving it into tonic? General scientific logic suggests an easy answer: the brain largely reacts to expectancy based on the cadence as an overlearned schema. Yet a deeper experience of music making may lead to the opposite explanation: Although it is the dominant that constitutes the cadence, its function exceeds that of marking the end of a musical statement. Dominants appear in the beginning and in the middle and at the end of a harmonic progression. Moreover, dominants constitute not only tonal definition but the modulation, the strategy opposite of the formation of the key. Even more, the overall direction of evolution of tonality in the nineteenth century can be labeled as ‘emancipation of the dominant’. In the late works of Scriabin the dominant function remains while its resolution is absent. Accordingly, the cognitive phenomenon of the dominant cannot be inscribed into schemata theory; it is heuristic, dynamic, and constitutive. This is another example of how general scientific ideas, hastily borrowed by musicians, can be misleading.

However, instead of switching the subject to “schemata theory” and “brain function,” it would be more practical to discuss how the resolution of the dominant to tonic would line up with the logic and syntax of the musical form. Yet the answers that theorists are receiving from cognitivists cannot satisfy the high demands of theoretical discussion of musical art. They are too elementary, descriptive, and misleading. In particular, since music theory has always been a prescriptive discipline, the goal of which was and is to guide a composer to the creation of new, never heard before, music. The books by A. B. Marx and Hugo Riemann, for example, contain very little abstract theorizing and analysis of existing scores. They provide musical examples only with the purpose of teaching a student the competencies needed to write music. Today, however, music theory is preoccupied with analyses of existing scores while the didactic element is reduced to a minimum.

Cognitive methodology appears to be even more deficient when viewed from the standpoint of contemporary continental philosophy. In the 1930s, Edmund Husserl

⁴ Of course, there is always an exception. Just as Aristoxenus in the Greek Hellenistic period, also the end of the nineteenth and the first half of the twentieth century has brought a magnificent maverick, the music theorist and psychologist Ernst Kurth. As he himself mentioned, his educational background was 50 percent in music theory and the other 50 percent in psychology. Therefore, his approach is truly interdisciplinary for that matter. Kurth, in contrast with post-WWII cognitivists, referred to the rich German tradition in philosophy and psychology. In particular, he lived in the era of phenomenology and psychoanalysis. His idea of energeticism in music describes best the late-Romantic harmony and can be stretched over all music history. Kurth was a psychologist who created a new approach to harmony, in particular that of Wagner. In contrast with contemporary cognitivists, Kurth did not adhere to the empirical method; his research was not based upon experiments. He and his passionate follower in Russia Boris Asafiev have created a hybrid discipline, a combination of music theory and psychology, called cognitive music theory, which the author is attempted to present in this chapter. Gustav Spet and Alexei Losev in Russia, Edmund Husserl, Roman Ingarden and Carl Gustav Jung in the West formed the context for Kurth's and Asafiev's thinking. However, both Kurth and Asafiev are dismissed today by Western and Russian traditions consecutively. For cognitivists their theory is too theoretical; for theorists it is too psychological.

declared a “crisis of European sciences” and confronted such crisis with the idea of phenomenology. Husserl criticized not only the method but the very goals of scientific research. In his seminal text *The Crisis of European sciences and the transcendental phenomenology* a subchapter 2 carries the title: “The crisis of sciences as the loss of its meaning for life” [8]. Remarkably, he mentioned that there had been a “lament about the crisis of our culture and the role here ascribed to sciences.” Husserl detected the “enigma of subjectivity” and the “enigma of psychological subject matter and method” [8]. These intuitions permeated nineteenth-century perception of the world through the medium of arts and humanities. Husserl suggested his own method—the *phenomenological reduction*. It deconstructs our habitual approach to reality defined by placing the results of immediate perception into a preexistent system of rationalization. These systems piled up in the course of millennial evolution of sciences. Scientists see not what is there but what they are trained and prepared to see. Logic and the language of terms also distort (or blur) the immediate perception. Husserl offered to clean that up. Husserl’s publication had overturned the field of European philosophy. It happened for the second time since Kant’s transcendental turn, compared to that of Copernicus. Husserl added to Kant’s initial idea that of phenomenological reduction, “going back to things themselves,” purified of any systematic imposition. It has been labeled ever since *the Copernican Turn* in philosophy. At this juncture, the most problematic aspect of an empirical approach to reality is the idea of observation. On the one hand, it seems simple and self-evident that, by observing things and processes, a human is connecting with them directly and immediately. In this regard, the best form of observation is a scientific experiment. However, what one seems to see directly is in fact contaminated by previous experiences, by the convictions of reasoning detached from experience, and by all kinds of beliefs and ideologemes. That is why phenomenologists and thinkers of that time avoided the methods of classical sciences of vulgar materialist kind and retreated to self-observation, or what the author of this chapter calls *gathering the circumstantial evidence*.

This issue cannot be disregarded, and the sciences cannot keep doing what they are doing without paying attention to this serious signal. If it proved to be urgent for the cluster of natural sciences, it must be even more important for the numerous fields within humanities.

In other words, both music theory and cognitive musicology may be enriched if placed in the true transdisciplinary realm governed by neither of them but relying on philosophical reasoning. Epistemological discussion of both fields, from an equidistant position, is the intention of the author of this chapter.

In the course of this text, three approaches to modulation in tonal common-practice style will be undertaken. First, Krumhansl’s encounter with modulation from chapter 9 of her book [4] will be offered; then, the music-theoretical concept of modulation from the nineteenth-century repertoire (Ernst Friedrich Richter and Hugo Riemann) will be provided. And lastly, the author will engage in the analysis of modulation in J. S. Bach’s *Allemande* from the French Suite no. 5 in G major as a hybrid of music-theoretical and cognitive approaches.

2. The core issue of cognitive musicology: Relationship with music theory

Cognitive musicologists rely heavily on the achievements of music theory. However, upon a closer examination, one can notice an important detail: cognitivists refer only to North America music theory of the post-World War Two period.

In her book [4] Carol Krumhansl relies on the views on music theory by L. B. Meyer as a final and authoritative source, while the history of music theory is much richer than this source could provide.

Thus, cognitivists have a limited view on the history of music theory, reduced to the segment from 1945 to 2000s and presented only in the English language.

It may come as a surprise to these researchers that the task of music theory has always been the description of compositional conventions in existing music. From the early Patristic era, through the Middle Ages, Renaissance, Baroque, Classical and Romantic periods music theorists provided in-depth conceptual commentaries to compositional conventions. St. Augustin and Boethius described the compositional conventions of the church chants; A. B. Marx and Hugo Riemann provided the prescriptive theories of composition for the nineteenth century. These descriptions require significant preparation, skills, talent, and investment of time and energy. The results of such descriptions are precious: they are commonly used by composers, performers, and educated listeners, none of whom ever complained about that.

Therefore, the questions raised by cognitivists would better be answered if they had started not from the views on harmony by L. B. Meyer or Walter Piston but from the ancient Greek categories of *tonos* (ὁ τόνος) and *dynamis* (δυνάμεις) used by Aristoxenus in his *Elements of harmonics* [9]. This would also immediately clarify the problems with current cognitive terminology. For example, the English indigenous term “pitch” is unique and untranslatable to many other European languages. It bears similarities, however, to Greek *tonos* as a source of energy and tension.

Recent research in music cognition represents a staggering breadth of topics and methods. Not all cognitivists side with Schenkerian dogmatics. Also, the term “computational” is sometimes added to the cognitive domain. Approaches range from traditional psychological questionnaires to neurobiological studies with the equipment borrowed from medicine and physics. Music under discussion is also unlimited in style and geography. Cognitivists allow crossing fields without any concerns. Thus, Jamshed Bharucha discusses *melodic anchoring*, using the terms tonic and dominant on the examples from Indian music: “In Indian music, it is typified by the performer dwelling on the semitone on either side of the Sa (tonic) or Pa (dominant) and finally resolving to the Sa or Pa with dramatic effect” [10]. In the course of evolution, however, cognitive musicology identified its own specific topics, such as, say, the speed of recognition of popular songs, researched by Carol Krumhansl [11]. One of the most successful and promising directions is the use of the *implication-realization model*, introduced by Eugene Narmour [12] and followed in a number of publications in cognitive musicology. For example, E. Glenn Schellenberg provides a convincing account of melodic expectancy using Narmour’s method [13], while Elizabeth Hellmuth Margulis, in her article about melodic expectation [14], provides the practical results of application of Narmour’s theory. This idea matches very well with the millennial history of music theory and serves both music theory and cognitive science rather efficiently. Narmour’s starting point in his book is that there is a genetic code in music that allows for “prospective realization of melodic expectations” [12]. In particular, “A + A implies A; A + B implies C” [12]. In other words, there is a hidden layer of structure and process in music that works on the level of implication rather than on the level of real representation. This, indeed, allowed Narmour to claim an alternative to Schenker’s theory which is completely dependent on graphic representation. It is not a coincidence that the same idea of an implied level in music was introduced some 260 years earlier by Jean-Philippe Rameau:

“By using the word ‘to imply’ [sous-entendre] we must be warned that the sounds can be heard in chords where they are not found. And even, with regard to the Fundamental Sound, it must be imagined that it ought to be understood below the other Sounds, when it is said to be implied” [15].

This rather profound observation that puts Rameau into the category of one of the earliest cognitive musicologists opens a path which any music researcher should follow if someone desires to reach tangible results. This is music’s specificity: most of its structure and processes are hidden from direct observation and verbalization. Rameau suggests, literally, that some tones of the chords are there while being “below the radar” of hearing. This – Rameau’s discovery – was just the opening step in the direction of expansion into the *hidden* world of music. Later on, theorists realized that musical function is also implied. In other words, scale steps and chords are visible; functions are invisible.

Scale steps, intervals, and musical functions when discussed in their visible and invisible aspects, do not line up and do not fit into a single table or a graph. There is a cluster of phenomenologies, each describing a group of musical events. Each group is systematic within itself but heterogeneous and rhizomatic in relationship with other groups. The list goes as follows: (a) scale step 5 is visible and real; its dominant function is invisible and non-tangible; it forms a group of its own, together with the seventh chord on top of its base; (b) scale step seven exists as a notated symbol and as a frequency; the leading tone is implied. There are many more implied differences, for example, (c) the major and minor seconds belong to a domain which is categorically different from that of the perfect fifth, perfect fourth and the octave. It was noticed by Aristoxenus who suggested that all intervals smaller than the fourth are *diaphonous*, and the fourth and the fifth are *symphonous*. These precise definitions suggest that the seconds are passing (*dia*) and the fourth and the fifth harmonize with each other and can be present in the same context (*syn*). In this regard, Krumhansl’s graph in **Figure 1**, that presents all the tones of the diatonic scale on the same Cartesian two-dimensional plane, is simply inadequate:

In this diagram, the continuous lines tie together the scale steps and intervals that do not belong in one table and one Cartesian two-dimensional plane. In a certain sense it is true from the universal position of an observer with standard education, however, when considering the differences in phenomenologies mentioned earlier, each point on these graphs comes from a different place. Thus, it is reasonable to assume that the space that this graph tends to map is not homogeneous but heterogeneous. It may require not Euclidian and Cartesian abstract planes but a curved ambient space (terminology by Nikolai Lobachevsky [16]) or a manifold and tensors (terminology by Bernhard Riemann [17]). In general, interdisciplinary studies often abuse the geometric abstract space delineated by the axes of abscissa and ordinate, using it for purposes remote from the original mathematical context. There are even more implied differences: (d) an interval of a second has two ontologically different versions: melodic and harmonic. The former is euphonous and consonant (because a melody contains mostly stepwise motion); the latter is dissonant (tones taken simultaneously as an interval of a second, major or minor, produce the least euphonous sonorities). (e) Moreover, the difference between a major second and a minor second is not merely quantitative: the smaller second has the capacity to constitute the leading tone; the larger one lacks such a capacity. (f) The thirds and the sixths present a special category. Their function is comparable to that of a glue or a filler. All the cases from (a) to (e) are real species (genres) of scale steps and intervals. They were constituted throughout the history of music and in the evolution of theoretical

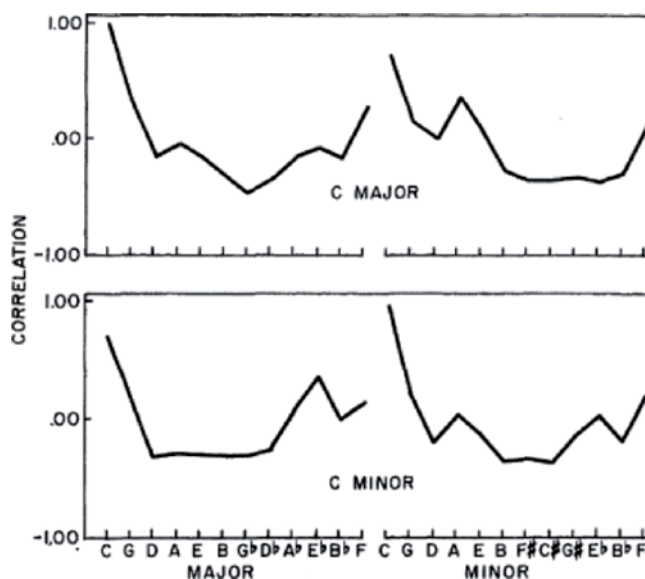


Figure 1.
 Krumhansl's graph of "correlations between the harmonic-hierarchy ratings in a C major key context."

views in different places at different times. The heterogeneous phenomenologies, listed from (a) to (e) are supported by various musical styles and compositional techniques. Further distinctions must be made in terms of (f); for example, the interval of the sixth in the era of fauxbourdon and in the nineteenth-century Romanticism carries different phenomenologies.

The list of implied structures and processes in music may be continued. One thing is remarkable: while these implied and subliminal aspects are carelessly rejected by science-oriented researchers, they constitute the second reality for well-trained musicians. They play the role of initial elements upon which a musician can build the structures and narratives of a higher level, such as musical emotion and the rhetoric and dramaturgy of performance. For a musician, trained in tonal-functional tradition, they are more real than notated signs in a printed score or numeric values of frequencies. Their existence is beyond questioning, while it is not communicable to humans with limited musical experience. It is important to emphasize that tonal-functional skill is the most expensive competency in professional music training, while the concepts derived from the background of general education, including some popular ones today, such as geometric (neo-Riemannian) and psychological, are rather cheap and cannot substitute for the latter.

Among all the variety of relationships within a cluster of adjacent fields, the regard of cognitive musicology on traditional music theory thus presents, perhaps, the most fundamental issue: Often, cognitivists treat music theory as the area of "common assumptions" while their own conclusions to analysis are considered final, objective, and rigorously scientific. However, upon a closer examination, the writings of cognitivists do not exhibit the novelty of approaches and methods and are pale in comparison with the achievements of "big science," say, those of Niels Bohr, Stephen Hawking, or Nicolas Bourbaki. This problematic condition necessitates not only an improved new approach to the old issues such as tonality, chord progression, and melodic fluency, but a great epistemological and methodological shakedown.

3. The problem of cognitive methodology: Considering the Husserlian warning of the crisis of European sciences

Cognitive musicology has made a step outside the language of music theory that went almost unnoticed. The change of the language of description, from the traditional musical to general scientific, was met with enthusiasm from the scholarly community. However, as already mentioned, philosophy in the twentieth century set forth a warning about the deficiency of the metalanguage of sciences. In this regard, noteworthy is Figure 1.1 from Carol Krumhansl's book [4] that lays the ground for her epistemological model:

Krumhansl suggests that there are two large categories in music perception, objective and subjective (I would rather use “more objective” and “more subjective” here). The objective one may contain musical *attributes* that are perceived and those that are not perceived, whereas the subjective category may contain psychological *experiences* that are related to music and psychological experiences that are not related to it, e.g. the listener's associations and memories. She specifies what cannot be perceived:

“There are undoubtedly features of music that can be objectively specified but cannot be perceived. This may be due to the listener's limited ability to discriminate along objective auditory dimensions” [4].

Evidently, these non-perceived elements are not related in any way to the idea of implied structures and processes proposed by Rameau and Narmour. Quite understandably, this model is then applied and replicated through the pages of her book as well as in a wealth of secondary literature on the subject. However, it remains unclear what these “attributes” and “experiences” are. As for perceived and not perceived, long before the emergence of cognitive science, the questions of perception were pondered by philosophers and aestheticians. Thus, perception occurs in steps, from the immediate phase to the more and more abstract. According to Immanuel Kant [18], there is perception and *apperception*. And the objects at hand are categorized by Kant accordingly: there are phenomena and noumena. So, returning to Krumhansl's diagram (see **Figure 2**) one may notice that she limits the non-perceived elements to the margin of errors of perception.

Direct and prompt perception of things is the core of Krumhansl's concept, its major assumption. Rather paradoxically, immediate perception has also been in the center of Husserl's and Merleau-Ponty's discussions. Yet what and how something is perceived immediately remains an open question. Contemporary cognitivists

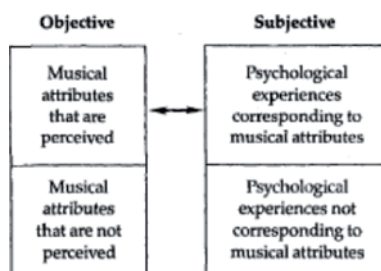


Figure 2.
Krumhansl's diagram 1.1.

returned to the earlier scientific beliefs, rooted in eighteenth-century experimental approaches in physics, chemistry and astronomy, in the possibility of direct observation. However, phenomenologists, and together with them the whole culture of the fin de siècle, were skeptical about such possibility. And the psychology of that time retained such an attitude. That is, although Carl Stumpf and his *Tone Psychology* (two volumes from 1883/1890) dealt with the aspect of perceiving single tone events as well as with perceiving consonances and dissonances for which he used the method of self-observation, he avoided placing a single tone into a system. Husserl, in his *Phenomenology of Time-Consciousness*, vehemently opposed to the idea of a system, when interpreting the perception of time and wrote about “moments of time” (*Zeitmomente*) that sink into the flux of time [19]. Sergey Prokofiev composed a piano cycle entitled *Fugitive visions*. He was influenced by the French impressionism in all its manifestations. This seemingly strange refusal to put perception into an iron-clad system was profoundly well-thought off. And, similar to Stumpf’s, the predominant method of psychological studies was ‘self-observation’ or ‘introspection’. This explains the principal difference between the empirical methods used by contemporary cognitivists, allowing gradual assessments in rating experiments using a 5- or 7-point Likert scale, and those used by phenomenologists.

4. The choice of the object of study: Well-trained professional musicians vs. untrained individuals

One can easily notice that Krumhansl and colleagues measure the reactions of listeners that are not properly identified. They write: “One subject reported having taken a ‘very basic’ introductory course in music theory, but the remaining subjects had no formal music theory background” [20]. In the reality of the musical scene in 1970s, undergraduate students in music “with no formal theory background” presented a true problem for the system of education. It is very difficult to imagine how these students had progressed in performance (Krumhansl suggests that they had on average 10 years of *performing* experience) without a formal music theory background. This condition that had to raise the brows of teachers is taken as an appropriate status for the study of perception. If the members of such a control group react to, say, the sound of scale step five, this may mean many things, from precise identification of the dominant function to a random pick of an unrelated tone, since it is not unusual to see among the respondents individuals with very low levels of aural skills. In other words, every recipient has a unique history and evolution of musical experiences and theoretical knowledge. For a group of musically untrained teenagers the sound of the dominant seventh chord can be perceived as an event of low level of importance. The formulation of the task in Krumhansl and Kessler is also rather unclear: “the subjects were instructed to rate on a 7-point scale how well, in a musical sense, each test tone fit into or went with the musical element just heard” [20]. First of all, the act of perception, an individual reaction to a phenomenon, is so fragile and transient that no *instruction* should interfere with it. Asking to *rate* a musical event is an intrusive and inappropriate question. The researcher forces his or her music-theoretical views and beliefs on the respondent. At least, such is the main objection of phenomenology to the methods of experimental psychology, as expressed by Husserl, Heidegger, and Merleau-Ponty.

Following this, the results of the so-called ‘reference group’ (or: control group) should be put into question. Who are the participants of experiments and what

kind of empirical data is expected to be gained from such experiments in cognitive musicology? In general, considering historic and cultural criteria that are commonly dismissed by many cognitivists, it would not be a surprise to assess the contemporary situation of hearing tonal music as rather deficient. Students at the conservatories today hear much less in tonal music than their peers at the end of the nineteenth century. Despite all these unanswered questions, cognitive musicology today applies the tools of measuring musical competencies by accepting the severe cognitive constraints of various control groups. Although actual experiments with living individuals for, say, the late eighteenth-century Italian tradition of *partimenti* (Italian eighteenth-century practice of composition and aural skills acquisition) are impossible, a method of analysis exists, termed *circumstantial evidence* in this chapter, by which the level of skills of a certain period *can* reliably be evaluated: In the absence of living experience of a certain kind, a scholar can engage in detective work searching for the traces and signs that may allow restoring such experience. Such residual evidence may remain in both the score *and* in the psyche of the researcher. It is the latter that Husserl, Stumpf, and Kurth used to define ‘perception based on self-analysis’. However, the traces of lost experiences can be found in the objects outside of the psyche. They can be discovered in the printed score, for example. This is what the author of this chapter calls *gathering circumstantial evidence*.

The difference between direct and circumstantial evidence used in sciences can be demonstrated with two examples. In 1831, the British physicist Michael Faraday experimented with copper wire and magnet. While moving the copper wire in vicinity of the magnet he discovered conductivity—the occurrence of electric charge. His experiment was based upon direct evidence. This is an example of a classical British empirical approach. A century later, Albert Einstein, John Wheeler and then Stephen Hawking discovered the black holes. They did it purely theoretically: they did not see them but deducted their existence from the general theory of relativity. In other words, black holes were discovered not because someone witnessed them in a direct empirical way, but because they must exist by judging the *contextual circumstances* provided by the new theories of space and time. In modern sciences, very few objects and events are visible and tangible. This indirect method is used in a number of sciences, including sub-atomic particle physics. The Higgs boson is imperceptible; however, it leaves circumstantial evidence after the collision with other particles in a hadron collider. The limits of the Universe are studied in astrophysics today while no direct empirical evidence is available for the objects that lie beyond the horizon of 13 billion years which is constituted by the constant speed of light propagating after the Big Bang. The method of *circumstantial evidence*, despite being the opposite of immediate observation, can bring solid results. The author will demonstrate the application of this method in Section 8 of this chapter, on the analysis of Bach’s Allemande.

5. Direct access to cognition by means of questionnaire vs. gathering of *circumstantial evidence*

Mental processes hide from classical scientific methods of experiment and observation. So, it would not be an exaggeration to suggest that *direct access to the mind is impossible*. As Heraclitus noticed, *physis kruptestai phylei*—nature likes to hide, and musical nature even more so. Music is, in essence, pure thought. On this, philosophers (such as Arthur Schopenhauer or Friedrich Nietzsche) and theorists (such as A. B. Marx, Igor Sposobin or Carl Dahlhaus) would be unanimous. German theorists

commonly use the term “thought” for music: *musikalische Gedanken* or *musikalisches Denken*. However, it is difficult to analyze thought by means of thinking. This paradox has been known to philosophers for centuries. However, just as for some processes in physics, music analysis can be done indirectly, by using *circumstantial evidence*.

It is very hard to tell what is happening in the mental apparatus (in psyche, or, in a narrower scope, in the brain) at the time an individual develops, for instance, a sense of tonality or a sense of modulation. However, a researcher may assume and reconstitute the mental processes that are taking place, first, in the psyche of a composer, then processed by the psyche of the performer, and finally, reaching the psyche of the listener, by observing the way how devices of certain kinds are used in the timeline of a musical composition, such as tonal-functional cycles embedded into metric-functional spans and motivic work. In this sense, music theory possesses the tools of verifiable cognitive analysis regarding mental processes during *creative processes*, on which the author will elaborate in detail in Section 8.

Noteworthy in this regard is Carol Krumhansl's setting for a *cognitive* experiment. She intends to gain objective results. For that she uses a questionnaire to compare the answers of a group of untrained musicians with those of a group of trained musicians. The questionnaire seems to be objective and direct. However, theoretically, every answer hides a long train of thought of the respondent that makes the results conditional and metaphoric. Such objection can be raised, all the more, when computerized equipment is used, making the respondent form his or her opinion in a split of a second, which is far from being accurate and authentic. The same objection applies to the significance of the results: it is related only to *passive* perception, while music is often termed as “making” rather than attending. The cognitive processes in the composer and the performer are also left beyond the inquiry.

Krumhansl's task (the instruction given to both groups) is to identify and further conceptualize the idea of tonality. However, the rating paradigm used in this experiment is simply insufficient for this rather complex task. In general, a questionnaire in psychological experiments is one of the most unreliable and deficient tools. When asked a direct question, a participant very often provides a quick and inadequate answer. Boris Teplov, the author of the seminal treatise *Psikhologiya muzykal'nykh sposobnostey* [*The psychology of musical competencies*] [21] chose a different approach. He conducted his research on similar matters with groups of highly trained students at the Specialized Music School, a branch of Moscow Conservatory. In this case, the existence of well-developed *active* skills in tonal music, including the aural definition of scale steps, intervals, keys, chords, chord progressions, and even modulations were known ahead of time. Teplov did not use the questionnaire method. Instead, he *observed* the performances of musicians in various age groups and made his conclusions using *circumstantial evidence*. The researcher was present during the regular classes and lessons and made his presence as inconspicuous as possible. For example, he noticed that musicians of the ages of five to eleven demonstrated formidable skills in recognizing and processing melody; at the same time, they could tolerate an untuned or erroneous harmonic accompaniment. Toward the age of eleven, they became aware of inadequate harmonic accompaniment and, thus, were attuned to both melody and harmony. This observation was made on the examples of hundreds of performances of young musicians and disciples at the Central Specialized School of Music in Moscow. It allowed Teplov to come up with formidable and convincing results: children of the ages five to eleven are capable of melodic hearing, and this competency develops at the highest rate; from eleven to fourteen, melodic hearing slows down when the harmonic skills start to develop. After that, the hormonal surge

wipes out most of these competencies, but the small percentage that is left allows a musician to progress efficiently at the college age. Teplov's conclusions regarding the age-specific phases in the development of musical hearing were proved by pedagogic practice and have been used in music-theoretical disciplines in Russia ever since.

In contrast to that, most of the experiments in cognitive musicology of the 1990s (and that was the peak of the development of this subfield) relied heavily on psychological questionnaires. From the epistemological point of view, such questionnaires force the participants to verbalize and quantify the parameters of music during immediate experience in continuous form, after which the obtained average values are put into tables and graphs. The very idea of perception is thus simplified and left beyond reflection. A participant, from this point of view, must formulate his or her reaction in a direct, precise, and objective way. Based on the results of such answers, the scholars make conclusions on the essence and inner structure of perception while many mental processes are left out of consideration.

In contrast to that, as seen in successful cases of cognitive studies, *circumstantial evidence* can indeed be gathered from the observation of the natural, age-dependent performances (such as the approach taken by Boris Teplov) and from the analysis of the score. Boris Teplov never *instructed* the responders and never asked them to rate anything.

6. The concept of tonality in music theory and the problem of tonal hierarchy

Analyzing the *circumstantial evidence from the observation of performance* has shown that cognitive musicology is in a position to make significant progress in the direction of understanding *tonality*, which is a category from ancient times. In their earnest intention to improve the knowledge about this, cognitivists used terms such as “anchoring” or “nesting.” However, these are not true scientific terms; rather, they are vivid metaphors that try to illustrate the interpretation of the theoretical category of tonality. Tonality is very well-defined and proven through the millennial history of music. Theorists are at an advantage here because their daily job is not so much the abstract contemplation as the practical task of educating composers and performers. From this rich practice, theorists derived the idea of tonality. The word exists in a cluster of modern European languages; its etymology is rich and powerful. Going back in time, one may associate tonality with *tenor*, the prevalent chanting tone in monodic chant, and earlier, to Greek *ho tonos*. The latter unites several meanings from different fields, such as physics (tonos as force) and music (tension of a string).

Another term that contemporary cognitive musicology tries to apprehend in the study of tonality is *hierarchy*. Since there has been a precedent before cognitivism, the graphic analysis of Heinrich Schenker (who himself never used the word hierarchy but it was ascribed to him by his followers), Fred Lerdahl worked with the same type of hierarchy in his publications [22], namely, top-down or tree (catalog) type. It is the most common type of hierarchy in general scientific fields and in social sciences. For music, it is applicable only in the presence of the completed work, or parts of it, in notated fixed form. However, from the standpoint of the creation of music, that of the composer or performer, the process that it takes is not visual assessment and description but discovery, a heuristic and constitutive strategy. For such a strategy, a different type of hierarchy must be used. It is known as bottom-up, ascending hierarchy or, simply, *emergence*. With few exceptions (including Eugene Narmour's already mentioned book), this idea was not implemented in

the English-speaking music theory and cognitive theory after World War Two. However, there is a wealth of literature on this subject [23], and it is unacceptable that the concept of emergence has not been touched by any recent trends in music theory. It opens a wide perspective and promises remarkable results. Musical form grows from the seed, from *Keim oder Trieb* [24] in terms of A. B. Marx (1897). Remarkably, Schoenberg's argument against Schenker relies on this difference: Schoenberg supported A. B. Marx's idea and suggested that the form grows horizontally from the *Grundgestalt*. Emergence is the creative type of hierarchy. Movement from the lower level to the higher involves obtaining a new quality (*quale novum*) and presents a quantum leap. Consequently, each new level contains elements of new nature, in comparison with the elements on the lower level. This may reveal a major mistake of Schenker, who thought that on all three levels the single principle of adjacency is governing the structure [25]. In this sense, the level of tones is based on the attributes of tones. The level of intervals is an entirely new level; here such new qualities as dissonance and consonance appear for the first time. The level of chords presents another quantum leap. The dominant seventh chord is not equal in its quality to the number of its components (this is the rule of hierarchy of emergence type).

The problem of tonal hierarchy according to Lerdahl occurs even with the first exchange of functions: T – D – T. In Schenker and in Lerdahl, the second tonic is deemed to be the same as the first (the tangible evidence of notation seems to provide this proof). However, the first tonic goes through a process and is transformed under the influence of the dominant. Thus $T_1 \neq T_2$. Apparently, a tree-type hierarchy (top-down kind) relies upon formal Aristotelian logic. However, in music this logic does not work all the time in all places. The lack of identity of the first tonic to the second precludes building a tree of hierarchy in Lerdahl's style:

According to Lerdahl [26], the geometry in **Figure 3** describes the formation of functions (metric, tonal, or, eventually, key regions) on a higher level. The higher tonic (T) controls the events at the lower level. Top-down hierarchy is, by definition, that of control from the top. This visualization may seem convincing on paper⁵ (for paper-theory or eye-theory—*Augentheorie*), but in real space of music it proves to be irrelevant.



Figure 3.
 Tonal hierarchy in Lerdahl's style.

⁵ On a side note, Lerdahl's tonal hierarchy is based upon binary branching. However, tonality cannot be constituted by using only two elements, tonic and dominant. A key element, the one that triggers the generation of the dynamic cycle that constitutes a key, is the subdominant. Tonic and dominant are equidistant; they generate a false key definition since the tonic is related to its dominant as the subdominant to tonic. For example, C is tonic in regard to G and, therefore, the G is the dominant; yet C can be the subdominant in regard to G if we assign the role of tonic to G. This becomes even more obvious when the two triads, C major and G major, are played on the keyboard without the given notation. The ear must choose between these two readings, which is virtually impossible without other elements. Only the actual presence of a subdominant is capable to break this vicious circle. In contrast with Schenker's interpretation that downgrades the subdominant, it has been always present in all concepts of tonality since Rameau. The subdominant is as important as the tonic and dominant; it must be located on the same level with them.

On the lower level, on which the tonic interacts with the dominant and loses much of its initial identity, the events are not just a concatenation of graphic objects but a real clash of forces with an unknown outcome. The result appears not by formal logic but by the cast of dice, as a heuristic choice. This dynamic process is best described in the language of Hegel, as a kind of *synthesis*: the first relation T-D negates and neutralizes its components, and the remaining T *emerges* as a new element on the higher level:

The event in **Figure 4** presents the path for the creation of new music. The change is irreversible, and the differences are non-resolvable. This pattern of emergence self-replicates on the level of smaller and then larger forms, where the large T must find a new element—a new large D, obviously the product of a different interaction on the lower level. Ultimately, it reaches the level of narrative and *musical dramaturgy* (the latter is the term of Viktor Bobrovsky, see the *Functional foundations of musical form* [27], derived from writings of Boris Asafiev and followed by Tatina Tchernova, Valentina Kholopova, and others).

Hugo Riemann suggests that the functional cycle of the dominant resolving to tonic is supported and emphasized by the metric function of light-to-heavy pattern. This adds even more complexity and diversity to musical structure and to the idea of tonality. Thus, the nucleus of tonal orientation is a strange combination of two processes that originate in incompatible fields, pitch/harmony and meter. One may add to these two a parameter of motive (motive-theme-dramaturgy). Thus, the defining factor of tonality (and of tonal hierarchy of the emergence type) is a triple heterogeneous fusion of *meter-harmony-motive*, all working into an organic cell-like element.

In his book *Systematische Modulationslehre* [28] Riemann first provides his definition of metric-harmonic interaction within the form, he called “symmetry in a succession,” leading to his metric postulate of light and heavy time, which is also valid for harmony. He thus concludes that “these two factors, harmonic design of a cadence and metric symmetry generate what is called form” [28].

Similarly, the Russian theorist Boleslav Yavorsky [29] developed a concept of *intonatsia*—analogous to Riemann’s metric-harmonic nucleus that is based upon *predictus-ictus* and tension-resolution models. Again, this unit is not based upon discrete quantifiable elements. Rather, it is a continuous process, akin to Riemann’s “dynamic shading” [*dynamische Schattierung*] [30], the ebb and flow of functional energy. Ernst Kurth’s ideas of linear development and *Spannung* (viscosity, stretchability of musical material) [31] are pertinent in this regard. Needless to say that such a definition of tonality is much more advanced than anything contemporary textbooks and treatises may offer.

This unit—an element with one strong beat and one functional cycle covering one motive—is prone to self-replication. There is an explanation of this

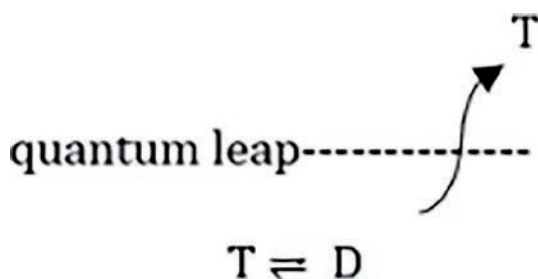


Figure 4.
Tonal hierarchy of emergence type.

phenomenon in Riemann [32]. Unlike Schenker's vague notion of a "biological urge to repeat" [33], Riemann's model carries a substantial explanation. In contrast with *chorée* (heavy—light pattern) that exhausts all the gained energy in its shallow and elongated second part, the iamb (light—heavy), since it ends abruptly, does not have time to expend all the gained energy. Therefore, an iambic cell carries an excess of energy. It is expended in the next cell, which begins with a higher level of energy than the previous. This way, the iambic model generates the line that unfolds with constant increase of energy. Hence, the urge to repeat and the dynamic character of Riemann's model. It contributes to the explanation of the "tonal hierarchy of the emergence type."

Thus, a tonal cell, or tonal sigh, has the ability to generate hierarchy. This hierarchy is produced by the even-odd pattern, derived from both, the breathing cycle 'arsis-thesis' and from the movement of human body (hence the most promising subfield of music theory is choreomusicology [34]). Thus, there is a tendency for hierarchy embedded in each level. Two patterns of two give one pattern of four. However, on each new level, the elements appear different. They gain a new quality. The light, mobile, dominant component on the basic hierarchical level relates to its own counterpart, the heavy, static component. On the higher level, the light, mobile and dominant component of the phrase of four (L), relates not to the first heavy component but to both light and heavy to the first cell; the same happens to the heavy component H on the higher level:

In other words, the higher-level light component (*L*) in **Figure 5** is in relationship with both heavy and light of the lower level. It implies an *L* to *l* and *L* to *h* correlation together with the correlation to a new entity, the synthesis of *l* and *h*. That changes the essence of relationships, in comparison with the simple *l* to *h* on the lower level.

When in a period form the first four beats line up with the second four, a new quality appears: the medial caesura, a cadence (provided that a musical theme consists of 16 bars). This makes the dominant of the third level function differently from both dominants of the first and the second level. After that, the cells do not simply replicate themselves mechanically. Therefore, the level of motives is based on one principle, the level of phrases—on another, and the level of smaller forms—again on another new principle. It is noteworthy that in this model, the beats, tonal-harmonic functions, bodily gestures and breathing cycles work together. This principle unites tonal hierarchy with that of dance movements and breathing cycles.

To summarize, the concept of tonal hierarchy is of ascending order, of emergence type. (A more detailed discussion of such hierarchy in music is available in my articles [35, 36]). It is based upon the levels with unique characteristics. The material on each level is comprised of elements of heterogeneous character. Their interaction within each cell is dynamic and processual. In this regard, simple quantifications and the application of top-down hierarchy do not serve music theory well. After all, composers, performers, and listeners want to know not how someone interprets an already composed and printed score but how music is being born, how to create music from

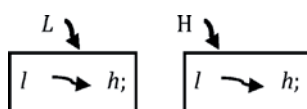


Figure 5.
 The hierarchy of emergence type in music, ballet, and breathing.

sounding elements. This creative task—the only justification of the existence of music theory—requires appropriate tools and methods.

7. Modulation viewed by music theorists and cognitivists

One of the most intriguing parts of Carol Krumhansl's research is her approach to *modulation* in tonal music. In regard to this, the author takes the liberty of adding the historic and theoretical background to this discussion. From a pragmatic point of view, as a method of teaching how to modulate and in complete disregard to the cognitive aspect, the act of changing key is rather simple and clear. As Ernst Friedrich Richter suggests in his textbook (one of the most commonly used in the second half of the nineteenth century):

“The concept of modulation: In the most ordinary sense, modulation is understood to mean the switching over from one key to another... Modulation occurs when a harmony foreign to the previous key occurs” [37].

Richter provides a simple example of such foreign harmony that switches the listener's attention from one key to another:

However, this description and the example in **Figure 6** are rather rudimentary. Here, one can witness how traditional music theory had been negligent to the complex psychological issues that accompany the process of modulation.

Hugo Riemann, the founding father of the interdisciplinary musical research and, together with Guido Adler, the creator of the systematic approach to music [38], provided a much more sophisticated interpretation of modulation. In his already quoted *Systematische Modulationslehre* he suggests the following definition:

“The will of the composer can intentionally force the tonality to shake or, even, for a short period of time, to become abandoned. If the composer will lead the dominant step on the heavy time, then the tonality of that chord (that will become temporarily the new tonic), will be emphasized” [28].

“The law of any modulation: a chord that occupies a certain moment in a cadence turns around (in German Umdeutung) and becomes a moment in a cadence in another tuning (key)” [28].

In other words, modulation is an intervention into a tonal-functional cycle T – S – D – T in which the penultimate D falls in the anacrusis and the last T is placed on the strong beat. If there is an intervention into this hard logic of musical syntax, akin to the generative grammar and the logic of verbal expression, the meanings of the



Figure 6.
Modulation (Ausweichung) in Richter's interpretation.

elements change and the whole system turns around. This aggressive event breaks the flow of syntax; it “shakes” the key or even forces the listener to abandon it. Rather paradoxically, the traditional music theory in its highest manifestation offers much richer material for cognitive studies than that used by the contemporary cognitivists themselves. For example, Riemann insists on the metric positions of harmonic events as the determinants of their meaning [30]. In the symmetry of succession, the parts are not equal in their “weight.” He also suggests that function in a cadence generates musical syntax and a sense of tonality. There is a discussion of the different roles of the dominant and subdominant in the system of tonal relationships. He also offers some insights into the phenomenon of *Umdeutung*, turning one meaning of a chord into another. All these truly perceptual and cognitive phenomena inherent in the theoretical realm await interpretation.

Riemann actively supports the inclusion of the study of perception and psychology into music theory:

“musical content is formed by melodic, dynamic and agogic rising and falling, those which carry the imprint of the spiritual conditions that generated them” [28].

Perhaps, this statement refers more to the condition of a composer and, to a lesser degree, to that of the listener. However, in his other writings and in the monumental plan for a systematic music scholarship, Riemann does not hesitate to delve into musical psychology and the perception of the listener. Thus, Riemann’s model of tonality and modulation is self-sufficient. His theory has been unfairly left in oblivion. It deserves revisiting.

8. Case study: Three approaches to modulation

All things considered, even with Riemann’s clairvoyant account of tonality and modulation, all the cases presented above in this chapter fall short of the analysis of modulation in a real musical masterpiece. They all, even the idea of tonal equilibrium and metric period, belong to categories of *theoretical* models, those that may or may not work in the application to a real musical composition. The following is a fragment from my paper entitled A Long Path to a Closely Related Key given at the European Music Analysis Congress (EUROMAC9) in Strasbourg, 2017 [39]. It goes without saying that any good musician, and especially a composer of tonal music, must study the theory of modulation in all its manifestations, presented above in this chapter. However, first and foremost, the following will be an example of the ‘hybrid’ approach that I called, in the title of this chapter, the “cognitive music theory.” Indeed, the analysis of modulation in a real composition must be able to demonstrate how the hybrid of cognitive musicology with the current state of music theory can bring about a new subdiscipline. The following discussion is intended to clarify the probable path of thought of Johann Sebastian Bach when he wrote the Allemande of the French Suite no. 5. He, for sure, knew how to render a modulation from G major to D major. In fact, it takes not more than a few chords to do it, according to the most rigorous textbooks:

A shown in **Figure 7**, in line with Riemann’s functional model of tonality, a modulation consists of three stages: first, establishing a key by means of a tonal cell, a cycle T-S-D-T; then the introduction of a chord that changes its meaning (*Umdeutung*); leading into another tonal cell, a new functional cycle T-S-D-T in a new key. Bearing in

G: I V_4^2 V_5^2 I vi_5^2

D: ii_5^2 $V_4^2 = V_7$ I

Figure 7.
A standard example of diatonic pivot chord modulation from I to V.

Figure 8.
J. S. Bach's Allemande from his French Suite no. 5. Nine gestures used for modulation from G major to D major.

mind this simple device, one may be surprised that it took J. S. Bach nine gestures and a page of music to accomplish just that – a modulation from tonic to the dominant:

This remarkable opening is rich in many parameters. For the sake of focusing on only one of them, the segment shown in **Figure 8** is viewed only as an extended harmonic progression. Modulation created by Bach is exemplary. Despite stylistic differences, greatest achievements of the same kind can be found in the music of the following periods.

The returning modulation from dominant to tonic is even more complex (see **Figure 9**). What caused Bach to apply this strategy? Neither a theoretical model of modulation nor empirical data can clarify this issue. The example above provides undeniable evidence that Bach was not preoccupied with devising a correct and functioning modulation. He was not trying to compose a modulation. Bach had a different problem at hand: the ear of the listener of his time. On the one hand, the French suite is included in the collection of keyboard exercises (*Klavierübungen*), thus one may assume that Bach intended it for learning by his students. On the other hand, as with Goldberg variations and other masterpieces, Bach undoubtedly expected to see among the listeners some connoisseurs and skillful musicians. A feature that distinguishes the eighteenth-century German listeners from those of our time is the deep knowledge of rhetoric. The listener of that time was attuned to the perception of rhetorical devices and gestures, mostly related to the Biblical text. Thus, music was

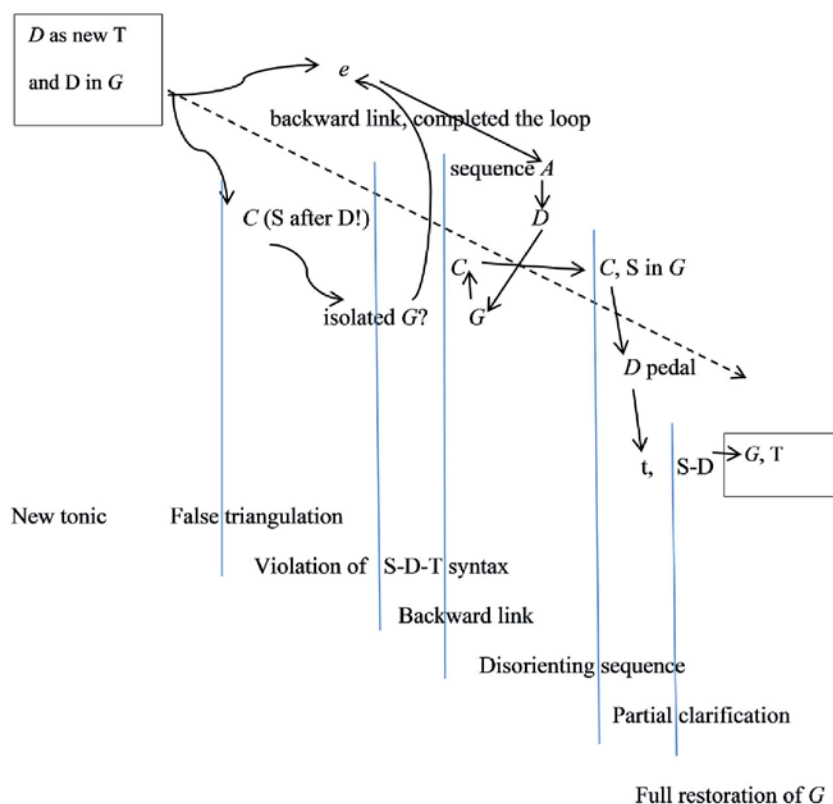


Figure 9. Schematic representation of the returning modulatory path in the Allemande. (The respective score is not depicted).

perceived through coding-decoding procedures. This was especially true for a learned musician, “musicus,” a term used by Johannes Mattheson (1681–1764). Also, music making was practical; listening was inseparable from composing and improvising; most of music was not written down and, in general, attending to well-established works in large concert halls in a passive listening mode was not in fashion. One way or another, the listening tradition of his time presupposed profound attention to the musical ideas. This may explain why Bach created these nine devices as nine suggestive *gestures* of the orator. In other words, the partitioning follows the limits of tonal-functional cycles and commonly known devices, such as sequences and cadences. Considering this, a theorist is capable of disclosing the intention of a composer: Bach may have prepared this score as a means to capture and manipulate the attention of the listeners.

Let us start the analysis of music in **Figure 8** with this strategy. As it is common to hear from the theory teachers, Bach begins by “showing the key.” That means, taking the listener through the tonal-functional cycle. Noteworthy is the fact that Bach uses the progression T-D-T-S-T-D-T in gesture 1, with the ten-ten-ten voice leading standard (I – V4/3 – I6 – IV – I6 – V4/3 – I)⁶ extended to seven chords. In this progression the subdominant is an absolute necessity. Without it, the triangle that allows to locate the tonic in the aural space cannot be formed.

Bach is a great master. He is also an excellent teacher, and that means patience and didactic consistency. He seems to understand that one tonal-functional cycle, although in theory sufficiently presenting the key, did not do it sufficiently for the listener who attended this music without having a score in hands. The repetition of the same syntactic pattern in Bach’s music is directly related to the studies in cognition, of learning processes in particular. Just as in the acquisition of a verbal text by ear, whereby even a native speaker may lose significant parts of information, in music the listener’s grammar proves to be much less advanced and flexible than the composer’s grammar. It seems reasonable, therefore, that Bach takes us through the “showing of the key” cycle again in gesture 2; this time by digressing into the dominant of the dominant, then bringing the dominant and tonic via resolution of V4/2. He does not hesitate to enrich each cycle with a particular voice leading solution. And yet, Bach’s listeners are still not fully convinced that the key is G major. Bach takes us through the cycle for the third time, in gesture 3. This may seem excessive for most trained musicians of the twenty-first century. However, Bach’s music has made its path to the hearts and minds of millions of listeners; he procured the most eloquent showing of the key, just to guarantee that even a less proficient listener will enjoy his music. These three presentations of a key instead of one reveal a secret of Bach’s composition: he knew what *cognitive constraint* means. As a successful pedagogue, he witnessed just how stubborn human perception is and how it proves to be tenfold worse when it

⁶ In the current North American music theory, the passing V4/3 is treated as “non-functional” and the whole progression is considered to be purely “contrapuntal.” This strange attempt to divorce counterpoint and tonal-harmonic function is anti-historical: both were in place in Bach’s time. That is, for a functionally trained ear of a musician raised in the European continental conservatory tradition, even the smallest passing case of the V chord presents the dominant function. And it is not a structure according to paper theory (*Augentheorie*) but a cognitive event: it appears on the level of sensations, akin to experiencing heat, cold, pain, etc. The dominant feels like a small piercing object, a shard of glass touching the skin. And in this case, the smallest dominant is the sharpest; it gives a strong impetus to harmonic progression. The fact that it is inserted into the most parsimonious voice leading amplifies this effect.

comes to recognizing musical structures and processes. The role of repetition in the acquisition of musical information has often been addressed in literature, for example by Peter Kivy [40].

Once Bach was done with showing the key, he engaged into the process of modulation. In order to start modulating, Bach needs to undo all the previous efforts on establishing the key. And that proves to be difficult: human perception is at a time stubborn and imperfect; at some other times it appears to be excessively efficient. A well-trained musician establishes the key in a way similar to a snapping mechanism. Once acquired, the evidence of an identified key is indelible from short-term memory. For such situations, by analogy and as the opposite of cognitive constraint, the author of this chapter suggests a new term: *cognitive excess*. While natural limitations of perception are well-documented, it may seem reasonable to touch upon the areas where cognition manifests the opposite case: ultimate capacity, exceeding the normative expectations. Such is the mechanism of immediately grasping an image. What has been seen cannot be unseen. The same applies to music: if the listener manages to hear and identify the key, it is extremely difficult to convince him or her to leave it. Hence, the modulation from tonic to dominant presents an ultimate difficulty for a composer: no matter how we try, the ear will always retain the hard link between tonic and dominant, and the new key will keep failing to present itself despite our vane attempts to ascertain a dominant as a new key, especially since *overlearned* processes regarding tonality and the cadence schema are effective.

This argument may explain what Bach does after showing the key. Instead of modulating directly to D major, he swerves from the path, going from tonic to submediant. However, from the standpoint of cognition, this common local key area in Baroque binary forms (going from tonic to submediant) needs an explanation. The job of a composer is to deceive the listener's image of tonic (to bypass the cognitive excess) and to disorient the listener in tonal space to the point that he or she would stop recognizing the simple scale step relationships. As an analogy, one may want to take someone to an adjacent room in a house without letting this person know that the room is adjacent. It makes sense to start by leaving the house through the window, traveling through the streets, and entering that room through a different door. A person guided this way would have an illusion that he or she is in an entirely different building.

The presentation in the key of E minor (gesture 4) is specific: it is an embryo of the future Subsidiary theme in Classical sonata allegro. It definitely takes the listener astray, not only key-wise but also by the change of mood and tenor. Here, if the composer wanted to be consistent, he had to return to section A (the rule of digression in Classical form: after the middle section one must return to the opening). Bach violates this principle and that allows him to disorient the listener. This is the work with real tonal space which is less abstract and mathematical than the model of tonal pitch space constructed by Lerdahl (see [41]). There are strange dark places in that space, and Bach took the listener to one of them. After that, in gesture 5, Bach adds the insult to injury: he takes us for a ride on a descending fifth sequence. In the beginning of the segment, we hear the resolution of the dominant to tonic in D major. This must have been the point of our destination. However, the cognitive strategies of a composer always deal with the proportion of time. Sometimes, a composer (and the listener) feels that something is happening too early; in other cases—too late. In this case, the appearance of D major is too early. It is not convincing and, evidently, it was not Bach's plan to stop the modulation here. By the inertia of sequential movement, circular rotation brings us for a moment to the key of G major. Hopefully, the

listener still remembers that this is the main key of the Allemande. He or she may see these two keys, initial and resulting, as the images through a side window of a car. The sensations of “home” and “not a home” mingle.

Gesture 6 offers a perfect authentic cadence in the resulting key. Here, the listener may give out a sigh of relief thinking that the journey is over. Yet, Bach, just as prescribed by Riemann, replaces the tonic on the heavy time of the symmetrical unit with the dominant (the tonic becomes the dominant seventh chord), and the modulating process is continued. At this juncture, a weary traveler does not even know what to think. The perspective is unclear; nobody knows where else this music will take us. At this moment, Bach reaches one of the most important goals in modulation: he succeeds in detaching the cognition of the listener from the anchoring structure of a key, thus defeating the “snapping to a grid” mechanism of identifying the key. The cognitive excess has been defeated: A listener proved to be able to unsee what he has seen and unhear what he has heard. And in order to keep his success, Bach takes us again to G major and E minor (gesture 7), as if returning to the very beginning of the Allemande, while preparing us for the last act in the drama of modulation. Bach kept the last turn in peripeteia in his pocket with the patience of a real magician. The unsuspecting listener is taken again to the dominant of dominant, but this resolves absolutely unexpectedly into D minor. Nobody could see this coming. This is a point of no return. It feels like a GPS playing the most undesirable trick on a tired driver at the end of a long ride. Bach steers out of this impasse. The harmonic pace accelerates to the limit; we get out of D minor using a same-name major and a triple dominant in G minor (G minor: E major – A major – D major). The key of D major appears as a long-desired figure; however, this desire has expired: the cognition is simply exhausted, and the listener perceives this key as a strange element of another game, unrelated to the roller-coaster journey of previous 11 measures. The long path to a closely related key can be presented in a graph (**Figure 10**; for a better comprehension the modulatory paths of the first and second parts are depicted separately, in the **Figures 9** and **10**).

There are more angles of view that a cognitive analysis must include. In fact, the analysis above is not intended to be of universal significance. Unlike the results of experiments in classical sciences (as in the case of Michael Faraday mentioned earlier), the analysis of musical work presents unique conclusions. Music theory simply follows the specificity of its object. In this regard, there are as many possible different analyses of the Allemande from Bach’s French suite no. 5 as there are professionally trained theorists.

Experiencing the formation of tonality and the escape into modulation is idiosyncratic. It depends on personal specificity and, to a greater degree, on the type of education. For example, many of the author’s colleagues are trained in the post-WWII tradition that canceled the fundamental role of tonal-harmonic function. For the author, in contrast, the functional cycle is not just an abstract concept, but the rule of perception of music. The author hopes that this rather specific interpretation may serve as one to foster future exchange with other analysts and other idiosyncratic readings. This analysis may be an example in favor of the method of introspection or self-analysis as used by researchers of the phenomenological tradition at the end of nineteenth century, discussed earlier in this chapter.

Bach and his listeners rely upon a complete array of tools offered naturally by cognition. Among them the most common are attention and memory. The gestures from 1 to 9 function regard these two faculties differently. In terms of memory, the most common types are used in this Allemande: working memory, long-term, and episodic. Thus, gesture 1 in **Figure 8** presents, rather obviously, fully concentrated

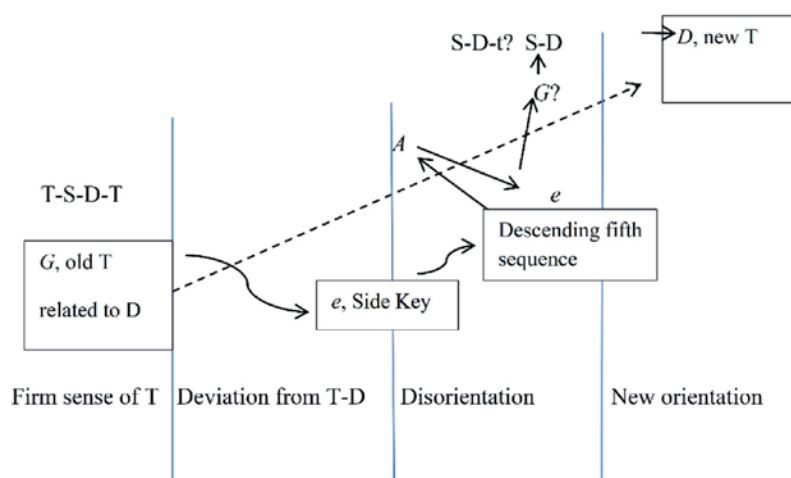


Figure 10.
Schematic representation of the modulatory path in the first half of Bach's Allemande.

attention and activation of working memory in a listener. Gestures 2 and 3, intended for establishing the key, are repetitions of the same tonal-functional cycle. As such, they require less directed attention. This is also beyond the time span of perceiving sounds in the immediate present. Attention here is relaxed a bit and the information gets stored in long-term memory. What happens next is much more eventful. Bach energetically interferes in the order of stored information by intentionally confusing it. In respect to memory and attention, gesture 4, a sudden digression into E minor, suggests Bach's intention to take the listener's memory away from the initial key. The listener, instead of closely following the line of the narrative, gets into the mode of "browsing" or "reeling," to use the jargon of contemporary clients of social networks. In the following gestures, Bach shows the listener only the fragments of both keys (G major and D major), so that the experience of perceiving them changes from true identification to superficial skimming through. The sense of key here is completely disoriented. A paradox of cognition, however, lies in the fact that the listener, through all this confusion, retains the sense of the initial key in long-term memory.

The composer keeps drifting away from the initial key until both the attention and memory of the listener hit a brick wall, with the sudden digression into D minor. Here the logic of syntax comes to a screeching halt, and the listener may experience a kind of cognitive collapse. The more trained the ear of a listener, the more acute the situation is with moving into a completely wrong place. Noteworthy is that for this type of structural listening, one must be experienced, and a musical piece should be presented several times. Such repeated listenings can happen in one session; more often, they are spread through the life of a musician, and their analytical resumes are piled up in memory. If so, a musician may listen to the Allemande only once (this time) and get all the necessary analytical results.

The modulatory path back from the key of the dominant to that of tonic is even more difficult (cf. **Figure 9**). Here, no matter how skillful the composer is in deceiving the attention of the listener, the hard link of 'dominant resolving to tonic' makes it virtually impossible to present D major to G major as a proper modulation. In most cases of the Baroque binary form, in the second half, one may notice a remarkable and persistent anomaly: after taking off with the dominant, the composer moves into

subdominant (G minor). This would be a grave syntactic error in any textbooks in harmony; yet it happens very regularly, as in the second half of the Allemande:

As shown in **Figure 9**, the path back is most convoluted. Bach intentionally introduces a false triangle I – V – IV – I in order to break down the natural and very strong perceptual logic of tonal syntax.

In terms of tonal hierarchy, the D major, reached in the middle of the Allemande, is entirely different from the dominant in the initial key of G major. Achieving this middle D took Bach all the skills in navigation through the murky waters of the tonal space. The D major is both, the dominant of G major in a larger-scale and a completely isolated auditory phenomenon. Only the *hierarchy of the emergence type* can explain (if possible, at all) the connection of the lower-level tonic and dominant to that highest degree of synthesis. The large-scale D does not have to be tied to the tonic by means of ‘voice leading’. It serves the highest level of musical logic, elevated from the surface.

This brief fragment of a longer analysis is brought up here to demonstrate the methodology. It is hard to call it “new,” since it relies on the achievements of European continental conservatory tradition in music theory, namely on *Kompositionslehre* which includes *Harmonielehre* (*Funktionstheorie*) and *Formenlehre*. This analysis did not provide an insight into the inner mental processes of the composer. The author simply observed numerous anomalies in comparison with normative textbook prescriptions. However, the novelty of this analysis and the method used should be seen in comparison with what traditional music theory is capable of. The latter mustered a voluminous corpus of knowledge on modulation, its functional logic and syntax. Traditional theory can explain the modulation in **Figure 7**. Still, for modulation in a real musical masterpiece as seen in **Figure 8**, traditional theory can only mark the event in E minor (gesture 4) as a “pivot chord.” However, logic calls to succeed the pivot chord with the cadence (functional cycle) in a new key. In Bach’s score, the “pivot chord” is abandoned and followed by a number of ultimately convoluted and twisted gestures that keep playing, back-and-forth, with initial and resulting tonics. (Regarding this, consider also the tone duration, i.e. the speed and acceleration of changes, effects that are only transient on the time-level of quarter notes—yet all these temporal aspects deserve a separate chapter). This obvious and massive phenomenon can only be interpreted in terms of *cognitive music theory*. This brings us back to the starting point of this chapter, the relationship between music theory and cognitive musicology. Perhaps, the point made here in this chapter reveals the discrepancy and acute necessity to unite (reunite) the forces of the two subdisciplines of music scholarships. What will be the result can be called *cognitive music theory*.

Thus, by using the method of *circumstantial evidence* applied to Bach’s work aligned with the listener’s perception, the author came up with notable conclusions. The author “questioned” Bach’s score on the subject of cognitive strategies. Bach took so much time and effort to modulate to closely related keys, because he composed not only a score of a musical masterpiece but also *a path for its perception*. And he was cautious and consistent in doing that. The overall appearance of this Allemande is brimming with excess: it is excessively embellished, excessively elaborate in modulatory strategies, and exhaustively didactic in regard of guiding the perception of the listener.

9. Conclusion

In this chapter, the author attempted to thread together some points that may give a new breath to two separate fields, cognitive musicology and music theory. By taking

the position of an independent observer, he was equidistant from both cognitive musicology and current North American music theory. What should have helped, undoubtedly, was the reference to philosophical reasoning. When two sides disagree on something, the role of the arbiter is commonly played by the philosopher. From the 1970s onward, musicologists with narrow professional training in this discipline endeavored to enter a new domain, that of experimental cognitive psychology. As mentioned above, this move proved to be costly: someone trained in one field cannot just start using the vocabulary of a different well-established field of knowledge without profound reflection on epistemological differences, even when using the terms from the metalanguage of science, such as 'system' or 'hierarchy'. Crossing the ground zero between musicology and cognitive psychology requires such reflection. In this respect, contemporary music cognitivists seem to overlook the phenomenological revolution of the 1930s (as shown above). The case of a preferential treatment of top-down hierarchy (TD type), while neglecting the bottom-up emergence type (BU), is also revealing. While top-down hierarchies are ubiquitous in scientific presentations, the bottom-up type, namely *emergence*, applies to creative, heuristic and constitutive conditions, thus also being appropriate for the art of music.

The author discussed in detail what path the listener's perception might take in regard of tonality and modulation which looked nothing like simply moving from tonic to dominant and back. One has to admit that the categorical apparatus of traditional music theory, its analytical discerning capacity, is capable of parsing the harmonic progression in Bach's Allemande according to the limits of tonal-harmonic cycles and modulatory techniques, such as sequences, digressions and pivot-chord modulations. However, while traditional music theory can help parsing the body of a musical composition it cannot explain the cause and the meaning of the sequence of events without the help from cognitive psychology. On the other hand, music cognition in its present condition cannot contribute much to this problem either. The cognitive processes in this Allemande are no easily available phenomena; music, indeed, likes to hide its deeper meaning and mechanisms. What is given in immediate experience, i.e., through gathering empirical data, cannot represent the deeper structure and processes in music. As was demonstrated on the example of scale steps and tonal-harmonic functions, their heterogeneity precludes positioning these elements on a single plane for further measuring and calculation. There are hidden, *implied* mechanisms that control the formation of a key and subsequent modulation. In this regard, cognitive musicology must make a step toward music theory and philosophy in order to reach feasible results in analysis and interpretation appropriate for a musical masterpiece.

The ontology of a musical work, its creation, performance, and perception dictate a retreat from the commonly used scientific methods toward self-study, introspection, and searching for information using *circumstantial evidence*. In the context of these methodological innovations, it seems reasonable that theory must join cognitive musicology in a more interactive and comprehensive way than today. Thus, a discipline of *cognitive music theory* should be established. This will be equipped for explaining the deep levels of musical structure, while teaching the composer, the performer, and the listener on how to deal with music in both speculative and pragmatic dimensions.

The analysis of Bach's Allemande, presented in this chapter, focused on the figure of the *composer*. It is possible to envision a similar discussion dedicated to the phenomenology of musical performance. Ultimately, both the composer and the performer are, first and foremost, listeners. So, in the third stage of creating cognitive

music theory one should return to developing a new understanding of the listener, by using a hybrid of music-theoretical, philosophical, and cognitive approaches.

One may come up with the following conclusion: in order to make music and music analysis more transparent, efficient, and intelligible, the composer must be preoccupied not only with his or her ideas and the technical concept, not with the execution of his or her skills, but also with processes directed at guiding the *perception of the listener*.

Author details

Ildar Khannanov
Peabody Institute, Johns Hopkins University, Baltimore, USA

*Address all correspondence to: ikhanna1@jhmi.edu

IntechOpen

© 2025 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Carpenter E. The theory of music in Russia and in the Soviet Union, ca. 1650-1950 [Ph.D. thesis, dissertation]. Pennsylvania: University of Pennsylvania; 1988. 1618 p
- [2] Jung C. Wagner und Hanslick. Kurze Geschichte einer Feindschaft [A short history of hostility]. Österreichische Musikzeitschrift. 2012;67(6):14-21
- [3] Popper K. The Logic of Scientific Discovery. London, England: HarperCollins; 1974. 480 p
- [4] Krumhansl CL. Cognitive Foundations of Musical Pitch. New York, NY: Oxford University Press; 2001. 318 p. DOI: 10.1093/acprof:oso/9780195148367.001.0001
- [5] Saussure F. Course in General Linguistics. Trans. Wade Baskin. New York: Philosophical Library; 1959. 264 p
- [6] Descartes R. Principia philosophiae (Online Resource). Available from: <https://www.gutenberg.org/cache/epub/4391/pg4391.html>
- [7] Deleuze G, Guattari F. Mille plateaux [A Thousand Plateaus]. Paris: Éditions de Minuit; 1980. 645 p
- [8] Husserl E. The Crisis of European Sciences and the Transcendental Phenomenology [Die Krisis der europäischen Wissenschaften und die transzendente Phänomenologie. Hamburg: Felix Meiner Verlag; 2012. 328 p.]. Evanston, Ill: Northwestern University Press; 1970. 405 p
- [9] Aristoxenus. Ἀρμονικά στοιχεία [Elements of Harmonics]. Transl. HS Macran. Oxford: At the Clarendon Press; 1902. 316 p
- [10] Bharucha JJ. Melodic anchoring. Music Perception. 1996;13(3):383-400. DOI: 10.2307/40286176
- [11] Krumhansl CL. Plink: "Thin slices" of music. Music Perception. 2010;27(5):337-354
- [12] Narmour E. The Analysis and Cognition of Basic Melodic Structures: The Implication-Realization Model. Chicago: Chicago University Press; 1990. 443 p
- [13] Schellenberg EG. Simplifying the implication-realization model of melodic expectancy. Music Perception. 1997;14(3):295-318. DOI: 10.2307/40285723
- [14] Margulis EH. Melodic expectation. Music Perception. 2005;22(4):663-714
- [15] Rameau JF. Traité de l'harmonie reduite à ses principes naturels [Treatise on Harmony Reduced to its Natural Principles]. Paris: Ballard; 1722. 494 p
- [16] Lobachevsky NI. Geometrical Researches on the Theory of Parallels. GB. Halsted (tr.). Chicago: Open Court Publishing; 1914. 52 p
- [17] Riemann B. Über die Hypothesen, welche der Geometrie zugrunde liegen [On the Hypotheses Which Lie at the Bases of Geometry]. Göttingen: Abh. Kgl. Ges. Wiss; 1868. pp. 133-150
- [18] Kant I. Critique of Pure Reason [Kritik der reinen Vernunft]. Gordonsville, VA: Palgrave Macmillan; 2003. 772 p
- [19] Husserl E. Texte zur Phänomenologie des inneren Zeitbewußtseins [The Phenomenology of Internal

- Time-Consciousness] (1893-1917). Hamburg: Felix Meiner Verlag; 1985. 255 p
- [20] Krumhansl CL, Kessler EJ. Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*. 1982;**89**(4):334-368. DOI: 10.1037/0033-295X.89.4.334
- [21] Teplov BM. *Psikhologiya muzikal'nykh sposobnostey* [Psychology of Musical Competencies]. In Russian. Moscow: Academy of Pedagogic Sciences; 1947. 335 p. Available from: <https://www.klex.ru/8nb>. French translation: Teplov BM. *Psychologie des aptitudes musicales*. Paris: P.U.F; 1966
- [22] Lerdahl F. Concepts and representations of musical hierarchies. *Music Perception*. 2015;**33**(1):83-95
- [23] McLaughlin B. Emergence and supervenience. *Intellectica*. 1997;**25**(2):25-43. Available from: https://intellectica.org/SiteArchives/archives/n25/25_01_McLaughlin.pdf
- [24] Marx AB. *Die Lehre von der musikalischen Komposition* [Theory and Practice of Musical Composition]. 5th ed. Vol. 1. Leipzig: Breitkopf und Härtel; 1897. 664 p
- [25] Schenker H. *Free Composition* [Der freie Satz]. Hillsdale, New York: Pendragon Press; 1977. 164 p
- [26] Lerdahl F, Jackendoff R. *Generative Theory of Tonal Music*. Boston: MIT Press; 1996. 384 p
- [27] Bobrovsky V. *Funktsional'nye osnovy muzykal'noi formy* [Functional Foundations of Musical Form]. Moscow: Muzyka; 1978. 332 p
- [28] Riemann H. *Systematic Teaching on Modulation as the Foundation of Teaching of Form*. In Russian. Moscow: Musical Sector; 1929. Original in German: Riemann H. *Systematische Modulationslehre als Grundlage der musikalischen Formenlehre*. Hamburg: J. F. Richter; 1887. 122 p
- [29] Yavorsky BL. *The Design of Musical Speech* [1903]. Bilingual Publication. Translated from Russian by Ildar Khannanov. Moscow: Kompozitor; 2018. 512 p
- [30] Riemann H. *Musikalische Dynamik und Agogik* [Musical Dynamic and Agogic. Paragraph]. In: Caplin W. *Theories of musical rhythm in the eighteenth and nineteenth centuries*. In: Cambridge History of Western Music Theory. Cambridge: Cambridge University Press; 2002. pp. 657-694. Ch. 21
- [31] Kurth E. *Romantische Harmonik und ihre Krise in Wagners "Tristan"* [Romantic Harmony and its Crisis in Wagner's "Tristan"]. Berlin: Max Hesse; 1920. 557 p
- [32] Riemann H. *Metrik* [Metrics]. *Musiklexikon*. Berlin: Max Hesse; 1909 [1654]. pp. 913-916
- [33] Schenker H. *Harmonielehre* [Harmony]. Vienna, Austria: Universal Edition; 1978. 460 p
- [34] Dorf SN, Minor HJ, editors. *Choreomusicology. Dialogues in Music and Dance*. Collection of Aarticles. New York: Routledge; 2025. 374 p
- [35] Khannanov I. Hierarchical structure in music theory before Schenker. *Res Musica*. 2017;**4**:92-111. Available from: https://resmusica.ee/wp-content/uploads/2017/04/rm3_2011_92-111_Khannanov.pdf
- [36] Khannanov I. *Music as Emergence: Non-Reducible Hierarchy in Music Theory*. Manuscript; 2016

[37] Richter EF. Lehrbuch der Harmonie
[Treatise on Harmony]. Leipzig:
Breitkopf und Härtel; 1860. 212 p

[38] Riemann H. Grundriß der
Musikwissenschaft [Comprehensive
Outline of Musicology]. Leipzig: Quelle
& Meyer; 1908. 115 p

[39] Khannanov I. A long path to closely
related key: Modulation in Bach's
Allemande BWV 816. In: Ninth European
Music Analysis Conference; June–July
2017; Strasbourg.

[40] Kivy P. Introduction to a Philosophy
of Music. Oxford: Clarendon Press; 2002.
296 p

[41] Lerdahl F. Tonal Pitch Space.
New York: Oxford University Press;
2004. 411 p

Section 3

Music Psychology

Chapter 3

Music and Emotion: Theory, Evidence, and Comparison with Literary and Visual Art

Keith Oatley

Abstract

Robin Collingwood, an English philosopher, proposed that a work of art typically begins with an emotion that a person externalizes into a musical, literary, or visual form, to understand it better. Another person who engages with the work can be affected by it emotionally and can also reach a deeper understanding of the issue that's involved. Music is a relatively ancient practice and can often occur in ways that foster an emotional connectivity, not just between the artist and those who engage with her or his works, but also among groups of people. It can prompt different emotions, such as happiness, sadness, anxiety, and anger; it can also enable changes to occur in listeners' personalities, probably by way of their emotions. As Steven Brown and colleagues have shown, some aspects of music—which they call prosodic—such as its loudness, pitch, and tempo, are shared with speech, while others, which they call tonal, such as major and minor scales are not shared in this way but can prompt emotions directly.

Keywords: music and the sister arts, emotions, understanding, connectivity, resonance

1. Introduction

“She played the first three of Bach’s six cello suites, and after a few bars ... I experienced one of those rare moments of absolute happiness that only music can give us” [1]. This quotation is by Ferdinand von Schirach, as he relates his experience-based story, “The cello.” It’s an apt beginning for this chapter because music—singing, dancing, playing an instrument, listening at a concert, and hearing musical accompaniments of films and television series—tend to prompt emotions.

In his book of 1938, *The principles of art* [2], Robin Collingwood proposed that a work of art is not produced by formula and has an outcome that’s not known in advance. It occurs when an artist is moved by an emotion that she or he knows is important but does not fully understand. This person then translates it into what Collingwood calls a “language” of music, words, or visual objects. In this way, the artist can articulate it and understand the emotion more deeply. The cellist in von Schirach’s story said about Bach’s suites that she played: “It’s the prelude to the first suite ... It’s all of life packed into three minutes” [1].

Collingwood continued his proposal by saying that those who engage with an artistic work do what the artist has done “over again for themselves” [2]. Specifically on music, Klaus Scherer and colleagues [3, 4] have done extensive research on how music can begin with emotions in human interaction with others, and how it evokes emotions in performers and listeners.

Starting with Collingwood’s theory, we move to the effects of music on interactions, and some of the processes involved. We then consider the pre-history and history of music, and how it can prompt emotions in humans, both indirectly and directly. We compare music with its sister arts—literary and visual—and evidence on what it is to be moved by artistic works.

In her book of 2019, *How art works*, Ellen Winner [5] says that typical requirements of works of art are “Skill and Versatility, Novelty and Creativity, Expressive Individuality, Emotional Saturation.” Art, including music, can be thought of as having four main aspects. The first is that proposed by Collingwood: art is generally founded in emotion and is capable of prompting emotions not just in the originating artists but also in others. The second, also from Collingwood, is that art is not mere amusement and that it is also unlike persuasion (as in advertising and political speeches) where someone wants others to think and do what that person wants. Instead, in engaging with a work of art, we are invited to think for ourselves. The third is that art is a mode of exploration. With it we can sometimes travel in our imagination into circumstances and modes where otherwise we would never have been. The fourth is that a piece of musical, literary, or visual art has been produced by an originator so that we, who engage with it, can also explore issues with which that the artist was concerned, that can concern us too. Having chosen to engage with a musical, literary, or visual work we can take it in, and make some aspect of it our own.

Pizzolante et al. [6] have added to these ideas the concept of *transformative experiences* that people find extraordinary, which prompt insights, and facilitate changes in people’s self-conception, views of the world, including their identity and social relationships. These researchers found art and esthetics as elicitors of transformation regardless of the type of artwork. Menninghaus et al. [7] say that people know that although a work of art is not real life, it is in the world, so they can separate it and hold it at a distance. In engagement with such works, we can feel both safe and able to embrace the new.

With a focus specifically on music Patrik Juslin and collaborators [8–10] have proposed that with music being a mode of abstract art that extends in time, it may seem difficult to see how it connects with emotions which are processes that often occur suddenly and have been essential for navigating human interactions and even survival. In 2025, Juslin [11] recapitulated his research and, drawing on an earlier proposal, suggested that the following processes occur in music, beginning each one with a capital letter as follows.

B. Brain Stem Reflexes: based on simple features such as loudness.

R. Rhythmic Entrainment: having people take part in rhythmic processes in a cooperative way, as in marching and dancing.

E. Evaluative Conditioning: alerts listeners to their wider learning history, their culture and sense of belonging.

C. Contagion: Starts with babies and lullabies; harmonizes moods during social ceremonies, mitigates feelings of loneliness.

V. Visual Imagery: enables people to visualize and feel close to absent relatives and friends.

E. Episodic Memory: enables people to re-experience significant past events and their bonding to loved ones.

M. Musical Expectancy: enables people to predict the musical patterns of other people, with a sense of shared goals and a similar musical history.

A. Aesthetic Judgment: prompts self-transcendent emotions such as awe, admiration, and gratitude, facilitating prosocial events such as helping and generosity.

2. Prehistory and history

Perhaps the most important psychological research of the Twenty-First Century has been that of Michael Tomasello and his research group (e.g. [12–14]). They have shown that a fundamental transformation in human evolution has been to *collaborate*. This means making arrangements with others and then fulfilling them, with “we” being able to become more important than “I”. Among the earliest indications of this is that we humans gather food together, take it to where family and friends are, and eat in groups. In 2020 a paper was published with evidence of the sharing of food from 170 thousand years ago [15]. In contrast, as Jane Goodall [16] has shown, although our closest primate cousins, chimpanzees, have hierarchies, look after infants, and make friends, they do not make new arrangements of the kinds that humans do. They roam together and, when they find a source of food, they grab what they can (with the alpha male generally getting first pick) then go off and eat alone.

Tomasello proposed that this new evolutionary change occurred in two phases. The first he called “joint intentionality”: making and carrying out arrangements with others. It included sharing of food. This phase, which probably began more than 200,000 years ago, also enabled the coming of human language, for instance in collaborative conversations that can occur while doing something else, such as gathering food or eating together. The second phase he called “collective intentionality”. It occurred when people came to live in groups and societies which developed what are now recognized as cultures in which people share a sense of what kinds of action are morally good and bad. This phase is a version of what Karl Popper (e.g. [17]) called “World 3,” which is distinct from “World 1”—the physical world— and “World 2”—people’s thoughts and emotions. World 3 is the products of thought such as scientific theories, stories, and myths. This world is created by human beings and changes continuously. It corresponds to the current state of human knowledge and culture.

In Tomasello’s collective intentionality, joint goals such as the good of the community, can become more important than individual goals. It was within societies that artworks, including music with its accompanying emotions, began to emerge, not just selecting and making of tools for practical purposes, but as a new kind of contribution to culture, able to reach not just those whom people knew personally but potentially many others.

The earliest surviving indications of art are visual: seashells drilled with holes to make beads for necklaces. These date back to more than 100,000 years ago [18]. Such necklaces were social; perhaps making a person more attractive. In terms of musical art, a flute has been discovered from 43,000 years ago [19]. From some 4000 years

ago archaeologists, for instance in Vietnam, have discovered instruments called lithophones, where “litho” means stone and “phone” means sound. These instruments are sets of stones in a line (not unlike modern xylophones) with each stone making a different sound when struck [20]. Contemporary with these, in the Old Kingdom of Egypt, musical instruments such as the harp were used along with the long-necked lute, which preceded the modern guitar. Bringing us towards modern times, Ada Lovelace who lived between 1815 and 1852 [21] realized that the machine on which she was working in collaboration with Charles Babbage—the “analytical engine” (the first computer; which was never actually constructed)—would be able to generate not just numbers but music. A helpful theoretical article on the history of emotions in relation to music is by Leonid Perlovsky [22] with a suggested split between speech and music. Speech tends to refer to objects, to what’s going on in the world, and to human interactions, whereas music is somewhat ambiguous and is perhaps more closely connected to emotions.

3. How music can prompt emotions

Johnson-Laird and Oatley [23] proposed that music can mimic and prompt emotions and moods in the following ways.

Happiness: medium tempo, loud, wide range of pitches in melody, major scale, consonant,

Sadness: slow tempo, soft, low pitch, small range of pitches, minor scale, mildly dissonant,

Anxiety: rapid tempo, moderate volume, low pitch, minor scale, dissonant,

Anger: rapid tempo, loud, high pitch, minor scale, dissonant.

When people watched early silent films in a cinema there was sometimes a pianist or organist who played music as the movie progressed: prompting audience members into these kinds of emotive states. When sound became available in movies, and then in television series, there was almost always musical accompaniment. When watching a film or a television at home, subtitles do not just let us know what characters are saying but often offer us a word about the kind of music that accompanies the action: “somber,” “eerie,” “foreboding,” and so on.

To find out more about how music affects people who watch pictorial images in cinemas or on their televisions Baumgartner and colleagues [24] used functional magnetic resonance imaging as a method to explore how music can affect people’s experiences when looking at pictures that had been selected because of their known properties of inducing fear or sadness. As compared with looking at the pictures without any sounds, which mainly produced activation in the prefrontal cortex, picture observation with congruent kinds of music activated many areas known to be associated with emotions, such as the amygdala, hippocampus, insula, and so on. The researchers concluded that combinations of pictorial and musical input are able to evoke strong emotional experiences.

A critical feature not discussed by Johnson-Laird and Oatley [23] has been found by Brown, Berry and Phillips [25] in a study in which they used functional magnetic

resonance imaging on a group of music professionals to identify brain areas that are activated with different kinds of music. The researchers' main finding is that there is a distinction between aspects of music that they call "prosodic" and "tonal". Prosodic features include modulations of pitch, height, melodic contour, tempo and amplitude. Music shares these features with speech. When, for instance, someone is angry she or he can speak quickly, loudly, with the voice sometimes at a higher pitch. Someone sad can speak slowly, softly, perhaps with the voice at a lower pitch. In contrast, however, music can also be tonal for instance with variations between a major scale, tending to be cheerful, and a minor scale, suggesting negative emotions. These modes — major and minor — do not occur in speech at all. They may have occurred with a further change in human evolution beyond the changes discussed by Tomasello and his group [14]. The distinctive areas that are activated with tonal music are the lateral frontopolar cortex and the dorsal anterior cingulate cortex. These areas both share activation features with distinct emotions, for instance happy and sad. Brown et al.'s article makes an important addition to the many neuroimaging studies of musically induced emotions which have received a recent meta-analysis by Fuentes-Sánchez and colleagues [26].

4. On being emotionally moved

Following from the proposal in the section (above) on "Prehistory and History" about humans having become collaborative, de Leeuw et al. [27] studied effects of listening to meaningful and pleasurable music on participants who listened to music for at least 2 hours a day. The researchers found that, as compared with listening to favorite pieces of pleasurable music, which participants had named in a questionnaire as making them feel happy and relaxed, listening to favorite pieces of meaningful music, which they had also identified—such as Chopin's Nocturne in E-flat Major—led to "stronger feelings of being moved. This included having a lump in one's throat and tears crying, more contemplation, a stronger motivation to seek what matters in life" [27]. When this happened there was a continuation of a sense of love for others with whom participants were close. These results confirmed several of the processes proposed by Juslin [11], which were presented in this chapter's Introduction, examples of which were: mitigation of loneliness, visualization of absent friends, and bonding with loved ones. De Leeuw and her colleagues also found that listening to both meaningful and pleasurable music led to a wide range of intense and blended emotions and a desire to connect with others.

In a different kind of study, Mehl et al. [28] investigated the role of emotion in the esthetic appeal of different kinds of art. Participants were presented with either a poem, a piece of music, or a painting. The researchers used participants' ratings of vividness of imagery, positivity or negativity of emotion, and how stimulating or relaxing the experience was. Vividness was the most significant predictor of the esthetic appeal of poetry, while ratings of how positive or negative they found the experience was a stronger predictor for the esthetic appeal of music.

Artists of all kinds work in relation to traditions. Some of these people work within a tradition, others extend a tradition a new way, yet others produce something completely different. Whatever the case such traditions are fundamental. Here's what Marcus Cicero said to a companion some 2000 years ago when they visited the grounds of the Academy where Plato had taught [29].

one's emotions are more strongly aroused by seeing the places that tradition records to have been the favourite resort of men of note in former days, than by hearing about their deeds or reading their writings...[Plato's] garden close at hand yonder not only recalls his memory but seems to bring the actual man before my eyes.

Similarly, in 1809, William Godwin (father of Mary Shelley, who wrote the novel *Frankenstein*, 1818) recommended that people visit graves of writers to keep company with them, to purify sentiments, and to connect with their cultural heritage [30].

In medieval times, there was a tradition of making religious pilgrimages to sites associated with saints. But visits of these kinds have been partly replaced by secular pilgrimages to places where musicians, writers, and visual artists, had lived or worked. Aslam et al. [31] did a study on this kind of journey. First one of the researchers talked informally with 29 acquaintances. Later, an online survey was conducted with a more representative sample of 91 participants (unknown to the researchers) each of whom had made a secular pilgrimage to a place where a musical, literary or visual artist had lived or worked. In both groups participants were asked to give a rating of how emotionally moved they felt when they reached their destination, using a scale of 0–10, with 0 meaning “not at all moved” and 10 meaning “the most moved I have felt in my life.” In the informal group, a professional violinist went to Prague to see where Wolfgang Mozart had lived. This person was “completely overwhelmed” and gave the experience a 10. Another participant visited Combray (now called Illiers), the place where the literary artist Marcel Proust had spent parts of his childhood and gave the experience a 9. She said she “felt Marcel’s heart within herself.” Yet another person visited the studio of the visual artist Frida Kahlo. She gave the experience a 9 and felt herself “there with her materials.” In the online study, 46 of the participants visited locations of musicians, 26 went to locations of writers, and the rest visited locations of artists of other kinds. In both the formal and the online studies the mean rating on the 0–10 scale of being emotionally moved was above 6. A paper with three empirical studies and some theoretical proposals on the psychological state of being moved was published by Menninghaus et al. [32]. These researchers found that being emotionally moved can occur with both sadness and joy, and sometimes a mixture of both.

5. Abstract and material art

Some distinguished commentators, such as Robert Simpson, have written that in the start of the recapitulation in the first movement of Beethoven’s Ninth Symphony there is “The sky ... blazing from horizon to horizon” [33]. Unlike most works of literary and visual art, however, music does not actually refer to anything in the material world. It is abstract. Meanings do, of course, occur but they are mental thoughts or reflections by individual humans.

Although when we look around with our eyes, it seems that there is a direct input to us from the world out there. But this is not the case. Instead, as we look outwards only a small patch about the size of an American quarter, or a one Euro coin, held at arm’s length, affects the retina’s rods and cones in high resolution. Everything else is vague. In 1882 Hippolyte Taine put it like this [34]:

So, our ordinary perception is an inward dream, which happens to correspond to things outside; and, instead of saying that a hallucination is a perception that is false,

we must say that perception is a hallucination that is of the truth. [translated by the chapter author; emphasis in original].

Visual perception occurs by the system sampling from rapid successions of fixations (which can occur as frequently as two to four times a second). These small patches of detail enable us to discern what are known as cues which select and guide what are known as mental models. These models are the bases of how the mind works. A clear explanation of how this occurs was made in 1943 by Kenneth Craik [35]. He proposed that when we encounter what he called a “problem”—something in which we do not quite know what to do or say—we translate it into terms of a mental model of the aspect of the world that seems problematic. We then work with that model, mentally manipulating it to find out what would happen if we did this, or said that. When we discover within the model something with an outcome that seems satisfactory, we back-translate from the model into an action or something to say.

An important distinction suggested by Brown et al.’s [25] research is that, as compared with music, the other sister arts—literary and visual—have both the artist and those who engage with artworks making mental models of what are depicted; they do this to project their inner understandings onto what they read, see, and hear. The models prompted by artworks are generally different from those that have occurred previously to people in their everyday worlds. They are, perhaps, unusual or surprising. That’s part of what makes new works of literary and visual art so engaging: both the artist and people who engage with the works may come to change how they think and understand some aspect of the physical or interactive social world, and/or of themselves. As explained in this chapter’s introduction, Collingwood [2] proposed that the changes are likely to have been prompted by emotions and can result in both the emotions and the mental models on which they are based being more deeply understood than before.

When we hear the aspect of music that Brown et al. [25] call prosodic, what happens is that our emotions are prompted in the same kind of way that occurs when we take part in conversations. We base our understanding on our mental models of the person we are talking with and sometimes of ourselves, to know what is going on in the interpersonal interaction. With a sense of identification, something similar can occur in our absorption in a story or film. Some of the brain areas involved here are the same as those for speech and/or facial based emotion recognition such as smiles or frowns. In this kind of situation, the emotions that occur are prompted by our understanding of what is happening in the conversation or in the story that we are engaged in. With the aspect of music that Brown et al. call tonal, no mental models are involved at all. There’s a direct input into parts of the brain in which emotions are mediated. In this kind of case, as von Schirach [1] suggests in the quotation that begins this chapter—without referring to anything in the outside world—music can sometimes prompt a strong emotion directly.

In a different kind of way, with literary works, there have been findings of emotion-based fluctuations in participants’ self-perceived traits of personality when they read stories that they rated as artistic, e.g. Djikic et al. [36]. These changes were different for each participant. One kind of change was feeling more outgoing, another was feeling more diligent, another was feeling more emotionally stable, and so on. This is very different from what occurs with persuasion in which, if it’s effective, everyone changes to think and act in the same way. A similar finding, of personality fluctuation, occurred from looking at pieces of visual art: Giotto’s “Seven vices” [37]. Again, the effects were idiosyncratic: not the same for everyone. They depended on personal

resonance [38]. See also the study by Salgues et al. [39] who asked participants to view pieces of visual art and found that resonance with the self, esthetic judgment, and emotional response are tightly intertwined. In an experiment on music, with participants who were English-speaking university undergraduates, Djikic [40] found that listening to Franz Schubert's "Ständchen, Zögernd leise," without being able to understand the German words, was able to prompt idiosyncratic fluctuations in participants' self-perceived personality, in a way that was similar to what occurred in the study of literary effects by Djikic et al. [36]. Reading an English translation of the lyric while listening to the music, or simply reading the lyric alone, had no such effect.

6. Art can reach within

In *Feeling Beauty* Gabrielle Starr [41] explored the emotional effects of engaging with art. In this thoughtful book she discusses literary, visual, and musical art, calling them "the sister arts." A central study here is by her and two collaborators [42] in which participants were asked to look at representational and abstract color paintings that were unfamiliar to them, while having their brains scanned by a functional magnetic resonance imaging machine. Here's what participants were told as an instruction: "The paintings may cover the entire range from 'beautiful' to 'strange' or even 'ugly.' Respond on the basis of how much this image *moves* you." They were asked to respond on a four-point scale from "least moving" to "most moving." They rated 16.7% of the images as "most moving." There was, however, little agreement among participants as to which these were. Feeling moved by a work of visual art is evidently idiosyncratic.

As the participants looked at the pictures, the researchers recorded activity in a circuit called the "default mode network", which is usually activated when people are musing, reflecting inwardly, thinking about themselves and about people with whom they are close. As participants gave ratings on how moving they found each picture, the default mode network became deactivated. This happens when a person is given a task in the outside world (such as making a rating) and therefore has to stop inwardly musing. With the pictures that a participant rated as "most moving," deactivation ceased quickly, and activation began in the default mode network. This did not generally occur with pictures that a participant found least moving. Starr and her collaborators concluded that a work of art that we find personally moving resonates with us, becomes part of us: it reaches within. In a comparable study of music, Taruffi et al. [43] found that listening to happy or sad music enabled people to withdraw into themselves and as this happened the default mode network was activated.

Also, with a focus on visual art, Maria Medved et al. [44] did a study of visitors to an art gallery. As they were leaving the gallery, these people were asked by an interviewer to choose a painting that they had found especially interesting. The interviewer asked them to talk about it as they would to a friend, and to say if any memories came to mind. They were also asked if it had evoked an emotion. Most of them said it did. Five weeks later, visitors who had agreed (about three quarters of them) received a phone call from the interviewer. They now reported experiencing the same emotion, although at a reduced level. On leaving the gallery, visitors' experiences had been piecemeal, but in the phone interview, they talked about the picture in ways that were more coherent and integrated, relating it to themselves. It would be helpful to our understanding if a study like Medved's et al.'s were done on music.

7. Conclusion

Emotions can be basic: examples are joy, sadness, and fear. When we experience these we do not necessarily know what they refer to. Emotions can also be complex: examples are envy and jealousy; these refer to ourselves and/or someone else [45]. An important emotion of either kind can occur to an artist who does not fully understand it. As Collingwood [2] proposed, to understand it more deeply, the artist can externalize it by translating it into a musical, literary, or visual form. When this artwork is out there in the world, a someone who engages with it can take it in and make some aspect of it part of her- or himself.

Whereas pieces of literary and visual art are based on mental models which are new, or even surprising, to those who engage with them, music is at least partly abstract. In its prosodic form, it does share emotion-based features with speech in conversations and suchlike. There are, however, aspects of music which occur in music alone; these aspects seem able to prompt emotions directly.

In ancient Hellenistic times, the Muses were goddesses who prompted artistic endeavors. There were nine of them, with four being associated with different kinds of music. The etymology of the word “music” derives from them. An artist not only strives to increase her or his understanding of an emotion but sometimes is close to a person who affects her or him deeply and experiences this person as a source of inspiration.

Acknowledgements

I am very grateful to Dr. Christiane Neuhaus who, in the course of this chapter’s preparation, has made very helpful and thoughtful suggestions, which have vastly improved the piece’s quality.

Author details

Keith Oatley
Department of Applied Psychology and Human Development, University of Toronto,
Toronto, Canada

*Address all correspondence to: keith.oatley@utoronto.ca

IntechOpen

© 2025 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] von Schirach F. *Crime: Stories*. New York: Knopf Doubleday; 2011. 257 p
- [2] Collingwood RG. *The Principles of Art*. Oxford: Oxford University Press; 1938. 347 p
- [3] Scherer KR. Which emotions can be induced by music? What are the underlying mechanisms? And how can we measure them? *Journal of New Music Research*. 2004;**33**:239-251
- [4] Scherer KR, Coutinho E. How music creates emotion: A multifactorial process approach. In: Cochrane T, Fantini B, Scherer KR, editors. *The Emotional Power of Music: Multidisciplinary Perspectives on Musical Arousal, Expression, and Social Control*. Oxford: Oxford University Press; 2013. pp. 121-146
- [5] Winner E. *How Art Works: A Psychological Exploration*. Oxford: Oxford University Press; 2019. 304 p
- [6] Pizzolante M, Pelowski M, Demmer TR, et al. Aesthetic experiences and their transformative power: A systematic review. *Frontiers in Psychology*. 2024;**15**:1328449
- [7] Menninghaus W, Wagner V, Wassiliwizky E, et al. What are aesthetic emotions? *Psychological Review*. 2019;**126**:171-195
- [8] Juslin PN, Västfjäll D. Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*. 2008;**31**:559-575
- [9] Juslin PN. Expression and communication of emotion in music performance. In: Juslin PN, Sloboda JA, editors. *Handbook of Music and Emotion*. New York: Oxford University Press; 2010. pp. 453-489
- [10] Juslin PN. *Musical Emotions Explained*. Oxford: Oxford University Press; 2019. 624 p
- [11] Juslin PN. Major theories of emotion causation and their applicability to music: The case for multi-level approaches. *Music Perception*. 2025;**42**:421-466
- [12] Herrmann E, Call J, Hernandez-Lloreda MV, Hare B, Tomasello M. Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science*. 2007;**317**:1360-1366
- [13] Tomasello M. *A Natural History of Human Thinking*. Cambridge MA: Harvard University Press; 2014. 192 p
- [14] Tomasello M. *Becoming Human: A Theory of Ontogeny*. Cambridge, MA: Harvard University Press; 2019. 392 p
- [15] Wadley L, Backwell L, D'Errico F, Sievers C. Cooked starchy rhizomes in Africa 170 thousand years ago. *Science*. 2020;**367**:87-91
- [16] Goodall J. *The Chimpanzees of Gombe: Patterns of Behavior*. Cambridge, MA: Harvard University Press; 1986. 673 p
- [17] Popper KR, Eccles JC. *The Self and its Brain*. New York: Routledge; 1999. 614 p
- [18] Vanhaeren M, d'Errico F, Stringer C, et al. Middle Paleolithic shell beads in Israel and Algeria. *Science*. 2006;**312**:1785-1788
- [19] Huron D. Is music an evolutionary adaptation? In: Peretz I, Zatorre R,

editors. *The Cognitive Neuroscience of Music*. Oxford: Oxford University Press; 2003. pp. 57-75

[20] Moran T. *Vietnamese Musical Instruments: A Monographic Lexicon*. National Library Board Singapore; 2020. 384 p

[21] Adams B. *Ada Lovelace: The world's First Computer Programmer*. Barnsley: Pen and Sword; 2023. 224 p

[22] Perlovsky L. Musical emotions: Functions, origins, evolution. *Physics of Life Reviews*. 2010;7:2-27

[23] Johnson-Laird PN, Oatley K. Emotions in music, literature, and film. In: Barrett LF, Lewis M, Haviland-Jones J, editors. *Handbook of Emotions*. 4th ed. New York: Guilford; 2016. pp. 82-97

[24] Baumgartner T, Lutz K, Schmidt CF, Jäncke L. The emotional power of music: How music enhances the feeling of affective pictures. *Brain Research*. 2006;1075:151-164

[25] Brown S, Berry M, Phillips E. Localizing music's "language of emotion" in the human brain: A functional MRI study of scale and emotion processing. *Psychology and Neuroscience*. 2025;18:40-59

[26] Fuentes-Sánchez N, Espino- Payá A, Prantner S, et al. On joy and sorrow: Neuroimaging meta-analyses of music-induced emotion. *Imaging Neuroscience*. 2025;3:425

[27] de Leeuw RNH, Janicke-Bowles SH, Ji Q. How music awakens the heart: An experimental study on music, emotions, and connectedness. *Mass Communication and Society*. 2022;25:626-648

[28] Mehl K, Gugliano M, Belfi AM. The role of imagery and emotion in the

aesthetic appeal of music, poetry, and paintings. In: *Psychology of Aesthetics, Creativity, and the Arts*. 2023

[29] Cicero M. *De Finibus (on the Ends of Good and Evil)*. Cambridge, MA: Harvard University Press; 54 BCE. 150 p

[30] Godwin W. *Essay on sepulchres*. In: Philp M, editor. *Political and Philosophical Writings of William Godwin*. London: Pickering and Chatto; 1993; 1809;6:1-30

[31] Aslam H, Oatley K, Djikic M. Secular pilgrimage to places associated with writers, musicians, and visual artists. *Journal of Tourism and Cultural Change*. 2023;21:501-515

[32] Menninghaus W, Wagner V, Hanich J, et al. Towards a psychological construct of being moved. *PLoS One*. 2015;10(6)

[33] Simpson R. *Beethoven Symphonies*. London: British Broadcasting Corporation; 1970. 62 p

[34] Taine H. *De l'intelligence*, Tome 2. Paris: Hachette; 1882. 496 p

[35] Craik KJW. *The Nature of Explanation*. Cambridge: Cambridge University Press; 1943. 123 p

[36] Djikic M, Oatley K, Carland M. Genre or artistic merit: The effect of literature on personality. *Scientific Study of Literature*. 2012;2:25-36

[37] Djikic M, Oatley K, Peterson J. Serene arts: The effect of personal unsettledness and of paintings' narrative structure on personality. *Empirical Studies of the Arts*. 2012;30:183-193

[38] Rosa H. *Resonance: A Sociology of our Relationship to the World*. New York: Polity; 2019. 576 p

[39] Salgues S, Jacquot A, Makowski D, et al. Self-reference and emotional reaction drive aesthetic judgment. *Scientific Reports*. 2024;**14**(19699)

[40] Djikic M. The effect of music and lyrics on personality. *Psychology of Aesthetics, Creativity, and the Arts*. 2011;**5**:237-240

[41] Starr GG. *Feeling Beauty: The Neuroscience of Aesthetic Experience*. Cambridge MA: MIT Press; 2015. 280 p

[42] Vessel EA, Starr GG, Rubin N. Art reaches within: Aesthetic experience, the self and the default mode network. *Frontiers in Neuroscience*. 2013;**7**:258

[43] Taruffi L, Pehrs C, Skouras S, Koelsch S. Effects of sad and happy music on mind-wandering and the default mode network. *Scientific Reports*;7(14396)

[44] Medved MI, Cupchik GC, Oatley K. Interpretive memories of artworks. *Memory*. 2004;**12**:119-128

[45] Oatley K, Johnson-Laird PN. The communicative theory of emotions: Empirical tests, mental models, and implications for social interaction. In: Martin LL, Tesser A, editors. *Striving and Feeling: Interactions among Goals, Affect, and Self-Regulation*. Mahwah, NJ: Erlbaum; 1996. pp. 363-393

Chapter 4

Music and Emotions: Spectrum of Possibilities in Esthetics Sports

Eliana de Toledo and Mateus Henrique de Oliveira

Abstract

This chapter examines the role of music in artistic sports, regarding the pedagogical, emotional, cognitive, and esthetic dimensions of training. Starting from literature in physical education, sport pedagogy, musicology, and psychology, the text explores how music functions not only as a performance backdrop but also as a structuring and expressive tool within training environments. Music is shown to enhance rhythmic perception, motor coordination, emotional engagement, and expressive sensitivity, especially in sports such as rhythmic gymnastics, artistic gymnastics, figure skating, and artistic swimming. Through a review of recent research, including qualitative analyses, the chapter highlights how the congruence between music and movement influences performance quality, as well as how music serves as a motivational and organizational device in training routines. The discussion extends to the training of coaches, advocating for the development of auditory sensitivity and musical literacy as pedagogical competencies. By framing coaches as esthetic mediators, the chapter emphasizes the need for musical education as part of coach development programs in esthetic-oriented disciplines. Finally, it problematizes the cultural devaluation of expressiveness and musicality in sports training, often seen as feminine attributes, calling for a broader understanding of music as an integral element of human movement and sporting excellence.

Keywords: music and emotion, esthetic judgment, choreographic composition, aesthetic sports, coach education

1. Introduction

The relationship between art and sport has undergone centuries of symbolic interpretation, cultural resignification, and epistemological tension. The debate about the relationship between art and sport dates back to the classical Greek ideal of integral human development, in which body, mind, and spirit were understood as an integrated whole. Art, understood as a sensitive and symbolic expression, and sport, as a performance activity requiring physical effort and skill, share values of creativity, beauty, and bodily communication [1, 2].

Karageorghis [3] systematizes decades of research on the effects of music on physical exercise and sports performance. His theoretical model considers three main factors that demonstrate the effectiveness of music: rhythmic response by the

athletes, extramusical associations, and cultural impact. These elements explain how certain music can improve mood, reduce perceived exertion, synchronize movements, and facilitate flow states.

In modern sport, this relationship is particularly evident in the artistic modalities, for which Best [4] coined the term “esthetic sports” due to their appeal regarding the esthetics of movements and their objectives within practice:

[...] There is a category of sports, which I call “aesthetic”, in which the aim cannot intelligibly be specified independently of the manner of achieving it. Included in this much smaller category are gymnastics, diving, skating, synchronized swimming, trampolining and surfing.

[...] The way in which the appropriate movements are performed is certainly not incidental to such a sport; on the contrary, it is central to and largely defines its character [4].

According to Toledo, Hess, and Oliveira [5], some of these esthetic sports can be: artistic swimming, figure skating, rhythmic gymnastics, and women’s artistic gymnastics (see Section 3.1 for a detailed discussion on the differences between them). These sports have an artistic character initially due to their relation with dance (and its typologies), which is considered part of the field of arts. Direct examples include some gymnastics modalities, which are explicitly derived from classical ballet or standard dance in their regulations [6], both mandatory in *routines* (a special term for “sequences of specific movements”). Indirect examples are those sports that use gestures and/or some elements of dance (classical, modern, folk, etc.) in their routines, often found in figure skating and artistic swimming.

These sports, with a strong expressive and esthetic component, are marked by the presence of music as a constitutive element of the performance.

When considering the role of music in these sports, we refer not only to its rhythmic function but also to its affective, communicative, narrative, and formative powers. More than just providing the rhythmic frame, it defines atmospheres, imposes themes, and offers pauses, accelerations, and intensities that are translated into movement, thus playing a fundamental role in the temporal and spatial organization of the performance. In sports such as rhythmic gymnastics, figure skating, and artistic swimming, this “music-movement relation” is a criterion for technical evaluation, with any dissociation between the two elements being penalized (a technical term used in the Code of Points and Regulations) [5].

Musicality, in this context, refers to the athlete’s ability to musically interpret their routine, demonstrating fluidity, expressive coherence, and rhythmic integration. The choice of soundtrack becomes strategic and is often tailored to the athlete’s physical style. This can determine the subjective impact of the performance on the judges and the audience.

Sborquia [7] emphasizes the term “choreographic construction,” as choreography is developed from the perspective that the subject – that is, the athlete, choreographer, or coach – constructs different worlds, distinct from their own. During this construction process, the subject has the opportunity to stimulate critical thinking, as he/she ceases to be merely a reproducer and receiver of his/her culture, becoming its creator. An example of this occurs when a choreographic theme is

chosen, and the search for songs related to it begins. In the process of musical selection, a series of feelings, memories, sensations, and desires can arise, and the subject will choose the music he/she prefers based on a critical stance regarding what they truly wish to express bodily.

In this same example, the subject can reflect on whether the musical style (classical, funk, popular, hip-hop, etc.) matches his/her culture, history, vision of human beings, and society. If the subject has a decolonial stance, he/she will certainly reflect that some styles and lyrics are not consistent with his/her position. Consequently, this will be reflected in the choice of movements and use of space, among other aspects. In all these processes, a critical and analytical stance is activated.

Thus, in addition to the individual impact of music on emotion, it is necessary to recognize that gymnastic disciplines are also shaped by cultural norms that regulate emotional expression. Snyder [8], in a study with collegiate gymnasts, demonstrated how the sporting experience is permeated by “feeling rules” that define which emotions may be displayed or suppressed in competitive contexts. This sociocultural dimension of emotional management shows that expressiveness in routines is not merely a spontaneous reflection but an affective work that is learned, negotiated, and adjusted to the expectations of the discipline and its audience.

In this context, the soundtrack of a performance can induce emotional states, reinforce symbolic gestures, give meaning to motor actions, and, above all, organize the esthetic experience lived by the athletes and spectators.

The estheticization of sporting gestures, especially in sports such as gymnastics and skating, means that many athletes are considered “artists of movement.” According to Melo [9], both sport and art are permeated by symbolic elements, rituals, and narratives that produce meaning for both observers and performers. By approaching sport as a popular art, the author also highlights its ability to mobilize collective emotions and identities. Indeed, some gestures are part of the collective memory of a given culture or sport, and when combined with the technical gestures of a sport, they produce distinct meanings. If, for example, in a choreography by a Black athlete, the “black panther gesture” is used in combination with a typical gesture of the sport (such as a balance, a jump, etc.), we will clearly have a production of meanings for spectators and participants. And, if this is further combined with a song with anti-racist lyrics, this identity dimension will be even greater.

Music, when incorporated into sports practice in a structured way, enhances complex emotional experiences. It organizes emotions, shapes atmospheres, and allows athletes to express their subjectivity. Lacerda and Mumford [10] argue that musical components foster creativity, intuition, and expressive freedom – qualities that transcend performance and open spaces for integral human development. By providing sensitive experiences, music can also be a pedagogical promoter in sports, contributing to the critical and creative development of athletes and coaches by fostering broader understandings of corporeality and symbolic communication in play, dance, and competition [11, 12].

The literature indicates that music, by inducing and modulating emotional states, organizes not only perception but also the quality of motor responses. Experimental studies with free dance show that emotions induced by happiness and sadness are not only subjectively experienced but also translated into specific bodily patterns. Van Dyck et al. [13] demonstrated that participants induced to feel happiness

exhibited faster, broader, and more impulsive movements, whereas the induction of sadness led to more restrained, slower, and contracted gestures. These findings reinforce the argument that the body responds to music in an integrated manner, articulating affective states and motor expression in line with esthetic principles. Thus, music is not merely an accompaniment to performance but an active element in producing meanings and atmospheres that are materialized in sporting gestures.

Corroborating this perspective, recent investigations with artistic gymnastics and trampoline athletes show that listening to music *before* competitions directly influences balance, emotional state, and perceived exertion, reinforcing the regulatory role of music in both preparation and performance execution [14]. These findings consolidate the idea that music functions as both an ergogenic and affective resource, optimizing motor parameters as well as psychological states.

It is in this secular scenario of sport that music vibrates and makes people vibrate, echoes, seduces, disturbs, and provokes. The objective of this text, therefore, is to describe the relationship between music and emotion in esthetic sports from different perspectives, with emphasis on athletes, coaches, and judges. This emphasis is important because the sports world encompasses a vast spectrum of professionals and social sectors, including members of the technical team (such as coaches, physical trainers, physiotherapists, nutritionists, etc.), businesspeople (sports product professionals, media professionals, etc.), sports management professionals (presidents and directors of clubs and federations, marketing professionals, communicators, legal counsels), and others.

2. Methodology

This chapter presents issues from a body of research developed by the authors and their partners on music in esthetic sports, combined with their life experiences as coaches, judges, former athletes, and university professors. Thus, it is within this mosaic of interpositions and coexistences that this text is constructed, bringing forth, in a unique way, a spectrum of possibilities regarding music and emotion in the sports context, in dialogue with other sources.

Table 1 contains links to documents in which rules and regulations applied in esthetic sports are described in detail. Unfortunately, these documents are not

Aesthetic sport	Regulation or code of points link
Artistic swimming	https://resources.fina.org/fina/document/2025/01/21/d9,948a65-9,295-44a0-ab96-bcfc22ade8b0/AQUA-AS_MANUAL_January-2025.pdf
Figure skating	https://isu-d8g8b4b7ece7aphs.a03.azurefd.net/isudamcontainer/CMS/Corporate-Site/Sports-Rules/Figure-Skating-Rules/Constitution-and-Special-Regulations/2024ConstitutionGeneralRegulationsFINAL0359870001741256957-1125.pdf
Rhythmic gymnastics	https://www.gymnastics.sport/publicdir/rules/files/en_1.1.%20WAG%20Code%20of%20Points%202025-2028.pdf
Women's artistic gymnastics	https://www.gymnastics.sport/publicdir/rules/files/en_1.1.%20RG%20Code%20of%20Points%202025-2028%20(mark-up)%20-%20valid%20from%201st%20April%202025.pdf

Table 1.
Documentary sample – regulations and code of points for esthetic sports.

always easily accessible, especially for nonexperts. Sometimes, there are different versions of the same document that are not official and, therefore, may contain errors and data distortions.

The research presented in this chapter is exploratory and of the documentary type, using as a sample the regulations of esthetic sports from the current Olympic cycle (2025–2028), with only the artistic swimming regulations following a different cycle (2022–2025). It also incorporates many articles (from indexed and open-access journals) that address the relationship between music and esthetic sports. The search period for articles was from 2021 to 2025, in national (Brazilian) and international (English-speaking) repositories (academic platforms). Special emphasis is placed on the relationship between music and emotion regarding the following themes: regulations, athletes, coaches, and judges.

3. Results – music and emotion: Mismatches in the sports context

3.1 The rule and the music

We initially present a brief description of some aspects of the regulations (or Code of Points [CoP]) in esthetic sports, aiming to bring out some of their specificities in relation to artistic aspects, including music. These regulations also show the demands placed on the athletes and how the artistic component is evaluated in the routines.

In this table, we identify how music is part of the artistic components of these sports, always in relation to another aspect. In other words, music is generally not evaluated separately, but rather in its relation to the movement, the theme, the costumes, the athlete's expressiveness, among other possibilities.

There are very similar assessment points between these sports, as they are all evaluated through choreography. However, the weight attributed to these artistic components in the different sports seems distinct.

[...] in each of these sports, the artistic component has a different weight in the routine's final score. That is, the regulations for each sport establish a greater or lesser value for the artistic component, which will proportionally impact the final score. In Artistic Swimming, for instance, the score a swimmer(s) receives for Artistic Impression is equivalent to 30 to 40% of the final score, depending on the stage of the competition. In Skating, the Artistic Component has much more diverse evaluation criteria than the Technical Components, proving to have greater relevance in the final score. And in Rhythmic Gymnastics, the maximum possible artistic deductions a gymnast can obtain represents a total of 25% of the total performance score [5].

Some regulations or the CoP have clear rules for the use of music, as well as for its duration (1 minute and 30 seconds, 3 minutes, etc.), whether it can be sung or not, or whether it can be original or mixed. These changes in the use of music occur in some Olympic cycles, based on specific demands of the sport and other demands aligned with social changes, especially among the youth (athletes and spectators). The mere fact that federations are sensitive to these tendencies highlights the importance of the relationship between music and emotion in the sporting context, as is the case with the use of sung music.

An example of this change in artistic sport, based on new trends in society, can be found in the records in the CoP of rhythmic gymnastics from the 1970s – (then modern gymnastics) – for example, music being played live, at which point preparation teams were allowed to bring their own pianists, who played the music for the gymnasts' choreography. This undoubtedly brought about a major change, both in training and in events, toward a more sensitive relationship between art and sport, since the artist (pianist) was no longer just part of training but also of events. Because of that, we assume that the emotions also changed, as they were felt differently by gymnasts, judges, and spectators during a live performance, triggering different experiences and senses.

The decade from 2010 to 2020 brought further significant changes to music. With the 2013–2016 Rhythmic gymnastics CoP (the specific term that refers to the regulations in gymnastics sports), sung songs were allowed to be used in athletes' routines, a notable milestone at the Rio 2016 Olympic Games [15]. Subsequently, the same occurred in figure skating at the 2018 PyeongChang Winter Olympics, resulting in an apparent improvement in expressiveness among gymnasts and a greater dialogue with the audience, thanks to the use of sung songs.

Artistic Swimming also began allowing the use of sung songs in its performances in the 2010s, with these songs being more prominent in some events and competition phases, such as solos/duets in Swimming and individual events in Skating, than in others. Of the aforementioned disciplines, the only one that does not allow sung songs in its performances is women's artistic gymnastics, which prioritizes only the melody and the beat of the music.

3.2 Athletes – harmonizing the body art in sports

Athletes in esthetic sports, as well as their coaches, follow the federative regulations (mentioned above), both for the experience of sport as a leisure activity and health promotion, and to achieve success in their sports careers, aiming to win a medal.

The training of the athlete for the development of artistic components must be understood as a process – from initiation to high performance – lesson by lesson, with music as an integrated component, that is, not separate from enhancing the skills in sports. Yürük and Asma [16], for example, demonstrated that coordination exercises with music, applied in gymnastics teaching for children, favored the development of balance and rhythmic skills, whereby the use of music brought elements of playfulness that contributed to affective engagement and motivation.

The relationship between body and sound was also explored by authors such as Krafte et al. [17], who discussed the interdependence between musical perception and motor organization, and by Mezei and Cristea [18], who pointed out the importance of integrating expressiveness and functionality into the technical training process.

In this context, the concept of emotional intelligence (EI) becomes particularly relevant for understanding the complexity of performance in gymnastic disciplines: Beyond the momentary induction of affective states through music, it also concerns the athletes' ability to recognize, regulate, and strategically express emotions in competitive settings. Recent studies in acrobatic gymnastics indicate that athletes exhibit high levels of EI in relation to age, experience, or position within the group, suggesting that regular practice in highly collaborative and expressive environments

fosters the development of this competence [19]. This finding broadens the discussion by emphasizing that EI not only contributes to the athletes' well-being and team cohesion but also enhances the expressiveness and esthetic communication in routines, bringing sports performance closer to a fully integrated artistic experience.

However, despite the fact that music is a component of the assessment, coaches tend to view it as a technical element rather than an artistic expression, in principle. This is common behavior in the training of esthetic sports athletes, considering any artistic expression as a means and not as the achievable goal of learning. In other words, artistic components generally seem to merely serve motor and performance goals from the perspective of skill development and code compliance.

As Pizzol et al. [20] warn, care must be taken to ensure that any component or type of art in sports should not merely be considered an “instrument” for achieving performance. They propose that art should be experienced by athletes as sensory, expressive, rhythmic, creative, esthetic, interpretive, and emotional, as an end in itself. Such experiences contribute to improvement in the sport but, above all, contribute to the human and esthetic development of individuals, both in the short and long-term.

It is therefore important for athletes to be in a musical environment that is rich in terms of rhythms, styles, melodies, musical instruments, ethnic roots, etc., which will certainly foster diverse emotions and unique performances. The regulations, or CoP, allow this musical diversity as well as, for the most part, suggest certain guidelines for use (see **Table 2**). These experiences can be had within training sessions, through educational activities, either individually or in groups, but also outside of them through visits to exhibitions or group musical programs (for shows, performances, orchestras, etc.), among other possibilities.

These experiences can be even more meaningful, as the athlete can then participate in proposals during the planning phase with greater autonomy. Toledo [21] emphasizes the need for sports education permeated by the pedagogy of autonomy [22] in a relationship where athletes and coaches share certain moments of classes or training sessions. In line with her proposal, athletes should have the opportunity to bring in music suggestions, whether for an activity or for the choreographic composition (or routines) they will compete in. This opportunity is mediated by the coach:

This sharing must be based on dialogue, in which teachers and students are considered as subjects who use their assumptions, experiences and knowledge to construct something new through collective maturation [21].

A recent example of this was the choice of the funk-style song “Baile de Favela” for medalist Rebeca Andrade’s artistic gymnastics floor routine at the Tokyo 2020 Olympic Games and the “Movimento da Sanfoninha” at the Paris 2024 Olympic Games [23–25] (**Figure 1**). These choices, made between the gymnast and choreographer Rhony Ferreira, were crucial to her medal victory, given the fact that it’s a musical style that’s part of the gymnast’s life, with which she’s familiar rhythmically and gesturally, and which comforts her emotionally.

In this case, it is worth mentioning that the use of this music symbolized, according to the gymnast, a decolonial and resistance aspect by performing to a typical Brazilian musical genre (and not to a North American or European one), often present in communities with high social vulnerability among many people of the same ethnicity.

AESTHETIC SPORT	DESCRIPTION	EVALUATION OF THE ARTISTICS COMPONENTS
Rhythmic Gymnastics 	Featured in the Olympic program since 1984, Rhythmic Gymnastics descends from the Swedish Method of gymnastics and has been recognized by the International Gymnastics Federation since its founding in 1921. Among the disciplines presented, it is the only one essentially practiced exclusively by women. In it, athletes must perform sets, combining elements of dance and choreography with portable equipment – hoop, ball, ribbon, and clubs – on a 12x12-meter stage. It is contested in two categories: Individual, in which each athlete must perform a set with each apparatus; and Group, in which a group of five athletes performs two different sets, one with a single apparatus set and the other with a set of two different apparatuses.	Two specific judges are responsible for evaluating each artistic component deductively, identifying artistic faults and classifying them as minor, medium, or serious. The total value of these artistic deductions is combined with the deductions applied for technical faults. Both types of deductions are subtracted from the initial value of 10.0 points to determine the Execution Score (E-score). This E-score is then added to the Difficulty Score (D-score) to obtain the gymnast's Final Score.
Artistic Swimming 	Artistic swimming – a term adopted by the International Aquatics Federation (FINA) in 2017 – was regulated as a sport in 1952 and became part of the Olympic program in 1984. The sport is contested in three categories: Individuals, Doubles (women's or mixed), and Teams of eight athletes. In this event, athletes must perform dynamically and artistically, using movements both in and out of the water, combining choreography with strength and flexibility exercises in a creative and synchronized manner.	The modality comprises two types of routines. In technical routines, five judges form Panel 2 – Artistic Impressions, which assesses the artistic components of the performance. These components represent 30% of the gymnast's final score. In free routines, the same judges evaluate artistic impressions, which correspond to 40% of the final score.
Figure Skating 	Currently divided into two distinct disciplines – Inline Figure Skating, regulated by the World Skate Federation since 1924; and Ice Figure Skating, regulated by the International Skating Union since 1892 – Figure Skating combines skating technique with elements of body expression, along with musical accompaniment. It is contested in three different categories: Singles (men's and women's), Mixed Doubles, and Dance Pairs. In each category, athletes must perform two types of routines: Technical Routines, which are shorter and require more rules and requirements to follow, and Free Routines, in which athletes must demonstrate greater creativity over longer time.	In total, nine judges are responsible for formulating the Presentation Score, which is the variable in the final score that includes artistic aspects. This variable is divided into aspects such as Skating Skills, Transitions (between choreographic and technical elements), Performance, Choreographic Composition, and Interpretation of the music, with a maximum value of 10 points for each of these items.
Women's Artistic Gymnastics 	A fixture on the Olympic program since 1928, Women's Artistic Gymnastics has been recognized by the International Gymnastics Federation since its founding in 1921. Among the disciplines presented, it is the only one with a gender division (Men's Artistic Gymnastics is a discipline with its own characteristics.) In it, athletes must perform series combining dance and choreographic elements with acrobatic elements on large apparatuses – such as the vaulting table, uneven parallel bars, balance beam, and floor platform. The balance beam and floor exercises include the so-called artistic component, as part of one of the variables of the gymnasts' performance score. The floor exercise also features musical accompaniment, which supports the execution of acrobatic movements and dance jumps.	In total, five judges are responsible for assigning the Execution Score (E) to gymnasts performing floor exercises. In addition to deducting technical errors for each movement, this group of judges is also responsible for analyzing so-called Artistic Faults in a deductive manner, seeking to understand what is different from the artistic standard of the sport in aspects related to the use of body parts, musical expression, music-movement relationship, and choreographic variety, thus composing one of the variables of the gymnasts' Final Score.

Source: Adapted from Toledo, Hess, and Oliveira [5].

Table 2.
The valuation of artistic components in esthetic sports.



Figure 1.
A typical gesture of Brazilian culture in Rebeca Andrade's Floor Exercise routine.

All this diversity of experiences promotes and raises a range of technical possibilities, but also esthetic and sensitive ones. They produce diverse emotions and indescribable meanings.

3.3 Coaches set the pace – music, emotion and coach training in esthetic sports

Music has thus been recognized as a fundamental pedagogical and expressive resource in the context of sports training, especially in artistic-expressive sports. Several studies indicate that its presence can contribute to improving physical performance, motor coordination, motivation, and training organization. Authors such as Karageorghis and Priest [25], Karageorghis [3], and Bateman and Bale [26] highlight the effects of music used synchronously – with movements coordinated to the rhythm – or asynchronously – as a background – on perceived effort, mood, and psychophysiological arousal. These effects make music an ally in creating more engaging and powerful training environments.

In aerobic gymnastics, for example, Mezei and Cristea [18] emphasize that artistic criteria, often undervalued in relation to technical ones, require more careful preparation: the appropriate selection of soundtracks, the coherence between music and movement, and the development of rhythmic sensitivity become essential elements for coaches and athletes. In this sense, Belenkaya [27] therefore proposes specific musical training for *coaches*, including knowledge of musical structure, style, dynamics, tempo, and expressiveness. Her emphasis on Dalcroze's eurythmy as a methodological tool reinforces the idea that pedagogical work with music should not be limited to intuition but requires technical mastery and esthetic sensitivity.

Music's contribution also extends to the neurocognitive field: Bredikhina et al. [28], using electroencephalography, showed that higher-performance athletes

exhibit more organized brain responses to rhythmic stimuli (in different frequency ranges like beta and theta), indicating that listening to music is directly related to neuromotor processing. This further increases the relevance of music in both athlete preparation and coach training.

This expanded perspective also includes the symbolic and visual domain. Kim and Chung [29], when analyzing figure skating costumes in relation to music, demonstrated how visual esthetics, intertwined with sonic choices, offer an additional dimension of interpretation for the coach. That is, it is the skaters' costumes, referring to the musical style and the characteristics of specific choreographies, using music from theater and cinema classics, such as 007, The Matrix, and The Phantom of the Opera, as well as tango and Spanish flamenco, that put additional emphasis on the artistic aspects of the performance. On the other hand, gender relations as conveyed through costumes, music, and expressiveness could be problematic. At least, Adams [30] points out that these visual-sonic practices are often associated with a tendency toward feminization, which distracts from technical recognition, thus focusing on other aspects like “sport related to media” or debates about how this can influence people to approach or distance themselves from it.

Of note is that, in addition to its expressive and formative roles, music can also serve as an organizing element, helping structure the training session (**Figure 2**). Studies, such as those by Terry et al. [31] and Karageorghis [32], indicate that the components of music (such as beats per minute, tonality, and style) can stimulate specific cognitive and emotional states – excitement, focus, or relaxation – thus being useful at different moments during training.

Musicality, therefore, is not merely an esthetic capacity: it underpins rhythmic perception, sense of timing, gesture construction, and training motivation, suggesting that listening to music favors the development of more refined and



Figure 2.
Spain's artistic swimming national team coach, Mayuko Fujiki, coaching elite artistic swimmers.

adaptive motor skills, in addition to enhancing creativity and playfulness – essential aspects in the early stages of development or transition.

Music, in this sense, doesn't just support the training – it transforms it. It acts as a language, a structure, a teaching tool, and an assessment criterion. For coaches, this involves developing attentive and sensitive listening skills, making them capable of perceiving sonic nuances, relating them to gesture, time, and space, and building from them richer, more expressive, and meaningful learning experiences. To train coaches who understand these (subtle) interfaces between sound and movement is to invest in more humane, coherent, and powerful practices in sports.

3.4 Music in the experience of sports judges: Analyzing body melodies

In esthetic sports, such as artistic gymnastics, rhythmic gymnastics, artistic swimming, and figure skating, judging transcends the purely technical dimension and takes on interpretative dimensions. In these contexts, music ceases to be a mere accompaniment and becomes a structuring element of the performance and, simultaneously, a perceptive and emotional challenge for the judges. Judging, in these cases, involves listening with the eyes and seeing with the ears – the sound begins to shape the scene, the emotion, and the narrative of the gesture (**Figure 3**).

Karageorghis and Priest [25] demonstrate that music can influence both performers *and* observers, affecting mood, focus, and psychophysiological states. In the case of *judges*, a specific soundtrack of a routine can trigger emotional reactions anchored in previous experiences and personal memories, which inevitably influence the interpretation and evaluation of the performance.



Figure 3.
Figure skater Adam Rippon from the USA competes in the PyeongChang 2018 Winter Olympic Games, being observed by a group of judges.

These interferences should not be understood as flaws but as inherent characteristics of artistic judgment. According to Plessner and Schallies [33], the judge operates in zones of normative ambiguity, where norms and sensibilities coexist. Judgment, then, requires more than technical mastery: it requires attentive listening, esthetic empathy, and an ethical stance of self-reflection.

Studies, such as those by Veit et al. [34], show that routines evaluated with soundtracks synchronized to body movements are perceived as more coherent and receive higher scores, both by judges and spectators. The musical beat, when compatible with the rhythm of the movement, amplifies the esthetic legibility of the routine and facilitates the decision-making process.

The judge's performance, in this sense, demands a heightened perceptive capacity. Karageorgis [3], Karageorgis and Priest [25], and Bateman and Bale [26] argue that listening to music activates brain areas related to motor memory, rhythmic anticipation, and esthetic empathy. This means that judges more familiar with the musical language are able to perceive gestural nuances, breaks in flow, and coherence between sound and movement with greater precision. However, despite these requirements, judge training is still predominantly focused on the normative aspects of the CoP, as highlighted by Oliveira and Toledo [35], neglecting the esthetic and expressive dimensions.

The subjectivity of the judges, therefore, is not a problem to be eliminated but a component to be recognized, discussed, and improved. Pajek et al. [36] demonstrate that there are significant variations among judges when evaluating artistic criteria in events such as the balance beam, highlighting the influence of individual repertoire. Suominen [37], studying groups of esthetic gymnastics in Finland, found that institutional ties and personal preferences also impact judging, often to the detriment of impartiality.

It is, therefore, important to consider robust esthetic-musical training for judges in sports to incorporate the artistic component into their evaluation, including practices related to music theory and exercises that highlight the relation between music and movement. Musicality here is not seen as a gift or ancillary skill, but as a technical-esthetic competence that can – and should – be developed through specific methodologies.

The role of music as a facilitator of judgment is also reinforced by neuroscience studies. Veit et al. [34] demonstrated that observing movements accompanied by sound activates neural areas related to action, suggesting that multimodal processing (vision and hearing) enhances the understanding of gestures. Music, in this context, acts as a guide for action, cognitively organizing the flow of routines and favoring more accurate and confident decision-making.

The emotional dimension is also present. Van Bokhorst et al. [38] demonstrated that the ability to recognize emotions is crucial for judgment accuracy, especially in situations with stimulus overload, since competitions can be long and permeated by different stressors, both for the athletes and for coaches and judges, for instance, in relation to nervousness and tension before, during, and after performing a routine. Maintaining the level of skill in this type of emotional recognition is a difficult task for judges throughout the competition period. Flessas et al. [39] observed that even experienced judges struggle to capture all technical and artistic elements simultaneously. Music, by providing rhythmic and temporal structure, can serve as a cognitive facilitator, promoting greater fluidity in reading the routine.

However, even with all this evidence, musical aspects are still undervalued in judge training. Fernandez-Villarino et al. [40] and Oliveira and Toledo [41] identified (by applying questionnaires to judges) that skills related to esthetic appreciation and musical perception are among the least considered by women's rhythmic and artistic gymnastics judges, despite being central to artistic evaluation. Leandro et al. [42] corroborate this finding by pointing out that criteria such as compositional unity and music-movement relation are perceived as the most subjective and, therefore, more difficult to assess consistently.

Yet, in judging practice, these elements related to how judges perceive and understand emotions are constantly mobilized. Stokowski et al. [43] observed that judges in college competitions in the US take into account the athletes' facial expressions, eye contact, and charisma – even though these aspects are not formally described in the regulation. Musical choice, while not directly altering final scores, impacts performance, audience connection, and emotional expressiveness, suggesting an implicit presence of music in judging.

Judging, therefore, involves a perceptive triad (technical, expressive, and musical) that challenges judges to produce quick, sensitive, and ethical summaries. The lack of objective criteria for evaluating music and expressiveness creates room for bias. Mercier and Heiniger [44] show that favoritism based on nationality, reputation, or artistic style occurs frequently, especially when there are no clear parameters to guide subjective evaluation. Recognizing these biases does not mean invalidating subjectivity, but rather understanding it as a constituent part of the judgment while investing in its improvement.

Regarding this issue, a study by Van Dyck et al. [45] yields interesting results. They demonstrated that spectators are able to identify, with a high degree of efficiency, some expressed emotions of happiness and sadness from merely observing body movements, even when presented in an amateurish way and without explicit musical reference. This data expands the understanding of the role of expressiveness in esthetic sports: the gesture is not only a consequence of an internal state but also a communicative signal that can be interpreted by the audience and, above all, by the judges responsible for the evaluation. By showing that bodily expressiveness is recognizable even in neutral conditions, this study supports the relevance of music as a stimulator of emotional atmospheres that enhance the esthetic power of performances.

In short, music in artistic sports does not only animate the gesture – it structures and qualifies it. For the judge, sound serves both as a reference and a challenge, increasing the complexity of listening and making necessary interdisciplinary training that articulates cognition, emotion, and esthetics. Evidence presented by authors such as Stokowski et al. [43], Veit et al. [34], and Oliveira and Toledo [41] reinforces the importance of rethinking training programs and evaluation criteria in sports judging.

Investing in sensitive listening, in creating more precise criteria for musical evaluation, and in promoting spaces for continuing education are paths toward fairer and more conscious judging, in tune with the complexities of contemporary sports. Sound, accompanied by gestures, should be understood not only as an esthetic ornament but also as a legitimate criterion of technical excellence – and, therefore, as a necessary and powerful instrument of judgment.

4. Conclusions – the orchestra in motion: Body, music, and sport

The relationship between sport and art is clear, and within it, music is a valuable artistic component. It fosters motion, constantly seeks harmony, and stimulates the senses. The intersection of music, movement, and emotion is one of the most complex and underexplored dimensions of the sporting phenomenon.

Music, when combined with movement, modulates physiological states. It also mobilizes emotions, symbolizes identities, alters motivation, and enhances esthetic experiences. Its use in sports goes beyond the functional dimension, proving to be a pedagogical, psychological, communicational, and cultural resource.

Competitive sports events include rules, athletes, coaches, judges, spectators, and many other actors in the sports scene. This chapter leads us to consider training, new paths, and ways of doing and being. It invites us to think about sport beyond performance, as a poetic language of the body in motion.

Music, and artistic components in general, are formative, expressive, and educational facets of contemporary sports. Some initiatives must be taken by sports institutions and professionals in the field to better understand and develop the relationship between music and emotion. The research presented, along with our experiences, has led us to put forward some concrete proposals, such as:

- The inclusion of the relationship between art and sport (addressing the topic of music) in professional training courses in Physical Education and Sports Science, within the university environment.
- The promotion of modules or training courses by sports federations, especially those involving esthetic sports (gymnastics, artistic swimming, and figure skating).
- Greater qualitative investment in the development of sports regulations, addressing artistic aspects.
- More and better training proposals for coaches and judges regarding artistic components, as well as the promotion of discussion and debate among them before and after competitions.
- The development of interdisciplinary research on sport, involving neuroscience, art, and music.
- Fostering greater exchange of research and knowledge among different professional profiles in the sports context.

Musicality, sensitivity, and emotion should be seen as strengths, not obstacles to high performance in sports. After all, it is through art that the sporting gesture becomes humanized in a distinctive way, touching everyone involved and resonating in all other senses throughout life.

For researchers, coaches, and educators, recognizing and exploring these dimensions can significantly broaden the understanding of the meanings of human

movement in contemporary times. More than just a backdrop or accessory, music is an integral part of the sporting experience – and perhaps one of its most human forms of expression.

The relation between music and aesthetic sports..

It's not ancillary, but structuring.

It's not just a rule, it's a meaning.

It's not just reason; it coexists with emotion.

It's not just evaluation, it's formation:

Human, cultural, critical, ethical, and aesthetic.

Author contributions

First Author (**Eliana de Toledo**): Conceptualization, Writing – original draft.

Second Author (Mateus Henrique de Oliveira): Writing – review, and editing.

Data availability statement

Source data not available for this chapter.

Conflict of Interest

The authors declare no conflict of interest.

Author details

Eliana de Toledo^{1,2*} and Mateus Henrique de Oliveira^{1,2}

1 Laboratory of Research and Experiences in Gymnastics (LAPEGI), Sport Sciences – School of Applied Sciences, University of Campinas, Limeira, Brazil

2 Gymnastics Research Group (GPG), School of Physical Education, University of Campinas, Campinas, Brazil

*Address all correspondence to: etoledo@unicamp.br

IntechOpen

© 2025 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Welsch W. Esporte–visto esteticamente e mesmo como arte [Sport – Seen aesthetically and even as art]. *Ética E Estética [Ethics and Aesthetics]*. Rio de Janeiro: Jorge Zahar. 2001;2:142–165
- [2] Graham G. *Philosophy of the Arts: An Introduction to Aesthetics*. London: Routledge; 2006
- [3] Karageorghis CI. *Applying Music in Exercise and Sport*. Champaign: Human Kinetics; 2016
- [4] Best D. Art and sport. *Journal of Aesthetic Education*. 1980;14(2):69–80
- [5] Toledo E, Hess CM, Oliveira MH. Em cena o esporte técnico combinatório: O caso da ginástica rítmica, do nado artístico e da patinação artística [On scene, technical combinatorial sports: The case of rhythmic gymnastics, artistic swimming and figure skating]. In Reverdito RS, Galatti LR, Scaglia AJ, editors. *Pedagogia Do Esporte: Ensino, Vivência E Aprendizagem Do Esporte Na Educação Física Escolar. [Sports Pedagogy: Teaching, Experience and Learning of Sports in School Physical Education]*. Cárceres: Editora Unemat; 2022. p. 17–34
- [6] Furtado LNR, Toledo E, Antualpa KF, Carbinatto MV. Ballet movements in rhythmic gymnastics routines: An analysis from the last two code of points (2013–2016 and 2017–2020). *Science of Gymnastics Journal*. 2020;12(3):395–406
- [7] Sborquia SP. Construção coreográfica: O processo criativo e o saber estético [Choreographic construction: The creative process and aesthetic knowledge]. In: Paoliello E, editor. *Ginástica geral: Experiências e reflexões [General Gymnastics: Experiences and Reflections]*. São Paulo: Phorte; 2008. p. 145–166
- [8] Snyder EE. Emotion and sport: A case study of collegiate women gymnasts. *Sociology of Sport Journal*. 1990;7(3):254–270
- [9] Melo VA. O esporte como uma forma de arte: Diálogos entre (duas?) linguagens [Sport as an art form: Dialogues between (two?) languages]. *Movimento (ESEFID/UFRGS)*. 2007;11(2):111–129. Movement
- [10] Lacerda T, Mumford S. The genius in art and in sport: A contribution to the investigation of aesthetics of sport. *Journal of the Philosophy of Sport*. 2010;37(2):182–193
- [11] Santos JCE. *Ginástica Para Todos: Elaboração de Coreografias E Organização de Festivais [Gymnastics for All: Development of Choreography and Organization of Festivals]*. 2nd edn. ed. Jundiaí: Fontoura; 2009
- [12] Toledo ED, Antualpa KF. The appreciation of artistic aspects of the Code of Points in rhythmic gymnastics: An analysis of the last three decades. *Revista Brasileira de Educação Física E Esporte*. 2016;30(1):119–131
- [13] Van Dyck E, Maes P-J, Hargreaves J, Lesaffre M, Leman M. Expressing induced emotions through free dance movement. *Journal of Nonverbal Behavior*. 2013;37(2):175–190. DOI: 10.1007/s10919-013-0153-1

- [14] Dallas G, Dallas C, Stavropoulou E, Toli D. The effect of classical music on balance, emotional state, and perceived effort in precompetitive artistic gymnasts and trampoline athletes. *Human Movement*. 2024;25(3):10–18
- [15] Toledo E, Oliveira MH, Scarabelim ML, Assumpção B. The impact of the change in the musical attributes of the code of points (2013-2016) on the routines in rhythmic gymnastics of the rio 2016 olympic games. *Science of Gymnastics Journal*. 2018;10:421–435
- [16] Yürük S, Asma M. Music supported exercises in gymnastic education. *CBÜ Beden Eğitimi Ve Spor Bilimleri Dergisi*. 2022;17(2):450–463
- [17] Krafte D, Fernāte A, Lāriņš V. The compositional and musical components of special artistry in rhythmic gymnastics. In: *Society. Integration. Education. Proceedings of the International Scientific Conference*. 2019;4:183–197
- [18] Mezei M, Cristea O. Performance criteria in aerobic gymnastics – Impact on the sportive training. *Procedia - Social and Behavioral Sciences*. 2014;117:347–351
- [19] Tatsi V, Selemidi E, Ypsilanti C, Venetsanou F. Emotional intelligence in acrobatic gymnastics: How important is the athletic profile? *Facta Universitatis, Series: Physical Education and Sport*. 2022;20(2):123–131
- [20] Pizzol V, Oliveira MH, Antualpa KF, Toledo E. Ballet in rhythmic gymnastics: Considerations about technique and instrumentation for the sport. *Science of Gymnastics Journal*. 2021; 95 :49–367
- [21] Toledo E. Aspectos Pedagógicos do Ensino da Ginástica Rítmica . In: Nista-Piccolo VL, Toledo E, editor. *Abordagens Pedagógicas do esporte - modalidades convencionais e não convencionais*. Vol. 1, 1st edn. Campinas: Papirus; 2014. pp. 345–361, Pedagogical aspects of teaching Rhythmic Gymnastics
- [22] International Olympic Committee. Rebeca Andrade fala sobre mundial, Paris e representatividade [Rebeca Andrade talks about the World Championships, Paris and representation] [Internet]. Olympics.com; 2025. Available from: <https://www.olympics.com/pt/noticias/rebeca-andrade-ginastica-mundisal-paris-jogos-exclusivo> [Accessed: 2025-July-30]
- [23] Olympics. Rebeca Andrade’s mesmerizing Floor Routine at Tokyo 2020! [Internet]. Youtube; 2021. Available from: <https://www.youtube.com/watch?v=us2zpccecdc> [Accessed: 2025-July-09]
- [24] Olympics. Rebeca Andrade - Apresentação no solo (Final Ginástica Artística) [Rebeca Andrade - Floor Presentation (Artistic Gymnastics Final)] #Paris2024 [Internet]. Youtube; 2024 Available from: <https://www.youtube.com/watch?v=kmOGe5Hv-mg>. [Accessed: 2025-July-09]
- [25] Karageorghis CI, Priest DL. Music in the exercise domain: A review and synthesis (Part I). *International Review of Sport and Exercise Psychology*. 2012;5(1):44–66
- [26] Bateman A, Bale J, editors. *Sporting Sounds: Relationships between Sport and Music*. London: Routledge; 2008
- [27] Belenkaya I. Musical training of coaches in aesthetic-oriented sports.

Slobozhanskyi Herald of Science and Sport. 2016;5(3):5–8

[28] Bredikhina YP, Kapilevich LV, Shunkina A. Music influence on athletes' sportsmanship (from the example of synchronized figure skating athletes). *Journal of Economics and Social Sciences*. 2018;(12):1–5

[29] Kim SE, Chung HS. An analysis of figure skating costumes in relation to the musical accompaniment. *The Research Journal of the Costume Culture*. 2011;19(3):579–596

[30] Adams ML. *Artistic Impressions: Figure Skating, Masculinity, and the Limits of Sport*. Toronto: University of Toronto Press; 2011

[31] Terry PC, Karageorghis CI, Curran ML, Martin OV, Parsons-Smith RL. Effects of music in exercise and sport: A meta-analytic review. *Psychological Bulletin*. 2020;146(2):91

[32] Karageorghis CI *The Scientific Application of Music in Sport and Exercise*. [Place unknown]: [Publisher unknown]; 2008

[33] Plessner H, Schallies E. Judging the cross on rings: A matter of achieving shape constancy. *Applied Cognitive Psychology*. 2005;19(9):1145–1156

[34] Veit F, Veit J, Heinen T. The influence of music on judges' evaluation of complex skills in gymnastics. *European Journal of Sport Science*. 2022;1(5):1–7

[35] Oliveira MH, Toledo E. Diferenças entre compreensão e aplicação de regras do código de pontuação na visão de árbitros(as) brasileiros(as) de ginástica

artística. [Differences between understanding and applying Code of Points rules from the perspective of Brazilian artistic gymnastics judges]. *Pensar a Prática*. [Thinking Practice]. 2023;26:1–26

[36] Pajek MB, Kovač M, Pajek J, Leskošek B. The judging of artistry components in female gymnastics: A cause for concern? *Science of Gymnastics Journal*. 2014;6(3):5–12

[37] Suominen S. Judge bias in aesthetic group gymnastics. *Science of Gymnastics Journal*. 2023;15(3):441–452

[38] Van Bokhorst LG, Knapová L, Majoranc K, Szebeni ZK, Táborský A, Tomić D, et al. "It's always the Judge's fault": Attention, emotion recognition, and expertise in rhythmic gymnastics assessment. *Frontiers in Psychology*. 2016;7:1008

[39] Flessas K, Mylonas D, Panagiotaropoulou G, Tsopani D, Korda A, Siettos C, et al. Judging the judges' performance in rhythmic gymnastics. *Medicine & Science in Sports & Exercise* 2015;47:640–648

[40] Fernandez-Villarino MA, Bobo-Arce M, Sierra-Palmeiro E. Practical skills of rhythmic gymnastics judges. *Journal of Human Kinetics*. 2013;39:243

[41] Oliveira MH, Toledo E. Competências "práticas" de árbitros(as) brasileiros(as) de Ginástica Artística: Um panorama diagnóstico. ["Practical" skills of Brazilian Artistic Gymnastics judges: A diagnostic overview] *Rev Bras Educ Fís Esporte*. *Brazilian Journal of Physical Education and Sport*. 2023;37 (Esp):e37nesp215406

[42] Leandro C, Ávila-Carvalho L, Sierra-Palmeiro E, Bobo-Arce M. The

evaluation rules in the view of the rhythmic gymnastics judges. *Journal of Sports Sciences*. 2016;**4**:232–240

[43] Stokowski S, Corr C, Wells R. Does music matter? Examining genre and tempo in collegiate women's gymnastics floor routines. *Journal of Sports and Games*. 2025;**7**(1):21–29

[44] Mercier H, Heiniger S. Judging the judges: Evaluating the performance of international gymnastics judges. *arXiv [Preprint]*. 2018. Available from: <https://arxiv.org/abs/1807.10021>

[45] Van Dyck E, Vansteenkiste P, Lenoir M, Lesaffre M, Leman M. Recognizing induced emotions of happiness and sadness from dance movement. *PLoS ONE*. 2014;**9**(2): e89773. DOI: 10.1371/journal.pone.0089773

Section 4

Music and Neuroscience

Chapter 5

Hyperscanning and Mirror Neurons: Musical Ensemble Playing from a Neuroscientific Perspective

Christiane Neuhaus

Abstract

Musicians performing in “small size” – as a duo or a quartet – are often admired for their elegance and skillfulness. This chapter explores the reasons why. The perspective is that of cognitive neuroscience, whereby processes on the micro and macro levels are considered. Part one takes a closer look at the special “mirror neurons”. The way they function within certain regions of the brain enables the players to quickly understand the movements of others and react precisely on time. In part two, a new and promising measuring method – called “hyperscanning” – will be introduced and discussed from all sides. Hyperscanning with EEG allows the recording of brain activity from two or more musicians simultaneously while playing in a concert. These inter-brain EEG patterns can then be analyzed for their spectral content and phase information, from which conclusions can be drawn about the state of brain synchrony and its change over time, or about the strategies employed during performance.

Keywords: ensemble playing, string quartet, social neuroscience, mirror neurons, hyperscanning, EEG

1. Introduction

Playing in a string quartet is high art – one of the most cultivated forms of making music. For many concert-goers, it is the epitome of elegance combined with reasonableness and mutual respectfulness. Even for Goethe, it was worthy of remark. For him, it was “a conversation among four intelligent people” (1829) [1].

What seems to attract is the way the string players deal with musical motifs: they pass them on – from the first violin to the second, then to the viola, the cello, and back – by repeating, modifying, and contrasting them, all done in a sophisticated and well-balanced manner. It is not surprising that string quartet literature has been (and still is) a rich field for music historians and analysts.



Figure 1.
String quartet playing as a model for teamwork effectiveness. The Goldmund Quartet, Munich, founded in 2009.

Haydn, for example, composed 68 pieces for this group, Mozart wrote 23, and Beethoven 16.

Today, string quartet playing has also been discovered by researchers in the fields of movement science, sociology, and even neuroscience. They consider it a model to investigate teamwork effectiveness and the dynamics within a group (see **Figure 1**). So, the focus has shifted to a certain extent: from analyzing the scores in terms of musical structure toward the “process” of playing, for which neural mechanisms are essential.

This chapter discusses musical ensemble playing from both the micro and macro perspectives of neuroscience. Part one takes a closer look at the so-called human mirror neuron system (MNS). The way it functions enables string players to react precisely on time and adjust themselves to the group. In part two, a new and promising measuring method – called “hyperscanning” – will be introduced and discussed from all sides. Hyperscanning allows researchers to record brain activity from two or more musicians simultaneously during concerts or rehearsals. Before delving into the matter, let us start with what has been discovered about string quartet playing on the behavioral level.

2. Ensemble playing – the behavioral side

The secret that makes a string quartet operate as a skillful work group lies in its subtle non-verbal mechanisms. Social cues like eye contact, facial expressions, or micro gestures help group members improve their play and/or deal with the performance situation. This has mainly been studied with three behavioral methods. The first: independent researchers analyze the performance behavior offline, provided that concerts have been video- and audiotaped. The second is the same, based on self-assessment, in which group members evaluate the quality of their play [2]. And the third consists of analyzing motion capture data, for which the string players must wear special suits with attached markers to precisely capture the movements of arms, hands, or the bows [3]. Based on these methods, the string quartet has been described as a “resilient system” [4], because group members are able to quickly react to disruptive influences like “being out of time” to re-establish

the previous state within a few seconds. Volpe et al. [5] point out that such compensating adjustments are grounded in special cognitive skills, in particular in divided attention and the ability to anticipate. In other words, ensemble musicians must be able to oversee and foresee both their own movements and those of the others to maintain their coordinated play. Each string player is equally responsible for the performance's outcome, making a string quartet act as a "self-managed team" [4].

Some interesting details have been found in relation to the movements of the head: head movements of musicians are considered as markers of expressiveness, since, obviously, they are fairly independent of instrument-specific movements [3, 5, 6]. Glowinski et al. [3] studied the head movements of two first violinists (one of whom was a member of the renowned Quartetto di Cremona) by using motion capture, video cameras, and stereo microphones. Both musicians had to repeat some short sections several times (taken from a string quartet by Schubert), either by playing the parts alone or together with the quartet. Their head movements were significantly more regular and foreseeable when playing along with the group than in the solo condition. In addition, the Cremona players reported that they align to a subjective and invisible common reference while playing – a "center of gravity" in front of them, called the "Ear" [3, 7] (also see **Figure 2**). To what extent this behavior is idiosyncratic or common practice among quartet members in general must be clarified elsewhere.

To put it in a nutshell: Sensorimotor cues such as eye contact, facial expressions, or head movements are essential means of communication, particularly when performing in real time on stage. Volpe et al. use the term "sensorimotor conversation" [5], which is, more or less, the action-based variant of Goethe's bon mot of "four intelligent people having a conversation with each other" [1].

A propos: What is correct when describing such performance behavior? Should one speak of "communication" and "conversation", or better of "interaction"? Obviously, this is a matter of perspective. Behne [8], for example, argues in favor of the word "communication" by focusing on the message as such.

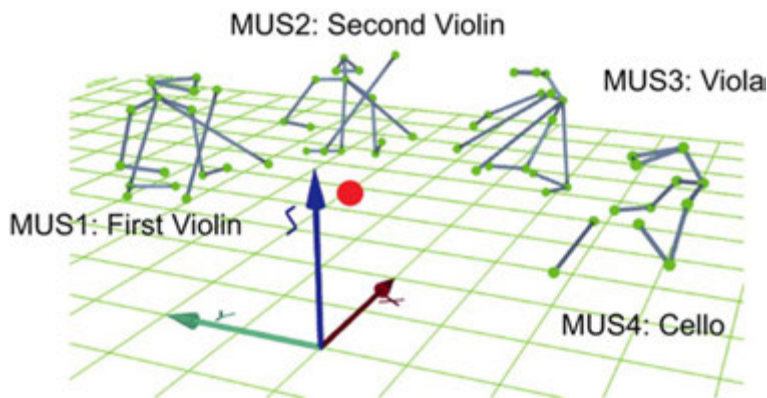


Figure 2.
 The "Ear" (red dot) – an imaginary reference point helping the quartet to keep track of the common goal [7] (red dot slightly enlarged by the author).

According to him, a musical message is based on certain rules (similar to language) with sounds, rhythms, and melodies as the elements. Even so, music does not aim at clearly informing the listener about something, as language does. Rather, it is an offer to the listeners (and also to the artists) to experience a certain mood or emotion through sound (associations included). Because of this vagueness (or ambiguity) of what is conveyed, Behne concludes that music is a kind of “as-if-communication” [8].

Interaction, by contrast, is a term commonly used when people coordinate their movements to handle a situation [9]. Interactions are complex forms of tuning-in and alignment, of making things work effectively as a whole. They vary in the number of participants and the type of non-verbal behavior [10]. However, all kinds of interactions share certain features: a common goal, mutuality in regard to doing (complementing each other) and feeling (“empathetic attunement” [11]), as well as in regard to temporal concurrence and attention directed to both group members and the object as such [10, 12, 13]. A good example is group improvisation in jazz. However, on closer examination, improvisation among two, three, or four people could also move the debate back in the direction of “conversation” or “discourse”, as musical ideas are spontaneously created on the basis of the generative power of the mind, similar to speech-based dialogs in language. The hypothesis that “improvisation is analogous to spontaneous speech” [14] was tested by a study by Donnay et al. [15] using functional neuroimaging as a method. In this study, a sort of “call-and-response”-situation was created in which professional jazz pianists improvised in pairs on small keyboards – one musician inside and one outside the scanner – by passing four-bar phrases to and fro, which is a playing technique called “trading fours” [15]. While doing so, functional brain activity was measured, showing that “call and response” activated the “classical” areas for language processing – which are left-hemispheric – (and also their right-hemispheric counterparts), providing evidence that “improvisation is [indeed] analogous to spontaneous speech” [14].

Let us quickly touch on the subject of ethnomusicology. In many cultures outside the Western world, especially in the Arab countries and on the Indian subcontinent, it is common practice to perform “in small-size” too. Even more, small formations – in the size of a chamber music ensemble – are the predominant type. However, on closer examination, their performance practice reveals that their concept of music-making is entirely different from how chamber music is understood in the Western world. First, music in these countries is almost always improvisatory in character. It is based on certain melodic formulae called *ragas* (in India) and *maqamat* (in the Arab world) that are modified spontaneously without relying on a musical score. In addition, their playing practice makes visible that role distribution is not equal for the most part. Rather, music organization is hierarchical, whereby the leading musician takes responsibility for the performance outcome. Detailed observations were made by Bonnie C. Wade, an American ethnomusicologist [16], who studied the performance practice of art music in India, particularly from the North. According to her observations, Indian ensemble musicians can be divided into those who are soloists and those who have an accompanying role. The first produce the melody – like the famous sitar player Ravi Shankar – whereas the others play the percussion instrument (e.g., the North Indian tabla) or produce the drone, a constant nasal-metallic sound by plucking the tanpura. In some ensembles, there is also a second sitar player who, however, has no independent role but rather orients

himself (or herself) toward the leading improvising instrument. From this, one can conclude that performance styles vary throughout the world. Both can be found: leading and following, as well as teamwork style. D'Ausilio et al. [17] call this "autocracy versus democratic group organization".

3. Social neuroscience and its philosophical basis

How can all this be explained by neuroscience? The essential point is that, about three decades ago, cognitive neuroscience underwent a change (or better: it experienced an expansion) – from the individual brain to a "two-brain"-situation. While the conventional form still focuses on how stimuli are processed in the single brain – with aspects like perception, judgment, recognition, and/or decision-making – the new branch, called "social neuroscience", enables researchers to focus on "cohesive processes" between two or more brains during joint attention or joint action [18], which is a matter in itself but also the prerequisite to understanding motivation and the intentions behind. Social neuroscience is based on the premise that neural processes are entirely different in situations of "reciprocal exchange" compared to unilateral approaches, where persons merely "observe" some sort of social scenario from the perspective of a spectator [10, 19, 20]. Social neuroscience is thus concerned with neural mechanisms underlying active processes in "real-life situations", including gestures, mutual eye contact, and facial expressions [20].

Interestingly, the rise of this "second-person neuroscience" [19] was preceded by a similar fundamental change in philosophy approximately 180 years earlier. The starting point of this change in perspective was the language-theoretical lecture "About the Dualis", given by Wilhelm von Humboldt at the Berlin Academy of Sciences in April 1827. Since then, dialog has appeared as a subject of philosophical debates, that is, the alternating between a speaker – the "I" – and the addressee – the "You". Through this, the prevailing stance of "ego-centrism" was called into question for the first time, established two centuries earlier by René Descartes (1596–1650) and his famous saying "cogito – ergo sum" [21, 22]. Humboldt's view was taken up by a group of philosophers called "dialogue philosophers" (first half of twentieth century), of which the most famous was the Jewish philosopher Martin Buber (1878–1965). Buber set the focus on the two-person relationship itself, that is, on the mere constellation between the "I" and the "You". Through this, so Buber, a new category comes into being, he termed the "In-Between" (in German: "*das Zwischen*"). This "In-Between" exists per se, independently from conversation or interaction. It is also present when two persons remain silent [22]. According to Buber, the "You" is also important from an ontogenetic perspective when the developmental stages of a child are considered. That is, in a mother–infant relationship, a child first says "You" before he/she says "I". So, there is, first, the primary relationship with the mother from which the child starts to develop some sort of self-awareness (before, then, building connections with some other sort of "You"). In short: "You" (mother) – "I" – "You" (other). In Buber's words: "The 'I' comes into being because of the 'You'" (in German: "Der Mensch wird am Du zum Ich" [23]).

Interestingly, this "I–Thou" relation was also discussed by "contemporary phenomenologists", with Edmund Husserl (1859–1938) as one of its main representatives. Their focus, however, was on "empathetic perception", thus

shifting the focus toward the emotional aspects of human interaction, toward “emotion sharing” [24], “joint musical experience” [11], and “the affective aspects of engagement” [10]. Although there is no doubt that this makes second-person approaches more complete, the emotional aspects of interaction will not be considered here.

4. Ensemble playing under the microscope – what functions do mirror neurons have?

Back to social neuroscience and the string quartet as a model for studying non-verbal interaction. Of note is that the brain mechanisms that make a string quartet operate as a skillful work group are, for the most part, elementary (or low-level). More precisely, the reason for such behavior lies in the properties of a special type of neuron called “mirror neuron”. So, what occurs on this micro level? What is it all about?

Mirror neurons were discovered purely by chance while conducting research on the monkey brain. In those days, the spike activity of certain single nerve cells was studied in detail by inserting micro-electrodes into their axons. Through this method, an Italian group of neuroscientists led by Giacomo Rizzolatti (as the principal researcher) made an important discovery: specific neurons in area F5, the monkey’s ventral premotor cortex, discharge with exactly the same intensity when a movement is either observed or performed – for example, when the monkey grabs a nut or sees another monkey grab a nut [25]. These nerve cells, called “*mirror neurons*”, belong to the group of bimodal neurons (see **Figure 3**). They discharge when goal-directed movements are either observed or self-performed. Kohler et al. [26] demonstrated that a subset of these neurons is even trimodal. These neurons also respond to action-related sounds, such as when the monkey hears someone cracking a nut with a nutcracker. In short, mirror neurons discharge in response to actions that are either seen, heard, or self-performed. By contrast, emotional sounds or other non-action-related vocalizations do not trigger any discharge at all [26].

What does all this have to do with our string quartet, described as a skillful work group? Let me explain this step by step: The first point is that mirror neurons are also an essential part of the human brain. They can be found in high concentration in six “spots” located across the human cortex – making up the “human mirror neuron system” (MNS) – of which the ventral premotor region and the inferior parietal lobule are of interest here. Note that, due to evolutionary progress, the MNS of humans differs slightly from that of monkeys [25]. While the mirror neuron system of monkeys merely reacts to movements directed toward an object, for example, a nut or a banana, the system in humans also reacts to the intransitive form, to the gestures themselves. What we have here in both humans and monkeys is some sort of automatized switching, often called “mapping”, since what is observed – the motion – is automatically passed onto the observer’s motor repertoire by activating the respective movement representation [25, 28]. In this way, actions of others can be recognized immediately without reflecting upon them.

For whom is this automatism good? What kind of advantages does our string quartet gain from this mechanism? Automatized switching, based on mirror neuron properties, enables the players to decode the observed movements of hands, fingers, or the bow within fractions of a second, which allows them to react precisely on time.



Figure 3.
(a) The leading figure in mirror neuron research: Giacomo Rizzolatti from Parma University (Italy). (b) Due to advanced light microscopy, mirror neurons can be made visible; one is depicted here (modality type [bi- or trimodal] not further specified) [27].

In other words, string players are able to understand the movements of the other group members intuitively without thinking, since what has been seen is immediately mapped onto their own motor repertoire. Even more, this enables them to anticipate the movements and actions of their teammates merely by watching. Thus, professional musicians are able to act in advance and also to hear the sound result ahead of time. They can even automatically grasp the intentions behind the actions, since this can often be inferred from seeing the movement as such (see [29, 30] for further details).¹

Besides their important role in “understanding the actions and intentions of others”, more has been discovered about the MNS in connection with music: First,

¹ Van Overwalle and Baetens [31] have shown by a meta-analysis that, in addition to the mirror neuron system reacting quickly and automatized in regard to perceived movements, a second, more abstract, system – called “mentalizing system” – is active in the human brain, mainly comprising the temporo-parietal junction (TPJ) and the medial prefrontal cortex (mPFC). TPJ is activated on the level of transient and task-relevant goals, whereas in mPFC aspects are processed that are highly abstract and more enduring such as beliefs or goals in life. Accordingly, the understanding of goals is a hierarchical process for which two distinct but complementary systems exist – the mirror neuron system on one hand and the mentalizing system on the other. Both may slightly differ in their time course of activation.

this “seeing and doing” is the basic principle on which “imitation behaviour” (and “imitation learning”) are grounded – from birth onward. Meltzoff and Decety [32] demonstrated that 12-to-21-day-old infants are capable of “neonatal imitation” [32, 33]; they are able to imitate the facial expressions of an adult straight away, for example, sticking out their tongue. This illustrates how important the MNS is for thriving development in early infancy. Even more, since facial imitations have already been observed in 42-minute-old newborns [32], imitation is undeniably an innate capacity with the MNS as its underlying mechanism.

Back to music: Mirror neurons also play a central role in teacher–student interactions, particularly in situations of imitation-learning, such as when learning to play a musical instrument. Buccino et al. [34] have shown, using functional MRI, that activity in the right ventral premotor cortex (as part of the MNS) is significantly increased when hand actions are first observed and then imitated. In this fMRI study, beginners were tasked with closely watching the hand positions of a professional guitar player (shown via monitor) and immediately repeating these positions on a fingerboard that had been placed in the scanner.

Newman-Norlund and Bekkering [35] demonstrated that “complementary” hand grips, opposite to those previously observed, activate the MNS even more strongly during the preparation phase. This can be explained by the fact that two-thirds of mirror neurons are “broadly congruent”, whereas the remaining third are “strictly congruent”. Broadly congruent mirror neurons discharge when counterpart variants (or opposite movements) are planned, whereas strictly congruent mirror neurons discharge only when movements must be executed in exactly the same manner as previously observed [25, 35]. One may deduce from this that certain “call-and-response” situations during jazz improvisation [6] will activate the MNS more intensely than, let’s say, pure motif repetition, since when a counterpart has to be invented, the broadly congruent mirror neurons are active.

Even more interesting is the question of whether the MNS is open to plasticity, that is, if it can expand structurally and if its functioning can be optimized through training. Apparently, yes. Efficient training can lead to refined movement representations in the brain onto which the observed movements (from skilled others) are matched. Calvo-Merino et al. [28] therefore speak of an “influence of motor expertise on action observation”. They themselves validated this in a study with functional brain imaging: 19 professional dancers, either trained in capoeira or classical ballet, had the task of watching video clips showing capoeira or ballet movements [28]. Due to their own motor experience in one of these styles, their MNS was more strongly activated bilaterally when watching movements of their own style than when watching movements of the other style. Obviously, training has a backward effect on how actions are seen. From this follows that the extension of one’s own motor repertoire is beneficial for sophisticated action understanding, for distinguishing between the subtle movements of teammates when observing them. In other words, expertise shapes our MNS, which, in turn, makes us see and understand the subtle nuances of gestures made by others.

All this shows that the MNS strongly affects the interactions among group members and the quality of a performance. In other words, the success of a string quartet as a “self-managed team” [4] depends largely on the malleability (or plastic change) of the MNS resulting from experience and training. Accordingly, the skillfulness of a musical ensemble, that is, to what extent players are able to deal with

non-verbal cues and micro-movements, is largely a low-level mechanism based on the properties of the mirror-neurons.

5. Hyperscanning with EEG – points and problems

From the cause to the method: What does the word “hyperscanning” mean? “Hyperscanning” is an extended recording technique by which the activity of two or more brains can be measured simultaneously. Multi-brain recordings bring both into focus: intra- as well as inter-brain activity, that is, processes occurring within each brain as well as those occurring between the brains, for instance, when a string quartet rehearses for a concert [36–38]. This approach coincides with the general idea that “interactive *and* individual elements in social cognition” ought to be systematically evaluated [10, *italics added*]. Hence, hyperscanning is commonly regarded as a promising tool to substantiate what is called the “interactive turn in social cognition research” [39, 40].

Here is a short list of example paradigms for which hyperscanning is a suitable method (in particular, when using that type with electroencephalography; see below): (1) Tasks: Interactions through eye gaze [41]; two-person finger tapping [38, 42]; musical ensemble playing [43, 44]. (2) Performance strategies: Cooperation versus competition [45, 46]; leader versus follower [47]; score-based versus improvised music making. (3) Role-taking/turn-taking: Complementary – questioner and respondent; pair dancing [36]; directional: leader and follower [47]; imitation [48]. (4) Group size: Two persons [44], four persons [43], or classroom size [49]. (5) Musical instruments: All types, for example, guitars [44], pianos [50]; strings, saxophones [43].

The variety continues when having a look at the recording methods themselves [36–38]. Mainly, three of them are in use: hyperscanning with functional MRI (fMRI), hyperscanning with fNIRS (functional near-infrared spectroscopy), and hyperscanning with EEG (electroencephalography).

The fMRI type requires that a lab be equipped with two scanners. If so, two subjects can be measured simultaneously when both machines are linked to each other. (Note that some time ago, a special dual-coil system was developed by which synchronous recordings could be done with two persons lying in one scanner; see the study by Babiloni and Astolfi [51] for details.) One of the big advantages of doing hyperscanning with fMRI is the fine spatial resolution, by which functionally activated regions can be located with high precision. However, a severe limitation is that, because of lying in the scanner, participants are completely unable to move, which is exactly the opposite of what one would expect from experiments in which interaction is investigated. Although any form of gesture-based interaction is unfeasible, hyperscanning with fMRI still allows addressing issues of conversation, like turn-taking or role-taking.

In contrast to that, a wide range of interactive scenarios has been investigated with fNIRS hyperscanning and with the EEG technique. The fNIRS is a method of optical recording, using caps with optodes placed on the scalp, which allows for precise investigation of the brain’s topography, although it is merely restricted to the cortex (the brain’s surface). In comparison with EEG, fNIRS is less susceptible to movement/motion artifacts; however, temporal resolution is lower [37, 38, 42]. Since both systems are also available in portable

form, costs can be considerably reduced while, at the same time, ecological validity is high.

As the chapter proceeds, only those studies will be discussed in which hyperscanning has been done with EEG. The “good old EEG” has an excellent time resolution, so that even tiny changes – intra- as well as inter-brain – can be detected [52]. EEG hyperscanning is done with two or more electrode caps, each consisting of 32 (or 64) electrodes placed on the scalp according to a standardized schema called the 10–20 system (see the study by Jasper [53] for details). The most challenging point is to connect the equipment’s components in such a way that the delay, due to signal transmission, will be in the range of some milliseconds. Ideally, multi-brain EEG data are only minimally affected by temporal jitter. **Figure 4** (taken from the studies by Babiloni et al. [43] and Barraza et al. [54]) shows two professional solutions: on the left, we see the setup used in the (early) EEG hyperscanning study done by Babiloni et al. [43]. They work with some sort of “device duplication” (or: “device multiplication”), where data recording and transmission run separately before the signals converge into a single workstation. This way, an ensemble of four Italian saxophonists was “hyperscanned” while playing an arranged version of a piece by Domenico Scarlatti. On the right (see **Figure 4**), a more user-friendly version is shown, which simply consists of two components: the standard equipment for EEG mobile recording (with two [or more] electrode caps) and a special cable for connecting the recording sets to a trigger-sending computer by which the multi-brain data can be synchronized later, that is, offline (see **Figure 4**; further details in the study by Barraza et al. [54]). Note that for safety reasons, each EEG hyperscanning setup must include two (or more) separate ground electrodes (GND) to individually connect each participant to earth.

Hyperscanning data – how can they be analyzed? Which measures are used for argumentation? What is quantified, in principle, is coherence, that is, temporal concurrence and the degree of similarity. In other words: to what extent brains will be synchronized when persons interact with each other. Inter-brain synchrony is, for the most part, quantified by calculating phase relationship and/or by determining spectral content [37, 38, 42]. Regarding the first parameter – phase – EEG curves from two (or more) musicians are assessed for coincidence, that is, for reaching the same position within the oscillatory signals at the same time. Phase synchrony is achieved when peaks and troughs between inter-brain signals are aligned [42], as indicated by markers called “phase-locking value” (PLV) or “phase synchrony index” (PSI) [36, 37]. Regarding spectral content – the second feature – persons’ EEG curves are analyzed for similarities in the frequency domain. For this, EEG signals (taken from the same electrode positions) are split up into several frequency bands (delta to gamma) via Fourier analysis to compare them across the brains. High resemblance indicates that brain-to-brain synchrony is high, and vice versa [36, 37]. This type of coherence analysis goes back to the 1980s; however, it was merely used for single brain recordings at that time [52, 55]. In this way, intra- and inter-hemispheric coherence has been studied during processes of creativity, for example, while composing a short piece of music spontaneously (see the study by Petsche [55] for details).

But back to multi-brain recordings: Hyperscanning data can also be analyzed with a statistical method called “Granger Causality”² [36, 37]. Granger Causality is based

¹¹⁰Developed in the 1960s by Clive WJ Granger, a British econometrician.

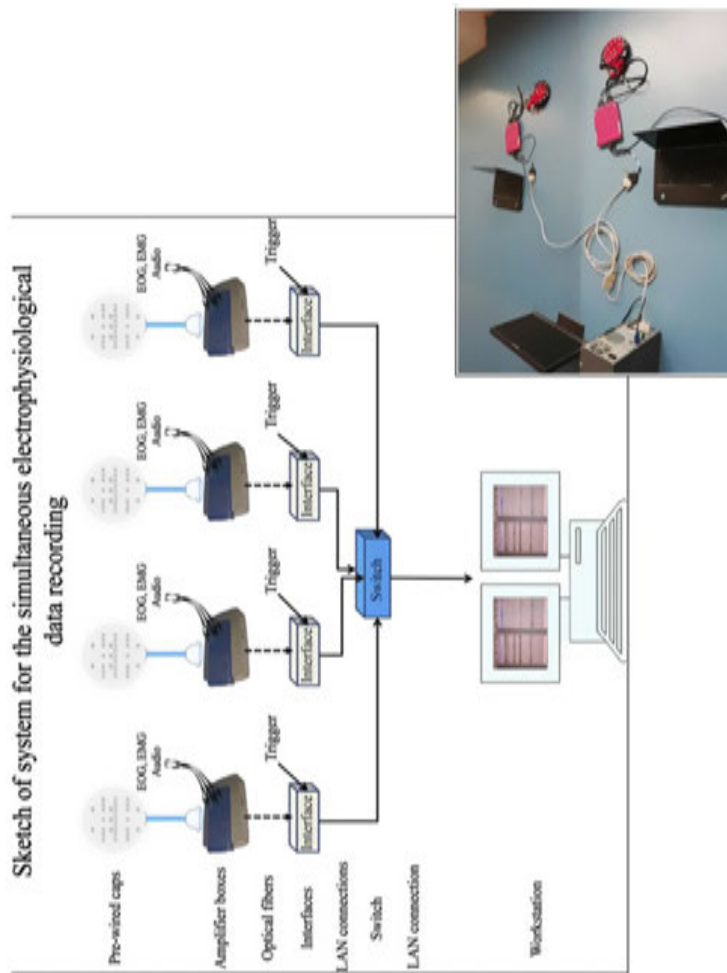


Figure 4.

Two professional setups used for EEG hyperscanning research [43, 54].

Left: Principle of “device multiplication”: Recording and amplifier devices for all saxophone players separately (four 32-channel electrode caps, four amplifier boxes). Amplifier units have additional input channels for electrooculographic (EOG) and electromyographic (EMG) signals used for artifact detection (e.g., eye blinks and muscle activity of the face and arm). Audio signals from microphone-recorded performances are also amplified to synchronize EEG data with the musical segments. The four interface devices in line receive the same external triggers to synchronize the EEG datasets with each other. The boxes converge to a single workstation for further signal processing (all datasets together).

Right: Two mobile EEG devices (two 64-channel caps, two amplifiers) are connected to two separate recording laptops. Both amplifier units are also connected to a trigger-sending computer (located on the left side) to synchronize the participants’ brain data.

on the assumption that data, in this case the EEG curves, are considered as time series. Given this, it is a proper method to prove whether a musical interaction is a leader–follower interaction, in that a soloist or a conductor – the leader – has a transient but direct influence on the accompanying musician or the orchestra, – in the sense of “cause and effect” [37, 42]. If so, a sort of time-delayed synchronization can be observed in the EEG data. More precisely, the leader’s specific EEG pattern at one electrode will, with a time lag, be roughly reflected in the EEG pattern of the follower. This (transient) process can also be shown for kinematic data. It is interpreted as an “index of successful leadership” [56].

However, things are more complicated when trying to define the relationship between brain activity and performance output, that is, when connecting the neural with the behavioral side. Is this relationship correlational, or is it causal? Regarding the first, brain synchrony could merely be a concomitant phenomenon, happening in coincidence with playing (see also below). Regarding the latter, brain synchrony must be the initial cause, an absolute prerequisite for social interaction, which “leads to changes in behavior” [57]. Novembre and Iannetti [58] point out that hyperscanning is, in principle, not designed for studying “cause and effect”. Obviously, this is also a matter of method. To solve this causality problem, they suggest the inverse: First, the brain tissues of two (or more) interacting persons are stimulated with transcranial stimulation (TMS) to then see what effects on social behavior this would have [58, 59].

Let’s put this aside – what are the pros and cons of doing hyperscanning research with EEG? Hyperscanning describes (and reveals) the “what is”, the coupling of brains during joint actions, for instance, when musicians play together (an example is depicted in **Figure 5**). First of all, one must appreciate the possibilities thus created: linking devices has opened up a new avenue for studying the simultaneous processes in the brains of people communicating or interacting with each other. A further advantage is the independence from lab-based testing, since hyperscanning devices

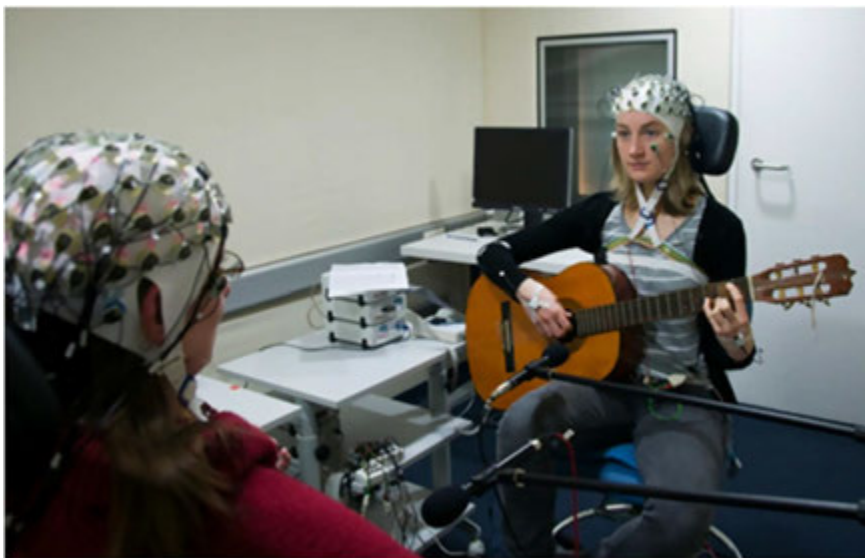


Figure 5.
An example paradigm in EEG hyperscanning research. Brain activity is recorded from two guitar players, each wearing an electrode cap [48].

are also available in portable form. This enables researchers to carry out experiments in situations of high ecological validity, for example, in a concert hall.

More importantly, hyperscanning offers the possibility to study the dynamic changes as the performance proceeds, that is, the altering patterns of intra- and inter-brain synchrony, for instance, when coming together after having been out-of-time. Such dynamic changes in brain synchrony have been extensively investigated by Müller, Lindenberger and colleagues [44, 60, 61] with guitar duos, guitar quartets, and other ensemble types. The results of their studies will be described in Section 6.

Let us turn to the disadvantages of conducting hyperscanning research with EEG. The first (of three points) is essential, as it touches on the problem of validity: Several researchers [36, 39, 62] pose the question of whether brain synchrony could merely be a by-product (or epiphenomenon) when musicians play together, meaning that it could already occur when two (or more) persons are simply listening, that is, paying attention to the same tone sequences they produce. According to this, a direct reciprocal person-to-person engagement is not necessarily required. In other words, brains could already synchronize when exposed to the same external factors, for example, when experiencing the same melodies, with no active form of interaction being necessary, for example, when playing in separate rooms. Holroyd therefore speaks of “interbrain synchrony on wavy ground” [62]. This point indeed touches on the problem of validity.

Second: Hyperscanning has been used for studying highly heterogeneous topics, ranging from finger tapping to teacher-pupil interactions in the classroom. For this, Holroyd used the expression “Wild West of task paradigms” [62]. What has been missing until now is a stepwise approach, where instructions are varied systematically and effects can be clearly disentangled from each other. Third: Hyperscanning reveals the “What is” when musicians perform together, so the focus is primarily on the coupling of brains during motor behavior (or joint action). However, many more cognitive (and emotional) processes are connected with this – think, for instance, of error correction.

One of the first attempts to disentangle the cognitive processes from pure sensorimotor activity was made by Gugnowska et al. [50]. They used a tempo-switching paradigm in which the new tempo had to be planned during short moments of silence. In this EEG hyperscanning study, pairs of amateur pianists were tested – one playing the melody, the other playing the bass line of Bach chorales, in which two-bar musical pauses had been inserted (see **Figure 6**). Visual cues signaled whether the second phrase – after the pause – should be played faster or slower than the first. In this way, brain activity (intra- and inter-brain) was studied while both musicians were planning the new tempo during those two bars of complete silence, in which no form of active playing (or joint action) was required (see **Figure 6**). However, the critical point is that both musicians were sitting in different rooms without any possibility to interact, for example, by eye gaze, which means that brain coupling was merely triggered by the cues themselves, whereas real synchrony, due to direct reciprocal exchange, did not exist.

6. Ensemble playing – results from EEG hyperscanning

Of course, one can raise the objection as to whether hyperscanning data are valid and meaningful at all. Yet, it is the only way to gain insight into the brains’

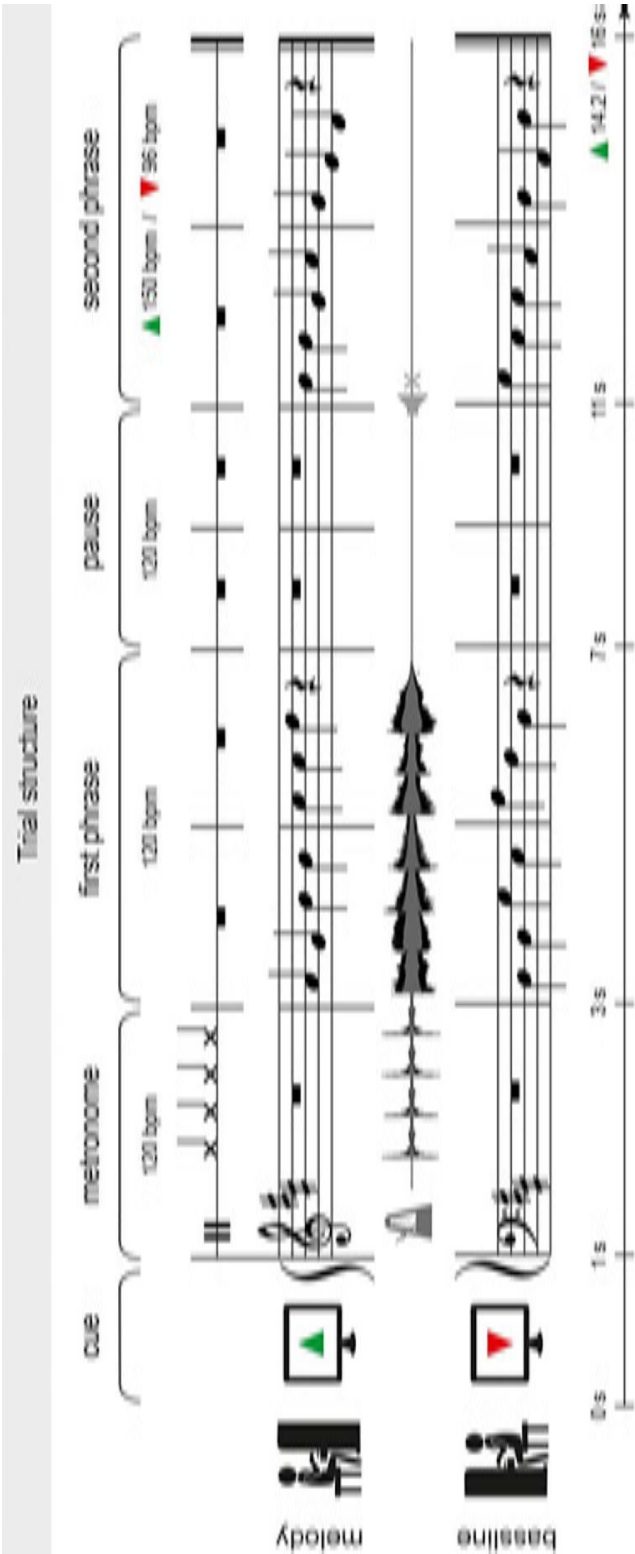


Figure 6. The tempo-switching paradigm used by Gagnouska et al. [50]. Both musicians were planning the new tempo during the musical pause in bars 4 and 5.

simultaneous processes beyond what can be seen from the outside. So, what has been found until now? From the many scattered results obtained with EEG hyperscanning, I would like to pick out four, starting with the brains' activity at the very onset of a musical piece while tuning-in. Each experienced concert-goer will sometimes draw his or her attention to this moment, especially when a musical piece is in a slow tempo.

This “swinging-in” of the brains becomes tangible by analyzing the time span prior to phase synchronization, as already done by Lindenberger et al. in 2009 [44]. In this EEG hyperscanning study, pairs of professional guitarists were asked to play a short jazz-like melody 60 times in unison, while particularly focusing on the beginning – the piece's start. Immediately before the onset, four metronome beats were given to facilitate temporal alignment. This way, EEG activity could be analyzed three seconds before and two seconds after the onsets of both guitars.

This study is noteworthy because of two aspects that also play a role in many other (EEG) hyperscanning studies. The first point is about brain topography: Lindenberger et al. [44] observed that brain synchronization mainly takes place in the prefrontal cortex (see also [47, 51]), which is convincing since action planning and cognitive control are located there, and because working memory, in general, has its seat in the dorsolateral prefrontal part. In other words, the coupling of brains is highest at frontal and central electrode sites.

The second point is based on the assumption that EEG curves are split into their spectral components, the frequency bands. A somewhat remarkable result is that the lower frequency bands, in particular theta oscillations, are more important than one might expect. That is, during temporal adjustment, phase alignment occurs primarily in the theta- and delta-bands [44], and this is true for both synchronization “within” and “between” the brains.

More importantly, however, are studies focusing on intra- and inter-brain coupling as such, that is, when phase synchronization has been put aside [60, 61]. These studies could show that the same joint activity, for instance, two musicians rehearsing for a concert, produces a sort of “frequency-based dissociation” in that inter-brain coupling occurs at lower frequency bands (theta and delta), while, at the same time, intra-brain coupling takes place at higher frequencies (8 Hz and higher), showing that different oscillatory networks are active. This frequency-based dissociation was found in eight pairs of guitarists improvising jazz together [60] and also in a quartet of professional guitarists who were playing tango pieces arranged for their group [61].

First attempts have also been made to explore the dynamic changes in brain connectivity over time, that is, the temporal aspects of synchronization. **Figure 7** shows the coherence maps (intra- and inter-brain) of the four guitarists (taken from literature [61]). These maps depict the connectivity in the delta range at three time points within a chosen musical sequence. (Note that in the study by Müller et al. [61], connectivity changes have also been analyzed in the range from theta to gamma.) The maps reveal that each guitarist develops his/her own patterns of within- and between-brain connectivity, which alter as the piece proceeds. Obviously, brain-to-brain coupling is both an idiosyncratic matter and one that adapts to the demands of the musical piece, which may include musical entries and/or leading parts.

Let us turn to another point – the performance strategies while making music. The first study in this respect was carried out by Sängers et al. [47]. They were interested in verifying a possible leader–follower relationship during performance

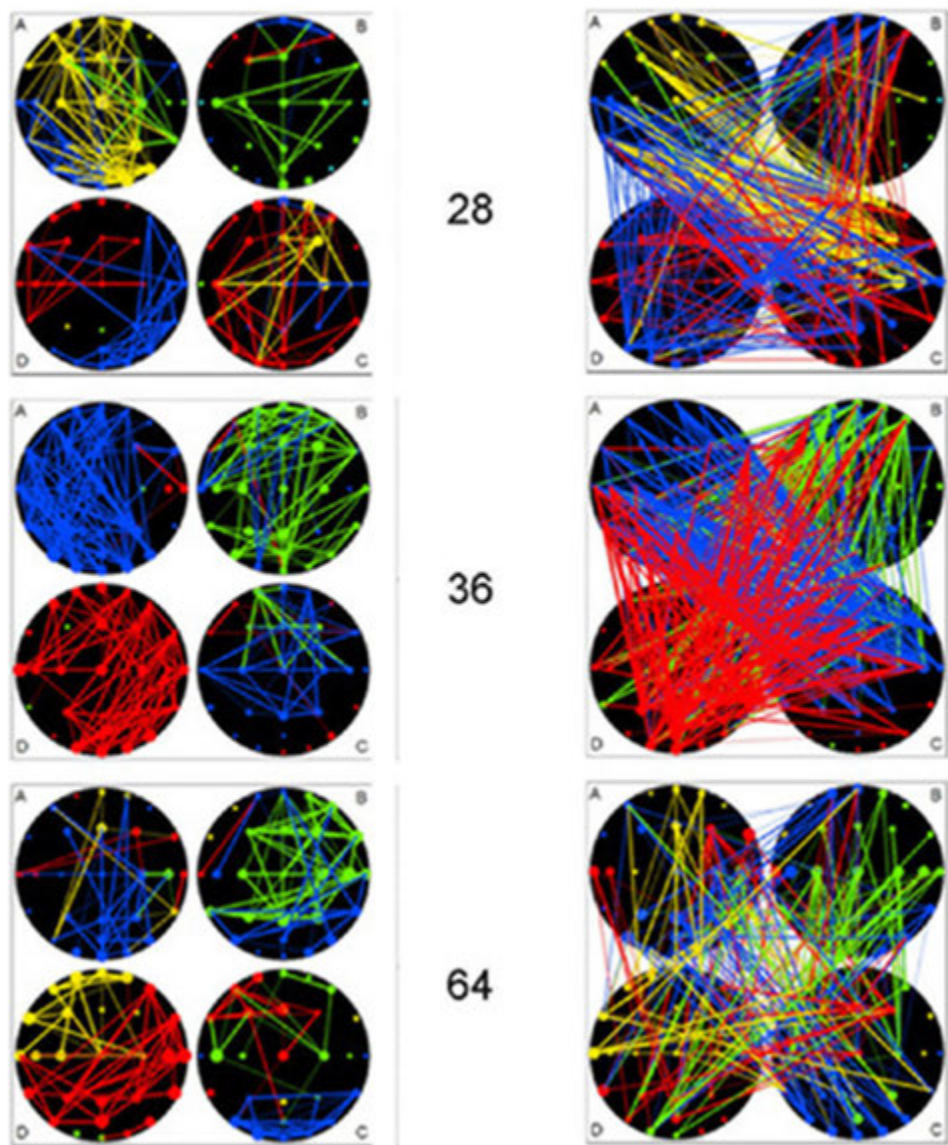


Figure 7.

Brain maps of four guitarists (A to D) playing in a quartet. Changes in connectivity strength across time. Intra-brain couplings (left) and inter-brain couplings (right); three time points (28/36/64) chosen from a musical sequence (note that time points are not equivalent to seconds). All maps show activity in the delta range (frequency of interest: 2.5 Hz). At time point 36, the density of intra-brain connections (left) and of inter-brain connections (right, players A, B, D) is high, indicating that functional couplings are high. Connectivity results differ considerably according to the piece's demands, the time point of analysis, frequency band, and the brains' specificities [61].

that could be proven through time-delayed synchronization, that is, through asynchronous coupling between the brains (note that for proving directionality, a different method than Granger was employed here; see above). In the study by Sanger et al. [47], the roles of leading and following were tested with guitar duos and

tempo re-adjustment, in which the leading guitarist had to take up the original tempo after a composed *ritardando*, that is, after some seconds of slowing down. While accelerating again, a time delay could indeed be observed in the fronto-central alpha and beta-bands of the follower in comparison to the leading guitarist, intra- as well as inter-brain. So, phase asynchrony does indeed exist between the brains of a leading musician and his/her follower where, again, the prefrontal regions have a decisive role in action planning – in this case, in re-adjusting the musical tempo.

Yet, in spontaneous music-making, as in jazz, it may also happen that group members compete against each other to stimulate their play, which is another type of performance strategy. Recently, the research question arose whether creative group activities, be it improvisation or brain storming, may gain more profit when subjects are in a “competitive” mood compared to a “cooperative” mood. To explore this with hyperscanning, a simple button-press experiment has been described in the study by Babiloni et al. [51], in which pairs of participants either had to press the buttons as fast as possible – which is an example of competition – or synchronize their button presses – as an example of cooperation. Using the fNIRS technique this time, strong inter-brain couplings could be observed in the right superior frontal cortex when “cooperating,” that is, when synchronizing the button presses, but not when competing against each other [51].

Similar results with fNIRS have been achieved by Lu et al. [45]. However, in their study, the task was more demanding: pairs of subjects had to perform the Guilford Alternative Uses Test, in which as many unusual uses for everyday objects as possible had to be invented. Again, inter-brain coupling was stronger during cooperation than during competition, activating both the right dorsolateral prefrontal cortex (rDLPFC) and the right temporoparietal junction (rTPJ). Obviously, these two regions are necessary for creative problem-solving. Pairs of less creative individuals profit from cooperation more strongly than pairs of highly creative persons. At least, inter-brain synchrony in the rDLPFC and rTPJ was stronger for the less creative persons than for the highly creative ones [63]. These studies by Lu et al. [45] and Xue et al. [63] are the first attempts to show that hyperscanning research can be extended away from studying joint actions like music-making in small groups toward exploring the mental efforts of a group when developing creative ideas.

New paths could also be pursued from the methodological point of view, for instance, through combining EEG hyperscanning with “event-related potentials” (ERPs). ERP is a special method of analyzing the EEG on the condition that, during recording, the same sorts of stimuli have been repeated many times [52]. ERP has the advantage that each half-wave (or component) of the final “grand average curve” indicates a specific step of information processing, which allows gaining insight into the cognitive processes beyond the “What is” of joint action. However, although the current technological state of research would allow working with this combined method, only single-brain processes have so far been analyzed using ERPs.

In pure ERP research (note: this is different from EEG hyperscanning), an interesting study about turn-taking has been carried out by Bögels, Levinson, and colleagues [64–66]. They were interested in the rapid switching between two persons having an everyday conversation – whereby one of them posed questions to which the other should give the right answer. Note that in this study, only one participant wore an EEG cap. Because of a clever design in which certain verbal cues could either appear at early or late positions in the questions, Bögels et al. [65] demonstrated that response preparation starts midway while comprehension is still ongoing. More precisely, it starts

as soon as the verbal cues used by the questioner have been grasped by the respondent. For this “midway-response-planning”, certain ERP components were visible in the respondent’s brain curves, with amplitude sizes depending on whether the verbal cues in the questions had appeared early or late. This paradigm could easily be transformed into one for using the combined EEG hyperscanning/ERP method to study turn-taking behavior in both dialog partners: the questioner and the respondent.

7. Conclusions

This overview chapter had the objective of describing musical ensemble playing from a neuroscientific point of view by considering both, processes at the micro and macro levels. The main results can be summarized as follows: First: The properties and functioning of mirror neurons are the secret to why a string quartet operates as a skillful work group. Mirror neurons discharge in relation to actions – either seen, heard, or self-performed [25, 26] – which is an elementary mechanism by which musicians can immediately understand the intentions and playing movements of others [28, 29]. The driving force for joint action can thus be found at the microanatomical level. Second: Hyperscanning – mostly with EEG – is a promising and expandable method by which brain activity can be recorded simultaneously from musicians playing together in real-time. Third: Intra- and inter-brain EEG patterns of musicians in interactive situations differ considerably in regard to frequency range [60, 61]. Intra-brain coupling takes place at higher frequencies, whereas inter-brain coupling occurs concurrently in lower frequency bands (theta and delta). Fourth: Both types of connectivity patterns also change considerably over time, partly due to the demands of a musical piece and partly due to idiosyncrasy [61]. Fifth: Causal leader–follower relationships among musicians, for example, between a soloist and accompanying musicians, can be confirmed on the neural level, provided that their EEG curves – considered as time series – reveal a certain time lag [47]. Sixth: Inter-brain coupling is stronger during cooperation than during competition, supporting the assumption that creative group activities, for example, musical improvisation, may gain more profit when subjects are in a cooperative mood compared to a competitive mood.

From the methodological point of view, many things could be improved toward systematization. For instance, the “classical” lab-based ERP-approach should be reconsidered to obtain hyperscanning results that are more consistent than at present, and variables such as “musical expertise” or “gender” could also be modified in a systematic manner.

An important new field is currently emerging in which hyperscanning (both EEG and fNIRS) is conducted with subjects who have deficits in social interaction, as is the case for individuals with autism spectrum disorder (ASD [67]). This branch, however, is still in its infancy.

All in all, hyperscanning is a suitable method to explore the “what is” of social life in all its facets. The main focus of interest throughout this chapter was on musicians interacting in small groups on stage. “Brain synchronization during joint action” and the “functioning of mirror neurons” are mechanisms that explain why musicians, performing in “small size,” do this so skillfully. However, the main thing with which musicians are preoccupied is still to convey the deeper meaning of a musical piece in an esthetically pleasing manner.

Author details

Christiane Neuhaus

Institute of Systematic Musicology, University of Hamburg, Hamburg, Germany

*Address all correspondence to: christiane.neuhaus@uni-hamburg.de

IntechOpen

© 2025 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Goethe JWv. Apercu. In: Grand Duchess of Saxony, S ed. Goethes Werke - Goethes Briefe [Goethe's Oeuvre – Goethe's Letters]. Weimar: Hermann Böhlau's successors. **46**; 1891
- [2] Davidson JW. Social and musical co-ordination between members of a string quartet: An exploratory study. *Psychology of Music*. 2002;**30**:186–201. DOI: 10.1177/0305735602302005
- [3] Glowinski D, Manchini M, Cowie R, Camurri A, Chiorri C, Doherty C. The movements made by performers in a skilled quartet: A distinctive pattern, and the function that it serves. *Frontiers in Psychology*. 2013;**4**. DOI: 10.3389/fpsyg.2013.00841
- [4] Glowinski D, Bracco F, Chiorri C, Grandjean D. Music ensemble as a resilient system. Managing the unexpected through group interaction. *Frontiers in Psychology*. 2016;**7**. DOI: 10.3389/fpsyg.2016.01548
- [5] Volpe G, D'Ausilio A, Badino L, Camurri A, Fadiga L. Measuring social interaction in music ensembles. *Philosophical Transactions of the Royal Society B*. 2016;**371**:20150377. DOI: 10.1098/rstb.2015.0377
- [6] Wöllner C. Call and response: Musical and bodily interactions in jazz improvisation duos. *Musicae Scientiae*. 2020;**24**:44–59. DOI: 10.1177/1029864918772004
- [7] Badino L, D'Ausilio A, Glowinski D, Camurri A, Fadiga L. Sensorimotor communication in professional quartets. *Neuropsychologia*. 2014; **1208**–104. DOI: 10.1016/j.neuropsychologia.2013.11.012
- [8] Behne K-E. Kommunikation oder Geste? [Communication or gesture?]. In Behne K-E, editor. *Gefühl als Erlebnis – Ausdruck als Sinn*. [Emotion as experience– expression as understanding.]. Laaber: Laaber Verlag; 1982. p. 125–143. DOI: 10.25656/01:24768
- [9] Vesper C, Abramova E, Büttepage J, Ciardo F et al. Joint action: Mental representations, shared information and general mechanisms for coordinating with others. *Frontiers in Psychology*. 2017;**7**. DOI: 10.3389/fpsyg.2016.02039
- [10] De Jaegher H, Di Paolo E, Gallagher S. Can social interaction constitute social cognition? *Trends in Cognitive Sciences*. 2010;**14**:441–447. DOI: 10.1016/j.tics.2010.06.009
- [11] Schiavio A, Høffding S. Playing together without communicating? A pre-reflective and enactive account of joint musical performance. *Musicae Scientiae*. 2015;**19**:366–388. DOI: 10.1177/1029864915593333
- [12] Sebanz N, Bekkering H, Knoblich GK. Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*. 2006;**10**:70–76. DOI: 10.1016/j.tics.2005.12.009
- [13] Moore C, Barresi J. The role of second-person information in the development of social understanding. *Frontiers in Psychology*. 2017;**8**. DOI: 10.3389/fpsyg.2017.01667
- [14] Sloboda JA. Music Performance. In Deutsch D, editor. *The Psychology of Music*. London: Academic Press; 1982. p. 479–496

- [15] Donnay GF, Rankin SK, Lopez-Gonzalez M, Jiradejvong P, Limb CJ. Neural substrates of interactive musical improvisation: An fMRI study of 'Trading Fours' in jazz. *PLoS ONE*. 2014;9. DOI: 10.1371/journal.pone.0088665
- [16] Wade BC. Performance practice in Indian classical music. In Béhague G, editor. *Performance Practice – Ethnomusicological Perspectives*. Westport, Connecticut: Greenwood Press; 1984. p. 13–52
- [17] D'Ausilio A, Novembre G, Fadiga L, Keller PE. What can music tell us about social interaction? *Trends in Cognitive Sciences*. 2015;19:111–114. DOI: 10.1016/j.tics.2015.01.005
- [18] Van Overwalle F. Social cognition and the brain: A meta-analysis. *Human Brain Mapping*. 2009;30:829–858. DOI: 10.1002/hbm.20547
- [19] Redcay E, Schilbach L. Using second-person neuroscience to elucidate the mechanisms of social interaction. *Nature Reviews Neuroscience*. 2019;20:495–505. DOI: 10.1038/s41583-019-0179-4
- [20] Schilbach L, Timmermans B, Reddy V, Costall A, Bente G, Schlicht T, Vogeley K. Toward a second-person neuroscience. *Behavioral and Brain Sciences*. 2013;36:393–462. DOI: 10.1017/S0140525X12000660
- [21] Burkhardt A. Der Dialog-Begriff bei Wilhelm von Humboldt und seine Folgen [Wilhelm von Humboldt's approach to dialogue and the consequences]. *Zeitschrift für Literaturwissenschaft und Linguistik*. [Journal of Literary Science and Linguistics]. 2018;48:717–740. DOI: 10.1007/s41244-018-0115-5
- [22] Jung T. Das Du-Bewußtsein – Zur Sozialontologie der Zweiten Person. [The Consciousness of the You – Socioontology of the 2nd Person]. 1st. Baden-Baden: Karl Alber; 2023. p. 128. DOI: 10.5771/9783495994948
- [23] Buber M. Das Dialogische Prinzip [The dialogical principle] 6th, Gerlingen: Lambert Schneider Verlag, 1992. p. 324
- [24] Zahavi D. Second-person engagement, self-alienation, and group-identification. *Topoi*. 2019;38:251–260. DOI: 10.1007/s11245-016-9444-6
- [25] Rizzolatti G, Craighero L. The mirror neuron system. *Annual Review of Neuroscience*. 2004;27:169–192. DOI: 10.1146/annurev.neuro.27.070203.144230
- [26] Kohler E, Keysers C, Umiltà MA, Fogassi L, Gallese V, Rizzolatti G. Hearing sounds, understanding actions: Action representation in mirror neurons. *Science*. 2002;297:846–848. DOI: 10.1126/science.1070311
- [27] Marshall J. Mirror neurons. *Proceedings of the National Academy of Sciences*. 2014;111:6531. DOI: 10.1073/pnas.1404652111
- [28] Calvo-Merino B, Glaser DE, Grèzes J, Passingham RE, Haggard P. Action observation and acquired motor skills: An fMRI study with expert dancers. *Cerebral Cortex*. 2005;15:1243–1249. DOI: 10.1093/cercor/bhi007
- [29] Iacoboni M, Molnar-Szakacs I, Gallese V, Buccino G, Mazziotta JC, Rizzolatti G. Grasping the intention of others with one's own mirror neuron system. *PLoS Biology*. 2005;3. DOI: 10.1371/journal.pbio.0030079

- [30] Kaplan JT, Iacoboni M. Getting a grip on other minds: Mirror neurons, intention understanding, and cognitive empathy. *Social Neuroscience*. 2006;**1**:175–183. DOI: 10.1080/17470910600985605
- [31] Van Overwalle F, Baetens K. Understanding others' actions and goals by mirror and mentalizing systems: A meta-analysis. *NeuroImage*. 2009;**48**:564–584. DOI: 10.1016/j.neuroimage.2009.06.009
- [32] Meltzoff AN, Decety J. What imitation tells us about social cognition: A rapprochement between developmental psychology and cognitive neuroscience. *Philosophical Transactions of the Royal Society B*. 2003;**358**:491–500. DOI: 10.1098/rstb.2002.1261
- [33] Simpson EA, Murray L, Paukner A, Ferrari PF. The mirror neuron system as revealed through neonatal imitation: Presence from birth, predictive power and evidence of plasticity. *Philosophical Transactions of the Royal Society B*. 2014;**369**:20130289. DOI: 10.1098/rstb.2013.0289
- [34] Buccino G, Vogt S, Ritzl A, Fink GR, Zilles K, Freund H-J, Rizzolatti G. Neural circuits underlying imitation learning of hand actions: An event-related fMRI study. *Neuron*. 2004;**42**:323–334
- [35] Newman-Norlund R, Bekkering H. The mirror neuron system is more active during complementary compared with imitative action. *Nature Neuroscience*. 2007;**10**:817–818. DOI: 10.1038/nn1911
- [36] Zamm A, Loehr JD, Vesper C, Konvalinka I, et al. A practical guide to EEG hyperscanning in joint action research: From motivation to implementation. *Social Cognitive and Affective Neuroscience*. 2024;**19**. DOI: 10.1093/scan/nsae026
- [37] Hakim U, De Felice S, Pinti P, Zhang X, et al. Quantification of inter-brain coupling: A review of current methods used in haemodynamic and electrophysiological hyperscanning studies. *NeuroImage*. 2023;**280**. DOI: 10.1016/j.neuroimage.2023.120354
- [38] Czeszumski A, Eustergerling S, Lang A, Menrath D, et al. Hyperscanning: A valid method to study neural inter-brain underpinnings of social interaction. *Frontiers in Human Neuroscience*. 2020;**14**. DOI: 10.3389/fnhum.2020.00039
- [39] Overgaard S, Michael JA. The interactive turn in social cognition research: A critique. *Philosophical Psychology*. 2015;**28**:160–183. DOI: 10.1080/09515089.2013.827109
- [40] Reybrouck M. A dynamic interactive approach to music listening: The role of entrainment, attunement and resonance. *Multimodal Technologies and Interaction*. 2023;**7**:66. DOI: 10.3390/mti7070066
- [41] Vesper C, Abramova E, Bütepage J, Ciardo F, et al. Joint action: Mental representations, shared information and general mechanisms for coordinating with others. *Frontiers in Psychology*. 2017;**7**. DOI: 10.3389/fpsyg.2016.02039
- [42] Cheng S, Wang J, Luo R, Hao N. Brain to brain musical interaction: A systematic review of neural synchrony in musical activities. *Neuroscience and Biobehavioral*

- Reviews. 2024;164. DOI: 10.1016/j.neubiorev.2024.105812
- [43] Babiloni C, Vecchio F, Infarinato F, Buffo P, et al. Simultaneous recording of electroencephalographic data in musicians playing in ensemble. *Cortex*. 2011;47:1082–1090.
- [44] Lindenberger U, Shu-Chen L, Gruber W, Müller V. Brains swinging in concert: Cortical phase synchronization while playing guitar. *BMC Neuroscience*. 2009;10. DOI: 10.1186/1471-2202-10-22
- [45] Lu K, Xue H, Nozawa T, Hao N. Cooperation makes a group be more creative. *Cerebral Cortex*. 2019;29:3457–3470. DOI: 10.1093/cercor/bhy215
- [46] Shemyakina NV, Nagornova ZV. Neurophysiological characteristics of competition in skills and cooperation in creativity task performance: A review of hyperscanning research. *Human Physiology*. 2021;47:87–103. DOI: 10.1134/S0362119721010126
- [47] Sängler J, Müller V, Lindenberger U. Directionality in hyperbrain networks discriminates between leaders and followers in guitar duets. *Frontiers in Human Neuroscience*. 2013;7. DOI: 10.3389/fnhum.2013.00234
- [48] Acquadro MAS, Congedo M, De Ridder D. Music performance as an experimental approach to hyperscanning studies. *Frontiers in Human Neuroscience*. 2016;10. DOI: 10.3389/fnhum.2016.00242
- [49] Dikker S, Wan L, Davidesco I, Kaggen L, et al. Brain-to-brain synchrony tracks real-world dynamic group interactions in the classroom. *Current Biology*. 2017;27:1375–1380
- [50] Gugnowska K, Novembre G, Kohler N, Villringer A, Keller PE, Sammler D. Endogenous sources of interbrain synchrony in duetting pianists. *Cerebral Cortex*. 2022;32:4110–4127. DOI: 10.1093/cercor/bhab469
- [51] Babiloni F, Astolfi L. Social neuroscience and hyperscanning techniques: Past, present and future. *Neuroscience and Biobehavioral Reviews*. 2014;44:76–93. DOI: 10.1016/j.neubiorev.2012.07.006
- [52] Neuhaus C. Methods in neuromusicology: Principles, trends, examples, and the pros and cons. In Schneider A, editor. *Studies in Musical Acoustics and Psychoacoustics*. Cham: Springer International Publishing; 2017:341–374. DOI: 10.1007/978-3-319-47292-8.ch11
- [53] Jasper HH. The Ten-Twenty electrode system of the international federation. *Electroencephalography and Clinical Neurophysiology*. 1958;10:370–375
- [54] Barraza P, Dumas G, Huanhuan L, Blanco-Gomez G, van den Heuvel MI, Baart M, Pérez A. Implementing EEG hyperscanning setups. *MethodsX*. 2019;6:428–436. DOI: 10.1016/j.mex.2019.02.021
- [55] Petsche H. Approaches to verbal, visual and musical creativity by EEG coherence analysis. *International Journal of Psychophysiology*. 1996;24:145–159. DOI: 10.1016/S0167-8760(96)00050-5
- [56] D’Ausilio A, Badino L, Li Y, Tokay S, Craighero L, Canto R,

- Aloimonos Y, Fadiga L. Leadership in orchestra emerges from the causal relationships of movement kinematics. *PLoS One*. 2012;7. DOI: 10.1371/journal.pone.0035757
- [57] Varela F, Lachaux J-P, Rodriguez E, Martinerie J. The brainweb: Phase synchronization and large-scale integration. *Nature Reviews Neuroscience*. 2001;2:229–239. DOI: 10.1038/35067550
- [58] Novembre G, Iannetti GD. Hyperscanning alone cannot prove causality. *Multibrain stimulation can. Trends in Cognitive Sciences*. 2021;25:96–99. DOI: 10.1016/j.tics.2020.11.003
- [59] Moreau Q, Dumas G. Beyond correlation versus causation: Multi-brain neuroscience needs explanation. *Trends in Cognitive Sciences*. 2021;25:542–543. DOI: 10.1016/j.tics.2021.02.011
- [60] Müller V, Sängler J, Lindenberger U. Intra- and inter-brain synchronization during musical improvisation on the guitar. *PLoS ONE*. 2013;8. DOI: 10.1371/journal.pone.0073852
- [61] Müller V, Sängler J, Lindenberger U. Hyperbrain network properties of guitarists playing in quartet. *Annals of the New York Academy of Sciences*. 2018;1423:198–210. DOI: 10.1111/nyas.13656
- [62] Holroyd CB. Interbrain synchrony: On wavy ground. *Trends in Neurosciences*. 2022;45:346–357. DOI: 10.1016/j.tins.2022.02.002
- [63] Xue H, Lu K, Hao N. Cooperation makes two less-creative individuals turn into a highly-creative pair. *NeuroImage*. 2018;172:527–537. DOI: 10.1016/j.neuroimage.2018.02.007
- [64] Bögels S, Casillas M, Levinson SC. Planning versus comprehension in turn-taking: Fast responders show reduced anticipatory processing of the question. *Neuropsychologia*. 2018;109:295–310. DOI: 10.1016/j.neuropsychologia.2017.12.028
- [65] Bögels S, Magyari L, Levinson SC. Neural signatures of response planning occur midway through an incoming question in conversation. *Scientific Reports*. 2015;5. DOI: 10.1038/srep12881
- [66] Levinson SC. Turn-taking in human communication – Origins and implications for language processing. *Trends in Cognitive Sciences*. 2016;20:6–14. DOI: 10.1016/j.tics.2015.10.010
- [67] Zhou X, Wong PCM. Hyperscanning to explore social interaction among autistic minds. *Neuroscience and Biobehavioral Reviews*. 2024;163. DOI: 10.1016/j.neubiorev.2024.105773

Section 5

Music Therapy

Chapter 6

Music Therapy in Classical Ayurveda: A Reconstruction from Ancient Sanskrit Sources

Abirlal Gangopadhyay

Abstract

This chapter describes the historical development and theoretical basis of Indian music therapy from the sixth century BCE to the sixteenth century CE, along with its modern resurgence. During all times, music was subtly integrated into Ayurvedic practices, particularly in internal medicine (*kāyacikitsā*), pediatrics (*kaumārbhṛtya*), and aphrodisiac therapy (*vājīkaraṇa*), and additionally in alchemy and iatrochemistry (*rasaśāstra*). The chapter examines foundational Ayurvedic texts and aesthetic treatises such as the *Nāṭyaśāstra* to reconstruct the *svara-rasa-doṣa* (SRD) framework, demonstrating how specific musical notes, scales, and their associated emotional states were mapped to balance the three somatic humors Vāta (air), Pitta (bile), and Kapha (phlegm). It also describes various therapeutic applications, including managing coma, fever, and supporting perinatal care, highlighting instances where ancient physicians prescribed music and even included musicians in hospital settings. The chapter concludes by outlining the resurgence of Indian music therapy from J. Pāl's 1930s model to current clinical research, while debunking modern myths.

Keywords: music therapy, Ayurveda, kaya chikitsa (internal medicine), coma, aphrodisiac therapy

1. Introduction

The *Atharvaveda*, one of the four holy texts about ancient Indian knowledge systems, is traditionally regarded as the Veda most closely linked to health, healing, and the rituals surrounding well-being. Unlike the other three Vedas – *gveda*, *Sāmaveda*, and *Yajurveda* – which focus primarily on hymns, music, and sacrificial procedures, the *Atharvaveda* places particular emphasis on healing practices, protective rituals, ethical disciplines, and the use of mantras for both physical and spiritual health [1]. It is from this Vedic foundation that *Āyurveda*, India's classical system of medicine, emerges. This is not just a system of medicine but also a philosophical and psychological discipline. The term *Āyurveda* (literally meaning “knowledge of life”; *āyur*(ṣ) = life + *veda* = knowledge) is defined as understanding life (*jīvita*), its continuity (*anuvṛtti*), and sustaining force (*dhāri*) [1].

The historical development of Ayurvedic medicine may be broadly divided into two major phases: the *Samhitā* period (before the sixth century BCE to around the seventh century CE) and the post-*Samhitā* period (approximately the eighth century to the sixteenth century CE). The *Samhitā* period saw the emergence of foundational treatises – the *Caraka Samhitā* (CS), *Suśruta Samhitā* (SS), and *Aṣṭāṅga Hṛdaya* (AH) – collectively referred to as the *Great Triad* (*Brhatṭrayī*). These works laid down the core principles of Ayurvedic thought and clinical practice and continued to be regarded as the primary sources of authoritative knowledge throughout the tradition.

Importantly, music as a therapeutic modality is mentioned in all three texts of the Great Triad, although not as a separate branch. Instead, it is subtly woven into domains such as internal medicine, pediatrics, and aphrodisiac therapy, and its aesthetic and psychological effects are acknowledged only occasionally. Among these, the *Aṣṭāṅga Hṛdaya* (AH) offers some of the most direct references to music therapy and remains central to later commentarial interpretations, including a rare mention of *active* music therapy in the medieval exegesis by Aṇadatta (c. 1200s CE).

The *Suśruta Samhitā* (SS), primarily a surgical text (sixth century BCE) attributed to Suśruta (eventually regarded as the father of surgery) and revised by the physician Nāgārjuna (first–fifth century CE), integrates brief nods to music in clinical care. The *Caraka Samhitā* (CS), in its oldest layers (dated no later than the second century BCE), reflects Buddhist influences and early philosophical frameworks. The physician Dṛḍhabala (c. 200s–400s CE) later redacted and completed the text, shaping the version known today. While music is not a major theme in the CS, its occasional references align with its wide-ranging concern for holistic well-being.

From the eighth century CE onward, Ayurvedic literature became more diversified and specialized, leading to the formation of the *Lesser Triad* (*Laghutrayī*), which includes the *Mādhavanidāna*, *Śārṅgadhara Samhitā*, and *Bhāvaprakāśa* (BhP). These works reflect an increasing focus on disease-specific classification, pharmaceutical preparation, and practical therapy.

Among the three, only the *Bhāvaprakāśa* (BhP) – composed in the sixteenth century CE – carries forward the theme of music therapy, making it an especially significant text. The BhP offers a systematic cataloguing of diseases, treatment protocols, and an extensive materia medica. It also holds historical importance as the first Ayurvedic text to record syphilis (*phiraṅgaroga*). Its references to music reaffirm earlier notions of receptive music therapy, wherein listening to music forms part of the healing system.

By contrast, the *Mādhavanidāna* (c. eighth century CE), while crucial for its classification of fevers and disease pathology, and the *Śārṅgadhara Samhitā* (thirteenth–fourteenth century CE), with their detailed descriptions of pulse diagnosis and pharmaceutical methods, do not engage with music therapy. Their silence on the subject marks a narrowing of the tradition's engagement with aesthetic modalities like music.

Beyond the Great and Lesser Triads, other classical sources from the *Samhitā* period also attest to the therapeutic use of music. The *Kaśyapa Samhitā*, for example, focuses on pediatrics and is especially relevant in understanding the sensory and developmental applications of music in early childhood care.

Among post-*Samhitā* texts, only Sodhala's *Gadanigraha* includes references to music therapy, marking it as the sole non-triadic medieval text continuing this

thread. The *Gadanigraha* carries forward certain medical and iatrochemical traditions, making selective but meaningful mentions of music as a healing tool.

In all these sources, music therapy appears almost exclusively in its receptive form, where patients passively receive its effects through listening. The specific instances and contexts in which these references occur are discussed in greater detail later in this chapter (see Section 4). In addition, the *svara-rasa-doṣa* (or SRD) framework will also be examined, as it is the backbone of these music therapeutic treatments.

One of the key objectives of this chapter is to refer to the sources written in Sanskrit with meticulous care, which, for the most part, are unknown to the Western reader.

2. Psychological healing in Ayurveda

2.1 Basis for musical intervention

In the foundational *philosophy* of Ayurveda, Ether (ākāśa), Air (vāyu), Fire (agni), Water (jala), and Earth (prthivī) are considered the five primary constituents of all objects (CS, *Sūtrasthāna*:26.10)¹ [2]. Each of these elements inherently possesses one or more of the following three fundamental qualities – *Sattva* (light, clarity, harmony), *Rajas* (activity, passion, restlessness), and *Tamas* (inertia, darkness, ignorance) – which are subtle, qualitative principles that govern the nature of conscious beings (subjects) and matter (objects) alike. These qualities are not superimposed from outside but intrinsic to the nature of the primary constituents [3]. Ether is primarily Sāttvic, associated with clarity. Air, being mobile and subtle, carries both Rajas (activity, passion, restlessness) and Sattva (light, clarity, harmony). Fire, responsible for transformation and brilliance, embodies Rajas (activity, passion, restlessness) and Sattva (light, clarity, harmony). Water holds Sattva (light, clarity, harmony) and Tamas (inertia, darkness, ignorance), symbolizing cohesion and fluid responsiveness. Earth, the densest of the elements, is predominantly Tāmasic, characterized by inertia and structural stability (SS, *Śārīrasthāna*:1.20) (**Figure 1**).

As these quality-infused elements combine to form physiological patterns, they give rise to three *humors* (*doṣas*) – Vāta (air), Pitta (bile), and Kapha (phlegm).² These humors are not arbitrary categories but emerge from the five elements and their inherent qualities (*gunas*) (**Figure 1**). Vāta, rooted in Ether and Air, carries the innate qualities of lightness, movement, and subtlety, reflecting embedded Sāttvic and Rājasic traits. Pitta (bile), emerging from Fire, embodies intensity, sharpness, and transformation, shaped by a blend of Rajas (activity, passion, restlessness) and Sattva (light, clarity, harmony). Kapha (phlegm), arising from Water and Earth, holds the embedded qualities of cohesion, stability, and heaviness, reflecting Tamas

¹ References such as (CS, *Sūtrasthāna*:26.10) follow a standardized format used for classical Sanskrit medical texts. The numbers indicate the section (sthāna/khaṇḍa), chapter, and verse or prose number, respectively.

²Here, 26 is the chapter number, and 10 is the line or verse number.

¹²⁹Note that drawing an analogy to Hippocrate's four fluids – blood, phlegm, black bile and yellow bile – is difficult. Ayurveda did not consider blood as a humor. Hippocrate, on the other hand, mentioned two biles, black and yellow, but Ayurvedic consideration is just bile, without any black or yellow division.

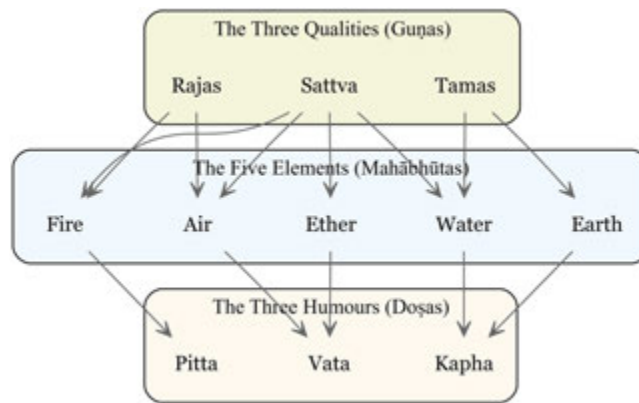


Figure 1.

Relationship between elements, psychological qualities, and humors (see SS, Śārīrasthāna:1.20 and AS, Sūtrasthāna:20.2). Note that due to their subtleties, the qualities (guṇas) are higher in ranking than the elements themselves.

(inertia, darkness, ignorance) and Sattva (light, clarity, harmony) (AS, Sūtrasthāna: 20.2).

Ayurveda is also concerned with *mental* disorders. They are not just an absence of well-being but disruptions of it. While often linked to imbalances of the humors, the primary focus is on the fundamental qualities of the mind. Disturbances arise from an imbalance of Rajas (activity, passion, restlessness) and Tamas (inertia, darkness, ignorance), which suppress Sattva (light, clarity, harmony), that is, the mind's inherent quality of balance (CS, Sūtrasthāna: 4.34). This understanding is central because Ayurvedic psychology views Sattva (light, clarity, harmony) as the mind's essence and stabilizing principle (prakṛti). This concept forms the crucial basis for the therapeutic process of regulating and uplifting the mind (sattvāvajaya), defining the Ayurvedic approach to mental well-being (CS, Sūtrasthāna: 11.54).

2.2 The structure of mind

According to the *Caraka Samhitā* (CS), the mind is a product of the ego (see **Figure 2**). It is also part of the 16 modifications of Nature (prakṛti). Accordingly, the mind is not part of the eightfold primordial constituents but rather arises subsequently along with the sense organs and the organs of action. In this framework, *Avyakta* has a double sense: as the ultimate self (avyakta/akṣara/ātman), it is unmanifest not because it is unconscious but because it is beyond speech and expression; and as unmanifest nature, *Avyakta* is the potential state from which manifest nature arises. This unmanifest self stands at the origin of the process (**Figure 2**) and is fundamentally distinct from the ego (ahaṁkāra), which is a later evolve within the manifest sequence.

The operations of cognition – such as perception, deliberation, emotional response, and the pursuit of self-awareness (not depicted in the present **Figure 2**) – are all mediated by the mind. While the senses provide the raw data of experience, the stability and engagement of the mind transform sensory input into meaningful knowledge. Thus, the mind is a subtle organ and the central coordinator of cognitive and affective processes. Hence, any successful intervention on the mind must

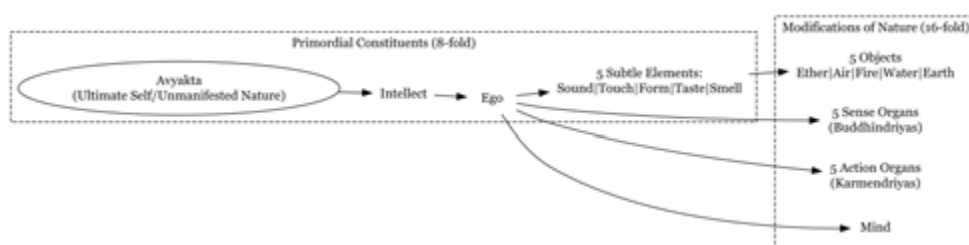


Figure 2.
 Diagram illustrating the evolution of cognitive and sensory faculties from primordial constituents and nature (*prakṛti*) (see CS, *Śārīrasthāna*: 1.61–66).

ultimately restore Sattva (light, clarity, harmony), directly or indirectly. This fundamental orientation opens the space for art forms – music in our present context – as tools of psychological restoration.

2.3 Sattva as therapeutic target

Sattva is thus regarded as the ideal and desirable quality in psychological states. It is the root of clarity, emotional harmony, memory, intuitive cognition (*buddhi*), and inner luminosity.

In Ayurvedic thought, Sattva can appear in pure (*śuddha*) form, but also in Rājasic or Tāmasic forms, with Rajas (activity, passion, restlessness) and Tamas (inertia, darkness, ignorance) existing on the same qualitative level as Sattva. Pure Sattva (light, clarity, harmony) is faultless and inherently oriented toward well-being. By contrast, Rājasic Sattva is tinged with agitation or anger, while Tāmasic Sattva is colored by delusion and inertia.

Now, it is essential to distinguish two layers of meaning: Sattva as a quality – a subtle constituent of mental function (see above) – and Sattva as the very name of the mind itself, as used by physicians in the classical texts [2, 4, 5]. As the *Caraka Samhitā* (CS) affirms, the mind or Sattva, devoid of Rajas (activity, passion, restlessness) and Tamas (inertia, darkness, ignorance), is known as Sattva (light, clarity, harmony). Rajas and Tamas are therefore considered two mental *doṣas* that disturb Sattva, the natural state of the mind. In detail, their manifestations – or *vikāras* – are: *kāma* (desire), *krodha* (anger), *lobha* (greed), *moha* (delusion), *īrṣyā* (envy), *mada* (pride), *śoka* (sorrow), *cintā* (anxiety), *udvega* (agitation), *bhaya* (fear), and *harṣa* (excessive joy), among others (CS, *Vimānasthāna*:4.6). These are not merely emotional states but also diagnostic expressions of a disturbed mind.

Furthermore, the exact text of *Caraka Samhitā* (*Śārīrasthāna*: 4.34) distinguishes between bodily and mental pathology: “There are three *doṣas* of the body – Vāta (air), Pitta (bile), and Kapha (phlegm) – which disturb the physical constitution; and two *doṣas* of the mind – Rajas (activity, passion, restlessness) and Tamas (inertia, darkness, ignorance) – which disturb Sattva (light, clarity, harmony). Disease arises only when both the bodily and mental domains are disturbed; without such disturbance, no pathology is produced.” This quotation again clearly affirms that mental health in Ayurveda is primarily a matter of maintaining the integrity of Sattva by regulating its intrinsic disturbances – Rajas and Tamas. Thus, cultivating Sattva (light, clarity, harmony) balances mental energy and acts as an internal medicine – a balm that restores the mind to its essential nature. This has direct

implications for the therapeutic use of music: sound must not excite Rajas (activity, passion, restlessness) or deepen Tamas (inertia, darkness, ignorance), but must elevate the psyche toward balance and harmony. The aim is not sensory pleasure but emotional purification through vibrational realignment.

Being so, music becomes medicine by modulating consciousness. This concept of purified Sattva (light, clarity, harmony) bridges Ayurveda with the emotional psychology of *Nāṭyaśāstra* (NS) (see below) [6].

3. Svara, rasa, and doṣa: Music, relish, and humor

3.1 Svara–rasa correlation

Before exploring the bridge between medical and aesthetic disciplines, especially *music*, it is essential to understand the music system contemporaneous with the early Ayurvedic physicians. In ancient times, the relish (rasa) of a musical performance was considered to be based on the predominance of a specific note (*svara*). With this in mind, we aim to identify the type of music ancient Indian physicians would have experienced.

First of all, a famous treatise has to be mentioned: The *Nāṭyaśāstra* (NS), regarded as the earliest foundational text on Indian performing arts and aesthetic theory. It describes a musical system known as the *Gāndharva tradition*. Its primary instrument was the arched harp, referred to in early Indian sources as the Vīṇā in its most ancient form (NS, 34.14ab) [7]. Secondary instruments included lutes (often made from gourds or tortoise shells), flutes, and drums (NS, 34.14 cd-15).

Such instruments and musical practices were also familiar to the early Ayurvedic physicians. Indeed, it is during the Saṃhitā period (sixth century BCE – sixth century CE) that Ayurvedic compendiums display a rich engagement with music as a therapeutic agent. This engagement appears to encompass not only instrumental music (such as arched harps, lutes, and drums) and vocal performance, but also the organized use of pitch and scales, whose possible therapeutic relevance – particularly in relation to emotional associations and svara–rasa correlations – will be explored in the following Section 3.1.1. In the context of the *Bhāvaprakāśa* (sixteenth century CE), such harps were no longer expected, as they had largely disappeared from the Indian musical landscape by the tenth–twelfth century CE.

3.1.1 Basic modes and their equivalents

The seven musical notes – Śadja (Sa), Rṣabha (Ri), Gāndhāra (Ga), Madhyama (Ma), Pañcama (Pa), Dhaivata (Dha), and Niṣāda (Ni) – originate in the ancient Gāndharva tradition. These note names have survived unchanged and continue to be used across all forms of Indian classical music. In modern performance, the scale is usually built upon a chosen Sa – for instance, if Sa is selected as D or C#, the rest of the notes follow accordingly. However, in the ancient system, these notes were assigned fixed frequencies [8]. Among them, Madhyama (Ma), the fourth note, held a particularly vital position. The *Nāṭyaśāstra* (28.65) explicitly states that Ma was an indispensable, non-negotiable part of any musical rendition, very much like the centrality of Sa in modern practice.

Interestingly, the consonance between Sa and Ma has remained consistent since antiquity. If we assign Ma to the natural C in Western notation, the corresponding Sa of the ancient system would fall a perfect fourth on G. Based on this mapping, the following scale – G A B $\flat$ C D E F (=Sa Ri Ga Ma Pa Dha Ni) – can be identified as the foundational scale (grāma) of the ancient Indian system. From this, seven primary modes (mūrcchanās) emerge, each with close parallels in both Western and Indian (Hindustani and Carnatic) classical frameworks:

1. Sa – G – Kāfi/Kharaharapriyā: G A B $\flat$ C D E F (equivalent to the G Dorian mode).
2. Ri – A – Bhairavī/Vakulābharaṇam: A B $\flat$ C D E $\flat$ F G (equivalent to the A Phrygian mode).
3. Ga – B $\flat$ – Kalyāṇ/Kalyāṇī: B $\flat$ C D E F G A (equivalent to the B $\flat$ Lydian mode)
4. Ma – C – Khamāj/Harikāmbhojī: C D E F G A B $\flat$ (equivalent to the C Mixolydian mode).
5. Pa – D – Āsāvārī/Hanumat-toḍī: D E F G A B $\flat$ C (equivalent to the D Aeolian/Natural Minor mode)
6. Dha – E – Lalit-Bilās (not very familiar)/(NA): E F G A A $\sharp$ C D (a distinctive scale, reminiscent of the Phrygian mode but with an augmented 4th and a missing perfect 5th).
7. Ni – F – Bilāval/Śaṅkarābharaṇam: F G A B $\flat$ C D E (equivalent to the F Ionian/Major scale).

The scale system outlined above – closely aligned with Western modes such as G Dorian, A Phrygian, and B $\flat$ Lydian – suggests a system of emotionally nuanced scales. While early Indian medical texts do not explicitly describe such modal usage, their references to music in therapeutic contexts, including aphrodisiac therapy, fainting, and coma (see Section 4), make it reasonable to cautiously infer that such scalar types may have been employed in practice. In consequence, these modes could have formed part of an early framework for music-based emotional modulation (see svāra-rāsa mapping in 3.1.2).

3.1.2 Predominant notes and rasas

From the time of the *Nāṭyaśāstra* (c. third century BCE to third century CE) up to texts from the thirteenth century CE, rāsa³ correlations are consistently described in terms of the seven musical notes – that is, as scalar frameworks, rather than in the sense of rāgas as melodic formulas for improvisation. In ancient musical practice, the starting note of a performance (vādin; usage in Western context: vādī) was considered the predominant or key note. Notes or intervals consonant with this

¹³³ rāsa is a core concept in Indian aesthetics that is difficult to translate with a single word; however, “relish” or “aesthetic relish” are example expressions of its meaning (see Section 3.2).

vādin were given second priority (saṁvādin). In contrast, semitones were typically considered dissonant and given lower importance, while the remaining notes and intervals received preference just after the consonant.

According to this ancient framework (NS, 29.11) [7]:

- Sa and Ri predominance was correlated with the Vīra (heroic), Adbhuta (wondrous), or Raudra (furious) rasas, while Ga and Ni predominance produced Karuṇa (pathetic) rasa.
- Ma and Pa were associated with Śṛṅgāra (erotic) and Hāsyā (comic) rasas.
- Dha predominance evoked Bībhatsa (odious) and Bhayānaka (frightful) rasas.

The correlation of Śānta (tranquil) rasa with specific notes remains a topic of historical debate, as the *Nāṭyaśāstra* offers no explicit association. However, three critical later sources do reflect such associations.⁴ In those texts, the predominance of Sādja (Sa), Madhyama (Ma), and Pañcama (Pa) is associated with the evocation of Śānta (tranquil) rasa [7].

3.2 Rasa–doṣa mapping

Ayurvedic physicians emphasized that music used for treatment must be pleasing to the ear (*śruti-madhura*) and delightful to the mind (*manorama*) [2, 9, 10], as will become clearer in Section 4. In therapeutic applications, where precision in emotional modulation is vital, some interpretive liberties, for instance, through Rasa Hāsyā (comic), risk compromising the effectiveness of the intervention and its alignment with the patient’s subtle needs. Therefore, the rasas truly relevant for healing are Śānta (tranquil), Śṛṅgāra (erotic), Karuṇa (pathetic), Adbhuta (wondrous), and Vīra (heroic).

3.2.1 Enduring emotions and psychophysical responses as therapeutic achievements

Yet, for a better understanding, it is useful to elaborate on the rasa concept in more detail. According to the *Nāṭyaśāstra* (Chapter 6), rasa – refined emotional relish – arises from the interplay of three elements: determinants (vibhāva), expressive responses (anubhāva), and transitory emotions (vyabhicāri-bhāvas) [6], whereby determinants are further divided into personal determinant factors and stimulating factors.

Taking Śṛṅgāra (erotic) rasa as an example: the lovers serve as personal determinant factors; the moonlit garden and music are stimulating factors; shy glances and soft gestures represent expressive responses; and emotions like bashfulness, longing, or joy are transitory emotions. Together, these generate a harmonized aesthetic emotion – rasa.

Alongside the triadic functions of determinants, expressive responses, and transitory emotions in aesthetic theory, the concepts of *sthāyi-bhāva* (enduring emotion) and *sāttvika-bhāva* (psychophysical response) hold critical

¹³⁴The *Viṣṇudharmottara Purāṇa* (3.18) (latest strata, seventh century CE), the *Bṛhaddeśi* (c. eighth century CE), and the *Saṅgītaratnākara* (c. twelfth–thirteenth century CE).

importance, especially in linking classical Indian aesthetics to Ayurvedic psychology [6].

Beings, by nature, possess enduring emotions such as tranquility (śama), love (rati), laughter (hāsa), sorrow (śoka), anger (krodha), enthusiasm (utsāha), fear (bhaya), disgust (jugupsā), and wonder (vismaya).⁵

When an aesthetic experience transforms individuals into an emotionally resonant and relishable state (rasa), which is an active and conscious process, it signifies the illumination of Sattva (light, clarity, harmony), temporarily suppressing Rajas (activity, passion, restlessness) and Tamas (inertia, darkness, ignorance), leading to internal purification. Thus, experiencing rasa is a realignment of one's inner state – an affective and psycho-spiritual transformation, evoking psycho-spiritual resonance. This is also the principle music therapy is based on: Its effectiveness is centered on the conscious experience of emotions and rasa states. A patient becomes aware of these states and, through that awareness, can consciously modulate them. This process transforms a passive emotional experience into an active, therapeutic one, allowing “inner purification” and realignment. Involuntary psychophysical responses (e.g., trembling, tears) confirm aesthetic relish, occurring only when consciousness is predominantly Sāttvic.

3.2.2 Toward a therapeutic framework: Svāra–rasa–doṣa

Disease, as noted in the *Caraka Saṃhitā* (CS), manifests from disturbances in both bodily and mental domains. Given the established correlations between principal qualities and *rasa*, and principal qualities and *humors*, we can now explore the critical relationship between *rasa* and somatic humors (**Figures 1 and 3**). This understanding is vital for reconstructing the logic of music therapy in classical Ayurveda, as compendia explicitly prescribe music for aggravated *doṣas* (bodily humors; see Section 2), while, at the same time, music directly operates through *rasa*. However, not all *rasas* are suitable for treatment. My analysis will focus on five specific *rasas* most appropriate for music therapy: Śṛṅgāra (erotic), Karuṇa (pathetic), Adbhuta (wondrous), Vīra (heroic), and Śānta (tranquil). Adbhuta and Śānta *rasas* are considered universally applicable for any *doṣa* imbalance, fostering clarity and harmony without aggravation. They effectively calm mental agitation and promote healing (**Figure 3**).

In Ayurvedic therapeutics, the principles that “like increases like” and “opposites balance” (in Sanskrit: Sāmānya-viśeṣa-siddhānta, SVS) are the foundational guides for all *doṣa* management (CS, Sūtrasthāna: 1.44). When applying this to the three fundamental qualities [Sattva (light, clarity, harmony), Rajas (activity, passion, restlessness), Tamas (inertia, darkness, ignorance)] inherent in various *rasas*, and considering the intrinsic fundamental qualities inherent in bodily humors (**Figure 1**), we arrive at a profoundly refined understanding of their therapeutic applications.

⁵ Further specification would lead to more marginal scholarly views. For example, one group thinks that all other emotions are modifications of tranquility, while another presents the view that tranquility has no criteria to fulfill as one of the *rasas*. Instead, they propose there are four basic *rasas* (love, anger, enthusiasm, and disgust) out of which the other four (laughter, sorrow, wonder, and fear) are modified (e.g., laughter is a modification of love, and so on), totaling eight *rasas*. Furthermore, some of the enduring emotions can also become part of the transitory emotions. Still, the information presented above provides the basic general structure of the theory.

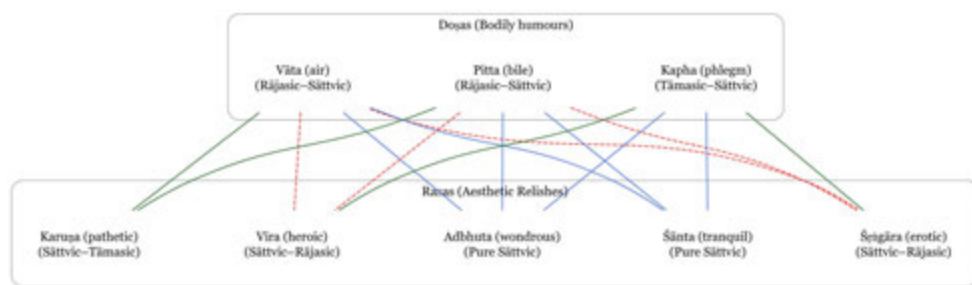


Figure 3.

Integrative model of rasas and doṣas [lines: green = therapeutic, red = contraindicated, and blue = universal balancing].

That is, for aggravated Vāta (air) and Pitta (bile) – the doṣas of the body, which are inherently Rājasic-Sāttvic – Śṛṅgāra (erotic) and Vīra (heroic), which are also Sāttvic-Rājasic, are less preferable (**Figures 1 and 3**). Their Rājasic component would increase the existing excess, contravening the pacification aim. Conversely, Karuṇa (Sāttvic-Tāmasic) is therapeutically appropriate. Karuṇa’s (pathetic) grounding directly counteracts intensity and heat, reducing excessive movement and fostering stillness (**Figures 1 and 3**). For aggravated Kapha (heavy, slow, Tāmasic), Śṛṅgāra (erotic) and Vīra (heroic) rasas are most suitable. Their inherent Rājasic energy directly opposes Kapha’s Tāmasic tendency, stimulating movement and dynamism to counter inertia (**Figures 1 and 3**). If all three doṣas are aggravated, the pure and universally balancing Adbhuta (wondrous) and Śānta (tranquil) rasas remain the preferred choice.

The discussion above shows that the ancient Indian approach to music and healing rested on the structured modal system (mūrcchanā) of the Gāndharva tradition. Thus, music was based on complete scales and not on single-note renditions, whereby each scale was defined by the predominance of a particular note (vādin). This predominant note was traditionally associated with a specific rasa, each embodying a characteristic composition of the three fundamental qualities – Sattva (light, clarity, harmony), Rajas (activity, passion, restlessness), and Tamas (inertia, darkness, ignorance). Since the bodily humors – Vāta (air), Pitta (bile), and Kapha (phlegm) – are likewise defined by fixed profiles of these fundamental qualities, the choice of mūrcchanā could be guided by the same principle of SVS as in Ayurveda (“like increases like”; “opposites balance”). Therapeutically speaking, the aim was to select a melodic mode whose rasa-based fundamental quality profile would counteract the qualities of an *aggravated* humor, while avoiding modes whose rasa might intensify the imbalance. This SRD framework⁶ explains the selective use of certain rasas for healing, the exclusion of others despite their artistic appeal, and the direct compatibility between the Nāṭyaśāstra’s aesthetic theory and the clinical logic of the Ayurvedic medical system. It is this framework that underpins the therapeutic applications of music, for which examples are given in the next section.

⁶ Note that in music therapeutical contexts, a rasa-doṣa correlation can never be established without the svara-link. In ancient texts, the rasas in music were always mentioned in conjunction with the svaras (musical notes). However, the correlation of one svara to one rasa does not mean that a single note could express a rasa. Rather, this was a sort of convention of ancient musicology: for example, mentioning Sa for Vīra, Raudra, and Adbhuta meant the melodic scale, where Sa was the vadin (predominant note). Similarly, Ga was correlated with Karuṇa, indicating a melodic scale where Ga was the predominant note.

4. Musical interventions in classical Ayurveda

Ayurvedic texts recommend music as an adjunct, that is, as part of an overall therapy. The following section provides detailed prescriptions regarding various circumstances and diseases, based on thorough readings and analyses of the respective Sanskrit texts.

4.1 Aggravated bile

The CS (Vimānasthāna:6.17) [2], AS (Sūtrasthāna:21.2) [11], and AH⁷ (Sūtrasthāna:13.6) [5] recommend music as a remedy in cases of aggravated bile or Pitta. Given Pitta's (bile) inherent Rājasic-Sāttvic nature, therapy aims to introduce opposing qualities. Karuṇa (pathetic) rasa, with its Sāttvic-Tāmasic profile, balances Pitta's intensity, while Adbhuta (wondrous) and Śānta (tranquil) rasas – being purely Sāttvic – offer universal calming effects. Musically, this emphasizes svaras like Ga, Ni, Ma, Pa, and Sa. Notably, while the detailed mapping of svara to rasa, especially for Śānta, became prominent by the time of AH and AS, that is, before 600 CE, the CS reference predates Drḍhabala's redactions, which happened around (c. second–fourth century CE). This implies that even in its earliest form, CS recognized music's therapeutic use based on *scales* for Pitta – likely favoring tonalities dominated by Ga and Ni. This approach finds a later cross-cultural parallel in the work of the medieval Jewish philosopher and physician Musa bin Maimun (Maimonides, 1135–1204 CE). He similarly prescribed music as a remedy for melancholy, an affliction he directly linked to an excess of black bile.

4.2 Coma

Music is recommended in coma-related treatment too, though with a different emphasis. SS (Uttaratantra:46.22–23) [12],⁸ explicitly prescribes the use of musical instruments (vīṇā, etc.) and soothing speech to break coma, indicating its active therapeutic role in reviving consciousness. By contrast, CS (Sūtrasthāna:24.51–53) and AH (Cikitsāsthāna:7.113–115) focus on patients who have already regained consciousness, recommending refined songs, instrumental sounds, pleasant sights, and memories to prevent the mind from slipping back into unconsciousness.

From the SRD model, coma reflects Tāmasic dominance. Thus, one should refer to Adbhuta (wondrous) (totally Sāttvic) and possibly Vīra (heroic) (Sāttvic–Rājasic) rasas for breaking this inertia. For relapse prevention (CS and AH), Adbhuta supports alertness, while Śānta fosters stability.

4.3 Pregnant women and newborns

In CS (Śārīrasthāna: 8.47) [2], music is prescribed for the newborn's well-being – songs (gīta), instruments (vāditra), and blessings are to fill the home. KS (Śārīrasthāna: Jātisūtriya Śārīra Chapter) extends this to prenatal care, advising melodies (brahmaghoṣa) with instrumental accompaniment for expectant mothers. The term vāditra in KS, likely referring to the vīṇā, suggests stringed or percussive

¹³⁷ Please see all abbreviations of AH, CS etc., in the Nomenclature table.

⁸ Likely redacted during Nāgārjuna's time around (first–fifth century CE).

instruments [13]. These early references imply that music held a supportive role in pre- and perinatal health. Such music – likely Sāttvic in nature – has been used to stabilize the mother’s or infant’s mind by enhancing Sattva (light, clarity, harmony). Note that in most Ayurvedic texts, the terms “gīta” (song) and “vāditra” (instruments) were used as a compound, “gītavāditra,” to generally refer to aspects of music. Gīta is a general term for both songs in classical and folk music. “Vāditra” is defined as the equipment used to play instruments, such as fingers or a plectrum. It is also a synonym for the term “vādyā,” which refers to four types of instruments: stringed, drum-like, wind, and metallic percussion [10].

4.4 Aphrodisiac therapy

Music is consistently prescribed in aphrodisiac therapy (vājikaraṇa) across all Br̥hatrayī texts, AS, and BhP, showing its sustained role from ancient through medieval periods. CS (Cikitsāsthāna:2/3.29–30) [9] links the music of the ancient Gāndharva tradition to enhanced sexual exhilaration, suggesting that its therapeutic use was culturally embedded [9]. SS (Cikitsāsthāna:26.8–11) [14], AS (Uttaratantra:50.133) [15], AH (Uttarasthāna:40.42–45), and BhP (Uttarakhaṇḍa:72.16–19) [16] echo this, describing pleasant music as enhancing potency. Within the SRD framework, this likely involves Śṛṅgāra (erotic) rasa (predominant Ma/Pa), which, with its Sāttvic-Rājasic profile, stimulates emotional warmth and vitality. Notably, this section was untouched by later redactors of CS or SS, affirming its antiquity. Both men and women are included as recipients of the therapy. However, whether this form of music therapy was exclusively for the rich and powerful or also accessible to the less wealthy is difficult to state with certainty. Most likely, the wealthy had greater opportunities to receive such treatment due to the resources and time involved.

4.5 Tuberculosis

The application of music in the treatment of tuberculosis (rājayaḥṣman) is first documented in CS (Cikitsāsthāna: 8.186–188, around the second century BCE), where pleasant music and instrumental sounds are said to prevent rogarāja, the “king of diseases.” Chakrapāṇidatta (c. eleventh century CE), the famous commentator of CS, interprets this as promoting a Sāttvic state of mind, which is essential for halting disease progression [17]. Significantly, this description predates Dr̥ḍhabala’s redaction, supporting its antiquity. SS does not mention music in this context, but both AS (Cikitsāsthāna: 7.117) [18] and AH (Cikitsāsthāna: 5.83) reiterate this prescription verbatim, as does GN (Kāyacikitsākhanda: 9.100–102) of Sodhala [19], who was active in the Yādava court of thirteenth-century Devagiri (in Karnataka, a state in South India).

Within the SRD model, Śānta (tranquil) and Adbhuta (wondrous) rasas – both Sāttvic – would be suitable, as they counteract the Tāmasic lethargy and emotional depletion caused by progressive tissue loss. Music in this context serves not merely as entertainment but as a subtle, rasa-based therapeutic tool for uplifting and stabilizing consciousness, providing psychological nourishment to patients enduring chronic wasting diseases.

4.6 Alcohol intoxication

The use of music in the treatment of alcohol intoxication (madātyaya) is notably referenced in CS (Cikitsāsthāna: 24.192–193), where “gāndharvaśabda” (musical

sound) is said to alleviate the condition. Since this section predates the redaction by the famous physician Dṛḍhabala, it likely reflects a pre-Gupta formulation. The intervention is explicitly applied in a type of alcohol intoxication (sannipātika), where all three doṣas are simultaneously aggravated. Chakrapāṇidatta (eleventh century CE) classifies this under “praharṣiṇī cikitsā” – exhilarative therapy – meant to restore mental joy (manahprīti) and physical calm (śarīrapraśama).

Within the SRD framework, Vīra (heroic) rasa, with its energizing Sāttvic-Rājasic qualities, appears therapeutically appropriate, as it cultivates clarity and enthusiasm. Though Suśruta (SS) and AH do not include this, the same idea recurs in AS (Cikitsāsthāna: 9.91) and BS (Cikitsāsthāna: 28.23) [20]. While Bhela’s (BS) current form may not be original, its agreement with Caraka (CS) suggests that music was part of a broader early therapeutic discourse in cases of intoxication.

4.7 Therapeutic emesis and purgation

The use of music also emerges as a specific therapeutic measure in emesis (vamana) and purgation (virechana) procedures, particularly when fainting occurs due to atiyoga – excessive or overly forceful action. In CS (Siddhisthāna:6.85), both comforting speech (sāma) and musical sound (gāndharvaśabda) are recommended for such episodes. Though this section belongs to the Gupta-period addition by Dṛḍhabala, a similar instruction appears earlier in SS (Cikitsāsthāna:34.12–13), where flute (veṇu), arched harps (vīṇā), and song (gīta) are used to revive a fainted patient [14]. AS (Kalpasthāna:3.23) [18] and AH (Kalpasthāna:3.39) [5] repeat this guidance.

According to the SRD framework, Adbhuta and Vīra rasas – which stimulate alertness – are most suitable, though Śānta (tranquil) rasa, affirmed in post-Gupta sources, may also apply. The continuity of this idea across the Great Triad, beyond 600 CE, confirms that music held a secure place in the classical medical treatment of consciousness loss.

4.8 Music artists in the hospital

Perhaps the most striking reference to music therapy in CS (Sūtrasthāna:15.7, c. second century BCE) is the inclusion of musicians and storytellers as part of the hospital infrastructure. The text recommends employing individuals skilled in singing, instrumental music, verses, epics, and mythical narratives to assist in patient care. This indicates an institutional awareness of mental well-being, centuries before similar practices appeared in Islamic medical contexts. The presence of such figures in a clinical setting suggests that music was not only therapeutic but also considered essential to creating a healing environment. It is thus plausible that individuals of this disposition were intentionally engaged in this work, reflecting a sophisticated integration of aesthetic psychology and public health within classical Indian medicine.

4.9 Companion to digestion

The KS (Khilasthāna: 5.55–56) describes music and storytelling as companions to digestion [13]. It recommends listening to songs accompanied by the vīṇā and veṇu (flute), dramatic performances, and engaging stories after meals to enhance strength and nourishment. This reflects a nuanced awareness that mental ease and sensory

delight promote better digestion. Notably, since the entire Khilasthāna is a later addition by the redactor Vātsya (c. seventh–ninth century CE), the credit for this refined digestive therapeutic model rightly belongs to him. It also reinforces the Ayurvedic view that *harṣa* (mental contentment) is vital to physical well-being. In lighter terms, the modern habit of listening to relaxing audio – such as a podcast or dramatic narration – during or after meals can be seen as a distant echo of this thoughtful prescription from the KS, as reshaped by Vātsya.

4.10 Hyperpyrexia

A unique reference to music as a remedy for burning fever or hyperpyrexia (*dāhajvara*) appears in Vṛddha Vāgbhaṭa's (c. sixth century CE) AS (*Cikitsāsthāna*: 2.77): *vallakimadthuram gītam. . . haranti dāham* – sweet music played on the *vallakī* (a synonym for *viṇā*) alleviates burning sensations [18]. This is the only known classical reference linking instrumental music directly to fever treatment. CS (*Cikitsāsthāna*: 3.259) discusses high fever (*dāha*) in detail, without mentioning music. Thus, Vṛddha Vāgbhaṭa is a significant contributor to the history of music therapy in Ayurveda. After the Caraka Saṃhitā, he appears to be the first physician to reintroduce original applications of music in clinical care.

4.11 Musical aids to alchemy

While *Rasāyana* – the rejuvenative therapy for strengthening bodily tissues – is one of the classical eight branches of Ayurveda, the development of alchemical and iatrochemical medicine (*Rasaśāstra*) marks a separate trajectory. By the fifth century CE, mercury-based preparations gained prominence within Śaiva and Buddhist tantric traditions. These traditions pursued the dual goals of *dehasiddhi* (the perfection and immortality of the body) and *lohasiddhi* (the transmutation of base metals into gold). This quest for corporeal perfection and alchemical transformation drove the complex methods of mineral purification and preparation that became the hallmark of this doctrine.

The *Ānandakanda* (twelfth–thirteenth century) [21], likely composed in Śrīśaila (Andhra Pradesh, South India), prescribes aesthetic stimulation, such as watching dance (*nartanālokana*) and listening to music (*gītaśravaṇa*) (1.6.91), as beneficial while taking iatrochemical medicine [22].

Similarly, Śālinātha's *Rasamañjarī* (6.111–112) (before the last quarter of the fifteenth century) [21] recommends *viṇā* music and songs before administering a compound of minerals (mica, copper, silver, lead) and herbs (*guggula*, *nāgakeśara*),⁹ used to treat *pāṇḍu* (anemia) and *kāmalā* (jaundice) [23]. This practice was aimed at calming internal heat and mental agitation.

4.12 Music in daily life

This section shifts from specific therapeutic contexts to the broader integration of music into daily life, as viewed through compatibility (*sātmya*) and

¹⁴⁰This alchemical medicine, which the author named *Pratāpalakṣeśvara*, was typically named after one of the many epithets of the Hindu deity Śiva (also known as Shankara or Mahadeva). In this iatrochemical tradition, the name is an inherited practice and serves as a devotional or symbolic label rather than a literal description of the specific herbs and minerals used in the preparation.

incompatibility (asātmya). According to Vāgbhaṭa (c. 600s CE), health depends on the balanced use of medicine, food, and lifestyle, adapted to the body–mind constitution (AH, Nidānasthāna: 1.6–7). Physician Aruṇadatta (c. 1200s CE), in his commentary on AH, includes music among the daily acts performed by speech, body, and mind. This suggests that music, whether practiced or passively received, contributes to maintaining mental balance and physical well-being [5].

This links directly to the concept of improper contact between the senses and their objects (asātmya-indriyārtha-saṃyoga) described in CS (Sūtrasthāna: 11.38–43). When appropriately chosen and practiced in moderation, music fosters psychosensory compatibility and helps prevent diseases caused by excessive (atiyoga) or inappropriate (mithyāyoga) stimulation. This Ayurvedic concept of asātmya-indriyārtha-saṃyoga applies to the contact between all five senses and their objects. Substance abuse is a form of mithyāyoga (inappropriate contact), particularly for the senses of taste and smell, as it introduces substances that are not conducive to health. Certain types of music can be both a form of atiyoga (excessive stimulation) if it is played too loudly, and mithyāyoga if the sound is jarring, dissonant, or unpleasant, which can agitate the mind and body. The same holds for visual stimuli; they can be atiyoga (e.g., prolonged exposure to bright screens or harsh sunlight) or mithyāyoga (e.g., exposure to frightening or disturbing imagery). This framework highlights that the quality, quantity, and appropriateness of all sensory input are crucial for maintaining health and preventing psychosomatic diseases. Accordingly, music is not merely entertainment but a mental–somatic regulator in daily life. Aruṇadatta’s inclusive vision hints at the potential for music to become a formal therapeutic discipline, though such a development was never fully materialized in Indian medical history, emerging for the first time from the early 2000s onward (see Section 5).

5. An overview of the modern era

After the sixteenth century, no further references appeared in Ayurvedic texts, marking the close of the classical discourse on music therapy.

Yet with the advent of the modern period in the twentieth century, the history of Indian music therapy entered a new phase. The earliest known instance appeared in 1938, followed by scattered efforts in the 1980s, mostly confined to the personal clinics of physicians or small-scale initiatives by musicians. These early endeavors, while notable, remained isolated and lacked the benefit of systematic evaluation or critical academic engagement. It was not until 2005 that the idea of therapeutic music began to gain explicit scholarly attention in peer-reviewed literature.

5.1 British India: 1930s

The early nineteenth century saw the establishment of formal institutions for mental health in British India, with a lunatic asylum opening in Munger (now in Jharkhand, India [East]) in 1795, followed by one in Murli Bazar, Dhaka (in Bengal Presidency, British India; now the capital of Bangladesh) around 1815. By 1922, these institutions were rebranded as “Mental Hospitals,” with *Ranchi’s facility* serving as a prototype example, founded in 1918. It became the Indian Mental Hospital (now known as the Central Institute of Psychiatry, Ranchi, Jharkhand, India [East]) [24].

While Dr. L.P. Varma¹⁰, writing in 1953, acknowledged the Dhaka asylum, he did not mention music therapy in that context, referring only to a 1926 proposal for a music room at Ranchi [24]. However, a 1984 article co-authored by Dr. Sridhar Sharma and Varma retrospectively claimed that Dhaka used music to calm patients, attributing this to Varma's earlier work [25]. Although this claim lacks supporting verification, it suggests a nascent interest in music therapy during the British Raj. Note, however, that there was no British or colonial influence on music therapy in India at that time.

After more than 350 years since BhP, a more concrete development emerged in the 1930s when J. Pāl¹¹ established the *Musico-Medical Institute on Chamberlain Road*, Lahore, British India (now in Pakistan). Our understanding of Pāl and his therapeutic methods primarily derives from his Hindi book, *Rāg Cikitsā*, published in 1938 [26].

J. Pāl's work first proposed a system correlating specific musical notes (svaras) with doṣa-dominant ailments, suggesting that the prominent note of a Raga could treat corresponding doṣa imbalances (the three bodily humors already discussed in Sections 2 and 3). For instance, he linked Ṣadja (Sa) to Pitta (bile) and Rṣabha (Ri) to both Pitta (bile) and Kapha (phlegm). His book *Rāg Cikitsā* (see **Figure 4**) also recounts numerous successes, including a notable case where Hakim Alam Shah utilized Sarangi players to calm a violently mentally ill patient, leading to recovery [26]. J. Pāl's book includes 20 testimonial letters from patients praising his musical cures, often after failed conventional treatments. A letter from July 2, 1937, by Vikramajit Singh, a lecturer at Forman Christian College, also attested to relief from insomnia through J. Pāl's Sitar music [26].

J. Pāl, a skilled sitar player, considered his instrument most suitable for therapy. While acknowledging earlier music applications by hakims (Unani doctors representing the Graeco-Arabic tradition of medicine), he aimed to establish a systematic, theoretical basis for music therapy within Ayurveda. His belief that a music therapist could discern deeper patient pain beyond conventional medicine reflected a distinct approach, suggesting that, for him, a traditional medical background was not a prerequisite for music therapists. In this way, Pāl, a sitarist and not a trained physician, advanced a musician-driven model of music therapy, rooted in practice rather than formal medical training. He was undoubtedly innovative in reviving a practice whose Ayurvedic precedents had largely faded centuries earlier.

5.2 1980s to the present decade

In 2002, reports from Maharashtra (a western state of India) highlighted the use of music therapy applications and consultations in several mental hospitals, including those in Pune, Nagpur, Thane, and Ratnagiri (all in Maharashtra, India [West]). Earlier, in collaboration with doctors, sitarist Shashank Katti had already been practicing music therapy since 1998 and subsequently established the *Sur Sanjeevan Music Therapy Trust* in Mumbai. Even further back, in 1987, Dr. Anil Patil in Dadar (Mumbai, Maharashtra, India [West]) used music as a supplementary treatment for physical and mental disorders [27]. Thus, between 1980 and 2002, music therapy wasn't common practice across India.

¹⁴² Dr. L. P. Varma was a pioneering Indian psychiatrist and leader of the Ranchi Institute, which gained international renown for psychiatric training and treatment under his guidance.

¹¹ J. Pāl: His first name, written in full, is not known to us till now.

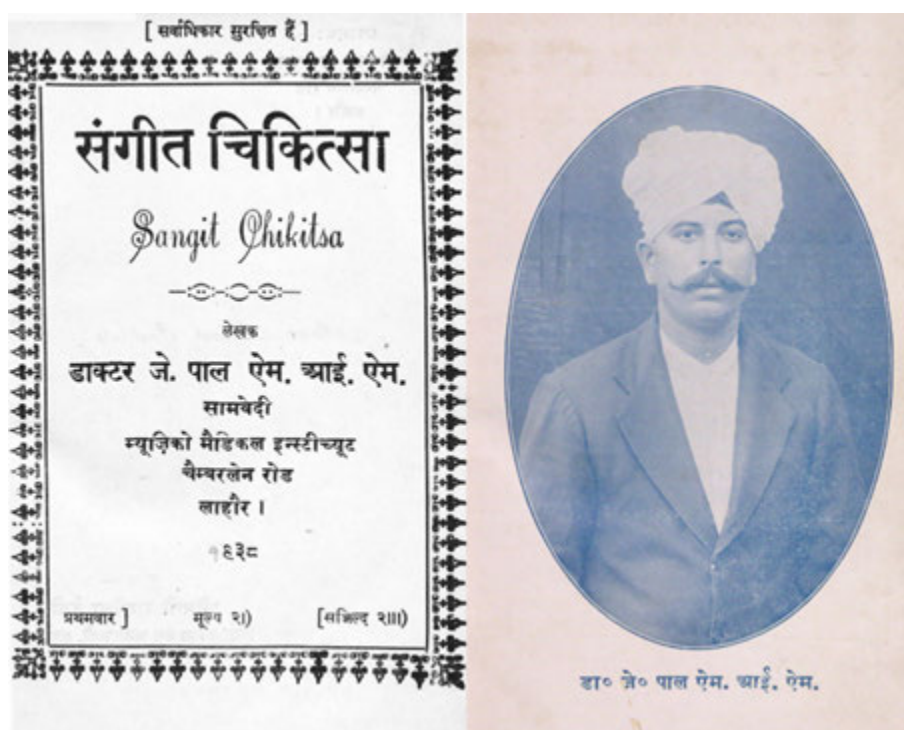


Figure 4.
 Front page and photograph of J. Pāl as printed in his book *Sangit Chikitsa*, published in 1938.

Despite that, public discussions surrounding music as medicine gained significant prominence in early 2001, with media outlets framing it as a holistic healing tool. However, an experimental initiative in Kolkata, reported in August 2001, aimed at reducing suicide rates through music underscored that music alone might not be a comprehensive solution [28]. Despite this caveat, optimism about music therapy persisted. By 2002, articles continued to highlight expert discussions on music's potential to support immune function; they also documented successful instances of anxiety reduction through sacred music achieved in mental hospitals [29, 30].

By 2005, music therapy began to be piloted in hospital settings in a more systematic manner, with reports about its use in Mumbai and the establishment of a dedicated raga therapy unit at *Chennai's Apollo Hospital* [31]. This period also saw significant institutional developments: the *Indian Association of Professional Music Therapists* (IAPMT) was founded in 2010 by Dr. Dinesh Chandra Sharma to train certified professionals, and in 2018, the *Indian Music Therapy Association* (IMTA) was established under T.V. Sairam's leadership, aiming to promote standards, ethics, and interdisciplinary research. By 2014, premier mental health institutions like the *Ranchi Institute of Neuro-Psychiatry and Allied Sciences* (RINPAS) had implemented raga therapy as a standardized healing procedure [32].

Early scholarly engagement with ragas began in 2005, with Uma Gupta and colleagues investigating rāga *Toḍī's* (Hindustani) impact on brain waves, depression,

and anxiety [33]. By 2009, broader music research emerged, examining pain relief during venepuncture, quality sleep in depressed patients, and, by 2014, *rāga Ānanda Bhairavī* (Carnatic) was assessed for post-operative pain relief [34–36].

By 2015, studies demonstrated that Indian classical music therapy can significantly reduce stress and anesthetic drug requirements both in preparation for surgery and during the procedure itself. It has also been used to effectively modulate blood pressure through autonomic function and elicit distinct emotional responses based on specific *rāga* structures and musical parameters [37–39]. Across these studies, *ragas* were pre-selected by the researchers, and music was delivered via headphones or earphones, ensuring a controlled auditory environment.

By 2019, research expanded to include music interventions for neurodevelopmental disorders. In addition, *rāga Bhūpālī* (Hindustani) was used for stabilizing attention in schizophrenia [40], alongside broader psychiatric research [41]. The early 2020s deepened the understanding of *rāga*’s psychological impact, with *rāga Kāpī* (Carnatic) being used for anxiety reduction [42]. Notably, while certain *ragas* like *Bhūpālī* are reported to help with specific conditions, there is little evidence showing that a *rāga* is effective for one condition but not for another. Such particularization – for example, “effective for A but not for B” – is largely absent in current studies.

From 2023 to 2025, research remains comprehensive. Studies from 2023 demonstrate that listening to Indian *ragas* boosts alpha brainwaves and attention [43], whereas lifelong Carnatic training enhances cognition and brain volume [44]. Current 2024–2025 research shows that music enhances brain connectivity and emotion-related activity, and reduces pain and anxiety in palliative care [45–50]. Together, these studies mark a growing scientific recognition of Indian classical music’s therapeutic potential across neurological, psychological, and medical domains.

5.3 Cakra–svara and *rāga*–roga myth

J. Pāl’s 1938 *Rāg Cikitsā* marks the first systematic attempt to relate specific *ragas* with diseases (*roga*), introducing the concept of *rāga–roga*¹² correlation. He also

¹² Proposed by Dr. J. Pāl, Hindustani *ragas* do have therapeutic effects that are highly specific: Insomnia – Basant, Khamaj, Marwa, Pilu, Sohini, Bhairavi; Hysteria – Tilak Kamod, Kafi, Multani, Khamaj, Darbari Kanada, Yog, Jogia, Sohini; Anxiety – Yaman Kalyan, Purvi, Pilu, Puriya, Tilang, Darbari Kanada; mental disorder due to air (Vāta) – Bahar, Bageshwari; mental disorder due to bile (Pitta) – Khamaj, Kedar, Lalit; Sattva (light, clarity, harmony)-dominant melancholic mental disorder – Asavari, Hindol; Rajas (activity, passion, restlessness)-dominant melancholic mental disorder – Bhairavi, Jogia; Tamas (inertia, darkness, ignorance)-dominant melancholic mental disorder – Sohini, Yog, Bhairav; Epilepsy due to bile (Pitta) – Darbari Kanada, Behag, Dhani; Headache due to air (Vāta) – Kamod, Multani, Paraj; Headache due to bile (Pitta) – Sohini, Todi, Bhairavi; Headache due to phlegm (Kapha) – Malkosh; Generalized Pain – Asavari, Bilawal, Bhairavi, Jaijaiwanti, Behag; Headache (old age/overexertion) – Bhairavi, Basant; Blood Pressure – Hindol, Puriya, Purvi, Todi; Grief – Lalit, Malkosh, Bhimpalasi; Tuberculosis – Tilang, Bageshwari, Behag, Puriya, Bilawal, Asa, Multani, Ramkali; Sexual Impotence – Basant, Sohrab, Khamaj, Sarang; Nocturnal Emission – Bhup, Bhairavi, Asavari; Heart Disease – Sohrab, Bhopali, Shankara; Pitta-dominant Pradar (abnormal vaginal discharge) – Sarang, Desh; phlegm(Kapha)-dominant abnormal vaginal discharge – Kafi, Jaunpuri, Mao, Kedar, Jogia, Tilak Kamod; air(Vāta)-dominant abnormal vaginal discharge – Malkosh, Bhimpalasi; Typhoid – Tilang, Behag, Multani, Malkosh, Bhimpalasi, Kalingda; Dysentery – Darbari Kanada, Bageshwari, Khamaj; Malaria – Todi, Pilu, Sohini, Marwa, Bhup, Hindol, Khamaj, Basant; Measles – Bhairav, Bham; Liver Problems – Purvi, Todi, Pilu, Sohini, Yaman Kalyan, Puriya, Shankara, Bhup, Dhani, Tilang, Khamaj, Behag, Bageshwari; Intermittent fever due to phlegm (Kapha) and bile (Pitta) – Desh, Sarang.

proposed a theory of svara-dōṣa correlation, where the dominant svara (vādin) of a rāga corresponds to a bodily dōṣa (humor). This theory, however, exhibits significant limitations. Though Pāl referenced medieval (c. thirteenth–fourteenth century CE) texts such as *Sanḡita-makaranda* (1.78–84), his interpretation – assigning svaras to colors and attributes such as dry/cold/warm – was closely based on Greek–Arab-influenced Unani theory rather than on Ayurvedic frameworks. For example, Ṛṣabha (Ri) is described as cold and dry, alleviating Pitta–Kapha diseases – a logic more aligned with Unani classification. Moreover, Indian classical music today involves 12 svaras, including komal and tīvra variants. Based solely on the seven primary svaras, Pāl’s theory offers no guidance for these newly evolved forms. In detail, no traditional or textual basis is provided if one asks about the rasa or therapeutic impact of komala ṛṣabha (flat Ri) or tīvra madhyama (sharp Ma). While in the ancient system, such svara-dōṣa mappings were based on fixed mūrccanā-based contexts (discussed earlier), the musical language has since transformed. Recent efforts to map rasa to all 12 svaras are ongoing but remain in an exploratory phase [51, 52].

Despite these issues, the rāga-roga linkage introduced by Pāl is frequently cited in contemporary Indian music therapy without acknowledging his contribution and is often mistakenly presented as an inherited part of tradition [53]. Similarly, popular claims linking individual svaras to specific chakras (subtle energy centers along the spine) lack a historical basis [54, 55].

6. Conclusions

The chronological survey of Ayurvedic literature reveals a deep and evolving engagement with music as a therapeutic remedy (**Table 1**). The SS, one of the earliest medical texts (c. sixth century BCE), provides foundational references to music within specific contexts, such as fainting during therapeutic emesis and purgation. However, it is the CS, dated to around the second century BCE, that offers the most diverse and conceptually developed applications of music therapy, including its use in aphrodisiac treatment, aggravated Pitta (bile), tuberculosis, alcohol intoxication, coma management, and prenatal care.

Later texts, such as the AS (c. 500s CE), AH (c. 600s CE), and KS (c. seventh–ninth century CE), reflect and elaborate upon these early formulations. By the thirteenth–fifteenth centuries, treatises such as the GN and BhP continued the tradition, while Aunādatta’s twelfth-century commentary on AH hinted at its *active* approach to therapy. Meanwhile, alchemical works such as the *Ānandakanda* and *Rasamañjarī* (twelfth–fifteenth century CE) incorporated music into ritualized iatrochemical practices. Yet, strikingly, the Ayurvedic treatise *Śārṅgadharma Samhitā* (Uttarakhaṇḍa: 4.40–44) (thirteenth–fourteenth century CE), while addressing complications such as fainting during emesis and purgation, remains silent on the use of music as a preventive aid [56] – an omission that contrasts sharply with the therapeutic role of music articulated as early as the sixth century BCE and signals a significant lapse in a once-integrated clinical art.

Spanning over two millennia, this trajectory illustrates how music, though never institutionalized as an independent branch of Ayurveda, remained a vital and adaptive therapeutic tool embedded in the subtle science of healing. However, this

Application	Chronology
Aphrodisiac therapy	Early SS (c. sixth century BCE), early CS (c. second century BCE), AS/AH (c. 500s–600s CE), BhP (c. 1500s CE)
Fainting during therapeutic emesis/purgation	early SS (c. sixth century BCE), redacted CS by Dṛḍhabala (c. 200s–400s CE), AS/AH (c. 500s–600s CE).
music artists in a hospital setting	early CS (c. second century BCE)
aggravated Pitta	Early CS (circa second century BCE), AS/AH (circa 500s–600s CE)
Tuberculosis	Early CS (c. second century BCE), AS/AH (c. 500s–600s CE), GN (c. thirteenth century CE)
Perinatal care (pregnant women and newborns)	Early CS (c. second century BCE), KS (c. seventh–ninth century CE)
Alcohol intoxication	Early CS (c. second century BCE), AS (c. 500s CE), BS (c. seventh century CE)
Breaking coma	Early CS (circa second century BCE), AH (circa 600s CE)
Preventing coma relapse	Redacted SS by Nāgārjuna (c. first–fifth century CE)
Hyperpyrexia	By Vṛddha Vāgbhaṭa (AS, c. 500s CE)
Companion to digestion	KS (c. seventh–ninth century CE, Vātsyā's addition)
Music in daily life	Arunadatta in Ayurvedic exegesis (c. 1200s CE)
Musical aids to alchemy	Ānandakanda (c. twelfth–thirteenth century CE), by Śālinātha (c. fifteenth century CE)

[AH = Aṣṭāṅga Hr̥daya, AS = Aṣṭāṅga Saṅgraha, BhP = Bhāvaprakāśa, BS = Bhela Saṃhitā, CS = Caraka Saṃhitā, GN = Gadanigraha, KS = Kāśyapa Saṃhitā, SS = Suśruta Saṃhitā.]

Table 1.
Chronology of applications of music therapy in Sanskrit texts.

chapter remains limited by its reliance on existing printed editions. A vast corpus of unpublished manuscripts and lesser-known regional or unprinted versions of Ayurvedic and alchemical texts remains unexplored. These sources may contain additional references to the therapeutic use of music – possibly in vernacular adaptations, ritual manuals, or localized commentaries – that have yet to be critically edited or catalogued. As such, a more comprehensive understanding of music's role in Indian medical traditions will require further philological work, manuscript study, and interdisciplinary collaboration across musicology, Ayurveda, and textual scholarship.

Nomenclature

AH	Aṣṭāṅga Hr̥daya
AS	Aṣṭāṅga Saṅgraha
BhP	Bhāvaprakāśa
BS	Bhela Saṃhitā
CS	Caraka Saṃhitā
GN	Gadanigraha
KS	Kāśyapa Saṃhitā
NS	Nāṭyaśāstra
SS	Suśruta Saṃhitā

Author details

Abirlal Gangopadhyay
Independent Researcher in Sanskrit Studies, Barasat, India

*Address all correspondence to: abirlalgangopadhyay@gmail.com

IntechOpen

© 2025 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Meulenbeld GJ. A History of Indian Medical Literature. Vol. 1A. Groningen: Egbert Forsten; 1999
- [2] Agniveśa C. Carakasamhitā [Caraka Samhitā text with English translation]. Vol. 1. Reprint ed. Varanasi: Chaukhambha Orientalia; 2010
- [3] Sastri VVS. Tridosha Theory. 5th ed. Kottakkal: Arya Vaidya Sala; 2009
- [4] Suśruta. Suśrutasamhitā [Suśruta-Samhitā with English translation of text and Ḍalhaṇa's commentary along with critical notes], Reprint. Sharma PV, editor. Vol. 1. Varanasi: Chaukhambha Visvabharati; 2004
- [5] Vāgbhaṭa. Aṣṭāṅgahrdaya [The Core of Octopartite Āyurveda, with the commentaries (Sarvāṅgasundarā) of Aruṇadatta and (Āyurvedarasāyana) of Hemādri], Reprint 9th. Parāḍakara H, editor. Varanasi: Chowkhamba Orientaria; 2005
- [6] Bharata A. Nāṭyaśāstra of Bharata Muni with the commmentary Abhinavabhāratī by Abhinavagupta, 2nd. Kavi MR, Sastri KSR, editors.. Vol. 1. Baroda: Oriental Institute; 1956
- [7] Bharata A. Nāṭyaśāstra of Bharata Muni with the commmentary Abhinavabhāratī by Abhinavagupta, 2nd Revised. Kavi MR, Pade JS, Kulkarni VM, Nandi TS, editors. Vol. 4. Baroda: Oriental Institute; 2006
- [8] Ratanjankar SN. Individual Notes and Specific Rasas. In: Aspects of Indian Music. Revised. New Delhi: Publication Division; 1970. p. 48–52
- [9] Agniveśa C, Ḍṛdhabala. Carakasamhitā [Caraka Samhitā text with English translation]. Vol. 2. Reprint ed. Varanasi: Chaukhambha Orientalia; 2008
- [10] Gangopadhyay A, Prasad JSR. Therapeutic elements of music in ancient India: a brief review in Br̥hatṭrayī. Indian Journal of History of Science. 2022;57(2):78–86. Available from: <https://link.springer.com/article/10.1007/s43539-022-00043-3> [Accessed: 2025-March-11]
- [11] Vāgbhaṭa (Vṛddha Vāgbhaṭa). Aṣṭāṅgasāṅgraha [Aṣṭāṅga Saṅgraha with Indu Commentory]. Panditrao DV, Pandeya A, editors. Vol. 1A. New Delhi: Central Council for Research in Ayurveda & Siddha; 1991
- [12] Suśruta, Nāgārjuna. Suśrutasamhitā [Suśruta-Samhitā with English translation of text and Ḍalhaṇa's commentary along with critical notes], Reprint. Sharma PV, editor. Vol. 3. Varanasi: Chaukhambha Visvabharati; 2005
- [13] Jīvaka V, Vātsya. Kāśyapasamhitā [The Kāśyapa Samhitā or Vṛddhajīvakiya Tantra by Vṛddha Jīvaka; revised by Vātsya]. Reprint ed. Varanasi: Chaukhambha Sanskrit Sansthan; 2010
- [14] Suśruta. Suśrutasamhitā [Suśruta-Samhitā with English translation of text and Ḍalhaṇa's commentary along with critical notes], Reprint. Sharma PV, editor. Vol. 2. Varanasi: Chaukhambha Visvabharati; 2005
- [15] Vāgbhaṭa (Vṛddha Vāgbhaṭa). Aṣṭāṅgasāṅgraha [Aṣṭāṅga Saṅgraha

- with Indu Commentary], Vol. 2. Panditrao DV, Pandeya A, editors. New Delhi: Central Council for Research in Ayurveda & Siddha; 1988
- [16] Bhāvamīśra. Bhāvaprakāśa, Reprint. Mīśra BŚ, editor. Vol. 2. Varanasi: Chaukhambha Sanskrit Bhawan; 2018
- [17] Cakrapāṇidatta. Carakasamhitā with the commentary of Cakrapāṇidatta. Hyderabad: National Institute of Indian Medical Heritage; 1941
- [18] Vāgbhaṭa (Vṛddha Vāgbhaṭa). Aṣṭāṅgasāṅgraha [Aṣṭāṅga Saṅgraha with Indu Commentary], Vol. 1B. Panditrao DV, Pandeya A, editors. New Delhi: Central Council for Research in Ayurveda & Siddha; 1998
- [19] Sodhala. Gadanigraha, Vol. 2. editor, Āchārya JT. Mumbai: Nirnaya Sagar Press; 1915
- [20] Bhela. Bhela Samhitā [The Bhela Samhita] [Internet]. Calcutta[now Kolkata]: University of Calcutta; 1921. Available from: <https://archive.org/details/in.ernet.dli.2015.284169> [Accessed: 2025-July-10]
- [21] Meulenbeld GJ. A History of Indian Medical Literature. Vol. 2A. Groningen: Egbert Forsten; 2000
- [22] Sastri SVR, editor. Ānandakanda [A treatise on mercury and other heavy metals - and their uses in medicine and alchemy (atomic phenomena)], Vol. 1. Tanjore: TMSSM Library Tanjore; 1952
- [23] Śālinātha. Rasamañjarī [Internet]. Kehamarāja Śrīkrṣṇadāsa; 1905. Available from: <https://archive.org/details/in.ernet.dli.2015.406254/mode/2up?view=theater> [Accessed: 2025-July-11]
- [24] Varma LP. History of Psychiatry in India and Pakistan. Indian Journal of Psychiatry. 1953;4(1):26–53. Available from: https://journals.lww.com/indianjpsychiatry/citation/1953/04010/history_of_psychiatry_in_india_and_pakistan.3.aspx [Accessed: 2025-March-11]
- [25] Sharma S, Verma LP. Psychiatry The State of the Art. In: Pichot P, Berner P, Wolf R, Thau K, editors. Vol. 8. 1st edition. Boston, MA: Springer; 1985. p. 45–52. DOI: 10.1007/978-1-4757-1853-9_6
- [26] Pal J. Sangit Chikitsa [Internet]. Lahore: J. Pal; 1938. Available from: <https://indianculture.gov.in/ebooks/sangaita-caikaitasaa-sangit-chikitsa> [Accessed: 2022-February-20]
- [27] Chatterjee S Music Therapy. Times of India [Internet]. 2002 June 1. Available from: <https://timesofindia.indiatimes.com/music-therapy/articleshow/11649538.cms> [Accessed: 2025-July-10]
- [28] TNN. Music therapy fails to stop suicide. Times of India [Internet]. 2001 August 30 . Available from: <https://timesofindia.indiatimes.com/city/kolkata/music-therapy-fails-to-stop-suicide/articleshow/247484768.cms> [Accessed: 2025-July-10]
- [29] PTI. Music therapy reduces anxiety for mental patients. Times of India [Internet]. 2002 October 22. Available from: <https://timesofindia.indiatimes.com/science/music-therapy-reduces-anxiety-for-mental-patients/articleshow/25941676.cms> [Accessed: 2025-July-10]
- [30] Dixit R Expert says Music can help body fight diseases. Times of India [Internet]. 2002 September 15.

Available from: <https://timesofindia.indiatimes.com/city/mumbai/expert-says-music-can-help-body-fight-diseases/articleshow/22296951.cms> [Accessed: 2025-July-10]

[31] Jaisinghani B Music ‘therapy’ to cure you. Times of India [Internet]. 2005 June 26. Available from: <https://timesofindia.indiatimes.com/india/music-therapy-to-cure-you/articleshow/1152826.cms> [Accessed: 2025-July-10]

[32] Sahuliyar AS Rinpas prescribes raga therapy - City lawyer-singer proves connect between music and mental health. The Telegraph [Internet]. 2014 March 6. Available from: <https://www.telegraphindia.com/jharkhand/rinpas-prescribes-raga-therapy-city-lawyer-singer-proves-connect-between-music-and-mental-health/cid/203707> [Accessed: 2025-July-11]

[33] Gupta U, Gupta BS. Psychophysiological responsivity to Indian instrumental music. *Psychology of Music*. 2005;33(4):363–372. DOI: 10.1177/0305735605056144

[34] kumar TS, Muthuraman M, Krishnakumar R. Effect of the Raga Ananda Bhairavi in Post Operative Pain Relief Management. *Indian Journal of Surgery*. 2014;76(5):363–370. Available from: <https://link.springer.com/article/10.1007/s12262-012-0705-3> [Accessed: 2025-March-11]

[35] Balan R, Bavdekar SB, Jadhav S. Can Indian classical instrumental music reduce pain felt during venepuncture? *The Indian Journal of Pediatrics*. 2009;76(5):469–473. Available from: <https://link.springer.com/article/10.1007/s12098-009-0089-y> [Accessed: 2025-March-11]

[36] Deshmukh AD, Sarvaiya AA, Seethalakshmi R, Nayak AS. Effect of Indian classical music on quality of sleep in depressed patients: A randomized controlled trial. *Nordic Journal of Music Therapy*. 2009;18(1):70–78. Available from: <https://www.tandfonline.com/doi/abs/10.1080/08098130802697269> [Accessed: 2025-March-11]

[37] Mathur A, Vijayakumar SH, Chakrabarti B, Singh NC. Emotional responses to Hindustani raga music: The role of musical structure. *Frontiers in Psychology*. 2015;6(APR). DOI: 10.3389/fpsyg.2015.00513

[38] Kar SK, Ganguly T, Saha Roy S, Goswami A. Effect of Indian Classical Music (Raga Therapy) on Fentanyl, Vecuronium, Propofol Requirements and Cortisol levels in Cardiopulmonary Bypass. *Journal of Anesthesia and Critical Care*. 2015;2(2)

[39] Kunikullaya KU, Goturu J, Muradi V, Hukkeri PA, Kunnavil R, Doreswamy V, et al. Music versus lifestyle on the autonomic nervous system of prehypertensives and hypertensives—a randomized control trial. *Complementary Therapies in Medicine*. 2015;23(5):733–740. Available from: <https://www.sciencedirect.com/science/article/abs/pii/S0965229915001272?via%3Dihub> [Accessed: 2025-March-11]

[40] Ahuja S, Gupta RK, Damodharan D, Philip M, Venkatasubramanian G, Hegde S. Effect of music listening on P300 event-related potential in patients with schizophrenia: A pilot study. *Schizophrenia Research*. 2020;216: 85–96. Available from: <https://www.sciencedirect.com/science/article/abs/pii/S0920996419305961?via%3Dihub> [Accessed: 2025-March-11]

- [41] Govindan R, Kommu JVS, Bhaskarapillai B. The Effectiveness of Nurses Implemented Music Add-on Therapy in Children with Behavioral Problems. *Indian Journal of Psychological Medicine*. 2020;**42**(3):274–280. Available from: <https://pubmed.ncbi.nlm.nih.gov/32612333/> [Accessed: 2025-March-11]
- [42] Sharma S, Sasidharan A, Marigowda V, Vijay M, Sharma S, Mukundan CS, et al. Indian classical music with incremental variation in tempo and octave promotes better anxiety reduction and controlled mind wandering—A randomised controlled EEG study. *EXPLORE*. 2021;**17**(2):115–121. Available from: <https://www.sciencedirect.com/science/article/abs/pii/S155083072030094X?via%3Dihub> [Accessed: 2025-March-11]
- [43] Babel S, Baral S, Srivastava A, Babel S, Baral S, Srivastava A. Impact of Listening to Indian Classical Music, or Rāgas, on the Electroencephalogram: A Meta-Analysis. *Cureus*. 2025;**15**(11). Available from: <https://www.cureus.com/articles/202935-impact-of-listening-to-indian-classical-music-or-ragas-on-the-electroencephalogram-a-meta-analysis> [Accessed: 2025-March-11]
- [44] Ghosh A, Singh S, Monisha S, Jagtap T, Issac TG. Music and the aging brain – Exploring the role of long-term Carnatic music training on cognition and gray matter volumes. *Journal of Neurosciences in Rural Practice*. 2024;**15**(2):327. Available from: <https://ruralneuropsychiatry.com/music-and-the-aging-brain-exploring-the-role-of-long-term-carnatic-music-training-on-cognition-and-gray-matter-volumes/> [Accessed: 2025-March-11]
- [45] Chand K, Chandra S, Dutt V. Raga Bhairavi in virtual reality reduces stress-related psychophysiological markers. *Scientific Reports*. 2024;**14**(1):24816. Available from: <https://www.nature.com/articles/s41598-024-74932-1> [Accessed: 2025-March-11]
- [46] Jethva DD, Patel BC, Sundar S, Patel JB, Vora HH, Sanghavi PR. Harmonising Hope: Impact of Music Therapy on Cancer Pain and Palliative Care. *Indian Journal of Palliative Care*. 2024;**31**(1):21. Available from: <https://jpalliativecare.com/harmonising-hope-impact-of-music-therapy-on-cancer-pain-and-palliative-care/> [Accessed: 2025-March-11]
- [47] Sukumar M, Thanneeru SK, Sutar RF, Sahu A, Agrawal A. Effect of Music Therapy on Quality of Life in Geriatric Population: A Systematic Review and Meta-Analysis. *Indian Journal of Palliative Care*. 2024;**30**(4):296. Available from: <https://jpalliativecare.com/effect-of-music-therapy-on-quality-of-life-in-geriatric-population-a-systematic-review-and-meta-analysis/> [Accessed: 2025-March-11]
- [48] Rajendran T, Patil A, Mohanta SK, Krishnadasa SN, Natarajan V. Classical Indian music for managing anxiety and pain among patients in a hospital setting: a systematic review and meta-analysis of randomized controlled trials. *J Public Health (Oxf)* [Internet]. 2025 May 7. Available from: <https://pubmed.ncbi.nlm.nih.gov/40331588/> [Accessed: 2025-July-10]
- [49] Kunikullaya UK, Pranjic M, Rigby A, Pallás-Ferrer I, Anand H, Kunnavil R, et al. The molecular basis of music-induced neuroplasticity in humans: A systematic review. *Neuroscience & Biobehavioral Reviews*. 2025;**175**:106219. Available from: <https://pubmed.ncbi.nlm.nih.gov/40412457/> [Accessed: 2025-March-11]

[50] Chawla T, Nangia T, Srivastava M, Kalra G, Langpoklakpam C. Comparative evaluation of effectiveness among various sensory distractors for reducing dental anxiety in pediatric population aged 6-9 years - A randomized controlled trial. *Journal of Indian Society of Pedodontics and Preventive Dentistry*. 2025;43(1):104–110. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/40159612> [Accessed: 2025-March-11]

[51] Gangopadhyay A. *Rasa Analysis of Sthūla-Pāśyantī id est Non-Lyrical Non-Percussive Indian Pure Music: A Pragmatic Approach in Therapeutic Context*. Hyderabad: University of Hyderabad; 2022

[52] Gangopadhyay A, Prasad JSR, Shamanna BR. Evaluation process of Rasa in the context of nonlyrical nonpercussive Indian pure music for possible application as music therapy. *Journal of Applied Consciousness Studies*. 2022;10(2):113. Available from: <https://www.jacsonline.in/article.asp?> [Accessed: 2025-March-11]

[53] Sinha J. *Rāga, roga va rogī : saṅgīta cikitsā*. 1st ed. New Delhi: Omega Publications; 2015

[54] McClellan R. *The Healing Forces of Music*. San Jose, New York: to Excel; 2000. p. 69

[55] Nagarajan K. *An Introduction to Indian Music Therapy*. Kindle New Delhi: Authorvine Publishers ed. 2021

[56] Śārṅgadharma. *Śārṅgadharma Saṃhitā*. Reprint ed. Varanasi: Chaukhamba Orientalia; 2010

Music and Psychoactive Substances: Cultural Practices, Adolescent Development and Music Therapy in Addiction Treatment

Jörg Fachner, Filippo Pasqualitto and Clemens Maidhof

Abstract

Music and psychoactive substances play a key role in forming identity during adolescence. Grasping the historical, neurobiological, and clinical aspects of this relationship is crucial for exploring how music therapy (MT) can support recovery from substance use disorder (SUD). In this chapter, the discussion begins by framing adolescent risk-taking, including substance experimentation, as a way of testing the self within a social context and its boundaries. Psychoactive substances are examined not merely as objects of pathology but as tools that historically co-evolved with music and popular culture, all playing a role in normalizing and contextualizing their use for young people. By analyzing what's "behind the scenes," the chapter examines the neurobiological isomorphism between drug-induced euphoria and the pleasure elicited by listening to music, both of which strongly involve the brain's dopaminergic reward system. This overlap provides the theoretical basis for using MT as a complementary intervention in the treatment of SUD. All in all, the chapter presents evidence of how MT modulates neural activity and contributes to symptom reduction, emotional regulation, and enhanced motivation, thereby supporting recovery in SUD treatment.

Keywords: music therapy (MT), psychoactive substances, substance use disorder (SUD), adolescence, rites of passage, craving, dopaminergic system, neurobiology, autobiographical memory, popular culture

1. Introduction

The music that we listen to when we become adults is linked to strong emotions and memories. Music therapy (MT) makes use of these autobiographical memories, for instance, when working with people living with dementia, allowing them to revive and remember the "good old days" of falling in love and discovering the world.

On the contrary, some drug-related memories cued by music may send a recovering addict back to the party nights when the music worked so well with the MDMA (ecstasy) at that rave party, or to weekends spent drinking and having fun with old friends. Music stirs up memories and gives us access to the strong emotions we experienced when we were young and on the way to becoming adults. James Dean, as well as Marco Leonardi playing the teenage boy Salvatore in Giuseppe Tornatore's fabulous movie *Cinema Paradiso*, allows us to identify with all the contradictions, dreams, and boredom an adolescent may go through and struggle with. For over 50 years, we have observed this "coming-of-age" genre of movies and books in popular culture, telling narratives about individuals and peers who are "becoming of age". The protagonists try to navigate through this uniquely wonderful and also difficult time of our lives, a time that lets us taste and remember a wealth of emotions and first-hand experiences during our adolescence.

In the following two sections of this chapter, we take a closer look at selected periods (1960s to millennium years) and aspects of popular culture, and how normal it became for adolescents to associate drugs and music together; furthermore, how the set and setting of taking psychedelic drugs became part of popular culture. After that somewhat historical focus, we will discuss how MT can help to overcome addiction and state-dependent recall of drug memories related to drug effects experienced with music.

2. No risk, no fun... adolescence, rites of passage, and popular culture

"Sex and Drugs and Rock'n Roll, that's all my brain and body needs..." With these lines, the English musician Ian Dury paraphrased the rituals of adolescence and its transition to adulthood within modern youth cultures, particularly during the "1960s, 1970s, and 1980s" of the last century. Liminal experiences have been the subject of institutionalized *rites of passage* since time immemorial [1]. In anthropology, liminality (from Latin *limen* – a "threshold") is the quality of ambiguity or disorientation that occurs in the middle stage of a rite of passage (during the transition from childhood to adulthood), when participants no longer hold their pre-ritual status but have not yet begun the transition to the status they will hold when the rite is complete. In other words, it describes the experiences of an adolescent during events like parties, initiation rituals, and in peer groups, but also in general. In short, the state of liminality becomes apparent in the experience of not yet being an adult nor being a child anymore. Young people test physical, social, and psychological boundaries to form their own social and personal identity, to differentiate themselves from others. Closeness and distance to like-minded and different-minded people are experienced, and romantic or intimate contact is sought. In the 1990s and the first decade of the 2000s, the "party" represented such activities in contemporary youth culture, where young people in their peer group were looking for themselves and for the other [2, 3]. From a relational perspective, such experiences may be seen as expressions of social identity, that is, the young adult is not isolated but shaped through interactions and shared contexts. Identity emerges here as a dynamic configuration, co-constructed in the interplay with others and with the surrounding environment [4]. Within this framework, alcohol often serves to reduce social inhibitions, facilitating contact and self-expression; likewise, it can be "chic," "hip," or "cool" to be seen as experienced and already grown up by

trying other psychoactive drugs. This is illustrated by an anecdotal statement from an American teenager who expressed his opinion on MDMA drug experiences in connection with rave parties during the MTV program *I'm on Ecstasy*, which was broadcast in Germany on October 4, 2005:

It's unfortunate when people can't handle it [MDMA], but they can recognize it [the experience of MDMA] as almost a rite of passage, that's how you discover your boundaries. I'm sure almost everybody on this continent had a couple of nights where staring into a toilet they swore off alcohol for the rest of their lives and I think it's about the same thing: once in a while, and particularly when you're young and you're growing up, you have to cross boundaries in order to find out where that boundary is [5].

Although this quote is from 2005, including the respective drug cultures and habits, it points toward an anthropological constant, as “adolescents testing boundaries” seems to be a timeless observation, and risk-taking may or may not include taking drugs. Insofar as risk-taking, peer acceptance, and a proclivity for sensation-seeking can be seen as typical adolescent motivational behaviors, it follows that adolescents will be more liable to engage in experimental substance use in the spirit of adventure and exploration and as a rite of passage *with peers*. Research suggests that peer group patterns indeed tend to influence social substance use in the early stages, with subsequent incremental stages of instrumental, habitual, and compulsive individual substance use — then denoting a more committed move from mainstream peer culture into a substance-using subculture [6]. The impacts of this shift are multiple, including moves toward a drug-using identity and a change in focus from peer motivations to personal impetuses for substance use, such as altering mood or seeking emotional relief. In social terms, progressive identification with a substance-using subculture carries risks of exclusion and alienation from mainstream culture, whilst in functional terms, increased frequency of substance use also correlates with reduced possibilities of a return to normal functioning.

Every decade of debating drugs has brought forward different arguments about why certain drug-taking behavior may need to be labeled as deviant and pathological or considered normal [7]. Some authors discuss psychotropic substance-seeking as a human evolutionary adaptation behavior [8], suggesting that external stimulation of the brain's reward system with plant toxins was important for stimulating its growth and separation from other species [9, 10]. Drugs are tools that humans have developed or discovered for various purposes, and, like all tools, they can be used or misused [11]. Using and misusing drugs during adolescence is an old phenomenon and seems to be one of the anthropological constants of mankind, like eating, drinking, sex, and war [10]. It has always been a social problem that the misuse of drugs influences some people more than others.

As Parker, Williams, and Aldridge pointed out, popular culture and its music play an important role for young people and adults in normalizing the recreational use of illegal drugs, because drug consumption and its meaning are integrated into a symbolically and historically mediated frame of reference for pop cultural practice [12]. In their research, they pay particular attention to the club culture and conclude that the recreational use of illegal drugs has become normalized in the sense of accepted social reality, because drugs have become culturally accommodated in society through popular music culture. Cannabis or “ecstasy”

(MDMA) no longer only occur in previously pathologized fringe groups but have become an accepted part of youth culture, and nowadays, consumption in several countries is legalized.

The neurobiologist Walter Freeman discusses how music and dance have been related to the cultural evolution of human behavior and social bonds. He sees a connection in the traditional knowledge about altered states of consciousness, which are triggered by chemical and behavioral forms of induction [13]. The trance states achieved in this way serve, on the one hand, to specifically break through learned habits and belief systems about reality, but, on the other hand, also to increase the willingness to absorb new information. Such deliberate changes may have led to the formation of “initiated” groups and trust in the transmission of important insights. Musical skills of those enabling insights for the recipients — via singing, drumming, or playing an instrument — seem to be important for the efficient transmission of knowledge. De Rios and Grob emphasize the increased suggestibility caused by drugs, which is specifically used by shamans for adolescents in ritual transition ceremonies to induce cross-modal intensification and regression processes that lead to more vivid associations and visions linked to the music [14].

2.1 Initiation rites, psychedelic drugs, mind-altering experiences, and partying

Ralph Metzner, an American psychologist, explains that in traditional systems of consciousness modification, such as shamanism, alchemy, or the Indian sadhus, psychoactive substances were always viewed as a sacrament, as a means of intensifying spiritual experiences within the respective culture. The white colonialists, who did not understand the religious practices and dismissed them as pagan, adopted the drugs only as a recreational product, but not the ceremonies in which they were used [15]. The ceremonies were about experiencing and treating the unity of religion and nature, of healing arts, and therapeutic effects on body, soul, and spirit in the community. After the ceremonies, the substances were also given out at the final celebrations and were, therefore, still part of the entire ceremony. Many traditions sought to combine mind-altering drugs, ritual spaces of action, and thus a combination of external and internal experience under the guidance of a teacher in a ritual context.

Music serves as a mental means of transportation that guides the “travelers” through the ritual context [16, 17]. Drugs affect the perception of time and metric frames, that is, they change the functioning of the inner clock regarding scaling and the internal mapping of auditory metrical units, resulting in a different experience of music, as if time slows down or speeds up, also resulting in a different perception of intensity levels and gestalt-background relationships [18]. Such interaction between music and drugs can be used for ritual-immanent purposes [16, 18–20]. In the following period, the post-ceremonial “party” aspects described by Metzner above were secularized by Western civilization in the profane use of drugs as stimulants. This shift in understanding became popular during the “psychedelic revolution” of the 1960s [21]. The rising subculture of hippies transferred core elements of psychedelic therapy into cultural symbols, and musicians went on stage to create the sound as an acoustic surrounding for the “pot- and acidheads” on their “trip” into inner and outer space. Here, music was mainly used as a guide during the hallucinogenic state to keep the acidheads “on track” [22].

3. Mind-altering drugs, the setting, and music in therapy

The psychiatrist Osmond introduced the term “psychedelic,” which means “expanding psychic reality.” Leary, Sandison, Grof, Leuner, Eisner, and other founders of psychedelic therapy utilizing LSD, psilocybin, mescaline, etc., discovered that *the setting* – realized via works of art, fairy tales, narratives, music, ritual, and selected interior – had a profound capacity to stimulate the unconscious, provoke imagery, and stir up emotions under the influence of psychedelics in guided and secure surroundings [21, 23–26]. Psychedelics evoke peak experiences in the right people under the right circumstances, as Abraham Maslow [27] describes in his book on values and peak experiences. This also points to an important concept in psychedelic *therapy*: optimizing the environment for the drug experience in order to facilitate access to unconscious material for psychotherapeutic purposes. Here, music [28], visuals, and art play an important role [25] in supporting peak experiences [29]. This aspect of setting has gained significant interest in both former and recent psychedelic research [30]. In general, psychedelic drugs are preferably consumed in a setting suitable for the interaction between the consumer and environmental cues to temporarily expand psychic reality [31]. Recent research on psilocybin and LSD has shown that vividness of imagery and emotion, along with music, are quite strong [32–34]. It is the psychedelic agent – mostly LSD, psilocybin, mescaline, MDMA, or even cannabis [35] – that intensifies the experience, changes the way we perceive the world around us, and weakens defense mechanisms [36, 37].

Moreover, the emergence of important imagery along with music yields strong imagery patterns with a noetic structure that may well yield a vast amount of important emerging issues that can be used for therapy [38]. Imagery enhancement under the influence of substances is indicated by increasing functional connectivity between the bilateral para-hippocampus (PHC) and the visual cortex (VC). Furthermore, for the interaction between music and LSD, an information flow from the PHC to the VC was observed. The latter result correlated positively with ratings of enhanced eyes-closed visual imagery, including imagery of an autobiographical nature [33].

3.1 Syncing context and mind–body experiences with music

While in the early days of hippie youth culture (1962–1965), the collective “drop-out” through shared psychedelic consumption was still a social experiment in a relatively protected and manageable ritual context [39], from 1966 onwards, the influx of pharmaco-adventurers and naïve spectators was no longer integrated. On October 6, 1967 (the first anniversary of the LSD ban in the USA), the hippies themselves publicly burned the popular “media hippie” in a performance on Haight-Ashbury Street in San Francisco, symbolically burying the hippie culture [40]. Through the popularization and marketing of youth culture, the drug-related ideas were diluted into clichés that ignored the connection between the situation and the meaning of drug use in the interests of expanding consciousness [41].

The era of large and small pop festivals began, where the consumption of legal and illegal drugs was part of the festival and party culture. Elements of the setting (muted light, relaxed atmosphere, art objects, music, comfortable furniture) of the psychedelic therapy-oriented procedures were now incorporated into the public

productions and were further developed artistically (see [25, 39]). Illumination, interior, and furniture were inspired by oriental tea rooms where opium was consumed [42]. The acid tests in San Francisco and the surrounding area, however, made these more private productions compatible with the masses.

In his book *The Electric Kool-Aid Acid Test*, Tom Wolfe [39] described the story of Ken Kesey and his traveling hippie commune – the Merry Pranksters – journeying in brightly painted buses across America. The Merry Pranksters experimented with light installations, light shows, strobes, large-format images with mystical-seeming motifs, sound objects, drama, artificial fog, and more. The aim was always to create synchronicity between external events and internal processes of thoughts, perception, and meaning, as Wolfe lets Ken Kesey explain as “The mysteries of the synch!” The individually significant synchronicity of external events and internal processes, which can be experienced in mystical ways, was specifically evoked in the multimedia contexts of parties, festivals, performances, and concerts. This element of targeted contextualization of the experience can also be found in the multimedia environments of psychedelic music performances of the 1960s, the chill-out rooms of rave culture, the video projections of Pink Floyd and their epigones, and, with light shows, visuals, and artificial fog, as a pop culture component of the ambient musical performance practice [22, 43].

The idea went beyond what would later be known as mixed-media entertainment, now a standard practice in ‘psychedelic discotheques’ and so forth [39]

The psychiatrist Ronald D. Laing, as a representative of the “Personal Growth Movement,” published a book entitled *The Politics of Experience* in the 1960s. In this book, he presented the value of experience for personal development and life perspective. For him, as for Maslow [27], evidence and liminal experiences were those deep personal experiences that gave people direction on their path and served as a reference for one’s own emotional and intellectual evaluations [44]. The goal of the “Growth Movement” was individuality, self-awareness, personal development, and also a look and a journey inward.

The question “Are you experienced?” could also be found on the cover of Jimi Hendrix’s debut album in 1967, which implicitly referred to this mental milieu [45]. For “progressive” young adults, having an experience meant more than cognizing, reproducing, or repeating learned content, because experience is about “trying over studying”; it’s about your own experience [46]. “We want the world, and we want it now,” shouted Jim Morrison of the Doors into the audience, taking aim at the hypocrisy of forever postponing experience through “straight” career planning, vacation policies, and retirement planning. “I want to experience the world now,” said Morrison, “and not when I retire.” This was also the opinion of all those who chose a lifestyle that focused on the present, on the “here and now” of experience at the center of their life planning.

3.2 Psychedelics and creativity

Masters and Houston [47], as well as Krippner [48], put a special focus on the effects of psychedelic drugs on creativity, particularly regarding the creativity of visual artists. Psychedelics produce an inner flood of images that the artists wanted to project onto their canvas. The authors also mentioned famous pieces from the

Beatles (Lucy in the Sky with Diamonds), The Byrds (Eight Miles High), or Grateful Dead (Dark Star), and other musicians who expressed their drug-induced experiences through a combination of text and music. Taeger [49] proposed that the combination of music and drugs in pop culture was not just born from a wish to make cash, but rather that the music and texts of pop artists revealed a spiritual search or longing. He underpinned this by exploring the interpersonal correlations between psychedelics and religious-mystical aspects in the counterculture of the 1970s. He found many indications of spiritual experiences and attitudes of musicians and artists on Long Player covers and in texts by pop artists of the 1960s and 1970s. Many images and symbols (for example, the “wise man on the Mountain,” a “stairway to heaven,” pyramids, chthonic animals such as snakes and scorpions) described in the texts revealed a mystical experience produced by psychedelics, as already described by C.G. Jung in his theory of archetypes, thus supporting the assumption that psychedelics provide access to the collective subconscious.

In the book *LSD, Spirituality and the Creative Process*, Marlene Dobkin de Rios presented work based on LSD research conducted by the psychiatrist Oscar Janiger from 1954 to 1962, during which more than 930 people in Los Angeles were given Sandoz LSD experimentally [50]. Included in the sample were artists and musicians. One of them described his experiences while listening to his preferred pieces of classical music:

My reaction to these pieces began with the conventional response but gradually took on a new character. It was as though the remaining ecstasy that flowed through me has washed away my patience with the exterior posturing of music. I felt that I saw directly into its heart and was interested only in what the music was really saying, remaining totally indifferent to how I was dressed. . . . The visual hallucinations were one of the more entertaining features of the afternoon. . . . I could not for a time distinguish between sight and sound. Later Mozart's melodic line was filling the room. Later woodwind harmonies released ethereal glowing purples and pinks in shafts of radiant light which streamed out from a picture in precise synchronization with the music. I felt that moment of incredible exaltation. I am, in this very instant, free from every petty negative emotion. [50]

Early research on music and drugs appeared as basic research on music perception, music production, and its therapeutic use [51, 52]. For instance, Weber's research in the 1960s focused on the use of psilocybin, a fungus with psychoactive ingredients [53, 54]. In Weber's research, drug-induced altered music perception was discussed as a model of functional regression to lower levels of cognitive development by promoting a “unity of the senses” experience. Discussing Piaget's steps of cognitive development, during which a diversification of sensory function and entities indicates progress in cognitive development, the psychedelic drug experience would bring back a former processing level that promotes synesthetic and coenesthetic experiences. In short, music would be experienced more holistically, in the way children would experience it.

3.3 Cannabis, cognition, and the experience of music

Since ancient times, cannabis has been used by Morocco's fishermen during the night as it increases night vision and visual depth perception, which was helpful for

the night shift [35, 55]. Also, in Tart's [56] book about 151 cannabis users, a complete chapter on their experiences with visual perception deals with temporarily increased vision and enhanced imagery when eyes are closed, especially when listening to classical, jazz, or pop music.

Tetrahydrocannabinol (THC), the psychoactive component of cannabis, appears to reduce activation in regions typically engaged under placebo, particularly the extrastriate visual cortex [57] – a finding that seems counterintuitive given reports of *enhanced* visual abilities. Similarly, fMRI studies indicate that cannabis decreases overall brain activity during music listening, dampening emotion and reward processing. However, increased functional connectivity in the auditory cortex correlates with pleasure ratings, suggesting a *complex* effect [58]. This is again in contrast to what cannabis users experience, reporting in recent research “significantly greater hearing sensitivity and state absorption in music” with altered cognitive processing and reframing of music experiences, emotional openness, and processes of immersion [59]. Importantly, reduced hemodynamic activation does not necessarily imply diminished information processing; rather, it may reflect altered mechanisms, such as shifts in selective attention to auditory input, evidenced by a reduced P300 electrophysiological response [60] and by changes in alpha and theta power topographies pre- and post-THC during music listening [61].

As testimonies from members of The Beatles or the band Fleetwood Mac confirm, listening to the sound mix of a newly recorded piece under the influence of cannabis is sometimes considered by some musicians to be a review strategy for a newly recorded song [62]. Descriptions of synesthetic effects by Baudelaire [63] and Tart [56], the THC-weakened filter function leading to changed visual depth perception [35], more creative interpretations of Rorschach's visual test patterns, and a transition to a divergent thinking style under the influence of THC — all point toward a change in the individual auditory perception style [18].

Globus et al. [64] demonstrated that cannabis influences subjective auditory processing by altering memory of loudness and the underlying phenomenon of state-dependent recall, where certain experiences occur only under intoxication, which will also be critical for understanding drug-related *memory* processes [65]. The next section will explore its implications regarding addiction treatment and music-evoked autobiographical memory.

Another aspect is time perception, which can also be significantly affected by cannabis consumption. Meck describes this effect as a reciprocal interaction between subjectively slowed time and an accelerated “inner clock” [66], accompanied by changes in cortical and cerebellar blood flow [67, 68]. The cerebellum, a key region for the temporal sequencing of motor activity, appears to play an important role in this process.

It seems as if cannabis temporarily allows one to focus more deeply on the acoustic sound space and the time structure of the music. Cannabis seems to have less impact on the perception of music *per se*, but rather acts similarly to an enhancer, exciter, or compressor from studio technology by altering the perception of musical time-space, changing the metric frame of reference for time perception, as well as changing the spaciousness of the acoustic events [18, 69]. In this sense, the impact of cannabis consumption on the perception of music can be described as a psychoacoustic quality enhancement, allowing one to “see” into “the space between the notes” [70].

4. Music therapy, craving, and drug memory

This section shifts the focus to the clinical realm. Here, we will examine how MT is applied to treat addiction.

As a healthcare profession, music therapists work with clients exposed to commercially-driven concepts of the self, conveyed by popular music culture. As discussed above, we know that adolescents have to pass through culturally mediated rites of passage [71, 72]. They must learn to handle the flowing and changing state of liminality regarding their being, their senses, and their mind, and they must integrate their first extreme bursts of emotional, sensual, and sexual experiences. Some juveniles seek extreme and risky sensations and experiences to get out of the domestic sphere. Limitations are set by adult society by checking the boundaries at parties, during binge drinking, heavy drug consumption, and other somewhat extreme lifestyle activities, which is a subject often discussed on MTV and the like [3, 22, 73]. Music therapists like Dijkstra and Hakvoort [74], and Punkanen [75], have shown what is left for some “trainspotters” after such experiences, and what MT can do to help restore, find, and maintain their identities lost and found in the “merry-go-round” of ontological uncertainty experienced in postmodern Western societies. Occasionally, some clients are even haunted by traumatic experiences, as Punkanen [76] outlined. In such cases, MT may provide a non-verbal channel for processing trauma by facilitating sensorimotor and emotional regulation. Other clients must be guided through unstable personality states on their biographic journey. But as time and therapy move on, persons with an addiction may learn not to obey the chants of the sirens, as Horesh [77] allegorizes, referring to the tale of Odysseus and his evoked hunger for relief from the sweet call of the sirens once upon a time in Greek mythology.

Even when some researchers are convinced that disorders are based on personality traits, as the discussion on premorbidity, genetic disposition, or acquired brain disease reflects [77, 78], the question of whether we are dealing in treatment with personality states or traits remains an open discussion. Recent research points toward the connection between brain development during adolescence and how we learn to interact with our social brain and regulate affect [79]. Music therapy addresses this developmental period by promoting dyadic, music-based experiences and connections in non-verbal social interaction between the therapist and the client. This active engagement enhances the capacity for affect regulation and emotional attunement. It also strengthens the neural networks associated with empathy and social cognition.

4.1 Processing strong emotions in the brain

Music can provide a means for promoting positive mood states, including euphoria [80, 81], and, with some caution on the choice of the musical material, it may help against the risk of relapse associated with negative mood states [82].

The concept of affect regulation is central to adolescence (and MT), denoting the capacity to control and master intense emotions “in adaptive ways in order to achieve goals, to act within social norms, and in ways that support rather than interfere with decision-making” [83]. Researchers further link this to the impact of situational factors on adolescents’ responsible decision-making by distinguishing between “hot cognition” (having to think when highly aroused and experiencing

intense emotions) and “cold cognition” (thinking when in a state of low arousal and feeling calm) [84, 85]. In other words, the same adolescent may make vastly different decisions based on either impaired (in the case of “hot cognition”) or rational (cold cognition) cognitive functioning, irrespective of their understanding of consequences. Most notably, peer influence, coupled with an intense desire for acceptance and hypersensitivity to exclusion, may play a role in adolescent decision-making and guide potentially risky behaviors founded on the fearful avoidance of actions that could lead to alienation from peers [6, 84].

Neuroscientific research on adolescent brain development provides a context to examine the impact of adverse childhood experiences on adolescent development. This neurological context is significant in that evidence clearly points to adolescence not only as a period during which specific changes in the brain provide some explanation for impulsive, risk-taking, and sensation-seeking behaviors, but also as a period of neuroplasticity [84], that is, with implications for well-timed support and interventions during the phase of adolescent identity exploration.

Psychoactive substances, as well as physical stimuli or behavioral patterns perceived as pleasurable, have reinforcing properties that can be ascribed to the neuronal reward system [85], which primarily involves structures like the ventral tegmental area (VTA), the nucleus accumbens (NAc) (part of the ventral striatum), and the prefrontal cortex. However, on the micro-level, the complex systems of neurotransmitters interacting with the nervous system to arouse euphoria are not yet satisfactorily understood [86]. The reward system theory postulates that reaching pleasurable or euphoric states is the major goal of drug users and persons with addictions. Psychoactive drugs, as well as specific behavioral patterns, can activate the reward system and are therefore used to close the cycle of motivation–search–fulfillment [87].

In pleasure, desire, and lust, the dopamine system is active in both a tonic and a phasic way. First, the tonic component of dopamine release in the brain’s prefrontal cortex regulates the readiness to react to stimuli. Second, there is an increase in dopamine release (phasic) when meaningful objects are in the focus of attention; the higher the personal meaning and valence of the object in focus, the more dopamine is released. Dopamine’s action thus supports feeling and distinguishing what, in this particular moment, is of personal importance and which current selection of perceptual and emotional content is consciously processed. All drugs of abuse increase dopamine release and therefore affect these two ways of dopaminergic functioning, altering the experience of pleasure [88].

4.2 Emotion and memory in music

The powerful effect of the reward system on behavior and emotion is not limited to substances; in fact, *music* shares notable neurobiological parallels with drug-induced euphoria. Research has shown that euphoric emotions elicited by beloved music reveal similar activation patterns to those induced by cocaine in the brain. Since Blood and Zatorre, in 2001, pointed out the underlying similarities between the pleasing effects of preferred music and euphoric drug action [89], there have been several investigations from the music sciences regarding corresponding brain mechanisms. For example, it has been noted that lesions in certain structures of the brain (i.e., the mesocorticolimbic network) impair the emotional response to music [90] and that dopaminergic transmission – manipulated with antagonists or

precursors in healthy participants – is *causally* involved in rewarding experiences elicited by music [91].

For music therapists working with addicted clients, this research stream suggests that enjoying musical activities in therapy would induce comparable euphoric neurotransmitter activities that may help to overwrite and reframe the craving-related pathways reacting solely to emotions raised by the drug action [65]. This *principle of isomorphism*, whereby the brain's reward system is either activated through music listening or through substance use, may reduce medication prescribed for addiction, depression, or anxiety. In other words, the parallel activation of pleasure mechanisms through both drug-induced and music-induced reward may support the hypothesis that therapeutic music engagement can partially substitute or complement pharmacological interventions [65].

Another approach that has been successfully employed in addiction treatment is through the altering of consciousness. For example, in drum circles, consciousness is altered during extended drumming, inducing the pleasure of playing and trancing in a group, shifting attention away from the drug and into experiencing the cohesive and mutual activity of musicking [92]. Metzner described the relationship between addiction and dependence as being related to the dynamics of consciousness in general [93]. In his model, addiction is represented by an extremely narrow, focused state within waking consciousness – but, on the other hand, taking psychedelic drugs occasionally results in broadened and, thus, altered states of consciousness. In these altered states, which are time-limited, the patterns of thought, feeling, and mood are changed in regard to the ordinary baseline conditions.

How the different drug actions are related to music perception cannot easily be generalized, and research into this field of music sciences is still in its infancy. A systematic investigation into the meaning of music for drug addicts has been carried out by Horesh [94, 95]. Conversely, the effects of drugs on music perception and production [18, 43] can also enrich the understanding of music. Keith Richards, for example, used to call some of his guitar work related to his time of being a heroin addict as “there I am playing my junkie riffs” [96]. What this may mean in terms of the musical elements used is left open. But as with Keith Richards and others who spiraled into addictions, drugs and music accompanied important and also difficult periods in their lives. Therefore, any music experienced during such periods will also relate to the *autobiographical memory traces* of the clients and their drug use experience [97]. This is what many music therapists working with addicted clients know [98, 99]. By doing so, therapists aim to modify the emotional valence of addiction-related memories, promote narrative coherence, and foster new, adaptive associations between music, the self, and recovery [65, 97].

4.3 Music therapy, brain processing, and craving reduction

Given the emotional depth often explored in music and MT, it is important to consider how emotions are related to one of the most challenging aspects of addiction: craving [100, 101]. Music can serve as a channel to access emotions that may be difficult to express verbally. In a therapeutic setting, music, therefore, provides an opportunity to engage with these emotions [65, 97, 99], to recognize and process them, so that the negative ones, such as anxiety or aggression, do not govern one's whole conscious life. Addiction, on the other hand, is a complex and multifaceted condition that significantly impacts

individuals' lives, with craving – defined as a powerful and persistent urge for substances – representing a central symptom [102]. Craving perpetuates the cycle of substance use disorders (SUD), affecting relapse rates and treatment outcomes, and is shaped not only by psychological factors but also by emotional and neurobiological mechanisms involving *memory and reward systems* [100, 101]. In this context, music may serve as a meaningful therapeutic channel to engage with the emotional and neurophysiological dimensions of craving. Indeed, MT added to standard care leads to moderate reductions in substance craving [99] through potentially modulating brain frequencies, in particular beta waves and alpha asymmetries involved in craving-arousal [103].

Music therapy, defined as the clinical use of music interventions by a certified music therapist within a therapeutic relationship [104, 105], has emerged as a promising complementary treatment for addiction [99]. The therapeutic effects of music are diverse and extensive, stemming from its profound capacity to engage the emotional and cognitive processing centers of the brain, in addition to the aforementioned reward system [81, 90]. Through tailored musical experiences – such as improvisation, songwriting, lyric analysis, and active listening – MT has been shown to facilitate emotional regulation, cognitive improvement, and craving reduction in people with SUD [103, 106]. These processes collectively support recovery by targeting craving and associated psychological symptoms, thereby enhancing individual motivation and resilience in overcoming addiction [99].

While considerable attention in former MT-related addiction research has been given to detoxification units, these treatment service environments primarily manage acute withdrawal symptoms through short-term medical interventions [107, 108]. For outpatient settings, by contrast, a more prolonged therapeutic framework is offered because those patients are regularly confronted with everyday stressors and potential relapse triggers. This setting also provides an ideal context for examining the sustained therapeutic benefits of MT [106]. Until now, however, research in outpatient environments remains limited, underscoring the importance of understanding MT's role beyond acute withdrawal management.

Recent systematic reviews highlight the positive effects of MT on emotional well-being, including reductions in depression, anxiety, and craving symptoms [99, 109]. For instance, studies have documented significant craving reductions immediately following individual MT sessions, as evidenced by participants' self-reported experiences [110, 111].

It is important to note that the mechanisms underpinning MT's effectiveness consist of several overlapping processes. On the neurophysiological level, MT can modulate brain activity associated with craving, as shown by changes in EEG patterns, in particular by reduced beta activity [104]. In other words, music interventions appear capable of attenuating the intense emotional reactions triggered by drug-related cues and memories – processes deeply tied to the brain's reward and memory circuits [65, 103]. Therapeutically, this involves retraining the emotional responses, thereby weakening the pathological cycle by which music is linked to memory and craving [112]. Additionally, improvisational MT, across diverse musical genres, can foster the improvement of cognitive functions, such as inhibitory control, further supporting clients in managing impulsive behaviors and cravings [103].

Moreover, the therapeutic relationship between patient and therapist, established through MT, significantly contributes to its effectiveness [113]. Music therapy

leverages the inherent relational and social dimensions of music-making, facilitating trust, emotional expression, and introspection. Participants frequently report feeling a deep therapeutic connection, which fosters motivation and supports their recovery journey. Qualitative research confirms that the therapeutic relationship, facilitated by music, helps clients confront and reinterpret past experiences by reducing the emotional intensity associated with drug-related memories and promoting positive behavioral change [103].

Innovative approaches, including EEG hyperscanning, allow researchers to simultaneously record and analyze the brain activity of both clients and therapists during MT sessions [114, 115]. This advanced technique provides unprecedented insights into the neurobiological underpinnings of therapeutic interactions, highlighting moments of emotional connection and mutual engagement. Such methodologies promise to deepen our understanding of the precise mechanisms through which MT facilitates healing and recovery.

4.4 Music therapy and processing negative emotions in recovery

As for the mechanisms of therapeutic change, it is important to stress that becoming aware of emotions and internal processes may lead to the experience of negative emotions. Experiencing negative emotions within a therapeutic journey is not necessarily a negative outcome. In a pilot randomized-controlled trial, we found significant differences at the post-intervention assessment stage between MT and a group that received standard interventions in a community setting [103]. The main difference referred to a frontal-brain asymmetry metric utilized as a marker of emotional and motivational regulation [116–118]. Results from the two groups were compared with a normative database containing the EEG signals from 625 healthy individuals [119]. The brain asymmetry scores – transformed from amplitude-squared units into z-score values – were closer to the values of the normative population after MT rather than after standard treatment, suggesting a shift toward a more balanced, bilateral pattern of frontal brain activity rather than a right-lateralized one [103].

Interpreting these findings in light of existing neuroscientific literature, the observed effect toward balancing out the lateralization patterns suggests that MT may have an influence on emotional and motivational regulation by reducing approach-related positive affect. Specifically, this effect may not reflect a reduction in positive affect per se, but rather a normalization of emotional engagement, including a decrease in excessive approach-related drive often associated with craving, impulsive behaviors, and a right-lateralized frontal brain asymmetry associated with withdrawal behavior, indicating the processing of negative emotions. The observed modulation in the prefrontal cortex (through EEG [103]) might be associated with a broader cortico-subcortical network involved in processing reward (see **Figure 1**), indicating the relative dominance of the withdrawal system, which is typically associated with the processing of negative affective states [116–118, 120]. For example, activation within this system has been linked to emotions such as sadness and self-reflection. In a therapeutic context, the recognition and expression – verbally and musically – of negative emotions, rather than their avoidance, can play a constructive role. Treating depression with MT, with improvisation and verbal expression, decreased anxiety and depression [121] and increased right frontal activity during the MT session [122]. Given that 70% of therapy activity contains

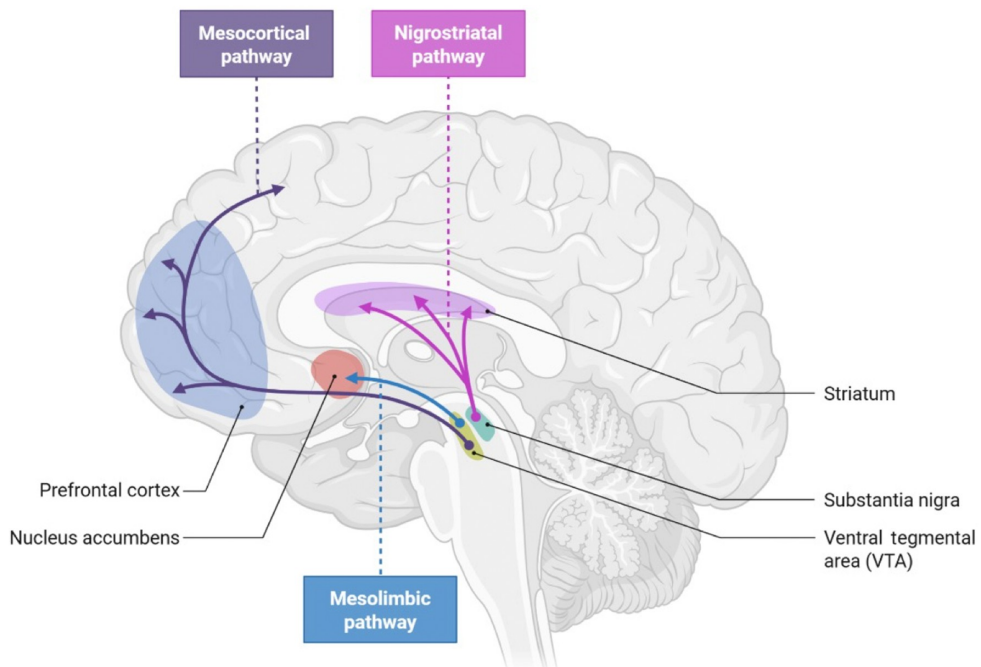


Figure 1.

Main dopaminergic pathways. Brain reward systems involved in the consumption of psychoactive drugs, including the mesocortical, nigrostriatal, and mesolimbic pathways (created with BioRender.com and adapted from insights presented in literature such as [85]).

talking about the music experience and current emotional states, the authors discuss how increased right frontal activity may also relate to right hemispheric changes in prosodic processing and the expression of emotion. A recent study examined the variability of prosodic features in verbalizations during MT as potential markers of change in depression [123]. Further, engaging with these emotional states during therapy may strengthen the therapeutic alliance and promote a more authentic and effective therapeutic process [124]. Importantly, psychological well-being does not depend solely on the presence of positive emotions or the absence of negative ones. On the contrary, research suggests that the acceptance and integration of negative or mixed emotions is associated with improved psychological and physiological health outcomes [125, 126].

5. Closing remarks

Regarding substance use and misuse, especially in adolescence, we can see that young people have sought to cross boundaries to find out where their own boundaries lie. This works well for most but not for all. For some young people, intense experiences can become strong autobiographical memories associated with the need to revisit them, which may spiral into addictions. While psychoactive substances have existed since the early days of humankind, shamanistic rituals that once helped young people strengthen their boundaries and integrate risky or extreme experiences during adolescence have largely disappeared. In the past, such rituals

were embedded within communities that provided guidance from experienced members during rites of passage involving these substances. The revival of psychedelic research and therapy approaches at the present time shows how the mind-altering properties of drugs can be used safely in guided sessions.

Nevertheless, the drug action of euphoric or sedative drugs, such as cocaine or heroin, may have an impact on emotion processing that requires therapeutic settings to regain the broad potential of emotional nuances that may have been lost through habituating a narrowed consciousness [93], since craving urges a repetition of a particular emotional body experience. Music therapy has been shown to support craving reduction and targets brain circuits of emotion processing during craving and emotion regulation. In MT, negative emotions can be expressed through musical activities, helping to relearn and retrain experiences that can also be linked to drug and music experiences, cueing (and obliterating) drug memory. Understanding how the interaction between drugs and music works is important for music therapists working in treatment centers as it allows for more targeted interventions in groups and single sessions with adolescents and adults.

In conclusion, MT is a powerful, multi-faceted intervention capable of addressing the complex emotional, cognitive, and social dimensions of addiction. While preliminary results from clinical trials are encouraging, further research – particularly in outpatient settings – is essential to fully understand MT's long-term impacts on craving and overall recovery. By integrating neurophysiological, psychometric, and qualitative methodologies, future studies can provide a comprehensive picture of music and MT's potential, ultimately enhancing treatment outcomes and the quality of life for individuals grappling with SUDs.

Author details

Jörg Fachner^{1*}, Filippo Pasqualitto^{1,2} and Clemens Maidhof¹

1 Cambridge Institute for Music Therapy Research, Anglia Ruskin University, Cambridge, UK

2 Department of Neuroscience and Rehabilitation, University of Ferrara, Ferrara, Italy

*Address all correspondence to: jorg.fachner@aru.ac.uk

IntechOpen

© 2025 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Turner VW. The ritual process: structure and anti-structure. Vol. **16**. New York: Aldine de Gruyter; 1995. p. 213
- [2] Collin M. Altered state: the story of Ecstasy culture and Acid House. Vol. **XII**. Updated rev. London: Serpent's Tail; 2009. 356 p
- [3] Rave RB. Communitas, and Embodied Idealism. Music Therapy Today. 2006;7(3):648–661. https://issuu.com/presidentwfmt/docs/mtt_7_3_2006
- [4] Aron L. A Meeting of Minds: Mutuality in Psychoanalysis. Hillsdale, NJ: The Analytic Press; 1996. p. 292
- [5] MTV True Life. Deutschland: MTV; I'm on Ecstasy. 2005. p. MTV
- [6] Chapman C, Rokutani L. Adolescents and substance abuse: what works and why? Vistas: compelling perspectives on counseling. Alexandria, VA: American Counselling Association; 2005. p. 25–28
- [7] Szasz TS. Ceremonial chemistry: The ritual persecution of drugs, addicts, and pushers. Vol. **xxi**. Rev ed. Syracuse, N.Y: Syracuse University Press; 2003. 290 p
- [8] Sullivan RJ, Hagen EH. Psychotropic substance-seeking: Evolutionary pathology or adaptation? Addiction. 2002;97(4):389–400
- [9] Sullivan RJ, Hagen EH, Hammerstein P. Revealing the paradox of drug reward in human evolution. Proceedings of the Royal Society B: Biological Sciences. 2008;275(1640):1231–1241
- [10] Blätter AE, Fachner JC, Winkelman M. Addiction – Dynamics and Relations to Altered States of Consciousness. In: Cardeña E, Winkelman M, editors. Altering Consciousness: A multidisciplinary perspective. 2 Biological and Psychological Perspectives. Santa Barbara: Praeger; 2011. p. 167–187
- [11] Moore P. Enhancing me: the hope and the hype of human enhancement. Vol. **14**. Chichester, England; Hoboken, NJ: Wiley/Dana Centre; 2008. p. 261
- [12] Parker H, Williams L, Aldridge J. The normalization of 'sensible' recreational drug use: Further evidence from the North-West England Longitudinal Study. Sociology. 2002;36(4):941–964
- [13] Freeman W. A neurobiological role of music in social bonding. In: Wallin NL, Merker B, Brown S, editors. The origins of music. Cambridge, Mass: MIT Press; 2000. p. 411–424
- [14] De Rios MD. Hallucinogens, cross-cultural perspectives. Vol. **17**. Prospect Heights, Ill: Waveland Press; 1996. p. 261
- [15] Metzner R. Molekulare Mystik: Die Rolle psychoaktiver Substanzen bei der Transformation des Bewußtseins [Molecular Mysticism: The Role of Psychoactive Substances in the Transformation of Consciousness]. In: Rätsch C, editor. Das Tor zu den inneren Räumen [The gateway to the inner spaces]. 1 ed. German, Südergellsen: Verlag Bruno Martin; 1992. p. 63–78
- [16] De Rios MD. The Role of Music in Healing with Hallucinogens: Tribal and

- Western Studies. In: Aldridge D, Fachner J, editors. *Music and Altered States - Consciousness, Transcendence, Therapy and Addictions*. London: Jessica Kingsley; 2006. p. 97–101
- [17] Fachner J. Music and drug induced altered states. In: Aldridge D, Fachner J, editors. *Music and Altered States - Consciousness, Transcendence, Therapy and Addictions*. London: Jessica Kingsley; 2006. p. 82–96
- [18] Fachner J. Drugs, altered states and musical consciousness: reframing time and space. In: Clarke E, Clarke D, editors. *Music and consciousness*. Oxford: Oxford University Press; 2011. p. 263–280
- [19] Maas U, Strubelt S. Polyrhythms supporting a pharmacotherapy - Music in the Iboga initiation ceremony in Gabon. In: Aldridge D, Fachner J, editors. *Music and Altered States - Consciousness, Transcendence, Therapy and Addictions*. 1 ed. London: Jessica Kingsley; 2006. p. 101–124
- [20] Shannon B. Ayahuasca and music. In: Clarke E, Clarke D, editors. *Music and consciousness*. Oxford: Oxford University Press; 2011. p. 281–294
- [21] Melechi A, editor. *Psychodelia Britannica*. 1 ed. London: Turnaround; 1997
- [22] Fachner J *Takin' it to the Streets...* Psychotherapie, Drogen und Psychedelic Rock: Ein Forschungseüberblick [Takin' it to the Streets... Psychotherapy, Drugs, and Psychedelic Rock: A Research Review]. Samples - Online publication of the German Society of the Study of Popular Music (ASPM) [Internet]. 2007-October-30. p. 6. DOI: 10.22029/jlupub-808
- [23] Eisner B. Set, setting and matrix. *Journal of Psychoactive Drugs*. 1997;29(2):213–216
- [24] Grof S. *LSD psychotherapy*. 2nd ed. Alameda, CA: Hunter House; 1994. p. 352
- [25] Leary TF. *Flashbacks: a personal and cultural history of an era: an autobiography*. 1st. Los Angeles; New York: J.P. Tarcher; Distributed by St. Martin's Press; 1990. 405 p. of plates
- [26] Leuner HC. *Die experimentelle Psychose [The experimental psychosis]*. Berlin Göttingen Heidelberg: Springer; 1962. 275 p
- [27] Maslow A. *Religions, Values and Peak Experience*. New York: Viking Penguin; 1970. 123 p
- [28] Kaelen M, Giribaldi B, Raine J, Evans L, Timmerman C, Rodriguez N, et al. The hidden therapist: evidence for a central role of music in psychedelic therapy. *Psychopharmacology (Berl)* 2018;235:505–519
- [29] Barrett FS, Robbins H, Smooke D, Brown JL, Griffiths RR. Qualitative and Quantitative features of music reported to support peak mystical experiences during psychedelic therapy sessions. *Frontiers in Psychology*. 2017;8(1238)
- [30] Carhart-Harris RL, Roseman L, Haijen E, Erritzoe D, Watts R, Branchi I, et al. Psychedelics and the essential importance of context. *Journal of Psychopharmacology* 2018;32:725–731
- [31] Zinberg NE. *Drug, set, and setting: the basis for controlled intoxicant use*. Vol. 15. 1ed. New Haven: Yale University Press; 1984. 277

- [32] Kaelen M, Barrett FS, Roseman L, Lorenz R, Family N, Bolstridge M, et al. LSD enhances the emotional response to music. *Psychopharmacology (Berl)* 2015;**232**:3607–3614
- [33] Kaelen M, Roseman L, Kahan J, Santos-Ribeiro A, Orban C, Lorenz R, et al. LSD modulates music-induced imagery via changes in parahippocampal connectivity. *European Neuropsychopharmacology* 2016;**26**:1099–1109
- [34] Barrett FS, Preller KH, Kaelen M. Psychedelics and music: neuroscience and therapeutic implications. *International Review of Psychiatry*. 2018;**30**(4):350–362
- [35] Emrich HM, Weber MM, Wendl A, Zihl J, Von Meyer L, Hanish W. Reduced binocular depth inversion as an indicator of cannabis-induced censorship impairment. *Pharmacology Biochemistry and Behavior*. 1991;**40**:689–690
- [36] Carhart-Harris RL, Friston KJ. The default-mode, ego-functions and free-energy: A neurobiological account of Freudian ideas. *Brain*. 2010;**133** (4):1265–1283
- [37] Carhart-Harris RL. How do psychedelics work? *Current Opinion in Psychiatry*. 2019;**32**(1):16–21
- [38] Bonny H. Music and consciousness: The evolution of Guided Imagery and Music. Gilsum, NH: Barcelona Publishers; 2002. p. 375
- [39] Wolfe T. The Electric Kool-Aid Acid Test. New York: Bantam Books; 1989
- [40] Doyle M. Death of Hippie. 1967. Available from: <https://www.ebsco.com/research-starters/history/death-hippie> [Accessed: 2026-January-13]
- [41] Lee MA, Shalin B. Acid Dreams. 2 ed. New York: Grove Press; 1992. 343 p
- [42] Hep-Cats JJ. Narcs and Pipe Dreams. 1 ed. Baltimore: Johns Hopkins University Press; 1999. 528 p
- [43] Fachner J. “. . . if you’re a viper.” Consciousness states in a social pharmacology of music. In: Cavanna AE, Nani A, editors. *Consciousness: States, Mechanisms and Disorders*. Hauppauge NY: Nova Science Publishers; 2012. p. 189–206
- [44] Laing RD. The politics of experience. Vol. 15. New York: Pantheon Books; 1967. 138 p
- [45] Hendrix J. Are You experienced? Long Player: Polydor International; 1967
- [46] Schamanismus RC. Techno und Cyberspace [Shamanism, Techno and Cyberspace]. Solothurn: Nachtschatten Verlag; 2001. p. 64
- [47] Masters REL, Houston J. Psychedelic art. London: Weidenfeld & Nicolson; 1968. 190 p
- [48] Krippner S. Psychedelic drugs and creativity. *Journal of Psychoactive Drugs*. 1985;**17**(4):235–245
- [49] Taeger H-H. Spiritualität und Drogen interpersonelle Zusammenhänge von Psychedelika und religiös-mystischen Aspekten in der Gegenkultur der 70er Jahre [Spirituality and Drugs: Interpersonal Relationships of Psychedelics and Religious-Mystical Aspects in the Counterculture of the

- 70s]. Markt Erlbach: Martin; 1988. p. 273S
- [50] De Rios MD, Janiger O. LSD, spirituality, and the creative process. Vol. 10. Rochester, Vt: Park Street Press; 2003. 246 p
- [51] Eagle CT. Music and LSD: an empirical study. *Journal of Music Therapy*. 1972;9:23–36
- [52] Bonny HL, Pahnke WN. The use of music in psychedelic (LSD) psychotherapy. *Journal of Music Therapy*. 1972;9:64–87
- [53] Weber K. Alteration of music perception in experimental psychosis (psilocybin). *Confinia psichiatria. Borderland of psychiatry. Grenzgebiete der Psychiatrie. Les Confins de la psychiatrie*. 1967;10(3):139–176
- [54] Weber K. Disturbances in time perception in relation to changes in experiencing music in experimental psychosis. *Confinia Psychiatrica*. 1977;20(2-3):79–94
- [55] Russo EB, Merzouki A, Mesa JM, Frey KA, Bach PJ. Cannabis improves night vision: A case study of dark adaptometry and scotopic sensitivity in kif smokers of the Rif mountains of Northern Morocco. *Journal of Ethnopharmacology*. 2004;93(1):99–104
- [56] Tart C. On Being Stoned, A Psychological Study of Marijuana Intoxication. 1 ed. Palo Alto: Science and Behaviour Books; 1971. 333 p
- [57] Winton-Brown TT, Allen P, Bhattacharyya S, Borgwardt SJ, Fusar-Poli P, Crippa JA, et al. Modulation of auditory and visual processing by delta-9-tetrahydrocannabinol and cannabidiol: an fMRI study. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology* 2011;36:1340–1348
- [58] Freeman TP, Pope RA, Wall MB, Bisby JA, Luijten M, Hindocha C, et al. Cannabis Dampens the Effects of Music in Brain Regions Sensitive to Reward and Emotion. *International Journal of Neuropsychopharmacology* 2017;21:21–32
- [59] Darakjian L, Glassman H, Lo CY, Russo FA. Exploring the interaction between cannabis and music. *Annals of the New York Academy of Sciences*. 2025;1551(1):140–158
- [60] Kempel P, Lampe K, Parnefjord R, Hennig J, Kunert HJ. Auditory-evoked potentials and selective attention: Different ways of information processing in cannabis users and controls. *Neuropsychobiology*. 2003;48(2):95–101
- [61] Fachner J. Topographic EEG changes accompanying Cannabis-induced alteration of music perception - Cannabis as a hearing aid? *Journal of Cannabis therapeutics*. 2002;2(2):3–36
- [62] Boyd J. Musicians in tune - Seventy-five contemporary musicians discuss the creative process. 1 ed. New York: Fireside, a Simon & Schuster imprimatur; 1992. p. 288
- [63] Baudelaire C. Les paradis artificiels: opium et haschisch [Artificial paradises]. Paris: Poulet-Malassis et de Broise; 1860. p. 304
- [64] Globus GG, Cohen HB, Kramer JC, Elliot HW, Sharp R. Effects of Marijuana induced altered state of consciousness on auditory perception.

Journal of Psychedelic Drugs. 1978;**10**
(1):71–76

[65] Fachner J. Music therapy, drugs and state-dependent recall. In: Aldridge D, Fachner J, editors. Music therapy and addictions. 1 ed. London: Jessica Kingsley; 2010. p. 18–34

[66] Meck WH. Neuropharmacology of timing and time perception. Cognitive Brain Research. 1996;**3**(3–4):227–242

[67] Mathew RJ, Wilson WH, Melges FT. Temporal disintegration and its psychological and physiological correlates changes in the experience of time after marijuana smoking. Annals of clinical psychiatry. 1992;**4**(4):235–245

[68] Mathew RJ, Wilson WH, Turkington TG, Coleman RE. Cerebellar activity and disturbed time sense after THC. Brain Research. 1998;**797**(2):183–189

[69] Fachner J. Out of time? Music, consciousness states and neuropharmacological mechanisms of an altered temporality. In: 7th Triennial Conference of European Society for the Cognitive Sciences of Music; University of Jyväskylä, Finland: UoJ: Department of Music for ESCOM; 12–16 December 2009. p. 103–109. <http://urn.fi/URN:NBN:fi:jyu-2009411245>

[70] Whiteley S. The space between the notes - Rock and the counter culture. London: Routledge; 1992. p. 152

[71] Gennep A, Vizedom MB, Caffee GL. The rites of passage. 2 ed. Chicago; London: The University of Chicago Press; 2019. p. xlvii

[72] Frith S. Performing rites. Oxford: Oxford University Press; 1998. p. 352

[73] Fachner J. CLEAN UP! “Heimatschutz”, Anti-Drogenpolitik und legislative Auswirkungen auf die Rave-Kultur nach dem 11. September 2001 [CLEAN UP! “Homeland Security,” Anti-Drug Policy, and Legislative Impact on Rave Culture after September 11, 2001.]. In: Helms D, Phleps T, editors. 9/11 - The world’s all out of Tune - Populäre Musik nach dem 11 September 2001 [9/11 - The world’s all out of Tune - Popular music after September 11, 2001]. ASPM - Beiträge zur Populärmusikforschung. Vol. 32. Bielefeld: Transkript Verlag; 2004. p. 81–98

[74] Dijkstra I, Hakvoort L. How to Deal Music? Music Therapy with Clients Suffering from Addiction Problems: Enhancing Coping Strategies. In: Aldridge D, Fachner J, editors. Music therapy and addictions. 1 ed. London: Jessica Kingsley; 2010. p. 88–102

[75] Punkanen M. Music therapy as a part of drug rehabilitation – from adhering to treatment to integrating the levels of experience. In: Aldridge D, Fachner J, editors. Music Therapy and Addictions. London: Jessica Kingsley; 2010. p. 123–131

[76] Punkanen M. On a journey to somatic memory: Theoretical and clinical approaches for the treatment of traumatic memories in music therapy-based drug rehabilitation. In: Aldridge D, Fachner J, editors. Music and Altered States: Consciousness, Transcendence, Therapy and Addictions. London: Jessica Kingsley Publishers; 2006. p. 140–154

[77] Blätter A. Kulturelle Ausprägungen und die Funktion des Drogengebrauchs. [Cultural characteristics and the function of drug use]. 1 ed. Hamburg: Wayasbah Verlag; 1990. p. 288

- [78] DuPont RL. The selfish brain: learning from addiction. Vol. 37. Center City, Minn: Hazelden; 2000. p. 505
- [79] Blakemore S.J. Development of the social brain in adolescence. *Journal of the Royal Society of Medicine*. 2012;**105** (3):111–116
- [80] Carlson E, Saarikallio S, Toiviainen P, Bogert B, Kliuchko M, Brattico E. Maladaptive and adaptive emotion regulation through music: A behavioral and neuroimaging study of males and females. *Frontiers in Human Neuroscience*. 2015;**9**:466
- [81] Koelsch S. Music-evoked emotions: principles, brain correlates, and implications for therapy. *Annals of the New York Academy of Sciences*. 2015;**1337**:193–201
- [82] Koob GF. Dynamics of Neuronal Circuits in Addiction: Reward, Antireward, and Emotional Memory. *Pharmacopsychiatry*. 2009;**42**(1): S32–S41
- [83] Dahl RE, Spear LP. Adolescent brain development: Vulnerabilities and opportunities. 1 ed. New York: New York Academy of Sciences; 2004. p. 472
- [84] Blakemore S.J. *Inventing Ourselves: The Secret Life of the Teenage Brain*. London: Doubleday; 2018. p. 256
- [85] Berridge Kent C, Kringelbach Morten L. Pleasure Systems in the Brain. *Neuron*. 2015;**86**(3):646–664
- [86] Koob GF. The dark side of emotion: The addiction perspective. *European Journal of Pharmacology*. 2015;**753**:73–87
- [87] Emrich HM, Schneider UFDS. Von der Neurobiologie zur Anthropologie. In: Von der Neurobiologie zur Anthropologie. [Facets of addiction. From Neurobiology to Anthropology.]. Frankfurt a.M: Peter Lang; 2006. p. 129
- [88] Yacubian J, Büchel C. The genetic basis of individual differences in reward processing and the link to addictive behavior. In: Dreher J-C, Tremblay L, editors. *Handbook of Reward and Decision Making*. 1 ed. Burlington: Academic Press; 2009. p. 345–360
- [89] Blood AJ, Zatorre RJ. Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences of the United States of America*. 2001;**98** (20):11818–11823
- [90] Koelsch S. Brain correlates of music-evoked emotions. *Nature Reviews Neuroscience*. 2014;**15**(3):170–180
- [91] Ferreri L, Mas-Herrero E, Zatorre RJ, Ripollés P, Gomez-Andres A, Alicart H, et al. Dopamine modulates the reward experiences elicited by music. *Proceedings of the National Academy of Sciences* 2019;**116**:3793–3798
- [92] Winkelman M. Complementary therapy for addiction: “Drumming Out Drugs”. *American Journal of Public Health*. 2003;**93**(4):647–651
- [93] Metzner R. Addiction and transcendence as altered states of consciousness. *Journal of Transpersonal Psychology*. 1994;**26**(1):1–17
- [94] Horesh T. Dangerous Music - Working with the Destructive and Healing Powers of Popular Music in the Treatment of Substance Abusers. In: Aldridge D, Fachner J, editors. *Music*

and Altered States - Consciousness, Transcendence, Therapy and Addiction. London: Jessica Kingsley; 2006. p. 125–139

[95] Horesh T. Drug addicts and their music - A Story of a Complex Relationship. In: Aldridge D, Fachner J, editors. Music therapy and Addictions. London: JKP; 2010. p. 57–74

[96] Roesberg D. Interview with Keith Richards. Fachblatt Musikmagazin; November. 1978

[97] Pasqualitto F, Panin F, Maidhof C, Thompson N, Fachner J. Neuroplastic changes in addiction memory. How music therapy and music-based intervention may reduce craving: a narrative review. Brain Sciences. 2023;13(2):259

[98] Aldridge D, Fachner J. Music therapy and addictions. London: Jessica Kingsley; 2010. p. 176

[99] Ghetti C, Chen X, Brenner A, Hakvoort LG, Lien L, Fachner J, et al. Music therapy for people with substance use disorders (Review). Cochrane Database of Systematic Reviews. 2022; (5):CD012576

[100] Koob GF, Volkow ND. Neurobiology of addiction: A neurocircuitry analysis. The Lancet Psychiatry. 2016;3(8):760–773

[101] Diagnostic and statistical manual of mental disorders: DSM-5. 2013

[102] Gauld C, Baillet E, Micoulaud-Franchi J-A, Kervran C, Serre F, Auriacombe M. The centrality of craving in network analysis of five substance use disorders. Drug and Alcohol Dependence. 2023;245:109828

[103] Pasqualitto F, Maidhof C, Murtagh D, Silva D, Fernie P, Panin F, et al. Music Therapy modulates Craving, Inhibitory Control, and Emotional Regulation: EEG, Psychometric, and Qualitative Findings from a Pilot RCT in a Community Outpatient Service. ResearchSquare preprints. 2025 submitted.

[104] Raglio A, Attardo L, Gontero G, Rollino S, Groppo E, Granieri E. Effects of music and music therapy on mood in neurological patients. World Journal of Psychiatry. 2015;5(1):68–78

[105] American Music Therapy Association. AMTA Official Definition of Music Therapy. Silver Spring, MD; 2006. Available from: <https://www.musictherapy.org/about/musictherapy/> [Accessed: 2026-January-13]

[106] Fachner J, Maidhof C, Murtagh D, De Silva D, Pasqualitto F, Fernie P, et al. Music therapy, neural processing, and craving reduction: An RCT protocol for a mixed methods feasibility study in a Community Substance Misuse Treatment Service. Addiction Science and Clinical Practice. 2023;18(36)

[107] Gossop M, Marsden J, Stewart D, Rolfe A. Treatment retention and 1 year outcomes for residential programmes in England. Drug and Alcohol Dependence. 1999;57(2):89–98

[108] James H, Morgan J, Ti L, Nolan S. Transitions in care between hospital and community settings for individuals with a substance use disorder: A systematic review. Drug and Alcohol Dependence. 2023;243:109763

[109] Hohmann L, Bradt J, Stegemann T, Koelsch S. Effects of

music therapy and music-based interventions in the treatment of substance use disorders: A systematic review. *PLOS ONE*. 2017;**12**(11)

[110] Silverman MJ. Effects of music therapy on change readiness and craving in patients on a detoxification unit. *Journal of Music Therapy*. 2011;**48**(4):509–531

[111] Farahmand SRE, Ahadi H, Kalhornia Golkar M, Sedaghat M. Comparison of the Effectiveness of Music Therapy and Cognitive Behavioral Therapy on Quality of Life, Craving and Emotion Regulation in Patients Under Methadone Maintenance Therapy. *Iranian Journal of Psychiatry and Clinical Psychology*. 2020;**26**(2):170–187

[112] Fachner J. Retraining of Drug reward, Music cues and State-Dependent recall in music therapy. *Music and Medicine*. 2017;**9**(1):8–14

[113] Silverman MJ. The Effect of Lyric Analysis on Treatment Eagerness and Working Alliance in Consumers Who Are in Detoxification: A Randomized Clinical Effectiveness Study. *Music Therapy Perspectives*. 2009;**27**(2):115–121

[114] Fachner J, Maidhof C, Grocke D, Nygaard Pedersen I, Trondalen G, Tucek G, et al. “Telling me not to worry. . .” Hyperscanning and Neural Dynamics of Emotion Processing During Guided Imagery and Music. *Frontiers in Psychology*. 2019;**10**(1561):80–102

[115] Maidhof C, Mueller V, Agres K, Lartillot O, Bloska J, Asano R, et al. Intra- and inter-brain coupling and activity dynamics during

improvisational music therapy with a person with dementia: an explorative EEG-hyperscanning single case study. *Frontiers in Psychology*. 2023;**14**

[116] Davidson RJ. EEG measures of cerebral asymmetry: Conceptual and methodological issues. *International Journal of Neuroscience*. 1988;**39**(1-2):71–89

[117] Davidson RJ. What does the prefrontal cortex “do” in affect: Perspectives on frontal EEG asymmetry research. *Biological Psychology*. 2004;**67**(1-2):219–233

[118] Harmon-Jones E, Gable PA, Peterson CK. The role of asymmetric frontal cortical activity in Emotion-related phenomena: A review and update. *Biological Psychology*. 2010;**84**:451–462

[119] Thatcher RW, Walker RA, Biver CJ, North DM, Curtin R. Quantitative EEG normative databases: Validation and clinical correlation. *Journal of Neurotherapy*. 2003;**7**(3-4):87–121

[120] Smith EE, Reznik SJ, Stewart JL, Allen JJB. Assessing and conceptualizing frontal EEG asymmetry: An updated primer on recording, processing, analyzing, and interpreting frontal alpha asymmetry. *International Journal of Psychophysiology*. 2017;**111**:98–114

[121] Erkkilä J, Punkanen M, Fachner J, Ala-Ruona E, Pöntiö I, Tervaniemi M, et al. Individual music therapy for depression - Randomised Controlled Trial. *British Journal of Psychiatry* 2011;**199**:132–139

[122] Fachner J, Gold C, Erkkilä J. Music therapy modulates Fronto-temporal

activity in the rest-EEG in depressed clients. *Brain Topography*. 2013;**26**(2):338–354

[123] Knight S, Spiro N. Tracing change during music therapy for depression: Toward a markers-based understanding of communicative behaviors. *Musicae Scientiae*. 2023;**27**(3):637–654

[124] Barnes M, Sherlock S, Thomas L, Kessler D, Kuyken W, Owen-Smith A, et al. No pain, no gain: Depressed clients' experiences of cognitive behavioural therapy. *British Journal of Clinical Psychology* 2013;**52**:347–364

[125] Ford BQ, Lam P, John OP, Mauss IB. The psychological health benefits of accepting negative emotions and thoughts: Laboratory, diary, and longitudinal evidence. *Journal of Personality and Social Psychology*. 2018;**115**(6):1075–1092

[126] Hershfield HE, Scheibe S, Sims TL, Carstensen LL. When feeling bad can be good: Mixed emotions benefit physical health across adulthood. *Social Psychological and Personality Science*. 2013;**4**(1):54–61



Edited by Christiane Neuhaus

This is a book about current topics in Systematic Musicology – about how music is experienced, performed, analyzed, and processed. Empirical work on these topics is interdisciplinary and extremely fast-paced nowadays, with new methods and software applications being developed almost every year. The current book aims to present the most recent state of research. It is a compilation of seven chapters, assigned to the following sections: Audio Technology, Music Theory, Music Psychology, Music and Neuroscience, and Music Therapy. The book's themes have been carefully elaborated by the authors, and its insightful chapters open new paths to deeper understanding and to opinion formation. The book deals with the following topics:

- Musical data sonification
- Cognitive music theory
 - Roots of artwork
- Music's role in artistic sports
- EEG hyperscanning, mirror neurons, and ensemble playing
 - Indian music therapy based on Sanskrit sources
- Anthropological and developmental aspects of drug use; music therapy in addiction treatment

Published in London, UK

© 2026 IntechOpen

© Alex Rodrigo Brondani / iStock

IntechOpen

