education sciences

MDPI

Article

Technology-Mediated Hindustani Dhrupad Music Education: An Ethnographic Contribution to the 4E Cognition Perspective

Stella Paschalidou

Department of Music Technology and Acoustics, Hellenic Mediterranean University, 74 133 Rethymno, Greece; pashalidou@hmu.gr

Abstract: Embodiment lies at the core of music cognition, prompting recent pedagogical shifts towards a multi-sensory, whole-body approach. However, the education of oral music genres that rely exclusively on direct teacher–disciple transmission through live demonstration and imitation is now undergoing a transformation by rapidly adapting to technology-mediated platforms. This paper examines challenges in embodied facets of video-mediated synchronous distance Hindustani music pedagogy. For this, it takes an ethnomusicological stance and showcases a thematic analysis of interviews featuring Dhrupad music practitioners. The analysis is driven and organized by the 4E Cognition principles, which stress the intimate relationship between body, mind, and environment. Findings indicate that while this adaptation aims to make music content more widely accessible, it comes at the cost of reducing opportunities for multi-modal engagement and interaction among participants. Results reveal limitations in transmitting non-verbal, embodied, multi-sensory cues, along with visual and acoustic disruptions of a sense of shared spatial and physical context, that hinder effective interaction and a sense of immersion, elements that are deemed vital in music education. They prompt concerns about the suitability of conventional videoconferencing platforms and offer key insights for the development of alternative technologies that can better assist embodied demands of the pedagogical practices involved.

Keywords: distance music education; music technology; motion capture technologies; 4E cognition; embodied music cognition; multimodality; oral music traditions; Hindustani Dhrupad music; ethnomusicology; thematic analysis

check for updates

Citation: Paschalidou, S. Technology-Mediated Hindustani Dhrupad Music Education: An Ethnographic Contribution to the 4E Cognition Perspective. *Educ. Sci.* **2024**, *14*, 203. https://doi.org/10.3390/educsci14020203

Academic Editor: Eleanor Dommett

Received: 30 December 2023
Revised: 31 January 2024
Accepted: 12 February 2024
Published: 17 February 2024

Copyright: © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Musicking [1] is a genuinely embodied process [2]. Therefore, in more recent years, music has been approached within the framework of multimodal discourse by embracing sensory modalities that extend beyond the acoustic, a concept anticipated as early as 1926 when Leo Theremin declared his intention to combine music with color, gesture, scent, and touch [3]. As a result, pedagogical music practices have been widely shifting their emphasis from viewing music as an object and process toward a meaning acquired through a multi-sensory, whole-body experience on a personal, social, and cultural level [4–6]. For instance, in singing, music-related gestures and heightened kinesthetic awareness are considered valuable tools for supporting memory, enhancing pitch accuracy, expanding vocal range, and mastering diverse vocal qualities [7].

Embodiment holds particular importance in oral music traditions like Hindustani (North Indian Classical), which, unlike Eurocentric music, does not rely on written notation and in which knowledge is instead (almost) exclusively transmitted directly from teacher to disciple through live demonstration and imitation [8]; crucially for this research, not only via sound but also vital multisensory corporeal information, encompassing visual and other physical means of interaction [9] (For this very reason, the term 'oral' has been criticized by Rahaim [9] for not encompassing the visual and kinaesthetic sensory channels involved in tacit knowledge transmission.). However, a contrasting pedagogy shift has

been observed in recent years in Hindustani music and its Dhrupad subgenre, a primarily vocal type of music that is essentially based on a court patronage system between teacher and disciple, with which this paper deals. Drastic changes in India's economic and social conditions [10], coupled with the increased popularity and uptake of Indian classical music in the West [11] and, most profoundly, an accelerated transition from in-person individual classes to online synchronous tutoring (happening in real-time, as opposed to asynchronous, that occurs at different times between instructor and student) prompted by recent technological advancements and the 2019 pandemic, has steered Hindustani music education away from its traditionally embodied, multi-sensory nature toward a rather disembodied approach.

The main premise of this paper is that gestures in Dhrupad are of paramount importance in the musical communication, teaching, and learning process. Based on this, it postulates a shift in teacher-student engagement within the instructional delivery followed by selected Dhrupad music teachers and their disciples through desktop videoconferencing, especially since the COVID-19 outbreak. It specifically highlights concerns about the potential disruption of the so-called 'interpersonal' space within technology-mediated interactions, a shared action space otherwise perceived as almost tangible, facilitating the multisensory exchange of information between teacher and student. By employing ethnomusicological methods (drawing on interviews) and an embodied music cognition perspective (organizing the analysis according to the 4E cognition principles), this paper explores the potential implications of this digital 'turn' [12] in the transmission of Dhrupad under the prisms of (a) pedagogy strategies and (b) technology issues. '4E cognition' (4E standing for Embodied, Embedded, Extended, and Enactive) is a relatively new field of interdisciplinary research, advocating that cognition is shaped and structured by dynamic interactions between the brain, body, and both the physical and social environments [13,14].

While it is difficult to assess the extent of this impact, it is legitimate to highlight the challenges associated with the embodied and experiential aspects of transmission and ask questions regarding the viability and appropriateness of common videoconferencing platforms that are not specifically suited for oral types of music. Given that these traditions might not be adequately supported by the currently offered online teaching solutions, it is also reasonable to emphasize the need for further investigation in the design of effective technologies that prioritize experiential and embodied knowledge over reliance on written notation. Hence, key inquiries of this paper include the following:

- The assessment of mainstream videoconferencing platforms in adequately meeting the demands of oral music instruction, as in Dhrupad;
- The exploration of the viability in effectively capturing, conveying, and transmitting salient movement information through these platforms;
- The understanding of the potential impact of the observed embodiment shifts on the transmission of oral music traditions and the risks of overlooking elements of the music style's heritage over time;
- The consideration of novel kinds of technologies and alternative methods for conveying embodied information that need to be developed to better support the educational needs of oral and underrepresented music traditions that do not rely on written scores.

Notably, there is a shortage of prior studies explicitly addressing concerns in contemporary practices of synchronous distance tutoring for oral music traditions, such as Hindustani [15]. This paper aims to bridge this research gap. This article is structured as follows. In the first section, the goals of this study are contextualized by outlining some relevant work that concerns both traditional and online pedagogical practices of Hindustani/Dhrupad music, as well as research work regarding recent technological advancements in network-mediated synchronous music education. The implicated methods and data are then outlined before presenting the interview analysis for the online Dhrupad pedagogy context. Findings are interpreted in the Discussion Section (Section 5), and finally, this paper concludes with some speculations on interesting future technological interventions.

2. Background

2.1. Traditional Pedagogy of Hindustani Dhrupad Music

In contrast to Western classical music, Hindustani music relies on the direct transmission of knowledge from the teacher to the disciple through demonstration, emulation, and repetition [15]. The so-called 'guru-shishya parampara' ('teacher–student tradition' in English) is often quoted as the ideal example of 'enculturation' [16], whereby the teacher (metaphorically, but also literally) embodies the tradition, which is slowly assimilated through a long and close relationship [17]. Teachers use hand gestures to guide students in reproducing challenging musical passages, as demonstrated in Figure 1 from a class of maestro Zia Fariduddin Dagar given to his disciple Nirmalya Dey. While gestures are usually not explicitly taught and teachers seldom focus solely on correcting 'incorrect' movements, students tend to adopt noticeable similarities in their movements compared to their instructors through gestural mirroring [9,18], albeit not as exact replicas. This physical engagement leads to the concept of a 'paramparic body' (from the word 'parampara,' i.e., tradition), a set of bodily dispositions developed over time [9], which is a concept that draws on Young's 'family body' [19]. Thus, not only the music style but also the movement style seems to be inherited.



Figure 1. Maestro Zia Fariduddin Dagar and his disciple Nirmalya Dey, Palaspe, India, 2012. Snapshot at 8:17 from video recording [20]. Reproduced with permission from Robert Caro, Youtube video: Ustad Zia Fariduddin DAGAR/The Dhrupad Legend; published on Youtube, 2015.

Dhrupad is characterized by smooth, slow melodic glides, whereby pitch is conceived as a continuum. In other words, while melody in Dhrupad involves the precise and subtle intonation of discrete notes (svaras), it is also conceived as the movement through paths, trajectories, and shapes in an imaginary pitch 'space' [21] that exists between scale steps highlighting the significance of the 'space between the notes' as often more important than the individual notes [22]. The smooth vocal navigation within this imagined pitch space is frequently mirrored by corresponding hand movements in the real, three-dimensional space. Findings from previous works reveal a systematic, even if idiosyncratic, relationship between hand movements and melodic phrases and patterns [9,18,23,24]. This connection goes beyond mere spatial geometry (the 3D Euclidean melographic representation in Hindustani music [9]), extending to the sensation of forces and effort [18]. Specifically, singers frequently engage with melodic ideas through manual interactions with imaginary objects (abbreviated as 'MIIOs,' ibid.), whereby motor imagery is 'materialized' through effortful physical actions directed towards such objects. Singers appear to engage with the

melody and the melodic modes by manually sculpting spaces as if they were physically tangible, posing a distinct sense of resistance. These spaces are conceived as consisting of pervasive substances, such as an elastic band, a ball, water, etc, that afford various interactions such as stretching, expansion, compression, pulling, throwing, etc. This process appears to enhance musicians' capacity to imagine musical sounds through the retrieval of motor programs and is a concept frequently employed by teachers for musical instruction. It is also observed in student–teacher interactions during teaching sessions ([9] and personal observations), which underscores the significant role of spatial perception and haptic sense in Dhrupad music pedagogy.

While traditionally dominated by the sense of sight, there is a growing acknowledgment of the implication of other senses as well in spatial perception and performance (haptic, auditory, and even the olfactory and gustatory sense), with Ratcliffe [25] and Pallasmaa [26] prioritizing the sense of touch. Touch, often referred to as somatosensation, seems to carry information about both an external object touching our skin and about our body itself [27]. Unlike other senses, it lacks a specific organ and includes the whole body, with its extremities functioning both as exploratory/sense and perfomatory/motor organs [28]. It entertains a range of sensory content from pain to texture (ibid.), and it is found to play a role even in intersubjective relations with others [29,30]. Even the act of hearing is not considered simply an aurally-based phenomenon; music can also be understood as a tactually oriented sensation [31,32].

Contrary to the conventional description of physical space as a continuous dimension, neuroscientific research reveals that the brain's perception of space is far from uniform, featuring multiple representations. They typically delineate three distinct spaces: the 'personal' space, coinciding with the body's surface; the 'peri-personal' space (PPS), i.e., the area surrounding and in close proximity to the body, reachable by the hands; and the 'extra-personal' space, extending beyond what is immediately within reach [33,34]. These discernible spatial sectors are often described in terms of distinct coordinates and 'frames of reference' [35] and are coded by a set of distributed cortical and subcortical neuronal structures [33,34] that rely on the input they receive from multiple sensory channels and modalities [36].

Làdavas & Farnè [37] have indicated the presence of a cohesive system governing visual and tactile inputs in the close proximity of the face and hand, which is functionally distinguished from the one overseeing visual information in the far-peripersonal space. Vision is supposed to offer external or environmental frames [36], while in the domain of touch, the frame of reference can vary according to the task and the individual's posture and may be described as either egocentric, a coordinate system centered on the body, or allocentric, determined by the position of an object external to the body [38]. These concepts coincide with the findings from an earlier study on Dhrupad singing gestures [39], whereby the classification of various MIIO types (such as stretching, pulling, pushing, etc.) was topologically based on the point of reference (arrival/departure point): either in proximity to the heaps (near the torso) with hands moving away from the body or a distant point in space with hands moving toward the body.

In light of these findings, it has been reasonable to suggest that Dhrupad singing involves sculpting interactions with spaces that exist both internally and externally in the body. Apparently, Dhrupad's vocal pedagogy effectively conveys such music-related space-sculpting interactions through mutual tactile experiences that are shared between teacher and student within a common physical space, the so-called 'we-space' [40] or 'interpersonal' space [41]. This is the action space that exists between participants and enables interpersonal multisensory exchange of information. However, little attention has been paid to the role of this shared action space within the realm of music education, particularly within the domain of oral music pedagogy, such as Dhrupad. Additionally, the recently observed shift from one-to-one classes to group-oriented learning sessions, as well as a heavy transition from in-person to screen-mediated communication, such as in videoconferencing, prompts concerns about potential changes or even disruptions in how

participants engage with each other within this shared haptic interaction space, which has not been extensively examined before. In the next section, I raise such concerns.

2.2. Observed Shift in Dhrupad Music Pedagogy

Dhrupad, often considered one of the oldest forms of Hindustani music and approached as a meditative practice (nada yoga or yoga of sound) [42], is characterized by a strikingly low pace and meticulous melodic development. Combined with the required rigor and precision of conveying subtleties in the intonation, practitioners often associate this with a profound sense of 'nowness' and 'hereness' (personal communication). Over the past 15 years, Dhrupad teaching has integrated into global yoga retreats, offering workshops that blend yoga, meditation, and singing. The rise in popularity of so-called world music [43] and of online group music workshops, including Dhrupad singing, has brought about not only financial (Financial: from 'disciple'—whereby students do not pay for classes—to 'tuition'—whereby students pay an agreed sum for a fixed class arrangement [44]) and structural (Structural: from a communal experience of embodiment to an individual focused, analytic and directions-based pedagogy [15]) changes in terms of pedagogy, but also a necessary physical displacement; a student will inevitably be required to sit a bit further away from the teacher to leave space for others in class, as illustrated by a snapshot from a recent Dhrupad music workshop in Figure 2. This transition has the potential to influence how students engage with the teacher's physical movements (visually and otherwise), a topic that the present paper intends to examine.



Figure 2. Pelva Naik and students, Dhrupad singing workshop, Labyrinth Catalunya, Spain, 2023. Recording by the author of this paper.

To further add to this shift, the 2019 pandemic has undoubtedly accelerated a digital 'turn' [12] towards online Hindustani music tutoring, which was only timidly being offered by music teachers in previous years, leading to a surge in online Dhrupad courses. Despite historical reluctance to use technology during music classes (even simple audio recordings) and while many have wrongly forecasted the demise of the tradition because of this, online instruction has admittedly contributed to making Dhrupad more easily accessible to a

global audience of students and has become part of the common tutoring practice. Dhrupad practitioners, notably the younger generation, now happily use the internet and widely available platforms for teaching, diverging from the traditional emphasis on in-person, face-to-face tutoring. However, there is good reason to suggest that the transmission of bodily dispositions may be affected or altered in a number of ways.

Most often, individuals will resort to standard videoconferencing or basic audio calls, thereby missing out on visual engagement. Among those who do use video, some may strategically position themselves in front of the camera, while others may not prioritize it. They may stay close to the camera, displaying only their head, or position themselves further away to make their upper or full body visible, as in Figure 3. It is, hence, only fair to suggest that this shift in two-dimensional visual engagement constitutes another pedagogical shift and may raise concerns about the effectiveness of transmitting salient gesture features in conveying intricate melodic patterns, rhythmic structures, emotional depth, and other stylistic elements. Such aspects are thoroughly examined in this study.



Figure 3. Snapshots of online Dhrupad singing teaching sessions by Meghana Sardar and her student Isadora Reig. Provided with written consent by the participants, 2023.

Screen-mediated communication, such as videoconferencing, has been portrayed as disembodied and emotionless [45], lacking the sensory richness of face-to-face interaction [46]. Criticisms include a diminished sense of social presence [47] in online pedagogy compared to in-person interactions [48] and an increased effort and fatigue attributed to the absence of non-verbal engagement [49]. Given that the peri-personal space (PPS) functions as a multisensory interface (tactile-auditory-visual) connecting the individual and the environment [50], it is worth considering whether technology-mediated online Dhrupad singing classes conducted through screen-based videoconferencing platforms might disrupt the music-related perception of the shared space between the teacher and student that is experienced in in-person classes.

As technological advancements in networked music performance have seen a significant acceleration, extensive work has been published on technical, perceptual, and aesthetic impediments and challenges restricting the overall performance experience [51,52], also including folk music [53]. The surge in distance learning applications and dedicated e-

learning platforms, especially prompted by the 2019 pandemic, has led to a significant body of scholarly literature on distance music education and related pedagogy strategies [12,54–57]. Nonetheless, despite recognizing the significance of gestures in music education, less emphasis has been given to the embodied aspects of the musical experience within the context of virtual education [58]. For instance, Duffy & Healey [57] have reported that the change in communication medium has resulted in the reduced availability of non-verbal cues and perception of relative spatiality.

There is also a very limited body of work on the impact it has had on the transmission and dissemination of oral music traditions in general [59] and specifically on classical Indian music [15,60], despite existing research on the embodied aspects of its education [61]. For instance, Roy [15] examined the structural elements of face-to-face Hindustani music pedagogy and compared them to synchronous online classes under a number of 'tenets' derived from interviews with students and teachers, which include, among other things, extra-musical materials such as gestures and verbal explanations. He observed a pedagogical shift, whereby online classes come structurally more suited for individual, analytic-based, and directions-oriented approaches of a coded music system (notation or verbal commands), moving away from the profoundly implicit, direct, experiential, and embodied approach encountered in the in-person context of the Hindustani oral music tradition. Nevertheless, the publication allocates only a small section to the exploration of bodily dispositions and multimodal aspects in music pedagogy, along with their transmission during synchronous distance tutoring.

In the current paper, I will be taking an embodied cognition stance. Specifically, I will be using the 4E perspective of cognition as a backdrop for structuring the analysis of the interview material and drawing relevant conclusions on the gestural aspects of music transmission in online synchronous Dhrupad music tutoring, in particular for singing. In contrast to the rather obsolete information-processing approaches to mind, which tend to discuss cognitive processes in terms of internal computations and representations, the 4E perspective views cognition as being dispersed throughout the entire body of a living organism and its immediate surroundings [62] and is further explained in the context of musicking in the section that follows.

2.3. The 4E Cognition Framework in Music Pedagogy

The last three decades have witnessed the development of a number of theories concerning what is plainly referred to as embodied cognition. Despite the diversity of perspectives, all these approaches depart from the Cartesian distinction between mind and body in a number of different ways and postulate the understanding that cognitive processes are distributed beyond the boundaries of the brain or the skull [63]. Hence, they emphasize the significance of an agent's physical body and its interaction with the environment in cognitive abilities [64,65]. Drawing on this, the human body has been acknowledged as an integral part of our experiences in music, too [66,67], leading Leman to coin the term 'embodied music cognition' [2].

Many of the more recent perspectives now fall under the rubric of '4E Cognitive Science', an umbrella term encapsulating different strands associated with the development of a 'new science of the mind' [68,69], the most radical form of the embodied cognition paradigm. In general terms, it advocates that cognition is shaped and structured by dynamic interactions between the brain, body, and (both physical and social) environment [63], guided by four main principles: Embodied, Enactive, Extended, and Embedded (ibid.). The Embodied aspect posits that cognition is fundamentally tied to the body, i.e., it cannot be fully described in terms of abstract mental processes and, instead, it is inseparable from and influenced by bodily conditions; the Enactive aspect suggests that it is not limited to mere contemplation but that it involves an adaptive two-way exchange with the environment; the Extended posits that it goes beyond the boundaries of the body to interact with external resources, both material and symbolic, for mental offloading purposes; finally,

the Embedded aspect suggests that it is not isolated, but situated within and supported by the environment [62,70,71].

This 4E cognition paradigm has also resulted in a more profound understanding of how the body plays a fundamental role in musical comprehension, sense-making, expression, communication, and creativity. Recent developments in the domains of music improvisation and participatory music making [72], as well as music education [71], have undergone a transformative shift influenced by the principles of 4E cognition. In accordance with the viewpoint put forth by Casas-Mas et al. [73], I aim to employ the 4E cognition framework in order to structure the thematic analysis of the interview material related to synchronous network-based Dhrupad music pedagogy. The subsequent section will delineate the methodology I employed in pursuit of this endeavor.

3. Materials and Methods

The present study employs a qualitative analytical approach, typical in ethnomusicological research [74], deemed suitable given the absence of an a priori theory from the existing literature, with the aim to develop a reasonable understanding of the implicated perceptual phenomena [75]. Specifically, it relies on semi-structured interviews to uncover common underlying concepts that are either explicitly stated or implicitly woven into cognitive thinking. Thematic analysis is applied to originally collected interview material, a qualitative data analysis method that entails reading through the dataset to identify, analyze, and interpret patterns and overarching themes in the meaning of the implicated data [76]. The objective is to gain a deeper understanding of the implicated sensory modalities and to identify links between musical attributes and other somatosensory elements involved in online teacher–student interactions. A distinctive aspect of this approach lies in organizing the emerging themes according to the principles of the 4E cognition framework. The conclusions are drawn not only from each of the overarching themes identified in the analysis but also from recognizing the interplay between individual themes and from situating them within the 4E cognition paradigm. Therefore, despite potential overlap, the analysis is presented according to the 4E principles. To enhance the narrative, representative quotes from participants have been included in the presentation of the analysis.

This study draws on material from a number of semi-structured interviews I conducted during 2023 with a cohort of Dhrupad practitioners, predominantly of the younger generation, namely Pelva Naik, Meghana Sardar, Cellin Wadier, Markus Schmidt, and Marianne Svašek, as well as their students Isadora Reig, Gauri Raje, and Aarya Patwardhan. Besides the interview with Naik that was held in person during a workshop at Labyrinth Catalunya in Cardedeu, Spain (5–9 April 2023), the rest were conducted online. Participants were selected based on their substantial engagement with both instructional settings of Dhrupad vocal pedagogy, in-person as well as videoconferencing, rendering them well-equipped to provide insights into both cases and offer comparative perspectives. For consistency reasons, all participants belong to the so-called Dagar gharana (stylistic school), indicating that they are practitioners of the musical tradition associated with the Dagar family. Whenever relevant, participants were asked about their experiences as both teachers and students. Table A1 in Appendix A compiles all pertinent interview data.

The interviews were conducted exclusively in English. Participants were only informed beforehand that the project focused on online Dhrupad music classes, while some were aware of my previous work on body movements in Dhrupad vocal music. The discussion was initiated by a rather general and open-ended question, such as inquiring about their experiences with Dhrupad singing or their level of expertise and years of training. Subsequently, follow-up questions were tailored based on the interviewee's responses and insights. All participants provided written informed consent and retained the right to discontinue their participation at any point during the interviews. No compensation was offered for their involvement.

The upcoming section delves into the structural elements of online synchronous Dhrupad vocal pedagogy, which are organized under the four principles that comprise

the 4E Cognition framework. In practical terms, the guiding tenets that were identified in the first part of this study concerning a similar approach to in-person Dhrupad vocal pedagogy were revisited and applied to the context of synchronous online instruction. The challenges associated with each spatial facet mentioned in the Introduction ('personal,' 'peri-personal,' 'extra-personal') are discussed separately under each individual principle of the 4E paradigm, while the comprehensive spatial perception of the online context will also be consolidated in the Discussion.

4. Interview Analysis—Results

> Marianne Svašek: 'I had a very deep connection as if we were sitting in one room'. (interview, online, 15 September 2023)
>
> Celinn Wadier: 'If I had discovered Dhrupad online, maybe I wouldn't have continued'. (interview, online, 15 September 2023)

From analyzing the interview transcriptions, two distinct "schools of thought" seem to emerge, revealing contradictive perspectives with respect to the online synchronous teaching of Dhrupad music. The two schools are nicely represented by the above-mentioned statements by the Dhrupad vocalist Marianne Svašek and her disciple Celinn Wadier, an experienced Dhrupad vocalist and teacher herself. One 'school' seems to emphasize the importance of content over medium, arguing that the quality of musical material and instruction is paramount, regardless of the delivery method. For instance, Marianne Svašek, Markus Schmidt, and Meghana Sardar contend that, with proper adaptation and utilization of technology, students can achieve similar musical proficiency as in traditional in-person classes. Proponents of this viewpoint suggest that technology can enhance the accessibility of singing classes, making them more inclusive, flexible, intimate, and even more 'accurate,' a term that was used by Svašek to refer to a heightened focus with minimal distractions. Svašek also reflects upon the following:

> 'One of my students said, "I want to see you online because I missed something". So, I started to think, "OK, what do you miss? What do you want from me more than material music and the whole vibe"? See, the interesting thing is you can have a very deep connection with the person, as if you are together, it is very much possible. If you're both very concentrated, you just have the communication. [. . .] So, it is a human thing in between two people. Not the matter, not the teaching, not the subject, not the material'. (Marianne Svašek, interview, Rotterdam, the Netherlands, 27 September 2010)

However, this perspective may overlook the experiential and sensorial aspects of singing education that go beyond mere content delivery.

Conversely, another 'school of thought' is concerned about the potential limitations of technology-mediated singing classes in terms of providing a rich, multimodal, sensorial, and embodied experience, which is considered to be as musically meaningful as the delivered content itself. The term 'experience' is employed to shift the focus away from the ideal and the ideological, directing it instead toward the phenomenal 'what people actually do, and what they feel, while engaging in music' [77]. Researchers within this camp argue that singing is not just about learning the material (notes, melodic treatment, and rhythms) but also about developing a deep connection between the body, voice, and emotions. For instance, Pelva Naik, Celinn Wadier, and Gauri Raje assert that the physical presence of an instructor during in-person classes allows for nuanced feedback and the embodiment of vocal techniques that may be otherwise challenging to replicate through screens. For them, technology may fall short in capturing the subtleties of body posture, facial expressions, and tactile sensations that are essential elements in the art of singing. For example, Wadier—Svašek's student and a Dhrupad teacher herself—posits the exact opposite opinion to that of her teacher:

> 'It depends actually on the person, on the student, but it's not so precise and easy online. [. . .] I can learn! You can appreciate the music without the movements,

> when you listen to a recording for instance. [But] if you want to sing, then it's good to see what's going on. You need to look and to be in the presence of the person to learn!' (Celinn Wadier, interview, online, 15 September 2023)

This perspective moves beyond common implications of video-conferencing platforms, such as temporal delay, poor audiovisual quality, and inadvertent disconnections [78], and mostly highlights the need to carefully consider the sensory and embodied aspects of singing education, that are rendered 'less precise' due to the technological mediation.

The forthcoming analysis will primarily concentrate on the second group of participants.

4.1. Embodied Aspect

In the realm of music education, the embodied aspect of 4E cognition reflects the notion that music learning and knowledge are deeply rooted in the physical and sensory aspects of the human body. Through the embodied principle, I aim to explore challenges in instructing and perceiving postures as well as covert aspects of embodied engagement in online classes, specifically those related to the mechanics of vocal production.

4.1.1. Challenges in Physical Engagement as a Rich Sensorial Experience

Multiple reports from participants indicate that the virtual environment of online lessons restricts the transfer of nuanced singing gestures, which are essential for conveying vocal techniques and expressions. They suggest that online learning lacks the multi-sensory richness of the in-person context, where various senses apart from the auditory, such as sight (peripheral too) and touch/sense of vibration, contribute to a deeper understanding of the music material. This limitation affects both the teacher's ability to effectively demonstrate and the student's capacity to accurately replicate these gestures.

Gauri Raje, a student of Dhrupad vocalist Pelva Naik, agrees with this notion:

> 'I didn't like it [online class] for the first maybe two or three months. It was very disembodied, the whole experience'. (Gauri Raje, interview, online, 18 September 2023)

Similarly, Wadier seems to recognize the significance of this multisensory transmission process:

> 'You see, [in-person] you have a picture much more precise of what's going on. So, I can see also what's going on with the jaw etc. [. . .] If I learn with Marianne [Svašek], or if I learn with ustad (guru, here referring to Zia Fariduddin Dagar), we were having this whole picture of the energy and gesture with the voice and the ragas. [. . .] You have the voice, you have the sound, you have the vibration. And you know, also for the breathing. This is something that I also need to share. Well, it's much better in person, because the sound, you hear much better'. (Celinn Wadier, interview, online, 15 September 2023)

Gauri Raji, a professional storyteller, appears to possess a deep comprehension of the complex multisensory engagement that occurs in face-to-face interactions. This engagement typically extends beyond direct eye contact when in close physical proximity but is inevitably replaced by exclusive visual cues in online environments.

> 'I think the visual engagement is slightly different for me when a person is live in front of me. [In-person] I don't need to look at the person in order to sense them. When I'm working with Pelva live, I'm in her house and learning from her, I'm not always looking at her. Because a lot of the times when I sing, my eyes are closed or they're looking elsewhere, they're not looking at Pelva. But when I'm online, that element for me is lost. I can't hear a person breathe. If they shift their bum slightly, I can't feel it. So, that source material is lost. And then visual has to become much stronger for me'. (Gauri Raje, interview, online, 18 September 2023)

The limited gestural perception complicates the identification of issues related to the mechanics of sound production, too, like body posture problems. These challenges are pertinent to both instructors and students. For instance, Dhrupad sitar player Markus Schmidt brought up concerns related to the visual perception of subtle details in gestures and postures, even in scenarios where the upper body is visible. They include the inability to observe individuals from multiple perspectives, as is the case in the real three-dimensional space, and the necessity for a selective visual reading of the student's gestures at each moment instead of the comprehensive visual engagement of the in-person context.

> 'You cannot see how he's sitting and so on. So, it's way more difficult to keep the whole picture. You have to really focus on different things at different times. And then the other thing is I cannot watch from different angles, [. . .] I just see the front side'. (Markus Schmidt, interview, online, 11 May 2023)

Aarya Patwardhan, a young student of Naik, expresses the same concerns from a student's perspective:

> 'Then my body posture, I found out that it was not good for Dhrupad. I didn't sit straight. That caused the voice projection that resulted. Madam corrected me when I was in the workshop. But in online, she couldn't notice that, I feel'. (Aarya Patwardhan, interview, online, 16 September 2023)

4.1.2. Challenges in the Embodied Conception of Personal Space

Conveying covert aspects of embodied engagement, such as the deliberate and active control of the abdominal muscles that are ideally involved in supporting singing, poses inherent difficulties even in traditional in-person instruction.

Wadier shares her perspective as an educator on the amplified intricacy of imparting such instructions during online sessions:

> 'Well, it's much better in person! Also about the gesture: It's not only the arms who are moving, it's also the heaps. [. . .] If I teach online, sometimes they can see how my hand moves, but then they don't see me doing anymore. . . [she lowers the camera to the abdomen]. There is something happening also with the heaps and how you are carrying yourself. So, it is more difficult to see how the student is carrying himself or herself online, because you don't have this full picture'. (Celinn Wadier, interview, online, 15 September 2023)

Her perspective is also echoed by students. Aarya Patwardhan recounted how her instructor was only able to recognize her voice projection technique issues when they had an in-person meeting. The limitation in identification appeared to stem from both the visual and acoustic shortcomings of online sessions compared to the in-person interactions:

> 'The voice projection and all, and the frequency is understood much better in person, I feel. Actually, I experienced this this year. In online, when madam teaches me, I'm singing what I feel right. But when I went to the workshop, I found out that I didn't have that voice throat; madam says that you have to build your core, so that your voice projection is much better. So, actually, I was singing all right, but then the projection that is needed for that particular music form, is not right. That's what I found out. In-person class is important, because I didn't know that in the online classes. I found out when I went to the workshop. That doesn't happen in online classes. (Aarya Patwardhan, interview, online, 16 September 2023)

4.2. Enacted Aspect

In the context of music education, the enacted aspect highlights the concept of action-as-perception [79], whereby knowledge is transmitted from teacher to student through gestural emulation. In this context, I aim to address challenges associated with both the

teacher's demonstration and the student's imitation of overt gestural engagement with the instructed musical material during online classes.

4.2.1. Challenges in Goal-Directed Manual Interactions

Multiple statements by participants underscored the impact of restricted gesture perception on the process of gesture inheritance, which is a fundamental aspect within traditional pedagogical contexts and is typically achieved through gestural imitation and teacher–student interaction.

Naik elaborates on her students' experiences, emphasizing that it is only during in-person face-to-face interactions with the teacher that they fully grasp the embodied aspects conveyed by their instructor and genuinely begin to actively connect with the melodic ideas through gestures:

> 'They [students] take much less of that [gestures in online classes]. Much less! Because I've seen there are some students who after, say, a year of online come and meet me in person and then they realize "Oh, wow, it's like this, you know?". And then they start doing this and that [she gestures in the air] and they are like "this has changed so much for me. I didn't know this for a long time". [...] I don't think so [that this gets across through the video]'. (Pelva Naik, interview, Cardedeu, Spain, 9 April 2023)

Similarly, Isadora Reig attests that bodily movements, which she acknowledges as important during in-person classes and even when watching videos of performances online, are most often not visible during online classes:

> 'We do classes with just WhatsApp video calls. So, most of the times I just see her face. So, in this sense, the online classes are really limited with that. I mean, I cannot see this whole thing, which, when I was singing with Pelva in this class [in-person] it was really amazing. And the other thing, [...] the two first years I was studying Dhrupad, I was watching a lot of videos of Uday Bhawalkar, for example, of the Dagar brothers as well. And it was very important to me to grasp..., it was one thing that I was really paying attention; what is the movement! Sometimes they are very subtle, they are just little movements. I have realized that, for example, [by] seeing Pelvis singing in concert or the meeting or the thing that you did with her and not the way that she was explaining things, I understood a lot of things there. Because I saw the rest of it'. (Isadora Reig, interview, online, 7 May 2023)

French Dhupad vocalist Celinn Wadier concurred with the aforementioned sentiments when expressing her perspectives:

> '[In the traditional in-person setting] we were like kind of imitating or interacting and that was a way of learning also. This is very, very important and that's for sure: online is not that easy to do. You have to guess!' (Celinn Wadier, interview, online, 15 September 2023)

The use of the verb 'guess' in this context implies the act of compensating for the missing information by incorporating various forms of other multimodal data, a process that will be discussed in more detail in the next section.

4.2.2. Indirect Inference of Missing Multimodal Information

Interestingly, when questioned about the body parts that contribute the most to her perception of gestures while observing her teacher in class, Raje acknowledged that she may even indirectly glean certain information about parts of the body not visible in the camera, such as the abdomen, from other cues:

> 'The other thing is that this part [shows shoulders and head], this much can be really expressive as well. So, there are times where you're trying to hit something, but you're actually going back. And you know there's that tension that is then

> conveyed. So, that I'm not just picking up on that, but I'm also picking up on how Pelva's neck is moving, how her shoulders are moving. Is it going forward, is it going back? So, I'm getting the sense of how her abdominal moves, for instance, from that'. (Gauri Raje, interview, online, 18 September 2023)

Likewise, Reig describes how she has incrementally become increasingly adept at discerning a range of information solely from facial expressions:

> 'I have learned to see a lot in the face about the singing. [. . .] I have become used to see a lot in the face, different notes'. (Isadora Reig, interview, online, 7 May 2023)

Nonetheless, she also asserts that this provides only a restricted perspective within her comprehension:

> 'Seeing Pelva singing in concert or the workshop I understood a lot of things, because I saw the rest of it. [In the online], I can see her face singing, so I can see somehow the movement reflected in the body, but very limited. But when I can see Pelva singing or in one-to-one class, it's like, "oh, wow, this whole thing, this other whole thing"'. (ibid.)

4.2.3. Sense of Presence and Connection

Most participants agreed that online teaching sessions are stripped of essential aspects of human social interaction. Recurring themes of online teaching challenges during interviews included the concepts of 'presence' and 'connection.' All participants agreed upon the fact that visual cues allowed for an—even if only shallow—feeling of connection.

The term 'presence' was used recurrently by her teacher, Pelva Naik:

> 'It is not just about the sound. That's something I've learned. So, it's also the "presence". So, the video helps only to make you feel that "ok, that person is present". Both ways. So, it's like a make-belief. You feel that the presence is there, but it's not the same as being in person. So, that's very distant'. (Pelva Naik, interview, Cardedeu, Spain, 9 April 2023)

Aarya Patwardhan, a student of Naik, attributes her limited comprehension in online teaching compared to in-person classes to a diminished sense of connection:

> 'I feel that when we have the classes in person, we understand more. The teacher also finds it easy to explain that particular topic and even the student understands much better than the online classes. [. . .] We get connected to the teacher when she is sitting just right in front of us. That connect between us, between me and madam makes me understand that particular concept much better'. (Aarya Patwardhan, interview, online, 16 September 2023)

Isadora Reig, Sardar's and Naik's student, attributes the sense of presence to the visual engagement with facial expressions when the full body is not visible:

> 'With the WhatsApp call I don't see the hands, I don't see the body. I don't see nothing. Just the face. But somehow the face is close. You are close. So, it was important'. (Isadora Reig, interview, online, 7 May 2023)

She contends that the limited feeling of immersion caused by the inadequate sensory exchange results in a restricted musical learning experience, making in-person interactions a necessity. Within this context, visual engagement with facial expressions provides at least a superficial sense of proximity.

> 'Watching a lot of recordings, a lot of concerts [of Dhrupad singer Uday Bhawalkar], has been really important for me, because I have been totally able to watch the movement, to study it and the note and the ragas and to get immersed. But with online classes no! It was more important to have this closeness. But the moment that I was seeing Meghana in these workshops, it was really making a difference. Like, "oh no, I need to be more close to it"'. (ibid.)

When I inquired about the potential to move farther away from the camera to ensure the entire body and associated gestures were visible to others, the majority of responses from participants expressed concerns about the limited visual clarity inherent in teleconferencing technologies compared to in-person interactions; many believed that such a move would only intensify the preexisting feelings of disconnection and distance brought about by technological mediation. For instance, Raje responded:

> 'In order to get the whole body, she [Naik, her teacher] would have to move away from the camera, and that gives it a kind of distance. I think what I appreciate is the kind of sense of closeness that I get, and with that, I'm picking up from this bit. I don't know if I would get anything more by seeing the whole body'. (Gauri Raje, interview, online, 18 September 2023)

Along the same lines, Wadier responded:

> 'Yes, maybe. Knowing that I have to make sure that I can hear them good. And then I would be a bit far, sort of. Because we are already separated like this on the screen. So, the energy is really difficult to feel. But I have to see the student also, that's the thing. If I'm going to sit far away, I don't see the student very well'. (Celinn Wadier, interview, online, 15 September 2023)

Finally, Sardar also agreed that the visuals are mostly there to keep the feeling of connection rather than offer an opportunity for meaningful engagement, similar to the liveliness of in-person experiences.

4.2.4. Visual Dominance, Fatigue and Stress

The interview material indicates a visual (and auditory) dominance over other sensory modalities in the online teaching context, which is attributed to the need to compensate for the missing sensorial information. Reports also suggest that this heightened reliance on visual cues may lead to increased levels of teacher and student fatigue.

> 'I think online over time I've learned to very quickly pick up clues visually, so that now it's almost second nature. But [. . .] I can't be online more than two hours, I just zone out. That's not to say if I'm in meetings in-person, I don't zone out. But part of my body is still working, and I will still recall things later on. It is more exhausting. Definitely'. (Gauri Raje, interview, online, 18 September 2023)

Additionally, according to Reig, the heightened sense of closeness conveyed through facial expressions during online sessions can become so intense that it becomes stressful and almost uncomfortable. At least, it is adjustable over time:

> 'And at the same time it's a thing... For example, first two years I was really struggling with it. She was looking me sometimes, because I was out of tune or I was trying, but I was really nervous. And it was difficult to have the face here [so close] because it was very clear. [. . .] Just the few moments that we needed to change to just audio call, I completely relaxed. It's like, OK, I don't have the face here [close to mine], like "what are you doing?" But this was at the beginning'. (Isadora Reig, interview, online, 7 May 2023)

4.2.5. Compensation through Increased Use of Verbal Communication and Visual Aids

To address the deficiency of embodied teacher-student interaction in the online context, teachers appear to expand the use of verbal communication and visual aids.

For instance, Pelva Naik, an experienced Dhrupad vocalist, mentioned a heavy reliance on linguistic metaphors and visual descriptors of musical imagery to counterbalance the restricted multisensory information exchange resulting from technology-mediated communication:

> 'In my own classes what I've started observing with myself is that I am double-explaining the visuals; much more than when I am in normal classes. Because

> there I feel that that will help. I really go into the whole metaphors, visuals and all that. I end up talking, explaining so much more. When you are in person, you don't really have to explain too much in words. You can exchange it with energy in space and just by singing'. (Pelva Naik, interview, Cardedeu, Spain, 9 April 2023)

Dhrupad vocalist Celinn Wadier expresses a similar sentiment in her reflection:

> 'I have a tendency to explain more online. I can also give explanations in the in-presence. But sometimes I have to put into words more online, because you need to. Though in-presence sometimes you don't need to, you can show more'. (Celinn Wadier, interview, online, 15 September 2023)

4.2.6. Challenges in the Enacted Conception of the Peri-Personal Space

Drawing on feedback from participants, it is commonplace for cameras during video calls to be set up in a way that they focus primarily on the face or the uppermost portion of the torso, just above the shoulders. This setup fundamentally leads to a restricted view of one another's body movements and gestures, rendering the enactive aspect associated with action-oriented intentionality less readily discernible to students in the online environment. 'Gestural confirmation' of musical concepts is therefore absent or limited, according to Gauri Raje, Naik's student, and the perception of the space surrounding the body (PPS) is obscured. Raje further explains:

> 'But online, usually it's this [she shows just the head], so you don't see the whole body. So, you don't know just how much of the bhav (Expression in Hindustani music) there should be, or just how much of the delicacy or just how quickly she is lifting something or how sharply she's going somewhere, that I think it's the gestures that kind of confirm things for me'. (Gauri Raje, interview, online, 18 September 2023)

Similarly, Isadora Reig, a student of Meghana Sardar who has also attended an in-person workshop with Pelva Naik, recalls that most often, online classes are limited to making only the head visible, thereby neglecting significant aspects of enactive and embodied understanding of music content. These aspects are fundamental in the in-person context, such as a workshop with Pelva Naik she had previously attended:

> 'Yes, it [paying attention at the gestures] plays a role, but with the online classes, WhatsApp video calls, most of the times I just see her face. So, in this sense, the online classes are really limited with that [gestures]. I mean, I cannot see this whole thing, which, when I was singing in this class with Pelva [Naik], it was really amazing'. (Isadora Reig, interview, online, 7 May 2023)

4.3. Extended Aspect

In the realm of music instruction and learning, the extended aspect of 4E cognition refers to the capacity of musicians to encompass external resources, such as musical instruments or tools, that allow the offloading of certain tasks and enhance music expression and performance. In Dhrupad vocal music, these tools can be stretched even further to entail imaginary objects that musicians are frequently observed employing during improvisation (MIIO: Manual Interactions with Imaginary Objects [18,39]). Thus, within the framework of the extended principle, I explore challenges associated with instances where singers interact with melodic ideas through gestures that appear to allude to the sensation of various resistive forces, such as elasticity through stretching–compressing, pressure through pushing, weight through lifting, etc.

4.3.1. Challenges in Manual Interactions with Imagined Objects (MIIOs)

The restricted ability to transmit the subtle nuances of singing gestures in online settings appears to impact the utilization of MIIOs and the sensation of forces as well.

While these gestures are still visible, they are not as clearly discernible as they are in face-to-face interactions. Hence, they prompt instructors to rely on verbal assistance for conveying these aspects, as in the case of Wadier:

> 'I try to transmit it as much as I can online, and it's for the comfort of the person. So, I speak about it and I try to show it a bit like. . . I use images, like "you imagine you have an elastic or you imagine you have some warm sun here or a cat purring like this and you just put your hands like this". I also tell it when I am in presence, but they can see also. They can see it better, and they can also see the way I move'. (Celinn Wadier, interview, online, 15 September 2023)

According to Raje, a student of Naik, gestures—and more specifically, MIIOs and the associated sense of forces—help her discern vital expressive musical components, like the nuanced handling of notes or melodic glides. These elements are often conceived of as almost tangible interactions and are not solely conveyed through visual means in genuine physical spaces. However, the restricted transmission of diverse sensory details in online settings prompts Raje to confine her interpretation of her teacher's MIIOs solely to visual cues:

> 'And I think online makes a big difference. Because when I'm live with Pelva, I can feel the heaviness or the gravity. In-person, even if Pelva didn't make the gestures, there would be a certain way in which the breath moves in her body, that if I'm in her presence, I sense it and I can see where she's going. You know, there's something about that body sensory communication, which sometimes gets filtered out online. I think online the gestures have really helped to communicate the sensory element. So, Pelva's gestures help me understand—say—the weight, the bhav that she gives somewhere or the kind of meend (Pitch glide/glissando) that she is trying to communicate, the delicacy of the meend sometimes, that with the gestures, becomes more present for me. [. . .] I can catch that a little bit more'. (Gauri Raje, interview, online, 18 September 2023)

4.3.2. Challenges in the Extended Conception of Peri-Personal Space

As participants reported that cameras most often capture only the face, resulting in the limited visibility of each other's gestures, it becomes challenging to perceive a sense of peri-personal space.

> 'Even though Pelva would be making gestures, I just wouldn't have a sense of things around her'. (Gauri Raje, interview, online, 18 September 2023)

This is also confirmed by Pelva Naik, Raje's teacher, who reports an interesting technology-reasoned shift that seems to take place in her pedagogy, whereby the visual engagement with her own video appears to affect her gestural instructions toward students online. Instead of occasionally engaging with the melodic content through manual interactions with imaginary objects and the sensation of forces in the wide three-dimensional space around her, she mostly seems to rely on two-dimensional trajectories or geometries in parallel to the computer screen.

> 'My way of teaching has changed a lot, because that [video] frame is there. I am looking at myself in the frame and so I use it as a board, so I am. . . [singing notes] 'Sa–Re–Ga' (1st–2nd–3rd degree of the scale in Hindustani music), so I am not doing this [MIIOs] any more when I am doing alap (improvisation) or when I am explaining, because I am looking at the person, but I am showing it, as if I am making a painting'. (Pelva Naik, interview, Cardedeu, Spain, 9 April 2023)

This may indicate a potential change of gestural engagement from three-dimensional MIIOs [18] to two-dimensional melographic representations [9] resulting from the constrained visual affordances of the two-dimensional videoconferencing platforms and affecting the use of the peri-personal space.

4.4. Embedded Aspect

In the realm of music education, the 'embedded' principle of 4E cognition can encompass various aspects. It may refer to the physical limitations imposed by an instrument, the societal and cultural norms that musicians are expected to follow, and even broader ecological factors such as the qualities of the physical or technologically-mediated environment, which include acoustic and visual elements. In this context, I will explore the challenges associated with the perception of extra-musical 'enculturation' influences [16], such as the level of familiarity with each other's personality and the perception of the surrounding environment.

4.4.1. Conceptual Adaptation

Several reports by the participants suggest that one's understanding relies on a complicated system of (multi-modal) information that, contrary to the in-person, does not become fully transmitted online. For instance, Markus Schmidt posits the following:

> 'Getting together with somebody in person is a very complicated system of all kinds of information and just a very little part of it gets through to your conscious mind. The other is absorbed unconsciously. And there are many things that you will miss when you just have a Skype conversation'. (Markus Schmidt, interview, online, 11 May 2023)

However, various interviews indicated that online participants' interactions and engagement improved gradually as they developed a deeper understanding of the other person and his/her contingent presence within the surrounding environment, even if through the virtual context. These reports suggest a contextual adaptation to technology mediation that helps alleviate the initial apprehension of participants regarding the use of online means for music education.

Raje serves as a notable example, whereby she explains that gaining a deeper understanding of her teacher's environmental context and personality over time has enhanced her ability to discern her teacher's music-related gestural cues, thereby contributing to her own musical comprehension and increasing her comfort in technology-mediated interactions with her teacher. Significantly, her gradual deeper understanding was largely shaped by her prior in-person experiences and the time they shared for activities beyond strict music instruction.

> 'In the beginning, when I started learning with her, I didn't have a sense of how she was, or who she was. I didn't know her at all. And then, now, over time, I have a better sense. Also, because I've been to her place now. It helped. Because I know this place she is in, from where she is, kind of. I think over time, there were times where we just talked in sessions we didn't sing, or sometimes we joked around, or we laughed. And that gave her body more of a materiality. I just get a sense of her body more and more. So, then it's not as strange anymore. So over time, even though I had not been with her in person, I had a sense of her as a person. I kind of have a memory of the weight, when she sits her whole body, what's the weight like. Or when she shifts a little bit. What is that shifting energetically? I just have a sense of it a little more now. Now, I'm used to Pelva, because I've worked with her for so long. So, I think I've arrived at a place where I'm comfortable with online'. (Gauri Raje, interview, online, 18 September 2023)

4.4.2. Challenges in the Embedded Conception of Extra-Personal Space

When Naik was questioned about the primary limitations she would like to see addressed in online classes, none of the common technical issues, such as sound quality, internet disconnections, lag, etc., were mentioned. Instead, her primary concerns revolve around space-related issues, particularly focusing on challenges posed by spatial constraints, which she associates with feelings of fatigue linked to visual dominance in videoconferencing.

> 'You know, emotions you can transfer through sound only. It's possible to do it even without visuals, [. . .] a Whatsapp audio call. But these other things work with space. And these other things don't happen [in online classes]. [. . .] There is a lack of the energy. So, there is a lot of effort, so it can be tiring, but the visuals help. [. . .] And you can't just go on teaching alap for 50 min or 1 h of class. But you can do that when you are in person somehow. Because there's this availability. It's vaster somehow. [. . .] It's like you are on a bigger vehicle, you know. And you can have that space. So, it feels more spacious working on alap and a raga in person than when you have one hour of [online]. So, that one hour is too much effort than two hours of just sitting and working on that rag in person'. (Pelva Naik, interview, Cardedeu, Spain, 9 April 2023)

Naik goes further, expressing her wish to expand the space between herself and her students to enhance interpersonal interactions. Specifically, she refers to the limited field of view resulting from the constrained frame of the video window that dramatically reduces the sense of surrounding space and the absence of three-dimensional vision, currently necessitating the imaginary reconstruction of the missing depth information. Finally, she refers to the discrepancy between visual and sonic cues, which renders the perception of space dubious and leads to a deprivation of the sense of proximity and the capacity of spatial localization. She concludes with her improvement requests in order of importance:

> 'What I would like to minus is the frame! It's like you are looking through prison window. So, maybe if I had a space where I can't see the frame? Just larger space! Where there is no [. . .] frame between the two people. Or at least there is an illusion that there is no frame. I mean my eyes are not disturbed by looking at the limited space between me and my student. [. . .] What I lack is the horizon! [. . .] It's not a problem, but you make believe a lot. You start imagining a lot. So, what you're doing is you are trying to go as close as possible to the in-person. [. . .] Doesn't matter if the sound is not as good also. [. . .] That was secondary. First, is this nearness. And it has to match the sound. So, if the sound is near, I want you near. Otherwise, my brain is starting "Is that really you? Or it's something else?" [improvement requests] Better sound that matches that ambience, definitely; that space! That larger space. And no frame!' (ibid.)

Feedback from various Dhrupad participants demonstrates an elevated awareness or heightened sensitivity to the audio-visual perception of space. Wadier attributes the inferiority of technology-mediated lessons to the spatial aspect and touches upon the lack of spatialized audio in the acoustic perception of the surrounding space. She links this also to the perception of gestures and our intricate somatosensory system, particularly emphasizing the sense of touch:

> 'There is the space involved, there is the acoustic also. When I speak with you now, my breathing is not that clear also. And the silence in between the phrases, which we can call space as well, it's also a part of the music, that is teached. [. . .] It's like I'm full ears online. And it's flat! It's very tiny. [. . .] Many times, actually, movements, they are quite subtle, and it's not that easy to see it, if you're not there actually directly. [. . .] So that was always much better [in-person]. [. . .] If she [her teacher, Marianne Svašek] is in front of me [in-person], it's three-dimensional. It's going many places and I can feel her movement, I can feel the emotion, I can feel some energy, and this [sense of space] I get. This I don't get on online, I don't get same, I mean. [. . .] And there [in-person] we share. . . It's not that obvious: We don't touch it, we don't touch each other, let's say, but even if it is from here (she points at her ear), this is obvious that there is a big difference. If I do this (she touches the screen), onto the screen, you don't feel anything. And this is also different in in-presence and in online. But this I can really feel more as a student. [. . .] I also wonder how you feel the presence of the person. . .'. (Celinn Wadier, interview, online, 15 September 2023)

When asked if this is a practical thing that could be technically corrected, interestingly, she reverts to discussing the concept of the shared physical extra-personal space that surrounds the two people interacting with each other when physically together:

> 'If technically it's perfect, like you have a big screen and they see you entirely and you see the person and you hear it good, for sure it's better. But the simplicity of the in-presence and the space between. . . the space between the people also, to me it's also part of the transmission. In *this* music'. (ibid.)

Finally, Reig raises similar concerns related to space constraints rather than solely focusing on sound. She emphasizes the significance of melding the perception of a broader space, unconstrained by framed video, with the precision of close visual engagement achievable only in real, physical environments. She immediately highlights that the concern is not merely about increasing the image size but rather about the sensation of proximity that can only be achieved by transitioning away from two-dimensional mediation:

> 'I could say I need maybe to see more space. But for example, Meghana, she uses to do three or four times a workshop with more students. And in this case she uses a big screen and you can see the whole body and you can see the movement. But because she needs to be with more distance from the camera, at the same time it's not that close as when I sing face-to-face, which is really. . . You can see the entire body, her sitting or playing the tempura maybe, you can see the movement. At the beginning I thought, "Oh, wow, that's cool, because I can see the movement". But then it was like, "Oh, no. But it is distant as well". So, this scene when face-to-face, I cannot see the movement, but there is something that is close'. (Isadora Reig, interview, online, 7 May 2023)

4.5. Future Technology Interventions

Towards the end of each interview, participants were questioned about their take on future technological interventions that could potentially enhance the online music education interaction experience. Most of the participants this paper has dealt with reported that they would embrace any promising technological advancements.

For instance, Isadora Reig emphasized the importance of embracing cutting-edge technologies that can offer both a sense of expanded space and high visual resolution concurrently; the former contributes to a profoundly immersive experience and a sense of proximity versus distance, while the latter ensures a clear visual resolution of nuanced bodily movements as is the case in close-up visual engagements in real spaces.

> '[I would need] something to be able to see the movements. But not seeing the face, not seeing in the camera... Something more like VR reality, for example. Something you can see the movements and you can see the thing, but you don't see it in one big screen. I think I will need additional space, to give more deep experience of it'. (Isadora Reig, interview, online, 7 May 2023)

Much like Reig, her teacher Pelva Naik ascertains that she would be in favor of new technologies if they could enrich the sense of presence and reduce the levels of mental effort required, which is only possible by removing the video frame, expanding the perception of space, moving from two- to three-dimensional visual engagement, and finally by achieving audio-visual congruency.

Gauri Raje's primary focus revolved around advocating for the integration of multi-sensory elements that are currently missing in technology-mediated communication. She considers the sense of smell to be the most essential among them, emphasizing that its absence leads to a disembodied learning experience.

Among the other participants, some stated that they do not miss anything with the currently existing technologies. Others expressed skepticism that new technologies would never measure up to the in-person experience, while a few asserted that the auditory channel suffices for conveying necessary information and that embodied aspects are only peripheral and not integral to the educational process.

5. Discussion

Two different perspectives have arisen from interviews with musicians who were queried about their experiences in online versus in-person settings, with one emphasizing the significance of content over medium, while the other is equally focused on the potential constraints associated with multimodal, sensorial, and embodied experiences. Despite the relatively small sample size of participants, it appears that senior musicians, including both teachers and advanced students, tend to align more closely with the first category (focus on content), while beginner students predominantly fall into the second category (stress on sensory experience). One plausible explanation for this phenomenon can be found in theories of embodied cognition, which propose that the perception and performance of music constitute active processes of exploration that draw upon an individual's accumulated "reservoir of sensorimotor skills" [79], developed over time through interactions with their environment. The subtly varied instances within these recurring patterns of sensorimotor experience contribute to the formation of cognitive structures [80]. Therefore, it stands to reason that the extensive interactions senior musicians have previously engaged in, whether with fellow musicians, mentors, or students, contribute to the development of robust connections between physical movements and musical sound, fostering a heightened understanding of musical affordances [81].

Consequently, senior musicians possess a more extensive toolkit for tapping into their previous encounters, enabling them to compensate for the deficiencies of particular sensory input and thereby mitigating the hurdles posed by technology utilization and geographical separation. In contrast, novice students exhibit a preference for grasping musical concepts in face-to-face settings as they struggle to replicate the absent sensorial and embodied dimensions online, stemming from their limited exposure to previous multimodal experiences. This postulated interpretation gains additional backing from recent discoveries [82], which highlight that engaging in multisensory learning enhances subsequent memory performance, even when considering distinct unisensory elements, such as those present in a teleconferencing environment. Ultimately, the ongoing debate underscores the importance of striking a balance between content delivery and the holistic singing experience according to the custom educational needs that fit every single student.

The paper has focused primarily on the second category. Findings are here outlined and discussed individually.

- Discussed under the embodied principle:
  - Reports from participants have highlighted a deficiency of multisensory information exchange in screen-based interactions compared to in-person contexts, where various senses beyond just auditory, including sight (also peripheral vision), touch, and even smell, contribute to a more profound, embodied understanding of the musical material. As a result, these digital environments restrict the transfer of covert embodied aspects. For instance, limited visual transparency and acoustic clarity hinder the identification of subtle details in the recommended postures and covert aspects of embodiment that enable effective sound production. Beyond communication asynchrony and differences in eye contact and directional gaze, crucial in social interaction [83], the limited visibility of the entire body during virtual interactions is known to impede quick and accurate emotion perception due to the absence of rich body language cues (ibid.).
- Discussed under the enacted principle:
  - Similarly, there is limited capacity to transmit even the overt embodied and enacted aspects of singing gestures in the online setting. Even when they remain observable, their clarity is diminished compared to in-person interactions due to the constrained transmission of multimodal information through the screen during videoconferencing teaching sessions. Sometimes, this information needs to be guessed.

- Certain information can be indirectly inferred from other (bodily) cues. However, this process provides only a restricted perspective of the entire gestural engagement.
- There is a wide consensus among participants that online teaching sessions lack crucial elements of human social interaction, with recurring topics centered around the ideas of a limited sense of 'presence' and 'connection.' Online presence plays a key role for both educators and students in creating a supportive environment for the effective transmission of knowledge [84]. The disruption of social presence in video-mediated communication has been previously attributed to the absence of tactile feedback and the limited transmission of non-verbal cues, such as subtle nuances of facial expression and body language [85,86]. Every participant concurred that visual cues play a significant role in this aspect, albeit only providing a superficial sense of connection.
- The interview material indicates a heightened reliance on visual cues during online teaching sessions, thought to be associated with increased fatigue levels. Biological anthropologists confirm that face-to-face interaction, prevailing for over 99% of human history [87], indicates an innate preference that is known as the media naturalness proposition in digital media [88]. It explains challenges, increased cognitive effort, and resulting fatigue in teleconferencing, negatively affecting collaboration satisfaction and effectiveness. Reports also corroborate the notion that stress is linked to an intense feeling of close scrutiny by the teacher, stemming from an intense gaze concentrated on the face alone due to the limited visibility of the body during videoconferencing.
- To counterbalance the deficiency of multi-sensory gestural teacher–student interaction in the online context, teachers appear to rely more heavily on verbal communication (such as linguistic metaphors) and visual aids (verbal descriptors of musical imagery).

- Discussed under the extended principle:
  - The limited visibility of each other's body (camera mainly focusing on the face) and the restricted capacity of videoconferencing platforms in effectively transmitting multimodal information appear to impact the transmission of MIIOs and the sensation of forces as well, i.e., the extended aspects of gestural transmission.

- Discussed under the embedded principle:
  - Several reports by participants suggest that one's understanding relies on a complicated system of (multi-modal) information that—contrary to the in-person—does not become fully transmitted online. However, various interviews indicated that online participants' interactions and engagement improved gradually as they developed a deeper understanding of the other person and his/her contingent presence within the surrounding environment, even if through the virtual context. These reports suggest a contextual adaptation to technology mediation that helps alleviate the initial apprehension of participants regarding the use of online means for music education.
  - Reports suggest that one's understanding relies on a complicated system of extra-musical information about the social and environmental context. However, they also affirm a contextual adaptation to technology mediation, based on a deeper understanding of each other's personality and their contingent presence within the surrounding environment, gradually leading to an improvement in teacher–student engagement.

- Challenges in the spatial perception of Dhrupad singing education:

Given the frequently reported deprivation of multisensory exchanges in screen-based instructional delivery, it is unsurprising that participants have also raised concerns about spatial perception and (social) presence issues. Various challenges associated with the

perception of spatiality in the online teaching context were raised by participants for all three concepts (personal, peri-personal, and extra-personal):

- Personal space, represented by the embodied aspect of the 4E framework: There is an even greater intricacy in conveying covert aspects of embodied engagement that are associated with the mechanics of vocal production, such as the deliberate and active control of the abdominal muscles, through the online medium than traditional in-person instruction. This intricacy is attributed to the limited exchange of multimodal information of screen-based interactions through sound, vision (peripheral too), touch, and even smell.
- Peri-personal space, classified under the extended and enacted aspects of the 4E framework: Its perception becomes challenging in online settings due to the fact that cameras most often capture only the face, resulting in the limited visibility of each other's gestures, including the utilization of MIIOs. Additionally, the mirroring of one's gestures on the screen appears to confine the potential for gestural interaction with melodic content mainly to two-dimensional geometric (melographic) representations. This is in contrast to the three-dimensional experience of haptic-related forces and the required effort exerted, commonly encountered during in-person teaching sessions and performances. This constraint is associated with the limited (two-dimensional) visual affordances of common video platforms and has the potential to lead to a shift in gestural instructions and engagement as well as the perception of peri-personal space, as audio-visual sensory deprivation is also known to degrade visuo-tactile peri-personal space [89]. It also aligns with findings from neuroscience about the effects of the brain's plasticity, whereby the neural representation of the peri-personal space (PPS) is expanded during real-world tool use but not when using tools in virtual environments [90].
- Extra-personal space, organized under the embedded aspect of the 4E framework: A number of participants have reported that their primary concerns about screen-mediated online teaching interaction are space-related and rely on our complex somatosensory system. They revolve around the disruption of the shared 'we-space' in which real interpersonal interactions take place, the constrained frame of the video window, the absence of three-dimensional peripheral vision, spatialized audio, and space-related senses, such as touch and smell, the diminished visual clarity regardless of zoom level that affects the awareness of each other's gaze, and finally, the audio-visual mismatch which currently renders the perception of space dubious. Interestingly, these reports surpassed commonly discussed technical issues, such as sound quality, internet disconnections, lag, etc.

In summary, beyond common technology-related videoconferencing problems, these space-related issues are attributed to the low capacity of screen-based communication in transferring numerous nuanced non-verbal cues. It refers to the limited capability in conveying and perceiving multi-modal information, including even the sense of touch and smell, the degradation of acoustic perception attributed to audio-visual incongruence and the deprivation of acoustical spaciousness and envelopment, and finally, the constrained transmission of various visual cues. The latter entails: the deprivation of direct and the lack of peripheral vision, which also causes the inability to simultaneously observe two things in distinct locations; the absence of a shared spatial reference frame, which is essential for the awareness of togetherness and of each other's gaze, and for establishing the relative positions and distances of the participants [91]; the limited field of view of the camera, especially when only the upper-body or just the head is visible, which dramatically reduces the available gestural space or so-called 'kinesphere' or 'personal space' [92], especially taking into consideration that there might be a tendency to perform larger gestures in the teaching context, at least of foreign languages [93]; the resulting challenges in the concurrent visual engagement with gaze and full-body aspects, sometimes even leading to hyper-gaze [49] or the paradoxical limitation of direct eye contact and ocular engagement through the camera [94]; finally, the lack of depth information and the low resolution that restricts

clear visual engagement and comprehension of postures and gestures. These results are in agreement with previous findings indicating that the spatial context communicated through an interface significantly affects the sense of social presence, co-presence, and gaze awareness [91]. As a consequence of all the above issues, non-verbal interactions that are effective when students and tutors are physically present in the same location can become impractical.

6. Conclusions

The current paper examined challenges related to the embodied elements of synchronous distance Dhrupad music education when conducted through common videoconferencing platforms. Adopting an ethnomusicological perspective, the research employed thematic analysis of interviews with Dhrupad practitioners. The analysis was guided by the four principles of the 4E cognition framework, highlighting the close relationship between the body, mind, and environment. This study has provided essential insights for advancing alternative technologies that can enhance the direct delivery of embodied instructions in a remote music educational context, better fitting the needs of Dhrupad pedagogy and of other oral music traditions than video-based communication.

The outcomes indicate that in Dhrupad vocal pedagogy, the constrained transmission of non-verbal, embodied, multimodal cues, along with the visual and acoustic disruption of a sense of shared spatial and physical context in video-mediated communication, restrict the potential for establishing an attunement between two individuals. This hinders effective social interaction and a sense of immersion, elements deemed vital in educational settings [46], encompassing music and the transmission of the Dhrupad music tradition style. Consequently, current video-conferencing technologies used in online education practices are mostly suited for the transfer of formal rules and procedures, such as score-based solutions [57], potentially lacking the required experiential and emotional engagement as well as nuanced multi-sensory, embodied, and contextual aspects [95] that are fundamental in oral music traditions, such as Hindustani. However, it is important to note that the issues mentioned previously have not been clearly attributed specifically to the two-dimensionality of screen-based communication [91], and even fundamental questions of whether web cameras enhance online communication compared to simple audio-synchronous interaction remains a topic of ongoing debate, at least in teaching foreign languages [96].

In spite of the foundational nature of this criticism, the limitation of embodied, multimodal interaction, as well as shared spatiality, can be reframed as a design challenge rather than a primary limitation of distance learning [97]. To improve these aspects—and hence the virtual music learning experience—and leverage the affordances of technologies used in distance music pedagogy, it is necessary to stimulate several or all the human senses by engaging users in an environment that perceptually surrounds them and allows intuitive and rich full-body interaction with other users and its content [98]. For instance, Nijs et al. [99] have developed an interactive multimodal system that can enhance the learning experience of a musical instrument through gesture-controlled color painting in space.

Additionally, the incorporation of novel technologies beyond the use of video has the potential to enhance the sense of co-presence in augmented learning environments [100] and also in music [101]. A collection of emergent technologies, such as three-dimensional marker-based or markerless motion capture, Virtual/Augmented/Extended/Immersive Reality (VR, AR, XR, IR), Immersive audio-haptic systems, holoportation, and holography, have been experiencing a resurgence in recent years. Combined with various real-time motion tracking technologies and sensors, such as BCI (brain-computer interfaces), head displays, and haptic devices, promise a deeply immersive, multi-sensory experience capable of bringing users together in a shared online space. Along these lines, the metaverse has gained strong interest in recent years as a fusion of both virtually-enhanced physical reality and physical-persisted virtual space [102] that aims at narrowing or even eliminating the gap between the virtual and the physical world.

The in-person meeting context remains the benchmark for all technology-mediated systems, and the sense of presence and immersion can serve as a measure of communication naturalness and satisfaction in all teleconferencing and collaborative virtual environments [103]. The key question revolves around the type of contemporary technologies and designs that can tackle the current limitation of full-body interaction in an otherwise joint physical space, as well as the type of new embodiment that arises from bodies being extended by modern technologies [104]. For instance, the study by Li & Lefevre [105] emphasizes the effectiveness of holograms, multisensorial elements, and artificial intelligence in videoconference design, enhancing participant engagement through stimulating senses. Holographic-type communication shows great promise in providing an unparalleled sense of presence through visual depth and realism, allowing users to perceive images from various angles and interact through touch or gestures. It may also enable direct touch of 3D holograms, emitting ultrasonic radiation for a tactile feel of resistance [106]. Further advancements in holography are foreseen, such as simulating temperature and scent, closely resembling real-life experiences and enhancing the sense of immersion [107].

However, while these technologies are observed to be progressively maturing, there is a relatively small body of literature that is concerned with using them as an alternative to simple video-conferencing for improving the experience of shared spatial perception, the sense of presence, and the transmission of multisensory embodied aspects in music education scenarios. For instance, although the metaverse [108,109] and holographic displays [110] have the potential to offer a new educational environment, systematic work on music pedagogy has been limited to date [111]. Drawing on the findings of this paper, it is a direction worthy of further investigation and points to priorities for future research. However, the ensuing inquiry emerges: within the context of frugal innovation and the democratization of technology, what types of custom-made, widely affordable, domestically accessible software and hardware technologies should be developed to enhance support for the teaching requirements of oral and underrepresented music traditions, most often lacking access to expensive state-of-the-art sensing and display technologies?

Funding: This publication (APC) is financed by the Project "Strengthening and optimizing the operation of MODY services and academic and research units of the Hellenic Mediterranean University", funded by the Public Investment Program of the Greek Ministry of Education and Religious Affairs.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Research Ethics Committee of the Hellenic Mediterranean University with protocol code 101 and date of approval 1.3.2023.

Informed Consent Statement: Written informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available within the article.

Acknowledgments: I want to express my gratitude to all participating musicians, as well as to the organizers of Labyrinth Catalunya for their support and assistance during the workshop.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A

Table A1. Information on original interview data.

Name	Type of Interview	Date
Pelva Naik	In-person; Labyrinth Catalunya, Cardedeu, Spain	9 April 2023
Isadora Reig	Online	7 May 2023
Meghana Sardar	Online	9 May 2023

Table A1. *Cont.*

Name	Type of Interview	Date
Markus Schmidt	Online	11 May 2023
Cellin Wadier	Online	15 September 2023
Marianne Svašek	Online	15 September 2023
Aarya Patwardhan	Online	16 September 2023
Gauri Raje	Online	18 September 2023

References

1. Small, C. *Musicking: The Meanings of Performing and Listening*; Music/Culture; University Press of New England: Hanover, NH, USA, 1998; ISBN 978-0-8195-2256-6.
2. Leman, M. *Embodied Music Cognition and Mediation Technology*; The MIT Press: Cambridge, MA, USA, 2007; ISBN 978-0-262-25655-1.
3. Smirnov, A.; Price, M. *Sound in Z: Experiments in Sound and Electronic Music in Early 20th Century Russia*; Koenig Books: London, UK, 2013; ISBN 978-3-86560-706-5.
4. Bremmer, M.; Nijs, L. The Role of the Body in Instrumental and Vocal Music Pedagogy: A Dynamical Systems Theory Perspective on the Music Teacher's Bodily Engagement in Teaching and Learning. *Front. Educ.* **2020**, *5*, 79. [CrossRef]
5. Simones, L.L. A Framework for Studying Teachers' Hand Gestures in Instrumental and Vocal Music Contexts. *Music. Sci.* **2019**, *23*, 231–249. [CrossRef]
6. Nafisi, J. Gesture and Body-Movement as Teaching and Learning Tools in the Classical Voice Lesson: A Survey into Current Practice. *Br. J. Music Educ.* **2013**, *30*, 347–367. [CrossRef]
7. Paparo, S.A. Embodying Singing in the Choral Classroom: A Somatic Approach to Teaching and Learning. *Int. J. Music Educ.* **2016**, *34*, 488–498. [CrossRef]
8. Kohut, D.L. *Musical Performance: Learning Theory and Pedagogy*; Stipes Pub. Co.: Champaign, IL, USA, 1992; ISBN 978-0-87563-415-9.
9. Rahaim, M. *Musicking Bodies: Gesture and Voice in Hindustani Music*; Music/Culture; Wesleyan University Press: Middletown, CT, USA, 2012; ISBN 978-0-8195-7325-4.
10. Jain, V. India Together: Changing Notes of Music Education—22 April 2005. Available online: https://indiatogether.org/music-society (accessed on 14 December 2023).
11. Farrell, G. *Indian Music and the West: Gerry Farrell*; Clarendon Press: Oxford, UK; Oxford University Press: Oxford, UK; New York, NY, USA, 1997; ISBN 978-0-19-816391-6.
12. Camlin, D.A.; Lisboa, T. The Digital 'Turn' in Music Education (Editorial). *Music Educ. Res.* **2021**, *23*, 129–138. [CrossRef]
13. Newen, A.; De Bruin, L.; Gallagher, S. *The Oxford Handbook of 4E Cognition*, 1st ed.; Oxford University Press: Oxford, UK, 2018; ISBN 978-0-19-873541-0.
14. Shapiro, L.A. (Ed.) *The Routledge Handbook of Embodied Cognition*; Routledge Handbooks in Philosophy; First Issued in Paperback; Routledge: London, UK; New York, NY, USA, 2017; ISBN 978-1-138-57397-0.
15. Roy, J. The Internet Guru: Online Pedagogy in Indian Classical Music Traditions. *Asian Music* **2016**, *47*, 103–135. [CrossRef]
16. Nettl, B. *The Study of Ethnomusicology: Thirty-One Issues and Concepts, Third edition.*; University of Illinois Press: Urbana, IL, USA, 2005; ISBN 978-0-252-03033-8.
17. Schippers, H. The Guru Recontextualized? Perspectives on Learning North Indian Classical Music in Shifting Environments for Professional Training. *Asian Music* **2007**, *38*, 123–138. [CrossRef]
18. Paschalidou, S. Effort Inference and Prediction by Acoustic and Movement Descriptors in Interactions with Imaginary Objects during Dhrupad Vocal Improvisation. *Wearable Technol.* **2022**, *3*, e14. [CrossRef]
19. Young, K. The Memory of the Flesh: The Family Body in Somatic Psychology. *Body Soc.* **2002**, *8*, 25–47. [CrossRef]
20. Robert Caro Ustad Zia Fariduddin DAGAR/The Dhrupad Legend. Available online: https://www.youtube.com/watch?v=FcGz-8ZxMrU (accessed on 14 December 2023).
21. Fatone, G.A.; Clayton, M.; Leante, L. Imagery, Melody and Gesture in Cross-Cultural Perspective. In *New Perspectives on Music and Gesture*; Gritten, A., King, E., Eds.; Ashgate Publishing: Farnham, UK, 2011; pp. 203–220.
22. Van der Meer, W.; Rao, S. What You Hear Isn't What You See: The Representation and Cognition of Fast Movements in Hindustani Music. In Proceedings of the International Symposium Frontiers of Research on Speech and Music, Baripada, India, 30 July 2006; pp. 12–20.
23. Martin, C. Time, Gesture and Attention in a Khyāl Performance. *Asian Music* **2007**, *38*, 71–96. [CrossRef]
24. Leante, L. The Lotus and the King: Imagery, Gesture and Meaning in a Hindustani *Rāg*. *Ethnomusicol. Forum* **2009**, *18*, 185–206. [CrossRef]
25. Ratcliffe, M. Perception, Exploration, and the Primacy of Touch. In *The Oxford Handbook of 4E Cognition*; Newen, A., De Bruin, L., Gallagher, S., Eds.; Oxford University Press: Oxford, UK, 2018; pp. 280–300. ISBN 978-0-19-873541-0.

26. Pallasmaa, J. An Architecture of the Seven Senses. In *Questions of Perception: Phenomenology of Architecture*; Holl, S., Pallasmaa, J., Perez-Gomez, A., Eds.; William Stout Publishers: San Francisco, CA, USA, 1994; pp. 27–37.
27. Serino, A.; Haggard, P. Touch and the Body. *Neurosci. Biobehav. Rev.* **2010**, *34*, 224–236. [CrossRef]
28. Warren, A.; Welton, M. On Touch. *Perform. Res.* **2022**, *27*, 1–6. [CrossRef]
29. Garner, J. *Kinesthetic Spectatorship in the Theatre: Phenomenology, Cognition, Movement*, 1st ed.; Cognitive Studies in Literature and Performance; Springer International Publishing: Berlin/Heidelberg, Germany; Palgrave Macmillan: Cham, Switzerland, 2018; ISBN 978-3-319-91794-8.
30. Zahavi, D. *Self and Other: Exploring Subjectivity, Empathy, and Shame*, 1st ed.; Oxford University Press: Oxford, UK, 2014; ISBN 978-0-19-959068-1.
31. Huang-Kokina, A. Touching through Music, Touching through Words: The Performance and Performativity of Pianistic Touch in Musical and Literary Settings. *Perform. Res.* **2022**, *27*, 48–55. [CrossRef]
32. Haynes, A.; Lawry, J.; Kent, C.; Rossiter, J. FeelMusic: Enriching Our Emotive Experience of Music through Audio-Tactile Mappings. *Multimodal Technol. Interact.* **2021**, *5*, 29. [CrossRef]
33. Cléry, J.; Guipponi, O.; Wardak, C.; Ben Hamed, S. Neuronal Bases of Peripersonal and Extrapersonal Spaces, Their Plasticity and Their Dynamics: Knowns and Unknowns. *Neuropsychologia* **2015**, *70*, 313–326. [CrossRef]
34. Rizzolatti, G.; Luppino, G.; Matelli, M. The Organization of the Cortical Motor System: New Concepts. *Electroencephalogr. Clin. Neurophysiol.* **1998**, *106*, 283–296. [CrossRef] [PubMed]
35. Berthoz, A. Reference Frames for the Perception and Control of Movement. In *Brain and Space*; Oxford University Press: New York, NY, USA, 1991; pp. 81–111. ISBN 978-0-19-854284-1.
36. Millar, S. *Space and Sense*; Psychology Press: London, UK, 2008; ISBN 978-1-135-42226-4.
37. Làdavas, E.; Farnè, A. Neuropsychological Evidence of Integrated Multisensory Representation of Space in Humans. In *The Handbook of Multisensory Processes*; Calvert, G.A., Spence, C., Stein, B.E., Eds.; The MIT Press: Cambridge, MA, USA, 2004; pp. 799–818. ISBN 978-0-262-26970-4.
38. Howard, I.P.; Templeton, W.B. *Human Spatial Orientation*; Wiley: Hoboken, NJ, USA, 1966; ISBN 978-0-608-11450-7.
39. Paschalidou, P.-S. Effort in Gestural Interactions with Imaginary Objects in Hindustani Dhrupad Vocal Music. Ph.D. Thesis, Durham University, Durham, UK, 2017.
40. Krueger, J. Affordances and the Musically Extended Mind. *Front. Psychol.* **2014**, *4*, 1003. [CrossRef] [PubMed]
41. Lloyd, D.M. The Space between Us: A Neurophilosophical Framework for the Investigation of Human Interpersonal Space. *Neurosci. Biobehav. Rev.* **2009**, *33*, 297–304. [CrossRef] [PubMed]
42. Cassio, F. Il Nāda Yoga, La Scienza Del Suono Nella Tradizione Musicale Indiana. In *Conscientia Musica. Contrappunti per Rossana Dalmonte e Mario Baroni*; Addessi, A., Macchiarella, I., Privitera, M., Russo, M., Eds.; LIM: Lucca, Italy, 2010; pp. 265–278.
43. Taylor, T.D. *Global Pop*; Routledge: London, UK, 2014; ISBN 978-1-135-25408-7.
44. Neuman, D. *The Life of Music in North India: The Organization of an Artistic Tradition*; University of Chicago Press: Chicago, IL, USA, 1990; ISBN 978-0-226-57516-2.
45. Baralou, E.; Tsoukas, H. How Is New Organizational Knowledge Created in a Virtual Context? An Ethnographic Study. *Organ. Stud.* **2015**, *36*, 593–620. [CrossRef]
46. Doumanis, I.; Economou, D.; Sim, G.R.; Porter, S. The Impact of Multimodal Collaborative Virtual Environments on Learning: A Gamified Online Debate. *Comput. Educ.* **2019**, *130*, 121–138. [CrossRef]
47. Lombard, M.; Ditton, T. At the Heart of It All: The Concept of Presence. *J. Comput.-Mediat. Commun.* **2006**, *3*, JCMC321. [CrossRef]
48. Cesari, V.; Galgani, B.; Gemignani, A.; Menicucci, D. Enhancing Qualities of Consciousness during Online Learning via Multisensory Interactions. *Behav. Sci.* **2021**, *11*, 57. [CrossRef]
49. Fauville, G.; Luo, M.; Queiroz, A.C.M.; Lee, A.; Bailenson, J.N.; Hancock, J. Video-Conferencing Usage Dynamics and Nonverbal Mechanisms Exacerbate Zoom Fatigue, Particularly for Women. *Comput. Hum. Behav. Rep.* **2023**, *10*, 100271. [CrossRef]
50. Serino, A. Peripersonal Space (PPS) as a Multisensory Interface between the Individual and the Environment, Defining the Space of the Self. *Neurosci. Biobehav. Rev.* **2019**, *99*, 138–159. [CrossRef]
51. Rottondi, C.; Chafe, C.; Allocchio, C.; Sarti, A. An Overview on Networked Music Performance Technologies. *IEEE Access* **2016**, *4*, 8823–8843. [CrossRef]
52. Wilson, R. Aesthetic and Technical Strategies for Networked Music Performance. *AI Soc.* **2023**, *38*, 1871–1884. [CrossRef] [PubMed]
53. Alexandraki, C. Experimental Investigations and Future Possibilities in Network-Mediated Folk Music Performance. In *Computational Phonogram Archiving*; Bader, R., Ed.; Current Research in Systematic Musicology; Springer International Publishing: Cham, Switzerland, 2019; Volume 5, pp. 207–228. ISBN 978-3-030-02694-3.
54. Biasutti, M.; Antonini Philippe, R.; Schiavio, A. Assessing Teachers' Perspectives on Giving Music Lessons Remotely during the COVID-19 Lockdown Period. *Music Sci.* **2022**, *26*, 585–603. [CrossRef]
55. Goodrich, A. Online Peer Mentoring and Remote Learning. *Music Educ. Res.* **2021**, *23*, 256–269. [CrossRef]
56. Koutsoupidou, T. Online Distance Learning and Music Training: Benefits, Drawbacks and Challenges. *Open Learn. J. Open Distance E-Learn.* **2014**, *29*, 243–255. [CrossRef]
57. Duffy, S.; Healey, P. A New Medium for Remote Music Tuition. *J. Music Technol. Educ.* **2017**, *10*, 5–29. [CrossRef]

58. Schiavio, A.; Nijs, L. Implementation of a Remote Instrumental Music Course Focused on Creativity, Interaction, and Bodily Movement. Preliminary Insights and Thematic Analysis. *Front. Psychol.* **2022**, *13*, 899381. [CrossRef] [PubMed]
59. Ward, F. Technology and the Transmission of Tradition: An Exploration of the Virtual Pedagogies in the Online Academy of Irish Music. *J. Music Technol. Educ.* **2019**, *12*, 5–23. [CrossRef]
60. Pudaruth, S.K. Online Teaching and Learning of Hindustani Classical Vocal Music: Resistance, Challenges, and Opportunities. In *Academic Voices*; Elsevier: Amsterdam, The Netherlands, 2022; pp. 147–159. ISBN 978-0-323-91185-6.
61. Pearson, L. Gesture in Karnatak Music: Pedagogy and Musical Structure in South India. Ph.D. Thesis, Durham University, Durham, UK, 2016.
62. Van der Schyff, D.; Schiavio, A.; Walton, A.; Velardo, V.; Chemero, A. Musical Creativity and the Embodied Mind: Exploring the Possibilities of 4E Cognition and Dynamical Systems Theory. *Music Sci.* **2018**, *1*, 205920431879231. [CrossRef]
63. Gallagher, S. *Embodied and Enactive Approaches to Cognition*; Cambridge Elements Elements in Philosophy of Mind; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2023; ISBN 978-1-00-920980-9.
64. Gibson, J.J. *The Ecological Approach to Visual Perception*; Erlbaum: Hillsdale, NJ, USA, 1986; ISBN 978-0-89859-958-9.
65. Johnson, M.L. *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*; University of Chicago Press: Chicago, IL, USA, 1992; ISBN 978-0-226-40318-2.
66. Cox, A. *Music and Embodied Cognition: Listening, Moving, Feeling, and Thinking*; Indiana University Press: Bloomington, IN, USA, 2016; ISBN 978-0-253-02167-0.
67. Reybrouck, M. *Musical Sense-Making: Enactment, Experience and Computation*, 1st ed.; Routledge, Taylor & Francis Group: London, UK, 2021; ISBN 978-0-429-27401-5.
68. Menary, R. Introduction to the Special Issue on 4E Cognition. *Phenomenol. Cogn. Sci.* **2010**, *9*, 459–463. [CrossRef]
69. Chemero, A. *Radical Embodied Cognitive Science, First MIT Press Paperback ed.*; A Bradford Book; MIT Press: Cambridge, MA, USA; London, UK, 2011; ISBN 978-0-262-51647-1.
70. Ryan, K.; Schiavio, A. Extended Musicking, Extended Mind, Extended Agency. Notes on the Third Wave. *New Ideas Psychol.* **2019**, *55*, 8–17. [CrossRef]
71. Schiavio, A.; van der Schyff, D. 4E Music Pedagogy and the Principles of Self-Organization. *Behav. Sci.* **2018**, *8*, 72. [CrossRef] [PubMed]
72. Ryan, K.J.; Gallagher, S. Between Ecological Psychology and Enactivism: Is There Resonance? *Front. Psychol.* **2020**, *11*, 1147. [CrossRef] [PubMed]
73. Casas-Mas, A.; Pozo, J.I.; Montero, I. Oral Tradition as Context for Learning Music From 4E Cognition Compared with Literacy Cultures. Case Studies of Flamenco Guitar Apprenticeship. *Front. Psychol.* **2022**, *13*, 733615. [CrossRef]
74. Hood, M. *The Ethnomusicologist*, New ed.; Kent State University Press: Kent, OH, USA, 1982; ISBN 978-0-87338-280-9.
75. Seaman, C.B. Qualitative Methods. In *Guide to Advanced Empirical Software Engineering*; Shull, F., Singer, J., Sjøberg, D.I.K., Eds.; Springer: London, UK, 2008; pp. 35–62. ISBN 978-1-84800-043-8.
76. Boyatzis, R.E. *Transforming Qualitative Information: Thematic Analysis and Code Development*; Nachdr.; Sage: Thousand Oaks, CA, USA, 2010; ISBN 978-0-7619-0961-3.
77. Clayton, M.; Dueck, B.; Leante, L. (Eds.) *Experience and Meaning in Music Performance*; Oxford University Press: New York, NY, USA, 2014; ISBN 978-0-19-936953-9.
78. Dejean, C.; Guichon, N.; Nicolaev, V. Compétences interactionnelles des tuteurs dans des échanges vidéographiques synchrones. *Distances Savoirs* **2010**, *8*, 377–393. [CrossRef]
79. Noë, A. *Action in Perception, 1st MIT Press paperback ed.*; Representation and Mind; MIT Press: Cambridge, MA, USA, 2006; ISBN 978-0-262-64063-3.
80. Varela, F.J.; Thompson, E.; Rosch, E. *The Embodied Mind: Cognitive Science and Human Experience*; 14. Print; MIT Press: Cambridge, MA, USA, 1993; ISBN 978-0-262-72021-2.
81. Menin, D.; Schiavio, A. Rethinking Musical Affordances. *Avant* **2012**, *3*, 201–215.
82. Okray, Z.; Jacob, P.F.; Stern, C.; Desmond, K.; Otto, N.; Talbot, C.B.; Vargas-Gutierrez, P.; Waddell, S. Multisensory Learning Binds Neurons into a Cross-Modal Memory Engram. *Nature* **2023**, *617*, 777–784. [CrossRef] [PubMed]
83. Kendon, A. Some Functions of Gaze-Direction in Social Interaction. *Acta Psychol.* **1967**, *26*, 22–63. [CrossRef]
84. Themeli, C. Tele-Proximity: The Experienced Educator Perspective on Human to Human Connection in Distance Education. Ph.D. Thesis, Lancaster University, Lancaster, UK, 2013.
85. Huisman, G. A Touch of Affect: Mediated Social Touch and Affect. In Proceedings of the 14th ACM International Conference on Multimodal Interaction, Santa Monica, CA, USA, 22–26 October 2012; pp. 317–320.
86. Nowak, K.L.; Biocca, F. The Effect of the Agency and Anthropomorphism on Users' Sense of Telepresence, Copresence, and Social Presence in Virtual Environments. *Presence Teleoperators Virtual Environ.* **2003**, *12*, 481–494. [CrossRef]
87. Boaz, N.T.; Almquist, A.J. *Biological Anthropology: A Synthetic Approach to Human Evolution*; Prentice Hall: Upper Saddle River, NJ, USA, 1997; ISBN 978-0-13-369208-2.
88. Kock, N. The Psychobiological Model: Towards a New Theory of Computer-Mediated Communication Based on Darwinian Evolution. *Organ. Sci.* **2004**, *15*, 327–348. [CrossRef]
89. Noel, J.-P.; Samad, M.; Doxon, A.; Clark, J.; Keller, S.; Di Luca, M. Peri-Personal Space as a Prior in Coupling Visual and Proprioceptive Signals. *Sci. Rep.* **2018**, *8*, 15819. [CrossRef]

90. Ferroni, F.; Gallese, V.; Soccini, A.M.; Langiulli, N.; Rastelli, F.; Ferri, D.; Bianchi, F.; Ardizzi, M. The Remapping of Peripersonal Space in a Real but Not in a Virtual Environment. *Brain Sci.* **2022**, *12*, 1125. [CrossRef]
91. Hauber, J.; Regenbrecht, H.; Billinghurst, M.; Cockburn, A. Spatiality in Videoconferencing: Trade-Offs between Efficiency and Social Presence. In Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work, Banff, AL, Canada, 4–8 November 2006; pp. 413–422.
92. Von Laban, R.; Lawrence, F.C.; Lawrence, F.C. *Effort: Economy of Human Movement*, 2nd ed.; Reprinted; Macdonald and Evans: Plymouth, UK, 1979; ISBN 978-0-7121-0534-7.
93. Tellier, M.; Stam, G. Stratégies Verbales et Gestuelles Dans l'explication Lexicale d'un Verbe d'action. In *Spécificités et Diversité des Interactions Didactiques*; Rivière, V., Ed.; Actes Académiques; Riveneuve éditions: Paris, France, 2012; pp. 357–374.
94. Bohannon, L.S.; Herbert, A.M.; Pelz, J.B.; Rantanen, E.M. Eye Contact and Video-Mediated Communication: A Review. *Displays* **2013**, *34*, 177–185. [CrossRef]
95. Dreyfus, H.L. *On the Internet*, 2nd ed.; Thinking in Action; Routledge: Abingdon, UK; New York, NY, USA, 2009; ISBN 978-0-415-77516-8.
96. Guichon, N.; Cohen, C. The Impact of the Webcam on an Online L2 Interaction. *Can. Mod. Lang. Rev.* **2014**, *70*, 331–354. [CrossRef]
97. Ward, D. What's Lacking in Online Learning? Dreyfus, Merleau-Ponty and Bodily Affective Understanding: What's Lacking in Online Learning? *J. Philos. Educ.* **2018**, *52*, 428–450. [CrossRef]
98. Hills, P.D.; Clavin, M.V.Q.; Tufft, M.R.A.; Gobel, M.S.; Richardson, D.C. Video Meeting Signals: Experimental Evidence for a Technique to Improve the Experience of Video Conferencing. *PLoS ONE* **2022**, *17*, e0270399. [CrossRef]
99. Nijs, L.; Coussement, P.; Moens, B.; Amelinck, D.; Lesaffre, M.; Leman, M. Interacting with the Music Paint Machine: Relating the Constructs of Flow Experience and Presence. *Interact. Comput.* **2012**, *24*, 237–250. [CrossRef]
100. El Raheb, K.; Soulis, A.; Nastos, D.; Lougiakis, C.; Roussou, M.; Christopoulos, D.; Sofianopoulos, G.; Papagiannis, S.; Rüggeberg, J.; Katsikaris, L.; et al. Eliciting Requirements for a Multisensory eXtended Reality Platform for Training and Informal Learning. In Proceedings of the CHI Greece 2021: 1st International Conference of the ACM Greek SIGCHI Chapter, Athens, Greece, 25–27 November 2021; pp. 1–8.
101. Van Kerrebroeck, B.; Caruso, G.; Maes, P.-J. A Methodological Framework for Assessing Social Presence in Music Interactions in Virtual Reality. *Front. Psychol.* **2021**, *12*, 663725. [CrossRef]
102. Zhao, Y.; Jiang, J.; Chen, Y.; Liu, R.; Yang, Y.; Xue, X.; Chen, S. Metaverse: Perspectives from Graphics, Interactions and Visualization. *Vis. Inform.* **2022**, *6*, 56–67. [CrossRef]
103. Biocca, F.; Harms, C.; Gregg, J. The Networked Minds Measure of Social Presence: Pilot Test of the Factor Structure and Concurrent Validity. In Proceedings of the 4th Annual International Workshop on Presence, Philadelphia, PA, USA, 21–23 May 2001; pp. 1–9.
104. Dall'Alba, G.; Barnacle, R. Embodied Knowing in Online Environments. *Educ. Philos. Theory* **2005**, *37*, 719–744. [CrossRef]
105. Li, N.; Lefevre, D. Holographic Teaching Presence: Participant Experiences of Interactive Synchronous Seminars Delivered via Holographic Videoconferencing. *Res. Learn. Technol.* **2020**, *28*, 2265. [CrossRef]
106. Reardon, G.; Dandu, B.; Shao, Y.; Visell, Y. Shear Shock Waves Mediate Haptic Holography via Focused Ultrasound. *Sci. Adv.* **2023**, *9*, eadf2037. [CrossRef] [PubMed]
107. Smart, J.; Cascio, J.; Paffendorf, J. A Metaverse Roadmap: Pathways to the 3D Web. 2007. Available online: https://www.w3.org/2008/WebVideo/Annotations/wiki/images/1/19/MetaverseRoadmapOverview.pdf (accessed on 14 December 2023).
108. Zhang, X.; Chen, Y.; Hu, L.; Wang, Y. The Metaverse in Education: Definition, Framework, Features, Potential Applications, Challenges, and Future Research Topics. *Front. Psychol.* **2022**, *13*, 1016300. [CrossRef] [PubMed]
109. Kye, B.; Han, N.; Kim, E.; Park, Y.; Jo, S. Educational Applications of Metaverse: Possibilities and Limitations. *J. Educ. Eval. Health Prof.* **2021**, *18*, 32. [CrossRef] [PubMed]
110. Themelis, C.; Sime, J.-A. From Video-Conferencing to Holoportation and Haptics: How Emerging Technologies Can Enhance Presence in Online Education. In *Emerging Technologies and Pedagogies in the Curriculum*; Yu, S., Ally, M., Tsinakos, A., Eds.; Bridging Human and Machine: Future Education with Intelligence; Springer: Singapore, 2020; pp. 261–276. ISBN 9789811506178.
111. Turchet, L. Musical Metaverse: Vision, Opportunities, and Challenges. *Pers Ubiquit Comput* **2023**, *27*, 1811–1827. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.