Asia-Pacific Journal for Arts Education

Special Issue

Unfold the Future of Music Education through Technology

Co-editors:

Prof. Bo Wah LEUNG

Dr. Cheung On TAM

Dr. Chi Hin LEUNG

Dr. Richard Guy WHITBREAD

The Education University of Hong Kong

http://www.ied.edu.hk/ccaproject/apjae/apjae.htm ISSN: 1683-6995

Volume 22 Number 1 December 2023

Seeking the Metaverse in Music Education

Xiao Qi

Nanjing University of the Arts

xq1177@126.com

Abstract

The metaverse, which exploded into the public consciousness in 2021, has been hotly debated since its appearance. With definitions still being contested, the various discussions have instead tended to center on its future impact. Building on its unique capacity to unite imagination and vision, and with the support of artificial intelligence, blockchain, 5G, virtual reality and other technologies, the integration of the metaverse and education is an area that has the potential to transform the way we view the world.

Concerns over the quality of education and diversified development have led to the gradual recognition that music learning occupies an indispensable position in the process of personal growth. How to enable music education to keep pace with the times has therefore become an urgent issue, the hope being that the unique opportunities for immersion, interactivity, and diversity offered by the metaverse may serve to stimulate the future development of music education.

In the metaverse of music education, learners' access to rich virtual experiences, diverse knowledge resources, decentralized knowledge creation and sharing, and powerful digital productivity can all serve to widen the limited boundaries of the real world. At the same time, generation of the metaverse can also help reconfigure knowledge construction, learning scenarios, and the environmental identity of music education teaching. Through theoretical and applied research, music education can, and should be, rethought and reimagined.

Key words

music education, metaverse, educational model

Introduction

In the fifteenth century, European explorers embarked on voyages, accumulating wealth and increasing national power in the course of their explorations. Nineteenth century Western European countries harnessed the value of mechanized production through a series of inventions, in the process bringing about an industrial revolution and becoming world capitalist powers. China at this time was still a feudal "face to the field, back to the sky" society "living at the mercy of nature", a situation that can, in part, be attributed to different perceptions of the unknown, change, and technology. In the present era of globalization, where everything is connected and interconnected, science and technology have become the leaders in every modern society's economic growth. Within this context of modern information technology, represented by the rise of, and reliance upon, computers, the waves of the technological revolution have significantly changed the ways music is created and the environments in which it is heard. The demand for music culture has become more individualized and differentiated, and music education has inevitably shifted in response to this influence. As the "metaverse" continues to integrate various cutting-edge technologies, its immersion, interactivity, and openness will undoubtedly bring to music education unprecedented new experiences, in the process opening up a "new era" for future generations.

The familiar past and an unfamiliar present

The Spring Festival is a traditional event and an essential symbol of Chinese culture. Likewise, the annual “Spring Festival Gala” on New Year’s Eve is considered essential TV viewing. As part of the 2022 CCTV Spring Festival Gala, virtual / augmented reality technology, holographic scanning technology, 8K naked eye 3D presentation technology, AI multi-modal motion capture technology, and other “Black Technologies” were combined with colorful performances. Together, these not only provided the audience with an unprecedented audiovisual feast, but also created a continuous and enthusiastic discussion on domestic and international social media platforms. Whether it is facial recognition, mobile payment, or driverless cars, scenes that once existed only in science fiction movies are gradually becoming a part of people’s everyday lives. We live in an era characterized by human-centeredness and individualism, in which knowledge, technology, and invention are gradually becoming the primary forces driving social development. Furthermore, education and the environment are interdependent. In discussing the relationship between the individual and the teaching environment, Zhong Qiquan cites R. H. Moos, an American professor of psychiatry and behavioral sciences, who argues that “Individual behavior is not an inherent characteristic of the respective person. In other words, it is not determined by personality and attitude, but is strongly influenced by the results of the environment.” (Zhong, 2021). In the ever-changing information age, the question of how both education generally and traditional teaching methods specifically have to change to keep pace with the times has become an urgent problem.

Technology is the engine of all things, and education is a powerful means of driving this engine. In 1350, Laurentius de Voltolina painted a picture depicting a crowded classroom at the University of Bologna. Ironically, despite being created more than six hundred years ago, there is much in the painting that would not look out of place in today's classrooms (see Figure 1).

Figure 1 *There are no essential differences between the delivery of music education as it occurs in today's classroom (right) and the organization of teaching more than six hundred years ago (left)*

What is no longer comparable is today's view of education, especially in terms of how it can and should prepare young talent for future development within the context of the kinds of technological, economic and cultural developments that are part and parcel of our everyday existence. With its emphasis on quality education and diversified development, music plays an indispensable role in personal development. Sukhomlinsky, a Soviet educational theorist, once said, “Music occupies an important place in the means of influencing the minds of young people, it is a powerful source of thought. Without music education, there can be no intellectual development that meets the requirements.” With the introduction of a Western music-centered curriculum and the establishment of modern new schools, school music education in China has entered a new era, complete with teaching models that aim to provide more students with the opportunity to study the subject. However, where “unified” music classes are typically the norm, the teacher is often the “leader,” and the students the “passive performers” who, under their tutors’ guidance, listen to and appreciate selected musical works in order to obtain the necessary aesthetic experiences. In this kind of “processing production” teaching system, students’ autonomy and their ability to define the meaning and purpose of learning music are often neglected.

Edgar Dale, an American audiovisual educator, proposed a theory called the “Tower of Experience’ whereby ten levels of abstraction, from direct, purposeful experience to verbal symbols, are employed. In the same way that words and language are related, the various notes and terms written on the music staff can be described as the symbolic expression of abstract music. Through the process of being musically educated, learners use their understanding of the symbols to engage in a variety of performance-related activities in order to gain the appropriate levels of perception, understanding, and aesthetic experience. As a result, teachers spend a great deal of time teaching students the meaning and function of various types of musical notation, and students spend corresponding amounts practicing repeatedly in order to grasp the intrinsic connection between notation and artistic expression. This kind of teaching model not only makes the learning process unnecessarily long, but also indirectly increases the difficulty of learning music.

In addition, the teaching process can be seen as one in which the teacher "encodes" the content and the students "decode" the information by listening to it. According to the "funnel principle" in the information transmission process, whether the teacher can deliver the information correctly and effectively, whether the students can extract the information they need in the listening process, and whether they can understand and master the information in the decoding process, all have a corresponding impact on the effectiveness of teaching. There is a common misconception that the process of learning and teaching is a unified whole, and that as long as the teacher demonstrates accurately the students will be able to grasp the relevant knowledge entirely and efficiently. However, due to limited teaching time, teachers are frequently called upon to utilize a range of teaching tactics and procedures in order to guarantee that students correctly grasp the information, meaning that the music class can all too easily turn into a performance class for the teacher. As a result, the uniqueness of learners is ignored, and music education increasingly resembles a “teaching product” that is “templated”, in much the same way that it was delivered in the fourteenth century. Indeed, as has been stated, “In the last 100 years, our listening style and environment have changed dramatically. But the current model of teaching still follows the methods and content of the old tradition, as if it were providing music learners oriented toward the new century with the skills needed by the masters of 18th century church music” (de Souza, 2020). Furthermore, the fact that, as a new generation of “digital natives”, the majority of students’ musical achievements occur outside of school, should not be overlooked. Thus, regardless of the enthusiasm of instructors for technology, new technologies of all types have become an integral component of the growth of music education among today’s youth.

Toward a “future paradigm” of music education

The rapid outbreak of COVID-19 in 2019 interrupted people’s everyday lives, turning each individual into “an island in the metropolis.” Individuals were forced to conduct many of their jobs and studies online, a situation that resulted not only in short-term changes, but stimulated a broader discussion about the nature of work and the different possibilities for the dissemination of information. In the field of music education, the majority of nations and regions chose online education to preserve the stability and continuation of education during the standstill induced by COVID-19. Nevertheless, although online education has become an essential part of teaching and learning, there is plenty of room for improvement. For example, students' sense of participation is frequently weak due to the lack of interaction, websites can be monotonous and boring, delays caused by network communications reduce learning efficiency, and the lack of sufficient learning data records makes it difficult to analyze and evaluate the effects of learning. As soon as the concept of the "metaverse" appeared, it was like a tsunami wave, causing heated discussions across various platforms. With the support of emerging technologies, such as virtual reality, 5G, blockchain, the Internet of Things, artificial intelligence, and cloud computing, can the metaverse lead education into a new network era?

Futurist Luke Shabro defines the metaverse as "a fuzzy, digitally hybrid reality with irreplaceable and infinite projects and roles, a cyberspace free from traditional physical limitations and constraints." (Hua & Huang, 2021) The Metaverse Development Research Report 2020-2021, published by the New Media Research Center of Tsinghua University, defines it as "a new virtual-real Internet application and social form resulting from the integration of multiple new technologies, which closely integrates the virtual world with the real world in terms of economic system, social system, and identity system, and allows each user to make content production and world editing". (Tsinghua University, 2021, p. XXX). As recently as thirty years ago, the famous scholar Qian Xuesen had a vision of virtual reality and the metaverse, even giving it the poetic name, "The Spiritual Realm." Compared with online education, the unique interactive, immersive, and participatory qualities of the metaverse have the potential to resolve the aforementioned defects of online education, while within the teaching context the metaverse can replicate the rich experiences between people and objects. With the assistance of 5G technology, the low-latency network environment satisfies learners' desire for quality immersion, enhanced engagement, and learning motivation in diverse learning scenarios, while with the help of Artificial Intelligence, 'Cloud' computing, and other technologies to plan knowledge content and utilize learning resources, learners can be helped to build three-dimensional knowledge systems. With the realization of more diverse, creative, and imaginative scenarios, the metaverse not only creates a "new world" where learners are free to play, but all the participants are contributors based on immersion, autonomy, and the co-creation of values. The metaverse's ideological core of "co-construction, co-creation, co-rule, and sharing" embodies the ecological, pluralistic, and decentralized concepts pursued by postmodernism.

In the music teaching environment of the metaverse, the realistic and fully immersive nature of virtual reality not only has the power to "reconstruct" the real world, but also to build up creative teaching scenes according to the user's independent imagination. Just as the teaching environment is no longer homogeneous, online learning is no longer an imitation and transfer of offline learning but instead an environment completely different from the real world. In this new and distinctive reality, exploration of music knowledge is an "embodied engagement" based on full perceptual stimulation. In this kind of open teaching environment, learners' perceptual experiences break through the boundaries of space and time, allowing the exercise of any form of creative and innovative thinking. Moreover, the use of multimodal learning resources blurs the boundaries between reality and the virtual world, so that the processes of creation and learning are simultaneously the processes of building and enrichment.

In addition to reshaping teaching scenarios and content, the teacher's identity as the dominant player will be reinterpreted. Firstly, each teacher in the virtual scenario will operate as an avatar, and "the avatar's characteristics will implicitly influence the user's cognition, attitude, and behavior", while the flexibility and freedom of virtual avatars to intermingle with various teaching scenarios will serve to better motivate learners. Secondly, virtual teachers shaped by intelligent technologies will also have a place in tutoring activities, significantly reducing the repetitive work of real teachers in terms of explaining knowledge, supervising the effects of learning, and processing learner learning data in real-time. The result of this "dual-teacher parallel" will be that the real teachers can focus more on content design, scene construction, emotional communication, and other curriculum materials. In addition, the inclusive and sharing nature of the metaverse creates value for the virtual world by allowing everyone to use their expertise and opening up an era of "teachers for all" in which everyone is given access to the learning resources they need. Learning from others in this way will have the effect of diluting teachers' absolute authority and building a harmonious educational environment where teaching and learning co-exist.

This change in the role of the teacher will also bring about changes in the teaching and learning process. The pursuit of knowledge in an interactive teaching and learning environment becomes an “adventure”, an embodied experience that sees learners not as passive recipients, but as active constructers to understand and apply information directly. It is anticipated that exploratory, project-based, and game-based teaching models will become the main forms of learning; in this scenario, the focus of the music education metaverse will be on developing learners’ sensory stimulation, and helping them develop in a more holistic way. The teaching environment of the metaverse not only provides students with a convenient practice channel for solving real-world problems, but also brings about the subversion of inherent cognition through a process of continuous investigation.

Enjoying music learning in the metaverse

Avoiding the “impossible triangle” of quality, affordability, and efficiency that represents one of the main restrictions to the balanced distribution of education resources by changing the educational scene and the presentation of rich knowledge, the metaverse also allows for high-quality educational resources to be replicated, while simultaneously disseminating and complementing the interactivity and practicability previously lacking in online education. In the music metaverse, the boundaries of time and space are broken; global learning becomes instantly attainable through the dramatic reduction of equipment costs and the continuous enrichment of resources, while with the efficient transmission of 5G technology, 4K HD presentation, and naked eye 3D technology, combined with various interactive means, learners are no longer spectators in front of the screen but personally and actively involved in various types of music learning and music performance activities. In this virtual world, access to any music content not only becomes easy, but high-quality, low-cost music resources provide more music-loving learners with opportunities to improve their abilities, in so doing promoting the equitable development of music education. In addition, the "sharing and building" environment provides a convenient and efficient platform for more capable and talented music creators and performers to display their talents. The idea of "letting the people with the ability to enhance others become the teacher", advocated by Confucius thousands of years ago, will be reflected in the metaverse.

In November 2021, the famous Canadian singer Justin Bieber held his first metaverse concert on the virtual music platform 'Wave'. Not only did it provide fans with a gorgeous view of the performance, but the virtual image was also particularly vivid (See Figure 2). The concert broke down many of the barriers between the singer and the audience by allowing the latter to move to different corners of the scene and change the viewing angle. Although there was not much interaction, the audience could use certain controls to customize their viewing experience, while the high speed and low latency network allowed the singer's real body and the virtual image to be synchronized in real-time, allowing the audience to see the performer's every move. This scene has since been widely used in music teaching, enabling learners to obtain implicit knowledge from multi-angle observation, build various scene environments independently, and learn more about the areas of continuous display, performance, and interaction.

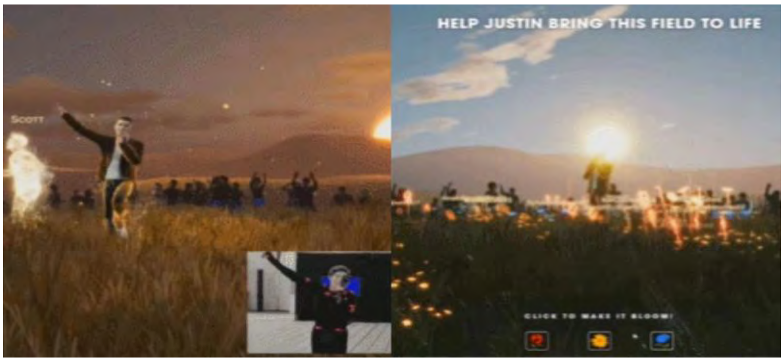


Figure 2 *The singer Justin Bieber performs via the virtual music platform, 'Wave'*

In the 2022 Jiangsu TV New Year's Eve concert, the appearance of the virtual persona Deng Lijun[1] generated lots of buzz (See Figure 3).

[1] Deng Lijun (1953-1995), one of the representative Chinese female singers in the 1970s and 1980s.

Figure 3 *Virtual persona Deng Lijun appearing on Jiangsu TV's New Year Concert*[2]

She can interact and talk as if she were a real person and sing the classic songs of the past exquisitely. The emergence of avatars in the metaverse not only replaces people as a means of completing the infrastructure, but also makes it possible to create a variety of them according to personal preferences. In the future metaverse of music education, avatars can serve as companions, learning assistants, and "human libraries", with a solid knowledge base to solve all kinds of problems faced by learners at any given time, in the process freeing teachers from the repetitive and mechanical tasks of teaching activities and allowing them to shift their focus to better motivating learners and designing new, high-quality and compelling content.

In addition to changing the traditional music teaching model, information and images in the metaverse environment will be integrated more appropriately in order to meet the demands of different types of knowledge. Theoretical knowledge that previously required the deliberate memorization of facts will be carried out in interactive scenarios through dialogue, interaction, and game activities, while the associated tacit knowledge can be experienced through task-based and breakthrough learning modes. Inner meanings and specific details of different types of knowledge and skills can be experienced through repeated challenges. Regarding practice and creative content, realistic scenarios can be projected into virtual environments, or new scenarios generated based on extensive data analysis to match students' learning preferences. For example, the 'Electronauts' music experience software developed by Survios, a game developer, can meet the various needs of creators through the real-time adjustment of parameters in order to obtain different auditory and visual effects. 'Foldit' is an experimental video game developed by the University of Washington. In 2011, 'Foldit' presented its users with a mission to find the protein structure needed to treat the AIDS virus, a problem that had plagued researchers for ten years; it was solved successfully within ten days by 60,000 players. In just the same way, future-oriented music teaching scenarios could well see participants from different backgrounds taking part in problem-solving processes together, and applying the results to the real world. Breaking down the boundaries and moving towards the integration of the virtual and the real in this way can result in a better and more just society for everyone.

Many of the existing rules will no longer apply in future music learning. Instead, new musical knowledge and experiences will be created through conversation, play, challenges, and collaboration, and music learning will be more participatory, autonomous, and creative. The diversity and inclusiveness of music education in the metaverse will also encourage participants to be more open to self-realization in different contexts.

Conclusions

The French literary scholar, Flaubert, once said, “The further we go, the more art becomes scientific, and at the same time, the more science becomes artistic, the more the two part from the foothills and reunite at the summit”. As the generation most deeply affected by technological changes, we should think about the future and face the current metaverse craze rationally and objectively. As music learners and teachers, we need to see it as an extension of the real world, not as a fantasy ‘utopia’ with technological support. Building the metaverse is a long-term commitment. Nevertheless, difficult as the journey may be, only sustained action can lead us to the destination. The future will undoubtedly usher in its own unique developments, enriching people’s music-learning lives in ways that are unimagined at present. The trumpet of the metaverse has been sounded. We stand on the threshold of a significant new era.

References

Bian, Y. L., Han, L., Zhou, C., Chen, Y. M., & Gao, F. Q. (2015). The Proteus Effect in Virtual Reality Social Environments: Influence of Situation and Shyness. *Acta Psychologica Sinica*, *3*, 363-374.

Coelho de Souza, R. (2020). In the Era of Computers, Internet and Multimedia, are we still Teaching Composers to become Chapel-Masters? Revista Vórtex, (1). 1-14.doi:10.33871/23179937.2020.8.1.1-14.

Hua, Z. X., & Huang, M. X. (2021). The Teaching Field Structure, Key Technologies and Experimental Study of Edu-Metaverse. *Modern Distance Education Research*, *6*, 23-31.

Liu, G. P., Wang, X., Gao, N., & Hu, H. L. (2021). From Virtual Reality to Metaverse: A New Direction of Online Education. *Modern Distance Education Research*, *6*, 12-22.

New Media Research Center of Tsinghua University. (2021). *Research Report on the Development of the Metaverse from 2020 to 2021*. Retrieved from: https://baijiahao.baidu.com/s?id=1714658282730354191&wfr=spider&for=pc

Zhong, Q. Q. (2021). *Deep learning.* Shanghai: East China Normal University Press.

Zhou, C. Z. (1983). *Creative psychology*. Beijing: China Youth Publishing House.

Asia-Pacific Journal for Arts Education

Special Issue

Unfold the Future of Music Education through Technology

Co-editors:

Prof. Bo Wah LEUNG

Dr. Cheung On TAM

Dr. Chi Hin LEUNG

Dr. Richard Guy WHITBREAD

The Education University of Hong Kong

http://www.ied.edu.hk/ccaproject/apjae/apjae.htm ISSN: 1683-6995

Volume 22 Number 2 December 2023

The Flexible-learning Approach in Higher Education Music History Courses

Dr. Pan-hang Tang

The Education University of Hong Kong

bphtang@eduhk.hk

Abstract

This study aims to investigate whether the flexible-learning approach may or may not foster student engagement in learning music history at the tertiary level. This approach provides students with a certain degree of autonomy in terms of deciding what to learn and how that learning should be assessed. Hence, it is highly learner-centred and deemed effective in cultivating students' learning interests. The flexible-learning approach has received global attention in recent years, particularly in Europe, where both policymakers and teachers in higher education have borrowed it for the purposes of curriculum planning and course design. To determine the impact of the flexible-learning approach on music history education, this study employed a retrospective post-pre survey in order to measure students' learning engagement. The results demonstrate that a flexible-learning approach can effectively enhance students' motivation in learning, implying that a single music history course does not need to stick to the traditional linear design, but should provide multiple trajectories for students to learn according to their interests and abilities.

Key words

flexible learning, music history, post-pre survey, retrospective evaluation, higher education

Introduction

Studying music history is essential for music students at the higher education level. However, mere lecturing can be boring. Various pedagogies have been developed as a means of enhancing students' engagement in learning. For example, Scarnati & Garcia (2008) adopted the "reverse chronological sequence" approach for jazz music history and theory education: allowing the students, as a first step, to learn from their favourite musical genres, succeeded in arousing their interests. Niren (2014) used a role-playing game, in which students portrayed various composers, in order to motivate them to learn the style and development of nineteenth-century Romantic programme music. This teaching approach had a great impact on the students, the participants returning to the school several years after graduation to tell the teacher that they had never forgotten the composers they had acted. Strandberg (2017) recommended the active learning approach for music history education, letting her students decide the repertoire of a concert, arranging the musical pieces, and writing the programme notes. She believed that students benefited from understanding the style and historical background of the pieces that they had investigated and analysed. Similarly, Rachlin (2017) found the active learning approach helpful, explaining that when students were looking for the necessary information, they could learn more efficiently by being directly involved. In short, "the more digging you do, the more appreciation you come away with for a particular subject" (p. 46).

Some researchers have suggested utilising computer technology to improve teaching and learning efficacy. Folio & Kreinberg (2009) discussed how they enhanced the quality of music history education through the computer-mediated collaborative learning approach. In their project, they assigned a reading assignment to their students. After reading, the students worked together in small groups to respond to five open-ended questions through 'Blackboard', the authors stating that "students thought deeply about what they had learned through this type of assignment" (p. 169). Gomes et al. (2016) adopted augmented reality (AR) technology to motivate students to learn twentieth-century music history, allowing the participants to access the AR learning content using a specially designed mobile application. Moreover, Neo et al. (2011) developed a computer programme to create a virtual learning environment to motivate students to learn jazz music history.

Perhaps not surprisingly, fostering students' motivation and enhancing their engagement in history education is an issue in other subject areas. To improve teaching effectiveness, Hicks et al. (2004) suggested training students like professional historians and equipping them with fundamental skills, including chronological thinking, historical analysis of cause and effect, discussion, debate, and persuasive writing. Julien et al. (2018) helped students learn labour history through social media platforms, the students investigating and reporting a self-selected historical event and sharing their findings on social media platforms, where peers could act as key stakeholders and interact with each other. Agus & Nara Setya (2017) advocated a more scientific approach, called hypnoteaching, which combines "quantum learning, accelerated learning, power teaching, neuro-linguistic programming, and hypnosis" (p. 4).

All of these teaching methods and approaches have one thing in common. They address students' study and professional development needs by creating a learner-centred environment, in which students act as active participants to construct the necessary knowledge and skills effectively (Lee & Branch, 2018).

Rationale of the Research

The flexible-learning approach

In recent years, I have been experimenting with the flexible-learning approach in order to enhance student engagement in learning. The flexible-learning method provides some degree of autonomy for students to determine what to learn (Collins & Moonen, 2001; Taylor & Joughin, 1997) and how to assess that learning (Wanner & Palmer, 2015). Given that it also transforms the teacher into a facilitator who encourages quality communication and interaction among students (Bostock, 2018), it is considered to be an approach that teachers and researchers in many areas can use to cultivate students' learning interests. For example, Lewis (1994) documented how a mathematics teacher adopted the flexible-learning approach in his undergraduate vector calculus course. His students were permitted to finish any learning materials and determine their study route based on their preferences, with class tests at specific times providing the necessary checkpoints for students to adjust their learning pace. Cornelius et al. (2011) detailed the flexible-learning approach in their Teaching Qualification (Further Education) programme. Students needed to finish a set of activities in which they had the choice of working either individually or in small groups. Furthermore, Irvine & Cossham (2011) reported how they improved the teaching efficacy in their Library and Information

Study undergraduate programme, whereby students could access the online learning resources and decide what, where, and when to learn.

At The Education University of Hong Kong (EdUHK), I teach a history course, *MUS2250 Traditions and Practices of Music I*, for the Bachelor of Arts in Creative Arts and Culture as part of a Co-terminal double degree programme. This course is compulsory for Year 1 students in the first semester. Most of them have just completed their senior secondary school education, while others are music graduates or have transferred from an associate degree or a higher diploma programme from other tertiary institutions. The 39-hour course aims to discuss the development of traditional Western music from the Ancient Greek period to the contemporary, 21st century. It is therefore a challenge to cater to the individual study needs of students with diverse musical backgrounds and mixed abilities, while simultaneously aiming to cover a wide range of musical styles. As a result, the course presents an ideal opportunity to let students decide (to some extent) what to learn with respect to their interests and abilities.

Figure 1 shows the generic structure I have designed for this course. After a main topic, several optional sub-topics are available, followed by a formative assessment to consolidate students' understanding at the end of each topic block. This process repeats until a summative assessment concludes the whole course.

Figure 1 *The generic structure of the* MUS2250 Traditions and Practices of Music I *course*

Figure 2 provides the alternative design of a topic block. For example, the instructor can deliver the optional sub-topics twice in a row or split the main topic into two parts and put the flexible-learning content between them. Moreover, the main topic can function as a self-sustained block without any optional sub-topics.

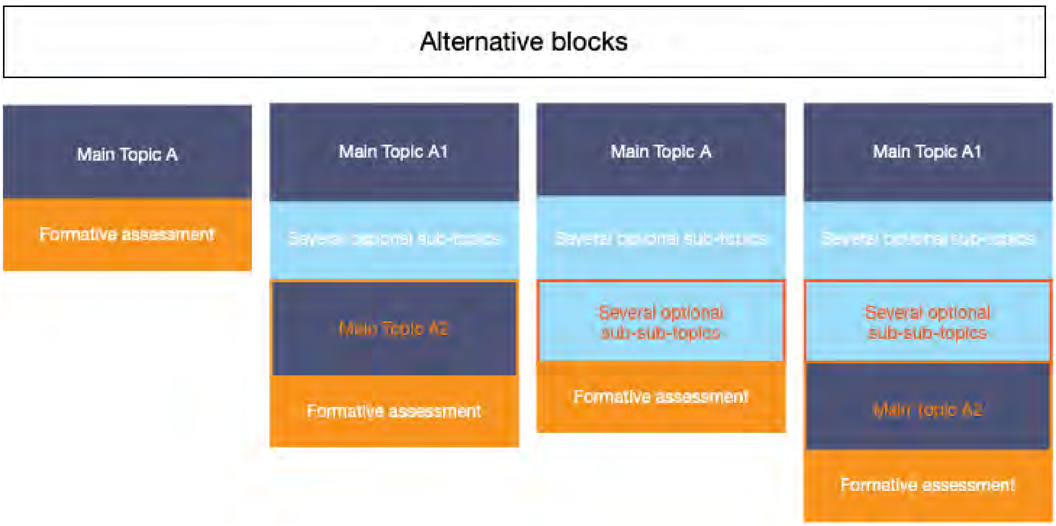


Figure 2 *Alternative topic block design*

Although this planning means that in theory the flexible-learning content can be delivered in different weeks, in practice he course takes longer to complete, an arrangement that is not ideal. Might it be possible to deliver the various sub-topics simultaneously in a way that is both cost-effective and time-efficient (Wild, 1994)? The magic of Information and Communication Technology reveals that the solution is actually quite simple. For example, Jackson (2020) investigated how the application of 'WhatsApp' might facilitate the process in a flexible learning environment. He concluded that 'WhatsApp' worked well in terms of fostering students' engagement in learning, while simultaneously enhancing interaction and discussion among teachers and students. The major advantage of using this app over a common Learning Management System, such as 'Moodle' and 'Blackboard', is that it is free to use for all parties, including the institute, the teachers, and the students.

Figure 3 is the course design of *MUS2250 Traditions and Practices of Music I*. The main topics occupy the middle line. There are a total of 13 of these, of which six consist of optional sub-topics. For example, after studying "The rise of humanism: the society in the Renaissance period", a student may choose "Renaissance instrumental music", "The Catholic tradition: Palestrina – The Prince of Music", or "The Religious Reformation and early Protestant Church music". The optional sub-topics either follow on top of or fall under the mainstream to indicate the different pathways towards the main "vocal / church music trend" and "instrumental music trend" topics. However, there are no restrictions limiting students to stick to a particular pattern. The only selection criteria are their own study needs and interests.

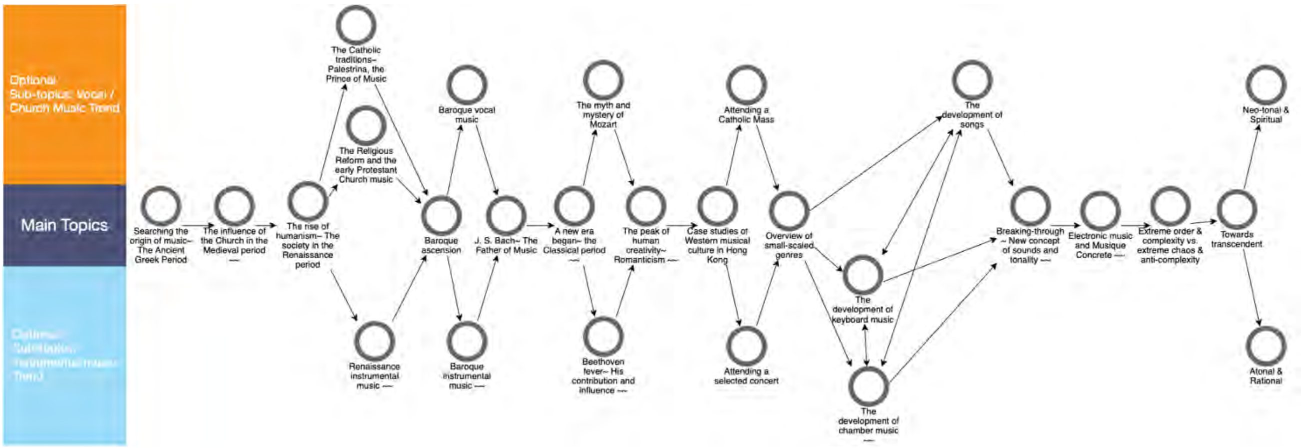


Figure 3 *Course design for* MUS2250 Traditions and Practices of Music I

In order to deliver different sub-topics simultaneously, the study materials and pre-recorded lecture videos were uploaded to ‘Moodle’ for students to access during the lessons. While the students were engaged with learning with the online materials, I supported them individually from the computer when they needed help. In order to assist with the consolidation of students’ understanding after each topic block, they were required to complete an online multiple-choice quiz, which was customised based on their choice of sub-topics. The ‘Moodle’ grading system helped monitor student learning progress by taking records and analysing the quiz results. Finally, a course essay – also personalised with regard to students’ learning interests – concluded the whole course as the summative assessment.

Research Questions

The above information and background led to the following research questions guiding the study:

- Can the flexible learning approach significantly enhance student engagement in learning music history?
- What other implications of the flexible-learning approach can we find in music history education at higher institutions?

Methodology

Definition of student engagement

To answer these two questions, we first need to define the meaning of "student engagement" and know how to measure it. O'Donnell & Reschly (2020) describe student engagement as a meta-construct containing three components in the form of behavioural, cognitive, and emotional engagement (Appleton et al., 2006, 2008; Reschly et al., 2014). Behavioural engagement is students' effort and persistence in learning (Birch & Ladd, 1997; Finn et al., 1995; Skinner & Belmont, 1993). Spending extra time learning and actively pursuing comprehensive knowledge outside class are examples of high behavioural engagement. Cognitive engagement means the value of striving for knowledge (Brophy, 1987; Newmann et al., 1992), e.g., students believing that what they have learned is essential to their professional and career development. Emotional engagement means students' reactions, such as finding pleasure in learning (Connell & Wellborn, 1991; Skinner & Belmont, 1993). For instance, students tend to enjoy listening to music they learn in class when they are emotionally engaged.

Research (Fredricks et al., 2004; O'Donnell & Reschly, 2020) tells us that it is possible to measure all three factors simply by conducting student self-evaluation.

The retrospective post-pre self-evaluation survey

To assess the degree to which the flexible-learning approach could enhance student engagement in my history course, I used a retrospective post-pre self-evaluation survey to measure changes in behavioural, cognitive, and emotional engagement. This post-pre approach stresses the changes in students' attitudes, knowledge, or skills regarding their perspectives (Hiebert et al., 2011), measuring the metamorphosis by comparing the extent of these attributes at the beginning and at the end of a learning process (Kanevsky, 2016). In the conventional pre-post procedure, the pre-test and post-test occur separately at a course's commencement and completion. However, even with noticeable improvement, the student may still rate themselves lower in the post-test because of a different standard of measurement after gaining the learning experience (Hiebert & Magnusson, 2014). The retrospective post-pre approach aims to avoid this inconsistency by putting the pre-test and post-test together at the end of the course so as to describe the actual changes more accurately (Bhanji et al., 2012).

Despite the criticism of subject bias (Nimon, 2014; Nimon et al., 2011), retrospective post-pre evaluation is highly regarded as a potentially effective tool for acquiring usable evidence of student learning (Coulter, 2012). Over the course of the past ten years, an increasing number of studies have adopted this approach in order to investigate student perspectives in learning. For example, Kanevsky et al. (2014) employed this tool to assess students' improvement after a two-year graduate diploma programme. Similarly, Stacey et al.

(2015) measured the impact of training on nurses' satisfaction using this evaluation method. Furthermore, Chen and O'Neill (2020) made use of this approach to measure changes in student engagement in a computer-mediated music composition activity. For this study, I came up with three descriptive statements in terms of students' behavioural, cognitive, and emotional engagement for every topic block listed in Table 1.

Table 1 *Title of the main topics and optional sub-topics*

Topic Block	Title of the main topic	Optional sub-topics
Topic 1	Search the origin of music: The Ancient Greek period	–
Topic 2	The influence of the Church in the Medieval period	–
Topic 3	The rise of humanism: The society in the Renaissance period	a. The Catholic tradition: Palestrina – the Prince of Music b. The Religious Reformation and the early Protestant Church music c. The Renaissance musical instruments and instrumental music
Topic 4	Baroque ascension	a. Baroque vocal music b. Baroque musical instruments and instrumental music
Topic 5	J. S. Bach: The Father of Music	–
Topic 6	A new era began: The Classical period	a. The myth and mystery of W. A. Mozart b. Beethoven fever: His contribution and influence
Topic 7	The peak of human creativity: Romanticism	–

Topic Block	Title of the main topic	Optional sub-topics
Topic 8	Case studies of Western musical culture in Hong Kong	a. Attending a selected concert b. Attending a Catholic Mass
Topic 9	Overview of small-scaled genres	a. The development of songs b. The development of keyboard music c. The development of chamber music
Topic 10	Breaking through: The new concept of sounds and tonality	–
Topic 11	Electronic music and Musique Concrète	–
Topic 12	Extreme order and complexity vs extreme chaos and anti-complexity	–
Topic 13	Towards Transcendent	a. Neo-tonal and spiritual b. Atonal and rational

Survey design

With regard to the design of the survey, the following descriptive statements guided students within each of the three forms of engagement. Participants were required to evaluate themselves for each topic block using a five-point Likert scale (0-4).

Behavioural engagement

I spend time reading about the information and listening to [music of the main topic and the chosen sub-topic(s), if applicable].

	Strongly Disagree →				Strongly Agree
	1	2	3	4	5
Before taking the course					
After taking the course					

Cognitive engagement

Understanding [music of the main topic and the chosen sub-topic(s), if applicable] is essential to my music study.

	Strongly Disagree →			Strongly Agree
	1	2	3	4	5
Before taking the course					
After taking the course					

Emotional engagement

I love listening to [music of the main topic and the chosen sub-topic if applicable].

	Strongly Disagree →			Strongly Agree
	1	2	3	4	5
Before taking the course					
After taking the course					

Results and Analysis

The data gathering process took place in the 2021-22 academic year, and the students took the post-pre survey on December 8, 2021, the date of the last lesson of the course. Due to suspension of the school caused by inclement weather, the lesson for Topic 9, "Overview of small-scaled genres", and the related optional sub-topics, were cancelled. Table 2 shows the results of the survey.

Table 2 *Results of the survey*

n = 35 Response rate = 92.11% Cronbach's alpha = 0.955			1	2	3	4	5	Mean	SD	Difference between the means	*t*	*p*	Effect size (Cohen's *d*)
Topic 1	Behavioural	Pre-	12	18	5	0	0	1.80	0.68	1.23	7.47	<0.001	1.72
		Post-	0	8	19	7	1	3.03	0.75				
	Cognitive	Pre-	5	13	10	6	1	2.57	1.04	1.03	6.59	<0.001	1.14
		Post-	0	1	16	14	4	3.60	0.74				
	Emotional	Pre-	8	11	11	4	1	2.40	1.06	0.94	5.43	<0.001	1.00
		Post-	1	3	15	15	1	3.34	0.80				
Topic 2	Behavioural	Pre-	10	16	6	3	0	2.06	0.91	1.03	6.38	<0.001	1.09
		Post-	2	7	14	10	2	3.09	0.98				
	Cognitive	Pre-	7	9	10	7	2	2.66	1.19	1.09	6.18	<0.001	1.05
		Post-	0	2	12	14	7	3.74	0.85				
	Emotional	Pre-	6	13	11	5	0	2.43	0.95	0.97	5.11	<0.001	1.03
		Post-	1	4	14	12	4	3.40	0.95				
Topic 3	Behavioural	Pre-	9	13	10	3	0	2.20	0.93	1.20	6.00	<0.001	1.30
		Post-	1	4	13	14	3	3.40	0.91				
	Cognitive	Pre-	6	11	9	9	0	2.60	1.06	1.03	4.97	<0.001	1.07
		Post-	0	2	15	12	6	3.63	0.84				
	Emotional	Pre-	5	9	14	6	1	2.69	1.02	0.89	4.99	<0.001	0.98
		Post-	0	3	12	17	3	3.57	0.78				
Topic 4	Behavioural	Pre-	6	7	13	8	1	2.74	1.09	1.09	6.02	<0.001	1.14
		Post-	0	1	11	16	7	3.83	0.79				
	Cognitive	Pre-	2	7	9	12	5	3.31	1.13	0.69	4.35	<0.001	0.71
		Post-	0	1	7	18	9	4.00	0.77				
	Emotional	Pre-	1	7	13	14	0	3.14	0.85	0.63	4.83	<0.001	0.74
		Post-	0	3	8	18	6	3.77	0.84				
Topic 5	Behavioural	Pre-	5	2	10	17	1	3.20	1.11	0.66	5.08	<0.001	0.69
		Post-	0	2	7	20	6	3.86	0.77				
	Cognitive	Pre-	1	2	9	18	5	3.69	0.90	0.51	4.34	<0.001	0.63
		Post-	0	1	3	19	12	4.20	0.72				
	Emotional	Pre-	1	6	9	17	2	3.37	0.94	0.49	4.36	<0.001	0.53
		Post-	0	2	10	14	9	3.86	0.88				
Topic 6	Behavioural	Pre-	0	3	13	13	6	3.63	0.88	0.54	3.93	<0.001	0.64
		Post-	0	1	6	14	14	4.17	0.82				
	Cognitive	Pre-	0	2	11	13	9	3.83	0.89	0.57	4.35	<0.001	0.73
		Post-	0	0	3	15	17	4.40	0.65				
	Emotional	Pre-	0	2	11	13	9	3.83	0.89	0.51	4.34	<0.001	0.62
		Post-	0	0	6	11	18	4.34	0.76				
Topic 7	Behavioural	Pre-	0	5	15	5	10	3.57	1.07	0.57	5.16	<0.001	0.59
		Post-	0	1	7	13	14	4.14	0.85				
	Cognitive	Pre-	0	1	12	14	8	3.83	0.82	0.51	4.34	<0.001	0.68
		Post-	0	0	4	15	16	4.34	0.68				
	Emotional	Pre-	0	2	11	12	10	3.86	0.91	0.34	3.43	0.002	0.41
		Post-	0	0	7	14	14	4.20	0.76				
Topic 8	Behavioural	Pre-	8	7	15	2	3	2.57	1.17	0.49	4.36	<0.001	0.44
		Post-	2	8	15	6	4	3.06	1.06				
	Cognitive	Pre-	0	7	16	11	1	3.17	0.79	0.63	4.61	<0.001	0.78
		Post-	0	1	13	13	8	3.80	0.83				
	Emotional	Pre-	1	6	14	8	6	3.34	1.06	0.49	3.87	<0.001	0.49
		Post-	0	2	12	11	10	3.83	0.92				
Topic 9	The lesson was cancelled due to Typhoon no. 8 bad weather signal.												
Topic 10	Behavioural	Pre-	9	11	9	5	1	2.37	1.11	0.86	5.37	<0.001	0.81
		Post-	0	9	14	7	5	3.23	1.00				
	Cognitive	Pre-	2	15	8	8	2	2.80	1.05	0.80	4.76	<0.001	0.83
		Post-	0	3	14	12	6	3.60	0.88				
	Emotional	Pre-	3	12	11	8	1	2.77	1.00	0.69	4.51	<0.001	0.69
		Post-	1	4	13	12	5	3.46	0.98				
Topic 11	Behavioural	Pre-	8	15	8	4	0	2.23	0.94	0.63	4.41	<0.001	0.65
		Post-	2	13	9	10	1	2.86	1.00				
	Cognitive	Pre-	5	9	13	8	0	2.69	0.99	0.54	3.63	<0.001	0.57
		Post-	1	6	14	12	2	3.23	0.91				
	Emotional	Pre-	8	8	16	2	1	2.43	1.01	0.54	3.09	0.004	0.53
		Post-	4	6	13	11	1	2.97	1.04				
Topic 12	Behavioural	Pre-	12	14	7	1	1	2.00	0.97	0.66	5.36	<0.001	0.67
		Post-	4	11	15	3	2	2.66	1.00				
	Cognitive	Pre-	10	14	7	3	1	2.17	1.04	0.77	4.42	<0.001	0.70
		Post-	4	8	13	6	4	2.94	1.16				
	Emotional	Pre-	7	16	8	2	2	2.31	1.05	0.71	4.42	<0.001	0.66

n = 35 Response rate = 92.11% Cronbach's alpha = 0.955			1	2	3	4	5	Mean	SD	Difference between the means	*t*	*p*	Effect size (Cohen's *d*)
		Post-	3	8	13	7	4	3.03	1.12				
Topic 13	Behavioural	Pre-	10	15	9	1	0	2.03	0.82	0.74	4.35	<0.001	0.83
		Post-	2	14	10	8	1	2.77	0.97				
	Cognitive	Pre-	7	13	9	6	0	2.40	1.01	0.86	4.66	<0.001	0.83
		Post-	2	6	12	11	4	3.26	1.07				
	Emotional	Pre-	8	17	6	4	0	2.17	0.92	0.69	4.88	<0.001	0.75
		Post-	2	10	15	7	1	2.86	0.91				

There were 38 students in the class, 35 of whom returned the questionnaire (N = 35; response rate = 92.11%). The alpha coefficient was 0.955, revealing a high level of internal consistency. The mean scores of the pre- and pro-test of all three factors of every topic were calculated, along with the standard deviations, in order to determine the data dispersion. The differences between the mean scores of the pre- and pro-test measured the extent to which student attitudes had changed. The results were larger than 0.0, indicating that positive growth occurred in all items after the course. With the exception of emotional engagement in Topic 7, all items changed significantly, with an increasing rate higher than, or near, 0.5. Some items, including all attributes in Topic 1, Topic 2, and Topic 3, as well as the behavioural engagement of Topic 4, even resulted in an intensively increasing rate larger than, or near, 1.0.

I also performed paired sample t-tests to compare the means of the pre- and post-test values of all the items. The *t-value* of every item in Topic 5 was consistently large (ranging from 3.09 to 7.47). At the same time, the two-tailed *p-value* remained extremely small (<0.001 in most cases), proving that the results were statistically significant. In addition, the effect size (Cohen's *d*) provides further evidence demonstrating the extent of the growth. Nineteen items recorded a medium value of >0.5, including the cognitive and emotional engagement of Topic 4, the behavioural and cognitive engagement of Topic 7, the cognitive engagement of Topic 8, and the emotional engagements of Topic 10 and Topic 13, as well as all the attributes of Topics 5, 6, 11 and 12. Fourteen items, including all the attributes of Topics 1, 2 and 3, the behavioural engagement of Topic 4, and the behavioural and cognitive engagement of Topic 10, as well as the behavioural and cognitive engagements of Topic 13, recorded a high value of >0.8. Only the emotional engagement of Topic 7, as well as the behavioural and emotional engagements of Topic 8, demonstrated relatively lower, but still substantial, values.

Accordingly, significant improvement occurred after the course across all items, apart from emotional engagement in Topic 7. This exceptional case resulted in the least extent of positive growth (0.34), the second weakest but still vital *t-value* (3.43), a relatively higher *p-value* (0.002), and the lowest Cohen's *d* (0.41). While implying that this attribute revealed the most trivial growth among the students, it should also be noted that a very high pre-test mean value of 2.86 was recorded for this attribute. As the highest in the course, it may have proved too difficult to garner any further improvement.

Table 3 lists the course average means and their differences in order to discuss the overall impact of the flexible-learning approach. Statistical significance analysis is unnecessary here because the course ratings cannot be insignificant when all the individual topic ratings are already statistically significant.

Table 3 *Course average ratings*

		Mean	Difference between means
Behavioural	Pre-	2.53	0.81
	Post-	3.34	
Cognitive	Pre-	2.98	0.75
	Post-	3.73	
Emotional	Pre-	2.90	0.66

Post-	3.55

Behavioural engagement started with the lowest pre-test mean value of 2.53, but jumped the most significantly (by 0.81) to reach 3.34 in the post-test, indicating that students' behavioural engagement changed from slightly disagreeing to slightly agreeing. On the other hand, cognitive engagement scored the highest mean value in both the pre- and post-test scores among all three aspects, with ratings of 2.98 and 3.73, respectively. The growth was 0.75, which is medium-high, meaning that the students, who were initially at a "Neutral" level in terms of their cognitive engagement, had strengthened by the end of the course to the level of "Agree". Lastly, emotional engagement achieved the second-highest pre-test mean value of 2.90. With the smallest, but still strong, increase of 0.66, this attribute rose to the second-highest post-test mean value with 3.55., demonstrating an observable improvement.

Conclusions and Implications

To summarise, students posted notable improvement gains in their behavioural, cognitive, and emotional engagement factors in the course *MUS2250 Traditions and Practices of Music I*, demonstrating in the process that the flexible-learning approach can enhance student engagement in music history education at the tertiary level.

Two decades ago, the flexible-learning approach received serious critiques (D. T. Chen, 2003), some researchers finding the quality of its learning outcomes to be unsatisfactory (Bell et al., 2001), while others questioned its cost-effectiveness (Ling et al., 2001). Positing different views, some researchers, such as Wild (1994) and Ratheswari (2018), had already foreseen the power of technology and its potential to overcome the previously-identified shortfalls. The flexible-learning approach now flourishes in higher education, especially in Europe, in countries such as Austria (Unger & Zaussinger, 2018), Finland (Moitus et al., 2020) and the United Kingdom (Brennan, 2021). In these locations, policymakers have brought the idea of the flexible-learning approach into curriculum planning in order to foster students' engagement and self-regulation. The flexibility of what, where, when, and how to learn transcends into a flexible-learning path. On the one hand, this offers students maximum control of their learning environment and routes. On the other, it promotes equity in education (Duarte et al., 2016).

Some teachers in tertiary institutes provide multiple study routes for completion of a single course, students being allowed to choose the best pathway to optimise their learning. For example, Duarte et al. (2016) put the flexible contents between two fixed-content lessons. After the first lesson, and before the next class, students were allowed to decide which learning tasks they wanted to take and in what order they wished to complete them. Meanwhile, Ibañez et al. (2022) suggested a six-sequence design. Each sequence consisted of its own set of worksheets, learning resources, and working tasks, students being given the freedom to select and deselect a learning sequence at any time during the duration of the course. Similarly, Rouhani et al. (2019) provided students with the flexibility to control their working progress based on the six modules making up their course, allowing them to choose those topics which were of most relevance to them.

In the same way, my history course provides students with the flexibility to decide the best study path to fit their learning needs. As such, a single music history course does not need to stick to the traditional linear structure, but can instead provide students with multiple trajectories based around their interests and abilities.

References

Agus, B., & Nara Setya, W. (2017). Hypnoteaching in History Lesson. *Journal Pendidikan Edutama*, *4*(2), 1-10. doi:10.30734/jpe.v4i2.50

Appleton, J. J., Christenson, S. L., & Furlong, M. J. (2008). Student Engagement with School: Critica Conceptual and Methodological Issues of the Construct. *Psychology in the Schools*, *45*(5), 369-386. doi:https://doi.org/10.1002/pits.20303

Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring Cognitive and Psychological Engagement: Validation of the Student Engagement Instrument. *Journal of School Psychology*, *44*(5), 427-445. doi:https://doi.org/10.1016/j.jsp.2006.04.002

Bell, T., Cockburn, A., McKenzie, B., & Vargo, J. (2001). Digital Lectures: If You Make Them, Will Students Use Them? Constraints on Effective Delivery of Flexible Learning Systems. *Interactive Multimedia Electronic Journal of Computer-enhanced Learning*, *3*(2). Retrieved from http://imej.wfu.edu/articles/2001/2/06/index.asp

Bhanji, F., Gottesman, R., de Grave, W., Steinert, Y., & Winer, L. R. (2012). The Retrospective Pre-Post: A Practical Method to Evaluate Learning from an Educational

Program. *Official Journal of the Society for Academic Emergency Medicine*, *19*(2), 189-194. doi:10.1111/j.1553-2712.2011.01270.x

Birch, S. H., & Ladd, G. W. (1997). The Teacher-child Relationship and Children's Early School Adjustment. *Journal of School Psychology*, *35*(1), 61-79. doi:10.1016/S0022-4405(96)00029-5

Bostock, J. R. (2018). A Model of Flexible Learning: Exploring Interdependent Relationships Between Students, Lecturers, Resources and Contexts in Virtual Spaces. *Journal of Perspectives in Applied Academic Practice*, *6*(1), 12-18. doi:10.14297/jpaap.v6i1.298

Brennan, J. (2021). *Flexible Learning Pathways in British Higher Education: A Decentralized and Market-based System*. Gloucester, United Kingdom: Quality Assurance Agency for Higher Education.

Brophy, J. E. (1987). Socializing Students' Motivation to Learn. In M. L. Maehr & D. A. Kleiber (Eds.), *Advances in Motivation and Achievement: Enhancing Motivation* (pp. 181-210). Greenwich, CT: JAI Press.

Chen, D. T. (2003). Uncovering the Provisos behind Flexible Learning. *Educational Technology & Society*, *6*(2), 25-30. Retrieved from http://ifets.ieee.org/periodical/6-2/3.html

Chen, J. C. W., & O'Neill, S. A. (2020). Computer-mediated composition pedagogy: Students' engagement and learning in popular music and classical music. *Music Education Research*, *22*(2), 185-200. doi:10.1080/14613808.2020.1737924

Collins, B., & Moonen, J. (2001). *Flexible Learning in a Digital World: Experiences and Expectations*. London: Kogan Page.

Connell, J. P., & Wellborn, J. G. (1991). Competence, Autonomy, and Relatedness: A Motivational Analysis of Self-system Processes. In *Self Processes and Development* (pp. 43-77). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Cornelius, S., Gordon, C., & Ackland, A. (2011). Towards Flexible Learning for Adult Learners in Professional Contexts: An Activity-focused Course Design. *Interactive Learning Environments*, *19*(4), 381-393. doi:10.1080/10494820903298258

Coulter, S. E. (2012). Using the Retrospective Pretest to Get Usable, Indirect Evidence of Student Learning. *Assessment & Evaluation in Higher Education*, *37*(3), 321-334. doi:doi.org/10.1080/02602938.2010.534761

Duarte, R., de Oliveira Pires, A. L., & Nobre, Á. L. (2016). *Increasing Adult Students' Learning Opportunities with Flexible Learning Pathways: Evidence from a Technology and Industrial Management Graduate Course.* Paper presented at the 2nd International Conference of the Portuguese Society for Engineering Education, Vila Real, Portugal.

Finn, J. D., Pannozzo, G. M., & Voelkl, K. E. (1995). Disruptive and Inattentive-Withdrawn Behavior and Achievement among Fourth Graders. *The Elementary school journal, 95*(5), 421-434. doi:10.1086/461853

Folio, C., & Kreinberg, S. (2009). Blackboard and Wikis and Blogs, Oh My: Collaborative Learning Tools for Enriching Music History and Music Theory Courses. *College Music Symposium*, *49/50*, 164-175.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School Engagement: Potential of the Concept, State of the Evidence. *Review of Educational Research*, *74*(1), 59-109. doi:10.3102/00346543074001059

Gomes, J., Figueiredo, M., Amante, L., & Gomes, C. (2016). Augmented Reality in Informal Learning Environments: A Music History Exhibition. *International Journal of Creative Interfaces and Computer Graphics*, *7*(2), 39-55. doi:10.4018/IJCICG.2016070104

Hicks, D., Carroll, J., Doolittle, P., Lee, J., & Oliver, B. (2004). Teaching the Mystery of History. *Social Studies and the Young Learner*, *16*(3), 14.

Hiebert, B., Bezanson, L., O'Reilly, E., Hopkins, S., Magnusson, K., & McCaffrey, A. (2011). *Assessing the Impact of Labour Market Information: Preliminary Results of Phase Two (Field Tests).* (Final report to Human Resources and Skills Development Canada). Canadian Career Development Foundation, Toronto. Retrieved from http://www.crwg-gdrc.ca/crwg/index.php/research-projects/lmi

Hiebert, B., & Magnusson, K. (2014). The Power of Evidence: Demonstrating the Value of Career Development Services. In B. C. Shepard & P. S. Mani (Eds.), *Career Development Practice in Canada: Perspectives, Principles and Professionalism* (pp. 489-530). Toronto: Canadian Education and Research Institute for Counselling.

Ibañez, J. S., de Benito Crosetti, B., García, J. M., & Carrió, A. L. (2022). New Flexible Designs and Modes of Organization in Higher Education: The Construction of Personal Learning Paths. *Pixel-bit: Revista de Medios y Educación*(63), 65-91. doi:10.12795/PIXELBIT.91739

Irvine, J., & Cossham, A. (2011). Flexible Learning: Reflecting on a Decade of Library and Information Studies Programmes at the Open Polytechnic of New Zealand. *Library Review*, *60*(8), 712-722. doi:10.1108/00242531111166728

Jackson, E. A. (2020). The Use of WhatsApp for Flexible Learning: Its Effectiveness in Supporting Teaching and Learning in Sierra Leone's Higher Education Institutions. *International Journal of Advanced Corporate Learning*, *13*(1), 35-47.

Julien, M., Stratton, M., & Clayton, R. (2018). History Is Not Boring: Using Social Media to Bring Labor History Alive. *Management Teaching Review*, *3*(3), 208-220. doi:10.1177/2379298117717257

Kanevsky, L. (2016). Assessing Students' Perceptions of the Effectiveness of Instructional Interventions with Post-Pre Surveys. Retrieved from https://www.sfu.ca/content/dam/sfu/istld/documents/Post-Pre_Kanevsky_handout.pdf

Kanevsky, L., Rosati, M., Schwarz, C., & Miller, B. (2014). *Post-pre Assessment of Effectiveness of a Two-year Graduate Diploma Program Offered in a Blended Format.* (Unpublished report). Simon Fraser University, Burnaby, BC.

Lee, S. J., & Branch, R. M. (2018). Students' Beliefs about Teaching and Learning and Their Perceptions of Student-centred Learning Environments. *Innovations in Education and Teaching International*, *55*(5), 602-610. doi:10.1080/14703297.2017.1285716

Lewis, P. (1994). A Flexibility Scheme for a First-year Mathematics Module. In W. Wade, K. Hodgkinson, A. Smith, & J. Arfield (Eds.), *Flexible Learning in Higher Education* (pp. 99-105). London: Kagan Page.

Ling, P., Arger, G., Smallwood, H., Toomey, R., Kirkpatrick, D., & Barnard, I. (2001). *The Effectiveness of Models of Flexible Provision of Higher Education*. Australia: Evaluations and Investigations Programme of the Department of Education, Training and Youth Affairs.

Moitus, S., Weimer, L., & Välimaa, J. (2020). *Flexible Learning Pathways in Higher Education: Finland's Country Case Study for the IIEP-UNESCO SDG4 Project in 2018–2021*. Finland: Finnish Education Evaluation Centre.

Neo, T. K., Neo, M., & Lim, M. (2011). The Use of Multimedia as an Innovative Method of Learning the History of Music: A Malaysia Experience. *International Journal of Instructional Media*, *38*(2), 187-196.

Newmann, F. M., Wehlage, G. G., & Lamborn, S. D. (1992). The Significance and Sources of Student Engagement. In F. M. Newmann (Ed.), *Student Engagement and Achievement in American Secondary Schools* (pp. 11-39). New York: Teacher College Press.

Nimon, K. (2014). Explaining Differences between Retrospective and Traditional Pretest Self-assessments: Competing Theories and Empirical Evidence. *International Journal of Research & Method in Education*, *37*(3), 256-269. doi:doi.org/10.1080/1743727X.2013.820644

Nimon, K., Zigarmi, D., & Allen, J. (2011). Measures of Program Effectiveness Based on Retrospective Pretest Data: Are All Created Equal? *American Journal of Evaluation*, *32*(1), 8-28. doi:doi.org/10.1177/1098214010378354

Niren, A. (2014). So You Think You Can Be a Composer: Pedagogical Strategies for Active Learning in Music History Courses. *College Music Symposium, 54*. Retrieved from www.jstor.org/stable/26574390

O'Donnell, K., & Reschly, A. L. (2020). Assessment of Student Engagement. In A. L. Reschly, A. J. Pohl, & S. L. Christenson (Eds.), *Student Engagement: Effective Academic, Behavioral, Cognitive, and Affective Interventions at School* (pp. 55-76). Cham: Springer International Publishing AG.

Rachlin, H. (2017). Use Band and Orchestra Compositions to Inspire Interest in Music History. *School Band & Orchestra*, *20*(9), 44.

Ratheeswari, K. (2018). Information Communication Technology in Education. *Journal of Applied and Advanced Research, 3*(1), 45-47. doi:dx.doi.org/10.21839/jaar.2018.v3S1.169

Reschly, A. L., Appleton, J. J., & Pohl, A. (2014). Best Practices in Fostering Student Engagement. In P. Harrison & A. Thomas (Eds.), *Best Practices in School Psychology: Student Level Services* (6th ed., pp. 37-50). Bethesda, MD:: National Association of School Psychologists.

Rouhani, M., Divitini, M., Vujosevic, V., Stai, S., & Olstad, H. A. (2019). *Design of a Programming Course for Teachers Supporting Flexible Learning Trajectories*. Paper presented at the Proceedings of the 8th Computer Science Education Research Conference, Larnaca, Cyprus. https://doi.org/10.1145/3375258.3375263

Scarnati, B., & Garcia, P. (2008). The Fusion of Learning Theory and Technology in an Online Music History Course Redesign. *Innovate: Journal of Online Education*, *4*(2), 1-6.

Retrieved from https://search-proquest-com.ezproxy.eduhk.hk/docview/61867160?accountid=11441

Skinner, E. A., & Belmont, M. J. (1993). Motivation in the Classroom: Reciprocal Effects of Teacher Behavior and Student Engagement Across the School Year. *Journal of Educational Psychology*, *85*(4), 571-581. doi:10.1037/0022-0663.85.4.571

Stacey, D., Skrutkowski, M., Carley, M., Kolari, E., Shaw, T., & Ballantyne, B. (2015). Training Oncology Nurses to Use Remote Symptom Support Protocols: A Retrospective Pre-post Study. *Oncology Nurse Forum*, *42*(2), 174-182. doi:10.1188/15.ONF.174-182

Strandberg, K. (2017). Music History Beyond the Classroom: Active Learning through Local History. *Journal of Music History Pedagogy*, *7*(2), 32.

Taylor, P., & Joughin, G. (1997). *What is Flexible Learning?* Queensland: Griffith University.

Unger, M., & Zaussinger, S. (2018). *The New Student: Flexible Learning Paths and Future Learning Environments*. Vienna, Austria: Institute for Advanced Studies, Vienna.

Wanner, T., & Palmer, E. (2015). Personalising Learning: Exploring Student and Teacher Perceptions about Flexible Learning and Assessment in a Flipped University Course. *Computers & Education*, *88*, 354-369. doi:10.1016/j.compedu.2015.07.008

Wild, P. (1994). Flexible Learning and Information Technology in Higher Education. In W. Wade, K. Hodgkinson, A. Smith, & J. Arfield (Eds.), *Flexible Learning in Higher Education* (pp. 35-45). London: Kagan Page.

Asia-Pacific Journal for Arts Education

Special Issue

Unfold the Future of Music Education through Technology

Co-editors:
Prof. Bo Wah LEUNG
Dr. Cheung On TAM
Dr. Chi Hin LEUNG
Dr. Richard Guy WHITBREAD
The Education University of Hong Kong

http://www.ied.edu.hk/ccaproject/apjae/apjae.htm ISSN: 1683-6995

Volume 22 Number 3 December 2023

Deep Neural Networks with Music Dereverberation for Technical Ear Training in Music Production Education

Manni Chen

City University of Hong Kong

manni.chen@my.cityu.edu.hk

Prof. PerMagnus Lindborg

City University of Hong Kong

pm.lindborg@cityu.edu.hk

Shuo Meng
City University of Hong Kong
shuomeng2-c@my.cityu.edu.hk

Abstract

Reverberation is an audio effect used in music production affecting the audio spectrum, as well as the timbre, of music. Technical ear training concerning the ways that reverberation works on music samples is important, since in order to achieve the desired texture sound engineers need to be able to perceive subtle audio changes. However, “dry” recordings, i.e. those that lack reverberation, are not universally available for music production students for the purposes of practice. In this paper, we propose a deep learning-based music dereverberation method to generate de-reverbed music samples for technical ear training in music production education. The experiment results show that based on various objective evaluation metrics, the proposed method can effectively realize dereverberation compared to other neural network-based methods.

Key words

(de)reverberation, audio spectrum, timbre, texture, technical ear training, music production

Introduction

Music production involves several workflow stages in the form of recording, mixing and mastering (Reiss et al., 2019). Recording typically refers to the technique of using microphones to collect the sounds of instruments, vocals, etc. to store into a media format. Mixing concerns the balance of individual musical tracks in terms of a series of audio features, including timbre and special location, while mastering deals with the final polish to the mix before distribution. During the music production process, audio effects, defined as "the controlled transformation of a sound typically based on some control parameters" (Wilmering et al., 2020, p. 791), are crucial tools for forming music texture. According to Wilmering et al. (2013), audio effects can be classified by their perceptual attributes as loudness, duration and rhythm, pitch and harmony, space, and timbre or quality. The "loudness" group can also contain audio effects (compressors, for example, possesses parameters including thresholds, attack time, release time, and make-up), while the "space" group includes reverberation, where pre-delay, reverb time, and wet constitute some of the more familiar parameters.

Reverberation is an audio effect that is widely used in music production applications. It affects the timbre of music and influences the perception of sound space (Zölzer, 2011). In this case, applying reverberation appropriately is essential in terms of creating the appropriate timbral and spatial attributes. However, adding reverberation is about tuning a series of parameters in order to control the audio transformation. Utilizing reverberation as a means of reaching the ideal perceptual target requires critical listening during the music production process, which means that sound engineers are required to make decisions on the setting of reverberation parameters based on their perceptual judgements. These perceptual judgements cover being aware of the indistinct musical discrepancies that exist when the parameters of audio effects are modified. In this respect, training in recognizing these almost unnoticeable changes is crucial for sound engineers. Consequently, the ability to understand how reverberation works on the perception of audio relies on a high level of technical ear training.

Technical ear training and reverberation

In traditional music education, ear training is a necessary skill for music composition and performance. In the field of music production, technical ear training is vital for work with audio signals through an ability to discern audio features. Corey & Benson proposed that technical ear training is a type of “perceptual learning focused on the timbral, dynamics and spatial attributes of sound” (2016, p. 5), and as such is an integral element of music production education.

Music production refers to the application of a series of technical tools to produce music. In addition to possessing a deep understanding of the theoretical background of sound production, sound engineers also need to be perceptually aware of music. In order to combine the scientific and artistic contexts, critical listening skills, based on the subjective decisions of sound engineers in response to what they hear, can serve as a bridge linking technical knowledge and musical aesthetics. Specifically, the operations involved in music production, including balancing the mix and adding audio effects, rely heavily on the ability to apply listening skills to achieve the targeted texture by tuning a series of parameters on audio software and hardware devices. A person without the necessary level of technical ear training is not capable of differentiating the subtle details of sound. For example, hearing the differences between quicker and slower attack times in a compressor as a precursor to making a decision about the most suitable parameter is acritical listening skill frequently called upon in music production. The ability to make judgements based on perceptual differences and translate these into technical modifications as they relate to audio signals, thereby attaining the required music production goals, is thus a requisite for every sound engineer. In this sense, the practice of acquiring critical listening skills through technical ear training is an inevitable and vital attribute.

1.1 Digital reverberation

With the introduction of digital technology has come the digitisation of audio effects for contemporary music production. Digital Audio Effects (DAFx) are transformations that blend audio signals, the modifications being conducted through a series of control parameters (Chourdakis & Reiss, 2017). Reverberation refers to the reflections of the delayed and attenuated copies of a direct sound (Zölzer, 2011); digital reverberation, as one of the DAFx, is a room acoustics modelling technique in which the technology is used to mimic those impulse responses. The responses are themselves divided into the direct sound, early reflections, and late reflections. The direct sound is the original sound triggered in the room, the reverberation occurring when the duplicates of the direct sound form as reflections. Defined by the directivity of the original sound and the physical properties of the surfaces, early reflections occur immediately after the sound event, while late reflections (perceived as decaying sound) are actually closely related to echoes. Accordingly, by mimicking the impulse responses in the room, digital reverberation, as adopted in music production, is able to bring the modifications of spaciousness, depth, and distances to the music, in the process creating different "environments".

Technical ear training can be helpful for music production students as a means of cultivating the ability to identify the changes brought about by reverberation. The learning materials for this ear training demand an ample number of music samples with reverberation effects (wet versions) and their corresponding dry recordings. However, it is unusual for the music production industry to release wet and dry version recordings at the same time (Martínez-Ramírez, 2022). Moreover, the recordings that *are* available are produced with added reverberation, which has led to a lack of (public) dry recordings. The result from the students' point of view is that merely listening to the reverbed version of a piece of music will not enable them to distinguish between, and pass judgement on, how reverberation works and what kind of reverberation suits a particular style of music. For instance, the amount and type of reverberation added to Baroque and Romantic pieces of music are likely to be very different, partly because of the sonic environments and reflecting surfaces in architectural buildings from these two very different historical eras. If the only accessible ear training materials are produced music, students will have no idea how to properly add reverberation after recording Baroque music, since they are unable to perceive the distance between un-reverbed Baroque recordings and the corresponding wet versions. Conversely, for ethnic music and instruments, dry and clean recordings for music production are more difficult to obtain compared to wet and coloured versions, leading to a lack of familiarity with the un-reverbed textures of individual instruments. The combination of the types of requirements required for reverbed and un-reverbed audio samples and the challenges of finding ways to remove reverberation effects in music in order to expand the number of learning materials for technical ear training consequently provided the impetus for the deep learning methods introduced within the design of neural networks outlined in this research project.

Existing deep learning methods

In recent years, neural networks have started to be introduced into audio applications. Mimilakis et al. (2016) utilized deep neural networks and the short-time Fourier transform (STFT) for the purposes of predicting a coefficient of dynamic range compression in music mastering applications, while the dilated residual network (DRN) (Grachten, 2018) and the convolutional neural network (CNN) (Ramírez & Reiss, 2018) have been applied to equalization in audio production. Oord et al. (2016) and Jang et al. (2021) proposed 'WaveNet' and 'UnivNet', respectively, for generating raw audio from text. However, the generative adversarial network (GAN) (Goodfellow et.al., 2020) has found the most favor among researchers in the field of audio application. Su et al. (2020) introduced 'HiFi-GAN', which combines 'WaveNet' and GAN for speech application, while Yamamoto et al. (2020) designed 'Parallel WaveGAN' for text-to-speech via the adoption of 'WaveNet' and a multi-resolution spectrogram. AlBadawy et al. (2022) proposed a neural vocoder which converts the spectral representations of audio signals into the generation of high-fidelity waveforms in real time. In an exception to the speech genre, GAN has also been utilized to inverse the reverberation caused by the car environment (Pepe et al., 2020), while Kumar et al. (2019) introduced 'MelGAN' for generating waveforms, including a synthesis of speech and music.

However, music production education in terms of technical ear training has been ignored by existing neural networks. The focus of this paper is on music dereverberation for the purposes of expanding the amount of technical ear training materials, based on the pre-existing condition that the construction of neural networks for music dereverberation can be applied to the same deep learning methods used in speech. Since the deep learning training process is closely related to the spectrogram of audio materials, the differences between the frequency content of speech (used here to refer to the “standard” means of human vocal communication based around a narrow spectrum range and less complexity in terms of resonance compared to an individual’s singing voice) and music deserve to be highlighted. However, the presence of important differences between speech and music brings its own set of difficulties in terms of designing deep neural networks for music dereverberation. Additionally, some of the difficulties associated with dereverberation also lie within unknown acoustic environments and moving sources, both of which involve time-varying reverberation effects (Nercessian & Lukin, 2019).

As a result, we propose a GAN-based model with a ‘UNet’ (Ronneberger et.al., 2015) to achieve high-quality music dereverberation. First, the reverbed music samples are converted to full-band Mel spectrograms (Stevens et al., 1937) as input to gain frequency information. Then, a multi-scale ‘UNet’-based architecture is adopted to generate non-reverb Mel spectrograms. Finally, a GAN-based vocoder (containing a generator with a series of transposed convolutions that produce input-sized waveforms and a multi-period waveform discriminator to grasp both spectral and temporal domain features) is designed to synthesize fine-grained waveforms as the output.

Proposed method

Due to high temporal resolution and the dependencies of audio data, modeling raw audio is particularly challenging. Most current methods therefore adopt an indirect approach through the modeling of a lower-resolution representation (Kumar et al., 2019) in order to reconstruct high-resolution audio. Inspired by 'UnivNet' (Jang et al., 2021) and 'U-Net' (Ronneberger et al., 2015), the overall architecture presented here contains two parts: a 'UNet'-based de-reverberator and a GAN-based neural vocoder, which contains a generator and a discriminator. The model 'learns' from paired examples using mono reverbed and dry audio samples. The block diagram of the proposed model is displayed in Figure 1.

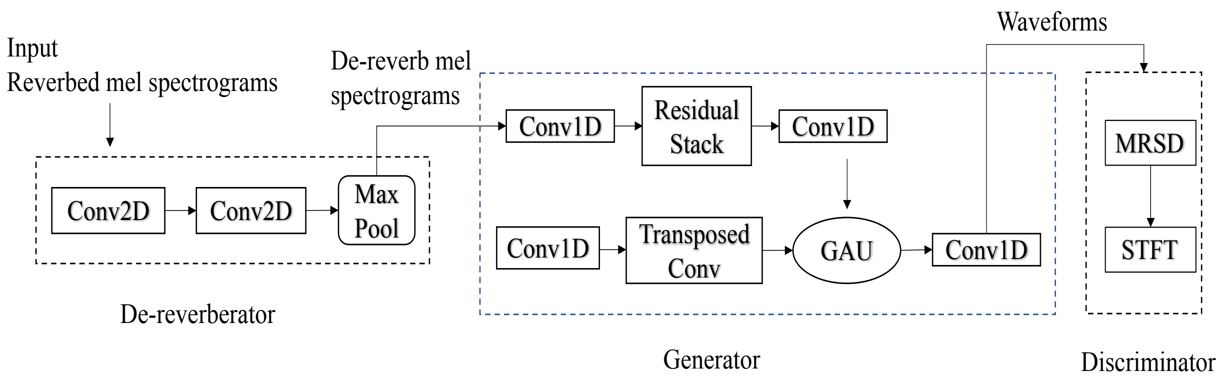


Figure 1 *Block diagram of the proposed model*

1.1 Architecture

'UNet' is widely used for image-to-image segmentation tasks (Ronneberger et al., 2015), dereverberation tasks (Ernst et al., 2018; Wang & Wang, 2020), and audio source separation tasks (Stoller et al., 2018). In this paper, we modify the 'UNet' structure for music dereverberation tasks; specifically, the architecture is a symmetric encoder-decoder network with skip connections which run on each encoder block and its mirror image between decoder blocks. As for the input data, we transfer the waveform to a Mel spectrogram rather than an STFT to feed into the neural networks, since an STFT contains more data points and a Mel spectrogram corresponds to human perceptions of sound, both of which facilitate the generation of de-reverbed music. Compared to using raw audio waveforms, a Mel spectrogram is easier to model and retains enough information to convert back to audio (Kumar et al., 2019). All waveforms were resampled to 22.5kHz using an STFT with 1024 FFT components, 1024 sample Hann window lengths, and 0-12kHz Mel spectrograms with 80 bands. Input signals were normalized using the mean and variance of the entire training set. The final output of the de-reverberator is the same size as the input in order to predict the de-reverberated Mel spectrogram.

Directly applying a 'UNet'-similar image-to-image architecture is hard to keep aligned with human perceptions. In this case, GAN-based neural vocoders are successfully applied to audio reconstruction tasks. Among them, 'UnivNet', one of the state-of-the-art neural vocoders, absorbed the advantages of various vocoders and used full band Mel spectrograms to obtain good speech reconstruction results. Consequently, this research adopted 'UnivNet' as the neural vocoder to conduct the music dereverberation tasks. The structure of the generator is the same as 'MelGAN', which used a background noise as the input and predicted the dry Mel spectrogram generated by the de-reverberator as the condition. Additionally, the location-variable convolution (LVC) and a gated activation unit (GAU) were added and a multi-resolution spectrogram discriminator (MRSD) (Jang et al., 2021), which employs a multi-spectrogram and a multi-period waveform discriminator to differentiate dry recordings and generate audio samples, adopted. The converted multiple spectrograms contain various temporal and spectral resolutions, while the multi-period waveform enhances adversarial modeling in the temporal domain.

1.2 Training losses

For the de-reverberator, the mean square error (MSE) loss was used, while for the neural vocoder, multi-resolution STFT loss (Yamamoto et al., 2020) as an auxiliary loss L_{aux}, and the commonly used least-squares GAN as the objective functions, were applied. The auxiliary loss L_{aux} combines the log STFT magnitude loss L_{mag} and spectral convergence loss L_{sc}. To improve the results, we added the MSE loss to the neural vocoder as another regularization term in order to compare the differences between the de-reverberated audio samples and the dry audio samples. All loss functions were defined as follows:

$$L_{MSE} = \sum_t \|x_t - \hat{x}_t\|^2$$

$$L_{sc}(s,\hat{s}) = \frac{\|s-\hat{s}\|_F}{\|s\|_F},\ L_{mag}(s,\hat{s}) = \frac{1}{S}\|\log s - \log\hat{s}\|_1$$

$$L_{aux}(x,\hat{x}) = \frac{1}{M}\sum_{m=1}^{M}\mathbb{E}_{x,\hat{x}}\left[L_{SC}(s_m,\hat{s}_m) + L_{mag}(s_m,\hat{s}_m)\right]$$

$$L_{GAN}(G,D) = \sum_t \left(\log D\,(Z_t, X_t) + \log\left(1 - D\left(Z_t, G(Z_t)\right)\right)\right)$$

$$L(G,D) = L_{GAN}(G,D) + \lambda L_{MSE}(G)$$

Experiments

1.1 Datasets

The dry audio samples were downloaded from the Chinese Bamboo Flute (CBF) dataset (Wang et al., 2019) and ‘The “Mixing Secrets” Free Multitrack Download Library’. The CBF dataset is a collection of 20 dry recordings of four Chinese bamboo flute pieces. As for the ‘Mixing Secrets’ dataset, in order to make sure that the audio samples contained more sustainable signals instead of silence as a means of facilitating the training process, multitrack excerpts, as opposed to the full multitrack, were used. Within the excerpt multitrack, we chose pop music and the dry recordings of songs from the multitrack, listening to every sample carefully in order to filter out the tracks that lacked continuous content, e.g., silence over three seconds, along with low-quality audio samples that included obvious artificial audio effects, such as pitch modification and lower clarity or breathing noise. The music samples were then cropped into blocks of three seconds each and the selected data separated into a training set consisting of 2400 audio samples and a testing set containing 600 randomly assigned audio pieces. Reverberation was added to the dry recordings via ‘Pedalboard’, a Python library containing the functionality of adding a series of audio effects. The built-in reverberation in ‘Pedalboard’, with parameters of a room size setting at 0.25, was adopted in order to make paired reverberated audio samples.

1.2 Evaluation metrics

In order to evaluate the performance of different models for dereverberation tasks, three objective metrics (Frequency-Weighted Segmental Signal-to-Noise Ratio (FWSegSNR), Log-Likelihood Ratio (LLR), and Signal to Distortion Ratio (SDR)) were employed. FWSegSNR is widely used for evaluating dereverberation performance (Kinoshita et al., 2016), defined as the average SNR ratio within a certain frequency band, the quality increasing with the value. LLR (Hu & Loizou, 2007) is also a conventional metric based on the LPC spectrum; suitable for the evaluation of reverberation effects, whereby those with smaller values are regarded as being of a better quality. SDR (Févotte et al., 2005) is a comprehensive index used to measure the sound quality of audio sources; the larger the value, the better the performance.

Results

The models 'MelGAN', 'Unet+MelGAN' and 'Unet+UnivNet-c32' were prepared for comparison. Since the original 'MelGAN' and 'UnivNet' act as neural vocoders, aimed at converting the Mel spectrograms into waveforms rather than dereverberation tasks, we added a cascade 'UNet' before them in order to realize the dereverberation. The cascade 'UNet' module took the Mel spectrogram of the reverbed audio as input to predict the dry Mel spectrogram. At the same time, we used 'MelGAN' to undertake the dereverberation tasks directly; this involved taking the Mel spectrogram of the reverbed sample as the input and predicting the dry audio waveforms directly. All models were based on the official implementations and the relevant hyper-parameters following the reference configurations of each one. To ensure stable sound quality, the consumption times were calculated on the same platform and all deep learning-based models were trained on NVIDIA RTX3090 GPUs until the loss tended to stabilize. The proposed model was trained with a batch size of 32 and an AdamW optimizer (Loshchilov & Hutter, 2018), with $\beta1 = 0.5$, $\beta2 = 0.9$ and a 1e-4 learning rate.

The results of the objective comparisons are summarized in Table 1. They show that using 'MelGAN' to generate de-reverbered audio samples directly failed to obtain ideal results, although it was the fastest model. By comparison, the neural network-based methods using 'UNet' cascaded resulted in a similar performance, illustrating that an indirect model is a requisite. The proposed method reaches the highest performance among the compared models in all objective metrics, except for speed, which can be tolerated in the real scenarios. Overall, the results demonstrate that the proposed method can realize high-quality music dereverberation while almost staying in real time.

Table 1 *Comparisons between different methods*

Model Name	**FWSegSNR**	**LLR**	**SDR**	**Speed**
'MelGAN'	6.57	1.03	7.80	**177.51**

'Unet + MelGAN'	9.09	0.51	9.17	142.86
'Unet + UnivNet'	10.58	0.44	9.14	118.11
Proposed Method	**12.31**	**0.29**	**9.67**	98.04

Note. The speed is calculated based on consumption time relative to real time.

Conclusions

In music production education, technical ear training is ineluctable when comparing un-reverbed and reverbed music. Moreover, the lack of un-reverbed samples has resulted in the obstacle of music production students failing to understand comprehensively how reverberation works. In order to tackle this problem, while simultaneously expanding un-reverbed music learning materials, this paper has proposed a GAN-based model to generate real time and high quality dereverberation music. The results of the different evaluation metrics indicate that our method is more capable of realizing music dereverberation compared to other neural network-based methods. Additionally, this implementation can be extended to other audio effects, such as delay, in the process exploring audio transformations as they pertain to technical ear training. Future explorations will include more universal, reality-sited dereverberation models based on unparalleled data, together with the extension of the applications to other audio effects in order to facilitate technical ear training within music production education.

References

AlBadawy, E. A., Gibiansky, A., He, Q., Wu, J., Chang, M. C., & Lyu, S. (2022, May). Vocbench: A Neural Vocoder Benchmark for Speech Synthesis. In ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 881-885). IEEE.

Corey, J., & Benson, D. H. (2016). *Audio production and critical listening: Technical ear training*. Routledge.

Ernst, O., Chazan, S. E., Gannot, S., & Goldberger, J. (2018, September). Speech dereverberation using fully convolutional networks. In 2018 26th European Signal Processing Conference (EUSIPCO) (pp. 390-394). IEEE.

Foote, C., Gribonval, R., & Vincent, E. (2005). BSS_EVAL toolbox user guide—Revision 2.0.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, *63*(11), 139-144.

Grachten, M., Deruty, E., & Tanguy, A. (2018). Auto-adaptive resonance equalization using dilated residual networks. arXiv preprint arXiv:1807.08636.

Hu, Y., & Loizou, P. C. (2007). Evaluation of objective quality measures for speech enhancement. *IEEE Transactions on audio, speech, and language processing*, *16*(1), 229-238.

Jang, W., Lim, D., Yoon, J., Kim, B., & Kim, J. (2021). UnivNet: A neural vocoder with multi-resolution spectrogram discriminators for high-fidelity waveform generation. arXiv preprint arXiv:2106.07889.

Kinoshita, K., Delcroix, M., Gannot, S. P., Habets, E. A., Haeb-Umbach, R., Kellermann, W., Leutnant, V., Maas, R., Nakatani, T., Raj, B., Sehr, A., & Yoshioka, T. (2016). A summary of the REVERB challenge: state-of-the-art and remaining challenges in reverberant speech processing research. *EURASIP Journal on Advances in Signal Processing*, *2016*(1), 1-19.

Kumar, K., Kumar, R., de Boissiere, T., Gestin, L., Teoh, W. Z., Sotelo, J., & Courville, A. C. (2019). Melgan: Generative adversarial networks for conditional waveform synthesis. Advances in neural information processing systems, 32.

Loshchilov, I., & Hutter, F. (2018). Fixing weight decay regularization in Adam.

Martínez-Ramírez, M. A., Liao, W.-H., Fabbro, G., Uhlich, S., Nagashima, C., & Mitsufuji, Y. (2022). Automatic music mixing with deep learning and out-of-domain data. In the 23rd International Society for Music Information Retrieval Conference (ISMIR). https://doi.org/10.48550/arxiv.2208.11428

Mimilakis, S. I., Drossos, K., Virtanen, T., & Schuller, G. (2016, May). Deep neural networks for dynamic range compression in mastering applications. In Audio Engineering Society Convention 140. Audio Engineering Society.

Nercessian, S., & Lukin, A. (2019, September). Speech dereverberation using recurrent neural networks. In Proceedings of the 23rd International Conference on Digital Audio Effects (DAFx-19), Birmingham, UK (pp. 2-6).

Oord, A. V. D., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., & Kavukcuoglu, K. (2016). Wavenet: A generative model for raw audio. arXiv preprint arXiv:1609.03499.

Pepe, G., Gabrielli, L., Squartini, S., & Cattani, L. (2020). Designing audio equalization filters by deep neural networks. *Applied Sciences*, *10*(7), 2483.

Ramírez, M. A. M., & Reiss, J. D. (2018, September). End-to-end equalization with convolutional neural networks. In 21st International Conference on Digital Audio Effects (DAFx-18).

Reiss, J., Stables, R., & De Man, B. (2019). *Intelligent Music Production: A Theoretical Overview*. Routledge.

Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Springer, Cham.

Stevens, S. S., Volkmann, J., & Newman, E. B. (1937). A scale for the measurement of the psychological magnitude pitch. *The Journal of the Acoustical Society of America*, *8*(3), 185-190.

Stoller, D., Ewert, S., & Dixon, S. (2018). Wave-u-net: A multi-scale neural network for end-to-end audio source separation. arXiv preprint arXiv:1806.03185.

Su, J., Jin, Z., & Finkelstein, A. (2020). HiFi-GAN: High-fidelity denoising and dereverberation based on speech deep features in adversarial networks. arXiv preprint arXiv:2006.05694.

Wang, C., Benetos, E., & Chew, E. (2019, November). CBF-PeriDB: A Dataset of Chinese Bamboo Flute Playing Techniques with Periodic Modulations. In 20th International Society for Music Information Retrieval Conference.

Wang, Z. Q., & Wang, D. (2020). Deep learning-based target cancellation for speech dereverberation. *IEEE/ACM transactions on audio, speech, and language processing*, *28*, 941-950.

Wilmering, T., Fazekas, G., & Sandler, M. B. (2013, October). Audio effect classification based on auditory perceptual attributes. In Audio Engineering Society Convention 135. Audio Engineering Society.

Wilmering, T., Moffat, D., Milo, A., & Sandler, M. B. (2020). A history of audio effects. *Applied Sciences*, *10*(3), 791.

Yamamoto, R., Song, E., & Kim, J. M. (2020, May). Parallel WaveGAN: A fast waveform generation model based on generative adversarial networks with multi-resolution spectrogram. In ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 6199-6203). IEEE.

Zhu, J.-Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. 2017 IEEE International Conference on Computer Vision (ICCV), 2242–2251. IEEE. https://doi.org/10.1109/ICCV.2017.244

Zölzer, U. (2011). DAFX: Digital Audio Effects, Second Edition (2nd ed.). Chichester, West Sussex, U.K: Wiley. The 'Mixing Secrets' Free Multitrack Download Library. Retrieved from https://www.cambridge-mt.com/ms/mtk/

Asia-Pacific Journal for Arts Education

Special Issue

Unfold the Future of Music Education through Technology

Co-editors:

Prof. Bo Wah LEUNG

Dr. Cheung On TAM

Dr. Chi Hin LEUNG

Dr. Richard Guy WHITBREAD

The Education University of Hong Kong

http://www.ied.edu.hk/ccaproject/apjae/apjae.htm ISSN: 1683-6995

Volume 22 Number 4 December 2023

Design of a College Teaching Model for Ethnic Vocal Singing Based on Artificial Intelligence Digital Algorithms

ZengYu Liu, MuXi Tang

Shenyang Conservatory of Music

tenorlou520@gmail.com

SiHao He

Shanghai University of International Business and Economics

sihaohe@126.com

Abstract

In recent years, there has been a surge in the development possibilities of ethnic vocal music performances and higher education teaching within the context of online education. The authors conducted a comprehensive investigation, taking into account the mixed drawbacks of conventional online education, in order to assess the current state of Internet technology development, a process leading to the development of a systematic model for summarizing the teaching system. In order to create a standardized model that integrates performance practice, teaching methods, and feedback mechanisms, the results of a statistical questionnaire survey conducted with 1657 individuals in the field of music in China were analyzed, as well as 22 Master's degree students in the Department of Ethnic Vocal Music at the Shenyang Conservatory of Music. The teaching mode is an integrated design of artificial intelligence, digital algorithm, vocal cord protection, vocal performance, applied music psychology, and vocal music teaching method, that supports the standardization and conscientization of teaching, as well as the visual display of evaluation records. To sum up, this model uses digital equipment, automatic management systems, and cyberspace as the carrier to naturally integrate national vocal music with the network using its software and extensive database, integrate resources in order to thoroughly improve teaching and performance practice, establish a professional ethnic vocal music teaching system, and address the drawbacks of online national vocal music teaching.

Key words

ethnic vocal music, vocal music performance, vocal music in higher education, vocal cord protection, artificial intelligence, digital technology

Introduction

Artificial intelligence (AI) has been used in many different industries due to the rapid expansion of the Internet. With its exceptional capacity for learning, AI has produced a large number of incredible artworks and melodies, including piano playing, in a short period of time (Civit et al., 2022). As a result of the multi-dimensional increase in the quality of teaching and singing, as well as the evaluation of teachers' instruction and the comprehension of course material by students, ethnic vocal music singing and instruction have also benefited from effective AI algorithms.

Modern online ethnic vocal music singing and instruction may be significantly improved by utilizing the real features of artificial intelligence. For example, AI can swiftly adapt to the performance and teaching environment by setting up an operating system that blends simulation, computation, learning, practice, feedback, and improvement. The flexible use of an offline teaching system, as discovered by Zhe (2021), is ideal for teaching vocal music in colleges and universities. Online instruction, on the other hand, is a relatively recent development that is still in the exploratory phase. The creation of a scientific, effective, and long-lasting online teaching model that can ensure the delivery of the educational material and even produce superior outcomes while utilizing resources sensibly (Wang, 2022) is thus an important first step.

The most effective approach to the transfer of knowledge is face-to-face, and vocal learning begins with imitation before moving on to secondary composition. Oral instruction, on the other hand, is less successful within the framework of online learning since students are unable to experience a genuine performance setting and are thus unable to completely comprehend what the instructor is trying to convey (Herbert, 2007). An AI-based education system thus offers advantages, such as the flexibility to study at any time and from any location, the capacity to provide expert services, and the assurance that outcomes obtained online can be reliably compared to those obtained offline (Lei, 2020).

The following research is presented in this paper:

(1) The teaching model, which supplements and adapts existing models, is based principally on the development of AI and is intended for Chinese folk vocal singing (Wei et al., 2022). This paradigm offers a strong basis for digitizing the instructional approach and course material for folk vocal music courses at universities, making it possible to create a “digital+” discipline and using it as a disciplinary instrument to support teaching and learning (Zhang, 2023).

(2) Comments on, and developments of, the features that users want from online teaching and learning tools, analyzing the opinions in a focused way. This feature involves the flexible use of AI (Yan, 2022), as well as the creation of easy-to-use instructional models for teachers and students.

(3) The development of particular concepts, the evaluation of findings from earlier studies, and an assessment of the viability of the ideas put forward in this study (Yu et al.,

2023).

The next four parts of the paper outline the research that was done. The first part covers the study's history and its guiding principles. The educational model design is described in the second part, together with outlining the prerequisites for performing and teaching folk vocal singing in schools, as well as describing the teaching methodologies that may be used for online delivery. It also explains the practical applications of vocal cord anatomy, music psychology, and vocal cord protection (Henrich et al., 2011).

Questionnaire analysis lays the foundation for AI design by examining what people are most interested in. The third part delves into the design concepts of artificial intelligence and an introduction to computers, while the fourth part summarizes the essential content of the text and introduces the teaching methods.

Related work

Since the development of the Internet, universities have experimented with different forms of teaching aids. For example, with the help of multimedia technology, the use of computers to produce class content, audio and video lectures, investigations into digital science as a means of managing tasks, and the development of small programs in chat software to assign homework (Webster, 2017), have all been explored.

In recent years, AI has made a splash in the arts, whether it be through conducting automatic piano playing or else writing a perfectly good opera excerpt in three seconds. In education, AI and the metaverse are also very much in the spotlight. As a result, there are a variety of opportunities for development within the realm of ethnic vocal singing and education.

Liu (2018) proposed the “twenty-four character formula” as the main design principle for teaching ethnic vocal music singing. Adopting this method to divide the teaching mode into dynamic, behavioral and cognitive modes, the voices of two tenors and two sopranos, a collection of singing data from national vocal music, and some music data from the piano and electric piano, were all collected as part of the research process. It was decided that these data would serve as the foundation for AI analysis and its design, the aims being to improve the audio and video quality of online instruction (Roads, 1985), examine users’ vocal cord maintenance, and compare the collected audio data as the standard reference for singing (Elghaish et al., 2022). On completion, it was hoped that AI could help teach national vocal music singing using its algorithm and imitation capability.

Liu et al. (2021) developed a comparative study between control and experimental groups based on the watch-summary-question (WSQ) vocal learning assessment model (Hsia, 2022), while Valle et al. (2016) found peer assessment (PA) to be a learning strategy that had the potential to engage students in knowledge building and skill development through understanding teachers’ guidelines for assessing, learning and reflecting on peers’ vocal practice. Aimed specifically at studying the impact of WSQ-based mobile PA on students’ vocal skills and learning perceptions, and founded on the hypothesis that utilizing electronic devices improves students’ abilities to better evaluate their peers’ vocal technical abilities (while also improving their own performance), vocal learning software was used to compare WSQ-based mobile PA and conventional mobile PA by focusing on the five learning units of vocal music (pitch accuracy, pitch centering, rhythm, expression or phrasing, and timbre quality).

The goal of a more recent study (Liu, 2022) was to determine whether or not mobile PA had any noticeable effect on students' vocal performances and learning attitudes, the ultimate aim being to guarantee that all students could engage emotionally, behaviorally and cognitively. This article focuses on the research methods used in that study while simultaneously using AI to design an effective and comprehensive vocal learning model.

Chinese ethnic vocal singing and teaching

A key singing concept for students majoring in ethnic vocal music is, "The articulation are driven by the cavity, enunciation can be driven by voice, then voice and prosody combination, and the voice is clear and strong." (Liu, 2021). Bel-canto perhaps comes closest in terms of style, but the singing effect is more natural than that of Chinese folk music, which favors "overtones" and whose timbre is more pronounced. Users' vocal cords are subjected to the collection of stroboscopic, perceptual, acoustic, and aerodynamic data in order to prevent vocal nodules (Fu et al., 2015).

Liu's "twenty-four character formula" stipulates that ethnic vocal music demands concurrent training in voice, singing exercises, pitch, and rhythm adjustments (Luo et al., 2020), underscoring the necessity for features such as automatic accompaniment. Moreover, it is crucial to extract noise based on the natural singing voice's vibration frequency and evaluate its potential impact on vocal cord health. The primary challenge in performing national vocal music lies in conveying "emotion". A variety of these, including "nationality, patriotism, love, labor, motivation, and positivity" (Liu, 2022), are integral to the genre, meaning that different song themes require specific, adept guidance in order to bring certain nuances to the fore.

Acoustic analysis and vocal cord protection

Breathing is very important for vocal cord protection (Valley, 2010), and is the prerequisite for vocal cord relaxation. In a study conducted by Hanyang University, Hwang et al. (2016) documented the effectiveness of basic breathing exercises in restoring laryngeal resonance, determining that during inhalation the strength of the breath and the correctness of breathing form the basis for determining the self-maintenance of the laryngeal vocal cords.

The auxiliary function of AI for ethnic vocal music learners is to serve as a voice pedagogue, defined as a teacher who has "scientific knowledge of the vocal mechanism as well as how to apply current research and valid techniques to maximize the efficiency of the voice" (Gutshall, 2006). The practical application of AI software in ethnic vocal music can effectively guide the user towards the correct vocalization method and singing rhythm. In a survey of the vocal cord structure of 376 college students, 20.75% of vocal music majors suffered laryngeal damage due to excessive voice use and poor voice practice methods, representing the largest source of laryngeal injury besides lesions. Among the symptoms of laryngeal cavity discomfort, hoarseness, throat discomfort, and air leakage during singing scored the highest proportions (Guo, 2006).

A mixed sample of 50 Chinese laryngeal cavities showed that through laryngoscope monitoring, the pronunciation frequency ranged from low to high. The dynamic laryngoscope automatic frequency measurement device showed that the frequency range was between 200 and 300HZ, and that low-frequency sounds increased the sliding degree of vocal cord mucosa (Lu, 1995). It was also determined that low-frequency sound exercises can play a role in vocal cord muscle massage.

"Glottal fry", also called "vocal fry" or "creaky voice" (Plexico, 2017), has been described as sounding similar to the "popping of corn" or a "motorboat engine" (Colton, 1996). Brought about by a low airflow rate related to the long-closed phase of the vocal fold vibratory cycle (Blomgren, 1998), and existing as a separate phonational register lying below the modal register on the frequency continuum (Hollien, 1968), it has a transparent feeling, giving the sensation of particles that are very dense but within which the gaps can clearly be distinguished. The texture is acceptable to the human ear and feels stimulating but not uncomfortable. Pulse register phonation or strohbass (Appleman, 2005) is a frequently-used exercise in ethnic vocal music that helps to rest the voice and restore singing function. It is characterized by low frequency and low decibel vibrations to make the sounds. "Opening the mouth wide" and "Grooving the tongue" simultaneously relax the entire facial musculature, while the weak airflow from the abdomen impacts the closed vocal cords, causing them to vibrate and produce a continuous bubbling sound (Guo, 2015).

The sound source tested the three processes of singing under fatigue, vocal fry recovery and high-quality singing. (See Figure 1.)

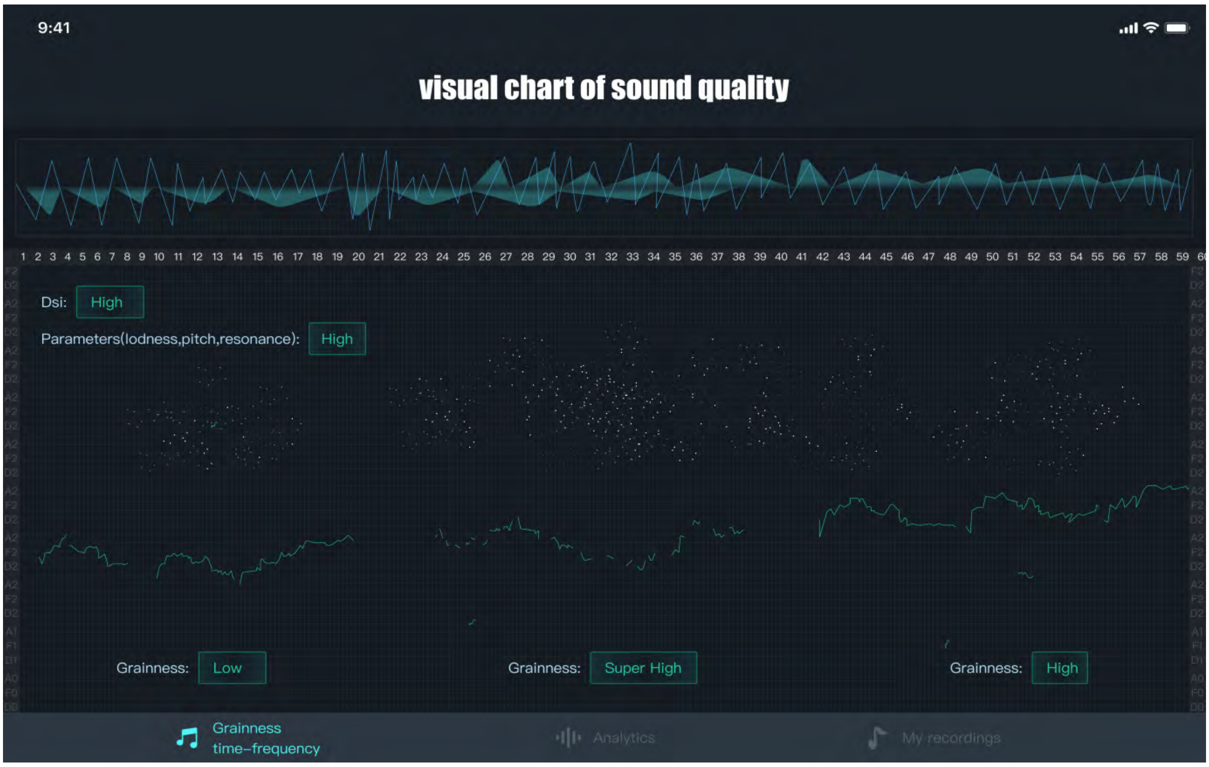


Figure 1 *Glottal fry data in AI application (Beta Version)*

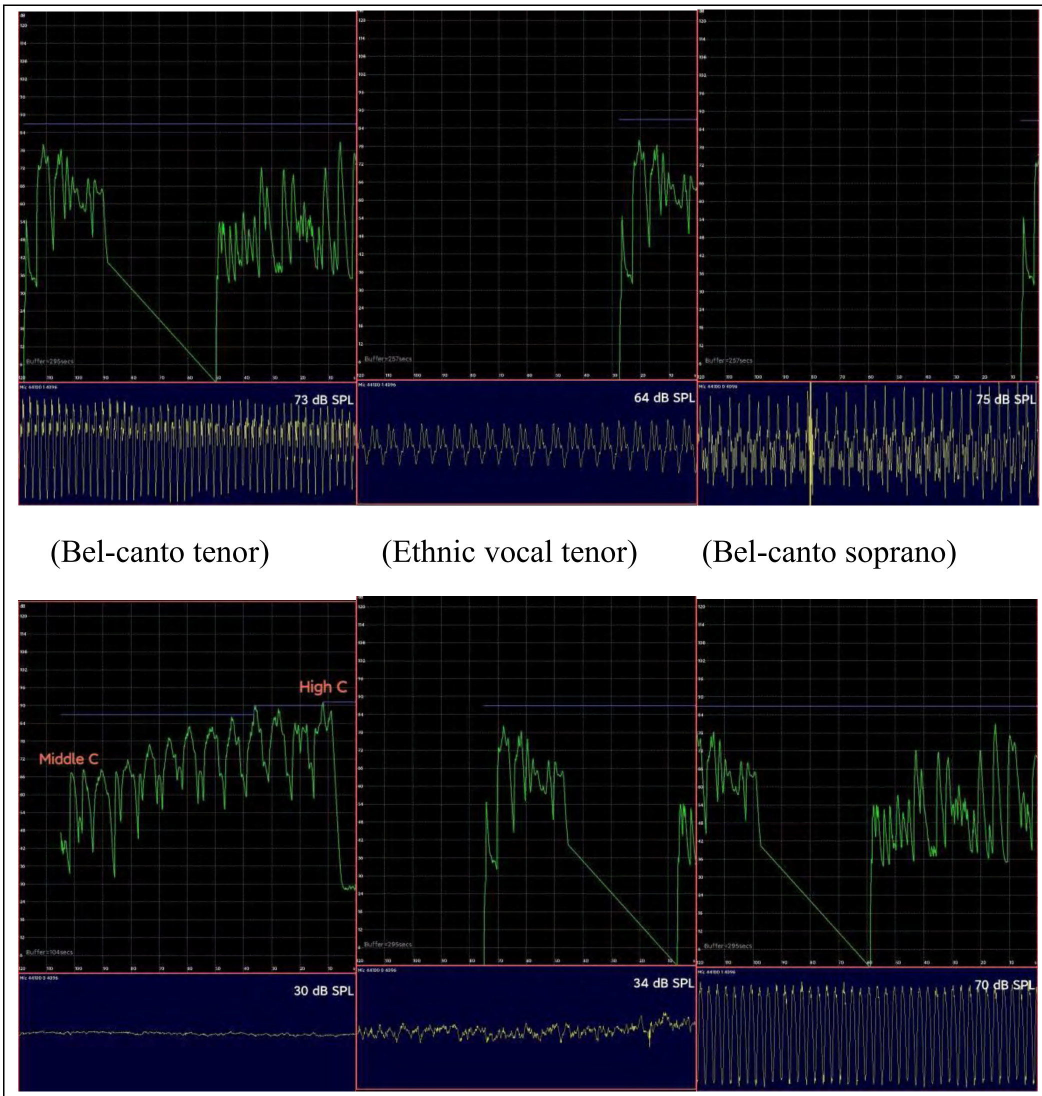
Vocal core and laryngeal cavity health monitoring are shown in Figure 2.

Figure 2 *Vocal cord and laryngeal cavity health monitoring (Beta Version)*

AI can analyze the hoarseness, roughness and breathiness in the voices of ethnic vocal music singers. In addition to the four standard acoustic parameters (F_0, Jitter, Shimmer, NNE), two additional acoustic parameters suitable for ethnic vocal music singing voices can be added: glottal spectrum frequency, and formant frequency perturbation (Huang, 2008). The glottal spectrum slope is a measure of the singer's voice signal in the high-frequency region and a quantity of intensity difference in the low-frequency region, thereby accurately monitoring the current voice state of the singer. AI also provides other exercises for vocal cord relaxation, such as producing the voice more naturally; warm-up and cool-down voice; and the relaxation of jaw and pharynx while producing nasal, vowel and humming sounds (Hazlett, 2011).

Audio processing software for online courses

As the output device typically employed for online teaching, it is a common experience for mobile phone users that the input voice is excessively loud. Liu et al. (2022) proposed a novel time-warping approach for pitch correction: shape-aware dynamic time warping (SADTW), which ameliorates the robustness of existing time-warping approaches, to synchronize the amateur recording with the template pitch curve. In order to improve the clarity of ethnic vocal singing on mobile devices, it is necessary to collect data on live vocals, musical instruments, and environmental sounds, and then make the necessary adjustments. Normally, when faced with mobile devices, the vocal effect will suffer a reduction in quality during the transmission process, or even be covered by current and ambient sound. A bel-canto tenor, a Chinese ethnic vocal tenor, a bel-canto soprano, and a Chinese ethnic vocal soprano were chosen as the four subjects to test their vocalization, examine their frequency of singing, and record and examine their singing data using "AudioTool" APP as the recording tool. (See Figure 3.)



(Ethnic vocal soprano) (Normal speech) (Ethnic vocal tenor scale practicing with piano)

Figure 3 *Vocal records on Hertz and dB SPL*

Distinct sound data is presented in Table 1 for the easy understanding of each singer's performance. This information aids the AI in tailoring the sound-carrying capacity of mobile devices to match the singer's current situation.

Table 1 *Sound data of individual singers*

Gender	Age	Vocal Part	Pitch Name	Herz	Decibel	Result
Male	23	Bel-canto tenor	G5	479HZ	75dB	High dB SPL,Low HZ,Voice Sound output is normal
Male	23	Ethnic Vocal Tenor	G5	511HZ	64dB	Low dB SPL,High HZ,Voice Sound output disappears
Female	23	Bel-canto soprano	A5	698HZ	73dB	High dB SPL,Low HZ,Voice Sound output is normal

Female	23	Ethnic vocal soprano	A5	783HZ	70dB	Low dB SPL,High HZ,Voice Sound output disappears
		Piano	C4	440HZ	54dB	When the chord above A5 is reached, the sound becomes weaker
		Normal talk		390HZ	31dB	The sound does not disappear

As can be seen, the online teaching software must modify voice input and output in keeping with the high decibel capacity of the ethnic vocal tenor and soprano, and the piano's sound amplitude above A5 (Chandna et al., 2020). By adjusting this fundamental setting, it is possible to guarantee that the folk song’s high voice will be adequately transmitted throughout the online lesson, while also depending on AI’s control and management to analyze the audio and make the necessary dynamic modifications. Such problems can be solved with a system capable of transmitting a low-bandwidth text / control stream over a network, or a similarly low-data-rate channel which facilitates the rendering of spoken or sung text at the receiving end of the transmission channel (Cook, 1993). Finally, the human voice, environmental sound and musical instrument sound data is transmitted to the device. After AI calculations, the device adjusts the input and the priority of the output sound data, so as to present the human voice completely and clearly.

The application of vocal psychology

Research has demonstrated that the importance value of vocal technology recognition increases with the level of improvement of psychological resources (Zhang, 2022). Some degree of harm can be done to the quality of a performance when pupils experience anxiety while singing, such as a rapid heartbeat, increased blood pressure, quick and shallow breathing, and dry mouth (Spahn, 2015), so teaching effectiveness needs to be optimized by ensuring that individuals are as emotionally “stable” as possible.

Information obtained via online instruction demonstrates the importance of motivation, encouragement, and the desire to study (Theiler & Lippman, 1995). Anxiety may be successfully managed within the context of the ethnic vocal performance process using slight, focused relaxation. During the 2022-2023 school year, the Shenyang Conservatory of Music's second-year graduate students used a combination of online and face-to-face engagement as part of their completion of a professional course in national voice music. The provide courses tuition in ethnic vocal music singing, including stage performance, opera, and national vocal music theory.

Under the author's guidance, 22 students convened to discuss how to address stage singing anxiety following the final national vocal performance examination. Of the participants, 17 expressed the view that, "Adaptability is crucial in mitigating anxiety related to stage performances." According to Zakaria (2013), the concept of "adaptability" not only bolsters self-confidence through an understanding of, and adjustment to, the song itself, but also extends to acclimating to the performance environment, keenly observing the audience's reactions and effectively appraising oneself psychologically, all of which ultimately contribute to the cultivation of a positive mental state.

Continued cross-modal research will undoubtedly provide further insights into the expressive aspects of vocal expression and music performance (Juslin, 2003). In the meantime, we may draw the conclusion that ethnic vocal music performance and training should be founded upon a psychological model that carefully takes into account various factors, as shown in Figure 5.

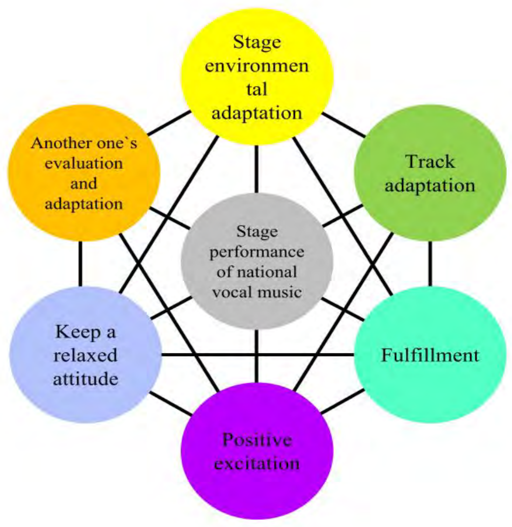


Figure 5 *Psychological model of national vocal music stage performance (Tang, XXX)*

Questionnaire investigation

Online surveys were distributed as the main data collection component of the present study, to be assessed based on predetermined criteria. While a total of 1,657 were completed, those with response times of less than 30 seconds were excluded. This led to 1,243 valid surveys being obtained, which translated to a successful recovery rate of 75%. According to the formula used to calculate individual user satisfaction, the outcome reflects a high level of satisfaction:

$$Sj = \frac{100}{1657}(\sum 10\ ) \quad (1)$$

The basic information regarding the 1,243 survey participants, as derived from the sample's primary data, is outlined in Table 2. The participants' ages range from 18 to 35, and predominantly comprise undergraduate and Master's degree holders (81.6% of the student body), a demographic characterized by a high level of knowledge and a strong openness to new concepts. Additionally, vocal music was the primary major for the majority of students.

Table 2 *Basic information of the sample*

User information	**The content**	**Quantity**	**Percentage (%)**
Age	Over 18	915	86.5
	Over 35	168	13.5
Identity	Student	950	76.4
	Teacher	293	23.6
Record of formal schooling	Junior college or below	125	10.1
	Undergraduate	673	54.1
	Master's degree	342	27.5
	Doctoral degree	103	8.3

Major	National vocal music	261	21
	Bel-canto vocal music	245	19.7
	Popular music	269	21.6
	Music education, Instrumental music performance, Musicology, Music therapy, etc.	250	20.1
	Other major	218	17.5

According to Table 3, a significant majority of individuals engage in daily vocal practice sessions lasting between one and three hours. This data reinforces the idea that most individuals find this duration comfortable, as well as aligning with the initial advice provided by the AI regarding recommended practice times. As a result, the author felt that recognizing the frequency of daily training was of paramount importance in shaping a model for the teaching and practice of ethnic vocal singing (LeBorgne, 2002).

Table 3 *Analysis of the current situation*

Sum of daily vocal training hours	Frequency	Percentage (%)
1 to 3 hours	875	70.4

Over 3 hours	368	29.6

In order to comprehend user intentions in terms of the employment of supplementary software functionalities, and with the aim of establishing a blueprint for the development of AI applications, the authors centered their attention on the “open throat vocalization” app developed by Beijing Xihang Technology Co., Ltd. This decision garnered strong endorsement from the software users, as the results of the survey results shown in Table 4 indicate.

Table 4 *Situational analysis on the use of the “open throat vocalization” app*

		Quantity	**Percentage (%)**
User satisfaction	Extremely satisfied	843	67.8
	Very satisfied	162	13
	Satisfied	148	11.9
	So-so	28	2.3
	Dissatisfied	62	5
Favorite function	Music scale training	260	20.9
	Intonation device	147	11.8

	Following songs	196	15.8
	Solfeggio & ear training	198	15.9
	Breath practice	165	13.3
	Virtual piano	72	5.8
	Metronome	115	9.3
	Another type	90	7.2
Would you like to use the app regularly?	Yes	1067	85.8
	No	176	14.2

The statistics reveal that most users are content with the software's user-friendliness, particularly favoring the "music scale training" feature. They also indicated a strong desire to use it in the future, in so doing underscoring the potential for the refinement of vocal practice software. The valuing of the incorporation of interactive elements and an expression of willingness to integrate tools that can enhance voice training in the future are both considered positive signs (Ng et al., 2013).

From an analysis of the data presented in Table 5, it is evident that users' expectations for each AI function are relatively consistent. Nevertheless, when viewed from a broader perspective, there is a significant surge in expectations for AI development, especially with regard to a demonstration of singing. As the most eagerly anticipated feature, this concept aligns with the ongoing endeavor to enhance human-AI interaction, particularly through advancements in human speech emotion recognition (Lawson, 2023).

Table 5 *Expected value of AI*

What functions do you expect from AI?		Response Number of cases	percentage (%)
1.	Auto accompaniment	992	18.7
2.	Course record	1095	20.6
3.	Singing evaluation	1055	19.9
4.	Suggestions for improvement	1062	20.0
5.	AI singing demonstration	1104	20.8

Teaching modes based on AI

Teaching in the fourth Industrial Revolution will aim at different ways of integrating advanced digital technologies into learning environments (Arel, 2010). Previous teaching models have been limited by time and space, making it difficult to meet the needs of different personalities. Using modern online educational technology to establish an efficient, fast and user-friendly application system has become an essential trend (Ma, 2021) as the technical basis of the “digital campus”. This study adopts the perspective of AI and information technology within the online teaching environment to subdivide and expand the tasks of the application scenarios in the teaching and personal learning of ethnic vocal music within universities. When intelligent technology is applied to various tasks, data consumers, in the form of students, teachers, and network platforms, can make more comprehensive and efficient use of data information, improve the quality of classroom teaching, and cultivate students’ music literacy.

In the field of music, the AI processing tasks are used mainly to identify and generate the logic and theory of music creation through the objective modeling of data and algorithm in the three main parts of music (Liang, 2022), performance (Li, 2021) and sound (Chen et al., 2019). In the acoustic model, the modeling needs to obtain fundamental frequency information, such as timbre (Pons et al., 2017), speaker information (Xue et al., 2021), pause information (Martin et al., 2022), language information (Schulze-Forster et al., 2021), facial features (Konstantinidis et al., 2018), breath control (Fukuda et al., 2018), sentiment analysis (Abburi et al., 2016), and other features. For different styles of music, the algorithm needs to analyze the composition and variation of the chords, rhythms, and instrumentation. Different from ordinary speech processing, deep learning not only needs to ensure sound quality

(Bochner et al., 2022), but also needs to pay attention to sampling rate (Gao et al., 2022), as well as continuity and fluency (Huang et al., 2022).

In the imitation and creation of songs, the characteristics of lyrics and melody can be simulated by a large number of data and models, but the weak relationship between them requires the strict monotony of tone and word pairing and connection (one-to-one to many-to-many), which in turn requires the quality assurance of data and the strengthening of prior knowledge. For example, if the tone of the lyrics is not consistent with the melody, the problem of homophone confusion will occur. The ability of models to align too closely does not create more diverse artistic music, especially in folk music (Savage et al., 2022). In fact, when it comes to expressing the emotion of the music more vividly and strongly, the rhythm of the lyrics and the melody are often exaggerated and over-dramatized. The melody corresponding to important lyrics will be enhanced by emotion (BAO, 2021), so emotion analysis needs to go beyond lyrics to incorporate melody and other factors. In the future, AI will cooperate with human beings to complete music in the ways outlined above, rather than replace their roles as artists (Feldman, 2017). As things currently stand, AI is widely used in aesthetic music applications. If AI music creation is also to be regarded as an information tool for producers, there are large-scale AI applications that can serve to promote the dissemination of musical works and enhance audience experiences. For example, the music recommendation system, as the mainstream personalized technology for audiences, can expand the number of music platforms and enhance the experiences of returning users. The combination of high quality and popular music platforms, in conjunction with the music recommendation system technology, requires an effective combination of audio characteristics, user interaction, user differences and commonalities, and user and platform annotation under music works (Liu & Liu, 2020). As a media for digital music communication, music platforms are more inclined to expand users' listening recommendations, thereby avoiding the "echo chamber" problem. At the same time, they also need to pay attention to copyright issues in the process of music transmission. In personal works, such as NFT digital products, privacy, security and other features need to be guaranteed by blockchain technology (Zhao & O'Mahony, 2018). For the audience, AI also requires research approaches that are multi-dimensional and multi-field.

Online education, as a teaching platform for universities in the information age, is typically composed of cloud classrooms and electronic resources; as such, it has become something of a mainstream teaching methodology, especially as a result of the pandemic. As the basis of online education, the network communication environment, sited between and where the devices of teachers and students are located, is crucial to ensuring the quality of cloud-based classroom teaching and the usage experiences of teachers and students. The cloud-based classroom should ensure a low-delay, highly reliable network communication environment (Bojović et al., 2020), as well as clear, three-dimensional conditions for high income recording with the removal of environmental noise. Even allowing for the establishment of a stable, cloud-based classroom, the cost and time limitations of university equipment, as well as the variations among students' equipment, should be taken into account in order to enhance specific user experiences (McRoy., 2020). Consideration also needs to be given to support students' ability to listen to lectures online in real time, to independently and flexibly select past courses, exchange discussions, and take online tests. At the same time, when applied to vocal music teaching, the cloud-based classroom needs to consider more diversified and three-dimensional forms of screen courseware that can effectively present the teaching content, as well as students' compositions, and the playing of musical instruments and singing (Fang, 2021).

A combination of AI and the digital curriculum system within the training cycle, with a musical instrument in the form of a harp from a specific teaching institution (http://s.immusician.com/static/web/index.html), is given here as an example. Its AI technology is mainly embodied in interactive video teaching based on voice and video recognition technology and customized personalized courses based on big data analysis, and the intelligent adjustment of learning paths based on data collected from users' actual practice. Students master the basic techniques of playing and singing in the early course; from the intermediate to the senior course, students can play and sing 80% of popular songs on the market without any difficulties. In the advanced course, students can master the principles of songwriting and try to create songs with the aid of voice and video recognition technology (called “audio-visual speech recognition”). In this form of deep learning, the recorded video and voice are used as the input data of the model for multi-mode feature fusion, extraction and coding. The construction of the algorithm model to achieve various specific tasks is shown in Figure 6 below.

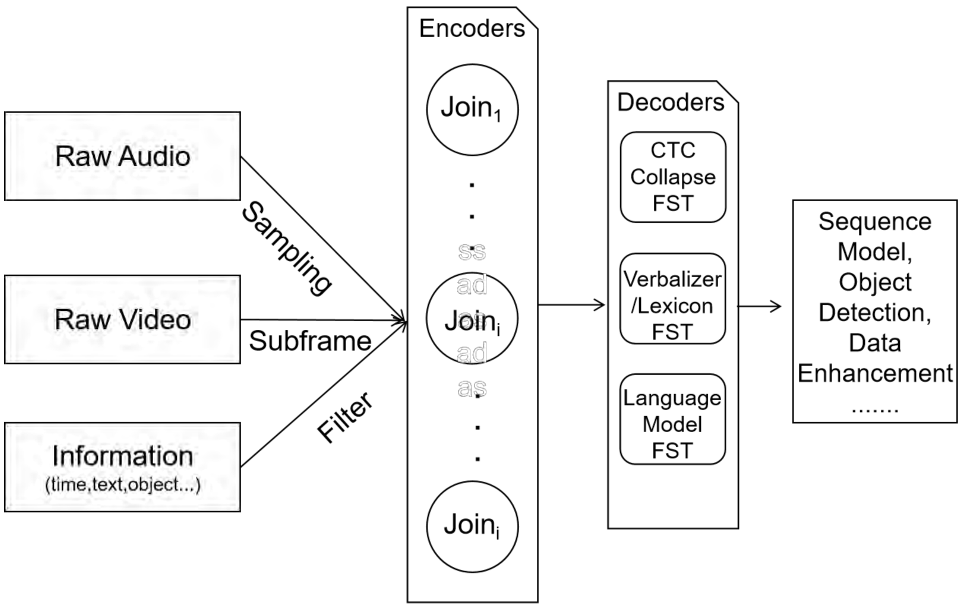


Figure 6 *Algorithm model*

In the university online ethnic vocal singing teaching model, AI can be subdivided into the following tasks, all of which complement each other:

1. According to the acquisition of students’ physiological signals (voice, ECG, eye movement, facial expressions, and body changes) in the process of learning and practicing, information feedback, evaluation, advice and emotional analysis can be carried out. In

addition, the relationship between a student's personality and music can also be explored. Combined with the capture and analysis of physiological signals, multimodal fusion technology is used to perform representation, fusion, conversion, and alignment of cross-modal data, the similarity, dissimilarity, freshness and other indicators being used to score, provide feedback, and evaluate students' singing and performance through combining the existing scores with the evaluation data. To further stimulate students' imagination, the sound of musical instruments and human voices are mapped into images in a visual way, i.e. the audio space is transformed into a visual space, thereby assisting the teaching process (Khulusi et al., 2020), as shown in Figure 7 below. Optical music recognition (OMR), which enables the detection of staff notation and the visualization of various musical elements (Calvo-Zaragoza, & Rizo, 2018), is one example. Audio-visual separation is used to distinguish the targets of different sound sources, while the detection and classification of acoustic scenes are closely related to multi-channel signals, e.g., microphone array multi-channel signals, conversion methods, e.g., Fourier transform and wavelet transform, and audio scaling. These processing elements correspond to the heterogeneous distribution and discretization of data features; group voice processing between teacher and student, or in a chorus, corresponds to speaker recognition (SR) in the speech domain (Li et al., 2022).

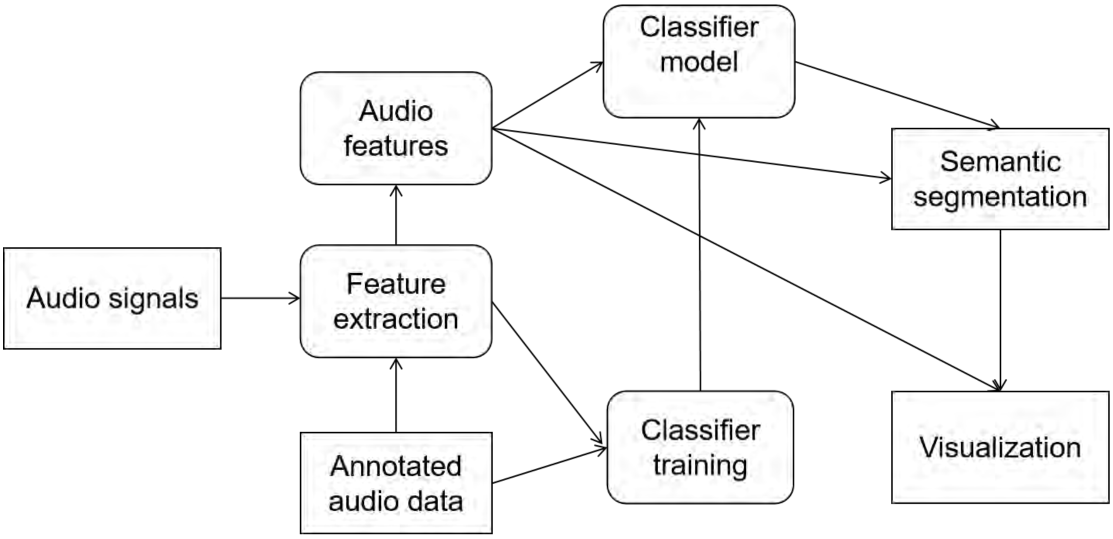


Figure 7 *Operation diagram*

2. Enhancing the experience of using multimedia intelligence tools, such as intelligent musical instruments and virtual musical teaching aids, the sound quality of online classes, and the immersion experience. The targeted, combined optimization of platform tools, such as audio and video recording, privacy processing, blackboard writing, and teaching content thought maps, are divided amongst teachers and students in line with the recommendation system. For the acquisition and overview of mainstream aesthetic works and information, “web crawler” and dynamic theme models can be used. Real-time and recorded data of teachers and students, along with distributed storage and privacy processing of computing using deep learning, together correspond to dynamic distributed messaging systems, blockchain and federated learning.

3. Exploring effective methods and strategies to improve students’ comprehension

quality and cultivate innovative consciousness in the development of electronic music creation and performance training and promoting the reform of electronic music online education. Teacher and student data can be used to judge learning models, and match professional knowledge, teaching content, progress and rhythm, and course and students' learning effect evaluations. For example, mapping the radiating function of wireless networks into online education and establishing the intimacy and immediacy of social learning systems can improve the effectiveness of the AI electronic music creation learning process. Scholars have analyzed representative online learning platforms, highlighting the problems and trying to use the structure of personalized learning systems and online learning content – constructed by using the improved optimization strategy, and automatically arranging sequences through a data-driven recommendation system, respectively – to build models.

4. The imitation and generation of musical scores and playing skills corresponds to the tasks in natural language processing (NLP), and assists students with opening their voices (performance preparation) and singing (performance). The generation of sounds and instrumental tones corresponds to speech understanding and speech synthesis in the field of speech, while generating a song according to one's own lyrics or melody can correspond to transfer learning and continual learning in the field of deep learning. It is also possible to generate models from text by giving arbitrary formats and templates in order to generate content that fits.

When a deep learning algorithm is applied to a specific scene, one of the necessary and particularly important problems is how to evaluate the effect of the model results. Subjective and objective perspectives should be combined to conduct a comprehensive evaluation, the aim being to make the effect of the algorithm more sustainable and practical. The judgement of musical works may be highly subjective, so the quality assessment mechanism should always guide the music processing direction and ability of the algorithm.

Teaching model

Among the methods of teaching vocal music, that of ethnic vocal music is characterized by being effective, intuitive, operational, open, and targeted. Focusing on the emotional, behavioral, and cognitive modes helps pupils learn based on their unique abilities (Kashina et al., 2020). The teaching model shown in Figure 8 may be constructed with the intention of assisting with the performance and teaching of ethnic vocal music (Wang, 2022); as such, it uses the performance techniques, instructional strategies, and artificial intelligence algorithm to collectively support this study's vocal teaching model.

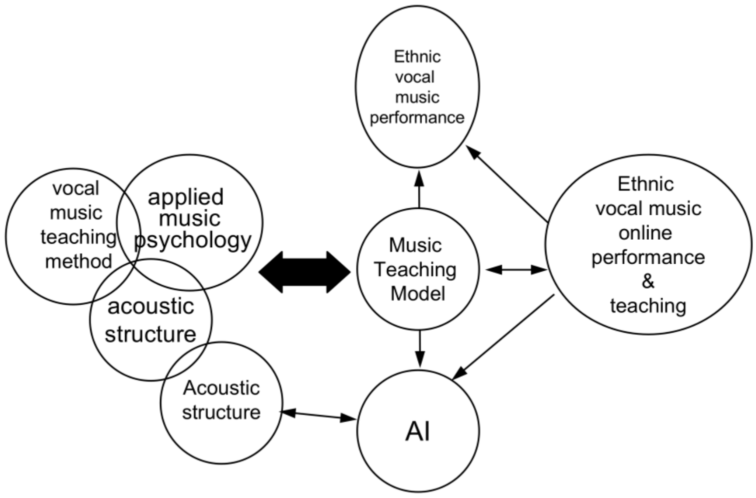


Figure 8 *Teaching model*

Table 6 summarizes the teaching mode of this article, As shown in the figure, the online performance and education of ethnic vocal music is based on other disciplines. At the same time, AI analyzes learning and teaching methods suitable for distance teaching based on data, in the process forming a complete teaching model (Su, 2022). Combined with the previous analysis, the WSQ model employs the observation method as a means of AI's fundamental assessment of vocal learners. Additionally, it utilizes the PA approach to facilitate a comparison between two software objects that do not occupy the same space (Body, 2021). This comparative evaluation yields valuable data, which in turn aids with individual assessment.

Table 6 *Model information summary*

The content of this teaching model	**Approach**	**Significance**
National vocal performance	Liu Hui "twenty-four character formula"	Plays a normative role in singing
Teaching methods of national vocal music	Emotional, behavioral and cognitive modes	Plays a normative role in teaching
Artificial Intelligence	User feature capture, course record, vocal music demonstration, scene	Assists the singing and teaching of ethnic vocal music in colleges and universities through digital algorithms

	analysis, etc	
Vocal cord maintenance	Tips Correct breathing	Protects students' voices
Applied music psychology	Psychological self adjustment	Spiritually guarantees performance and teaching
Acoustic structure	Adjust audio decibel and frequency	Guarantees normal online teaching

It can be inferred from the study presented in this article that ethnic vocal music and teaching are interconnected and that AI algorithms may be a useful tool for classroom instruction. The inclusion of different viewpoints and the feedback received from the study as it relates to its practice suggests that the teaching model is suitable for folk vocal singing and instruction.

References

Abburi, H., Akkireddy, E. S. A., Gangashetti, S., & Mamidi, R. (2016, January). Multimodal sentiment analysis of telugu songs. *In SAAIP@ IJCAI.*

Appleman, R., & Bunch, M. (2005). Application of vocal fry to the training of singers. *J Sing*, *62*(1), 53-9.

Arel, I., Rose, D. C., & Karnowski, T. P. (2010). Deep machine learning - a new frontier in artificial intelligence research [research frontier]. *IEEE Computational Intelligence Magazine, 5*(4), 13-18.

Bao, C. *Generating music with sentiments[J]*. 2021.

Blomgren, M., Chen, Y., Ng, M. L., & Gilbert, H. R. (1998). Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers. *The Journal of the Acoustical Society of America*, *103*(5), 2649-2658.

Bochner, J., Indelicato, M., & Konnur, P. (2022). Effects of Sound Quality on the Accuracy of Telephone Captions Produced by Automatic Speech Recognition: A Preliminary Investigation. *American Journal of Audiology*, 1-8.

Body, M. (2021). Flipped Classroom-Perfect Mix of Face-to-face and Online Learning. *Scripta*, 15.

Bojović, Ž., Bojović, P. D., Vujošević, D., & Šuh, J. (2020). Education in times of crisis:

Rapid transition to distance learning. *Computer Applications in Engineering Education*, *28*(6), 1467-1489.

Calvo-Zaragoza, J., & Rizo, D. (2018). Camera-PrIMuS: Neural End-to-End Optical Music Recognition on Realistic Monophonic Scores. *International Symposium/Conference on Music Information Retrieval.*

Chandna, P., Blaauw, M., Bonada, J., & Gómez, E. (2020, May). Content based singing voice extraction from a musical mixture. *In ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP) (pp. 781-785). IEEE.

Chartrand, R. (2016). Advantages and disadvantages of using mobile devices in a university language classroom. *Bulletin of the Institute of Foreign Language Education Kurume University*, *23*(1), 1-13.

Chen, X., Chu, W., Guo, J., & Xu, N. (2019, March). Singing voice conversion with non-parallel data. *In 2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)* (pp. 292-296). IEEE.

Civit, M., Civit-Masot, J., Cuadrado, F., & Escalona, M. J. (2022). A systematic review of artificial intelligence-based music generation: Scope, applications, and future trends. *Expert Systems with Applications*, 118190.

Colton, R. H., & Casper, J. K. (2001). Understanding Voice Problems: Voice Problems

Associated with the Pediatric and the Geriatric Voice.

Cook, P. R. (1993). SPASM, a real-time vocal tract physical model controller; and singer, the companion software synthesis system. *Computer Music Journal*, *17*(1), 30-44.

Elghaish, F., Chauhan, J. K., Matarneh, S., Rahimian, F. P., & Hosseini, M. R. (2022). Artificial intelligence-based voice assistant for BIM data management. *Automation in Construction*, *140*, 104320.

Fang, P. E. N. G. (2021). Optimization of music teaching in colleges and universities based on multimedia technology. *Advances in Educational Technology and Psychology*, *5*(5), 47-57.

Feldman, S. S. (2017). Co-creation: human and AI collaboration in creative expression. *Electronic Visualisation and the Arts* (EVA 2017).

Fu, S., Theodoros, D. G., & Ward, E. C. (2015). Delivery of intensive voice therapy for vocal fold nodules via telepractice: a pilot feasibility and efficacy study. *Journal of Voice*, *29*(6), 696-706.

Fukuda, T., Ichikawa, O., & Nishimura, M. (2018). Detecting breathing sounds in realistic Japanese telephone conversations and its application to automatic speech recognition. *Speech Communication*, *98*, 95-103.

Gao, Z., Tang, M., Li, A., & Chen, Y. (2022, April). An Audio Frequency Unfolding Framework for Ultra-Low Sampling Rate Sensors. *In 2022 23rd International*

Symposium on Quality Electronic Design (ISQED) (pp. 1-6). IEEE.

Guo, L., Zhang, L., & Li, Y. (2006). Analysis of 376 cases of voice diseases among vocal music students. *Music Exploration (Journal of Sichuan Conservatory of Music)*, *01*, 86-88.

Guo, X. (2015). The important role of bubble sounds in vocal singing. *Northern Music*, *35*(17), 64.

Gutshall, E. M. (2006). The twenty-first century vocal pedagogue: A mediator and translator for the vocal studio and its interdisciplinary colleagues. *ProQuest*.

Hazlett, D. E., Duffy, O. M., & Moorhead, S. A. (2011). Review of the impact of voice training on the vocal quality of professional voice users: implications for vocal health and recommendations for further research. *Journal of Voice*, *25*(2), 181-191.

Hebert, D. G. (2007). Five challenges and solutions in online music teacher education. *Research and Issues in Music Education*, 5(1), 1-10.

Henrich, N., Smith, J., & Wolfe, J. (2011). Vocal tract resonances in singing: Strategies used by sopranos, altos, tenors, and baritones. *The Journal of the Acoustical Society of America*, *129*(2), 1024-1035.

Hollien, H., & Michel, J. F. (1968). Vocal fry as a phonational register. *Journal of Speech and Hearing Research*, *11*(3), 600-604.

Hsia, L. H., Hwang, G. J., & Lin, C. J. (2022). A WSQ-based flipped learning approach to improving students' dance performance through reflection and effort promotion. *Interactive Learning Environments*, *30*(2), 229-244.

Huang, R., Cui, C., Chen, F., Ren, Y., Liu, J., Zhao, Z., & Wang, Z. (2022, October). Singgan: Generative adversarial network for high-fidelity singing voice generation. *In Proceedings of the 30th ACM International Conference on Multimedia* (pp. 2525-2535).

Huang, Z. M., & Wan, P. (2008). Research on the correlation between voice acoustic parameters and voice quality. *Journal of Clinical Otolaryngology-Head and Neck Surgery*, *2008*(06), 251-255.

Hwang, Y. S., Shim, M. R., Kim, C. J., Choi, Y. S., Kim, S. Y., Choi, C. H., & Sun, D. I. (2016). Comparison of Vocal Cord Motion and Voice Characteristics of Applied Music Singing Students before and after Singing Voice Therapy. *Journal of The Korean Society of Laryngology, Phoniatrics and Logopedics, 27*(2)*,* 114-121.

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological bulletin*, *129*(5), 770.

Kashina, N. I., Pavlov, D. H., & Li, H. (2020, May). Music-digital technologies as a means of developing creative independence of students of colleges and higher schools in China and Russia. *In International Scientific conference "digitalization of education:*

History, trends and prospects" (DETP 2020) (pp. 27-31). Atlantis Press.

Khulusi, R., Kusnick, J., Meinecke, C., Gillmann, C., Focht, J., & Jänicke, S. (2020, September). *A survey on visualizations for musical data. In Computer Graphics Forum* (Vol. 39, No. 6, pp. 82-110)

Konstantinidis, D., Dimitropoulos, K., & Daras, P. (2018, October). A deep learning approach for analyzing video and skeletal features in sign language recognition. *In 2018 IEEE international conference on imaging systems and techniques (IST)* (pp. 1-6). IEEE.

Lawson, F. R. S. (2023). Why can't Siri sing? Cultural narratives that constrain female singing voices in AI. *Humanities and Social Sciences Communications*, *10*(1), 1-11.

LeBorgne, W. D., & Weinrich, B. D. (2002). Phonetogram changes for trained singers over a nine-month period of vocal training. *Journal of Voice*, *16*(1), 37-43.

Lei, D. (2020, October). RETRACTED: Research on Network Teaching of Music Major Based on Artificial Intelligence Technology. *In Journal of Physics: Conference Series* (Vol. 1648, No. 4, p. 042094). IOP Publishing.

Li, H. (2021). Piano education of children using musical instrument recognition and deep learning technologies under the educational psychology. *Frontiers in Psychology*, 12.

Li, X., Liu, S., & Shan, Y. (2022). A Hierarchical Speaker Representation Framework for One-shot Singing Voice Conversion. *arXiv preprint arXiv*:2206.13762.

Liang, M. (2022). Music Score Recognition and Composition Application Based on Deep

Learning. *Mathematical Problems in Engineering*, 2022.

Liu, C., Wan, P., Tu, Y. F., Chen, K., & Wang, Y. (2021). A WSQ-based mobile peer assessment approach to enhancing university students' vocal music skills and learning perceptions. *Australasian Journal of Educational Technology*, *37*(6), 1-17.

Liu, H. (2011). Thoughts on the relationship between sounds and words in singing -24-character formula for the relationship between sounds and words in Chinese national vocal music singing teaching. *Songs of the Music Bureau (Journal of Shenyang Conservatory of Music), 04*, 9-10.

Liu, H. (2018). Inheriting beauty, expressing beauty, creating beauty, spreading beauty, and leading beauty: Thoughts on the development and construction of Chinese vocal music. *Chinese Music*, (01),7-14+193. doi:10.13812/j.cnki.cn11-1379/j.2018.01.002.

Liu, J., Li, C., Ren, Y., Zhu, Z., & Zhao, Z. (2022). Learning the beauty in songs: Neural singing voice beautifier. *arXiv preprint arXiv:*2202.13277.

Liu, N. (2022). Study on the Application of Improved Audio Recognition Technology Based on Deep Learning in Vocal Music Teaching. *Mathematical Problems in Engineering*, 2022.

Liu, S., & Liu, C. (2020). A review of music recommendation systems. *Journal of Guangzhou University: Natural Science Edition*.19(05):36-46+77.

Liu, Z. Y. (2022). My opinion on the improvement of national vocal music singing through musical aesthetics. *Cultural Monthly*, *09*, 163-165.

Lu, Y. Y., Wang, S. Q., & Ling, L. (1995). Study on the relationship between vocalization at different frequencies and vocal cord mucosal waves under dynamic laryngoscopy. *Zhejiang Journal of Integrated Traditional Chinese and Western Medicine*, *1995*(02), 34-35.

Luo, Y. J., Lin, Y. J., & Su, L. (2020). Toward expressive singing voice correction: On perceptual validity of evaluation metrics for vocal melody extraction. *arXiv preprint arXiv*:2010.12196.

Ma, J., & Feng, B. (2021). Integrated design of graduate education information system of universities in digital campus environment. *Hindawi Limited*.

Martin, V. P., Arnaud, B., Rouas, J. L., & Philip, P. (2022). Automatic Estimation of Pauses Duration and Location in Read Speech as Biomarkers of Sleepiness. *Submitted to ICASSP*, 2021.

McRoy, C., Patel, L., Gaddam, D. S., Rothenberg, S., Herring, A., Hamm, J., & Awan, O. (2020). Radiology education in the time of COVID-19: a novel distance learning workstation experience for residents. *Academic radiology*, *27*(10), 1467-1474.

Ng, S. C., Lui, A. K., & Lo, W. S. (2013). An interactive mobile application for learning music effectively. In Knowledge Sharing through Technology: 8th International Conference on Information and Communication Technology in Teaching and

Learning, ICT 2013, Hong Kong, China, July 10-11, 2013, Revised Selected Papers 8 (pp. 148-157). *Springer Berlin Heidelberg*.

Plexico, L. W., & Sandage, M. J. (2017). Influence of glottal fry on acoustic voice assessment: A preliminary study. *Journal of Voice*, *31*(3), 378-413.

Pons, J., Slizovskaia, O., Gong, R., Gómez, E., & Serra, X. (2017, August). Timbre analysis of music audio signals with convolutional neural networks. *In 2017 25th European Signal Processing Conference (EUSIPCO)* (pp. 2744-2748). IEEE.

Roads, C. (1985). Research in music and artificial intelligence. *ACM Computing Surveys (CSUR)*, *17*(2), 163-190.

Savage, P. E., Passmore, S., Chiba, G., Currie, T. E., Suzuki, H., & Atkinson, Q. D. (2022). Sequence alignment of folk song melodies reveals cross-cultural regularities of musical evolution. *Current Biology*, *32*(6), 1395-1402.

Schulze-Forster, K., Doire, C. S., Richard, G., & Badeau, R. (2021). Phoneme level lyrics alignment and text-informed singing voice separation. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *29*, 2382-2395.

Spahn, C. (2015). Treatment and prevention of music performance anxiety. *Progress in brain research*, *217*, 129-140.

Su, L. (2022, April). Analysis on the Application of Modern Multimedia Devices in the Current College Vocal Music Guiding from the Long Distance Signal Communication

Perspective. *In 2022 International Conference on Sustainable Computing and Data Communication Systems* (ICSCDS) (pp. 792-796). IEEE.

Theiler, A. M., & Lippman, L. G. (1995). Effects of mental practice and modeling on guitar and vocal performance. *The Journal of General Psychology*, *122*(4), 329-343.

Valle, C., Andrade, H., Palma, M., & Hefferen, J. (2016). Applications of peer assessment and self-assessment in music. *Music Educators Journal*, *102*(4), 41-49.

Valley, P. M. (2010). Generation Y and Vocal Fry.

Wang, X. (2022). Design of Vocal Music Teaching System Platform for Music Majors Based on Artificial Intelligence[J]. *Wireless Communications and Mobile Computing*, 2022.doi:https://doi.org/10.1155/2022/5503834

Webster, P. R. (2017). Computer-based technology and music teaching and learning. *In Critical essays in music education* (pp. 321-344). Routledge.

Wei, J., Karuppiah, M., & Prathik, A. (2022). College music education and teaching based on AI techniques. *Computers and Electrical Engineering*, *100*, 107851.

Xue, H., Yang, S., Lei, Y., Xie, L., & Li, X. (2021, January). Learn2sing: Target speaker singing voice synthesis by learning from a singing teacher. *In 2021 IEEE Spoken Language Technology Workshop (SLT)* (pp. 522-529). IEEE.

Yan, H. (2022). Design of online music education system based on artificial intelligence and multiuser detection algorithm. *Computational Intelligence and Neuroscience*, 2022.

Yu, X., Ma, N., Zheng L., et al. (2023). Developments and applications of artificial intelligence in music education[J]. *Technologies*, *11*(2), 42.

Zakaria, J. B., Musib, H. B., & Shariff, S. M. (2013). Overcoming performance anxiety among music undergraduates. *Procedia-social and behavioral sciences*, *90*, 226-234.

Zhang, Y. (2022). Cultivation and interpretation of students' psychological quality: Vocal psychological model. *Frontiers in public health*, *10*, 966628.

Zhang, Y. (2023). Optimizing Ethnic Vocal Performance Using Artificial Intelligence Technology.

Zhao, S., & O'Mahony, D. (2018, December). Bmcprotector: A blockchain and smart contract based application for music copyright protection. In *Proceedings of the 2018 International Conference on Blockchain Technology and Application* (pp. 1-5).

Zhe, T. (2021). Research on the model of music sight-singing guidance system based on artificial intelligence. *Complexity*, 2021, 1-11.

Asia-Pacific Journal for Arts Education

Special Issue

Unfold the Future of Music Education through Technology

Co-editors:
Prof. Bo Wah LEUNG
Dr. Cheung On TAM
Dr. Chi Hin LEUNG
Dr. Richard Guy WHITBREAD
The Education University of Hong Kong

http://www.ied.edu.hk/ccaproject/apjae/apjae.htm ISSN: 1683-6995

Volume 22 Number 5 December 2023

Integrating ubiquitous music ecologies into STEAM scenarios in music teaching-learning processes

Yannis Mygdanis
Ph.D. Candidate, Music Education (EUC)
yannis.mygdanis@icloud.com

Elissavet Perakaki
Specialized Educational Staff, Department of Music Studies (NKUA)
eperak@music.uoa.gr

Abstract

Technological progress over the course of the past few decades has transformed how children interact with sound and music, offering new and extended ways of expression, creation, and learning. This digital environment can provide opportunities for constructing a framework of sound perception, musical praxis, and creativity enhancement, the emergence of cross-platforms and a growing variety of hardware and software serving to support these developments through an emerging context of ubiquitous acoustic ecologies. The aim of this research was to involve educational scenarios in music lessons following such ecological perspectives through a pilot study for children aged 7 to 9 in a conservatoire setting in Greece. Actions for the current practical intervention have been designed following a STEAM project-based learning approach, which offers students cooperative activities, transdisciplinarity, game-based, augmented reality, playful learning, and authentic problem-solving experiences. Analysis revealed four distinct, emerging thematic categories that drew on the development of auditory perceptual ability, creativity development, computational thinking cultivation, and the shaping of digital and physical musical worlds. The results of the educational intervention underlined the fundamental role of ubiquitous music ecologies in planned actions, which served to widen students' musical horizons.

Key words

ubiquitous music ecologies, STEAM education, creative and computational thinking development, conservatoire setting, music teaching and learning

Introduction

Technological progress over the course of the past few decades has offered new and extended ways of expression, creation, and learning, transforming the way children interact and communicate with sound and music (Mygdanis & Kokkidou, 2021). Based on a ubiquitous computing perspective, this digital environment can provide opportunities for constructing a framework of sound perception, musical praxis, and creativity enhancement (Etmektsoglou, 2019). The development of digital cross-platforms and the growing variety of hardware and software have arisen in an emerging context of ubiquitous acoustic ecologies (Keller, 2020). Ubiquitous computing technologies cover a wide range of Do-It-Yourself (DIY) practices using browser-based platforms, mobile computing, interconnected distributed resources, the Internet of Things (IoT), low-cost hardware interfaces, microcontrollers, open-source software platforms, and programming languages (Lazzarini et al., 2020). These tools and procedures emphasize the social-cultural environment for collaborative creative processes through everyday life and objects (Dionysiou, 2019).

This research aimed to incorporate ubiquitous music ecology perspectives into music lessons through a pilot study for children aged 7 to 9 in a conservatoire setting in Greece. The focus was on examining students' learning processes and exploring their perceptions of musical acquisition. Actions for the current practical intervention were designed following a STEAM project-based learning approach, which offered students cooperative activities, transdisciplinarity, game-based, augmented reality, playful learning, and authentic problem-solving experiences. The methodology was based on qualitative educational research (content analysis). For data collection, multiple methods were utilized, including observations and field notes during the lessons, semi-structured interviews with the children, informal discussions inside and outside the classroom, and ubiquitous musical artifacts.

Acoustic music ecologies and sound-based pedagogy

Sound is everywhere and defines our relationship with the environment. Educational actions focusing on sound can provide additional value in music education by forming multiple meanings for students (Paynter & Aston, 1970; Schafer, 1977) and strengthening their environmental and cultural awareness (Etmektsoglou, 2019; Truax, 1996). In recent decades, there has been a strong interest in the shift from organized sound and music aesthetics to the soundscape in the musical educational process (Dionysiou, 2019).

Active listening is fundamental in sound-oriented learning environments (Westerkamp, 2011), connecting the listener with the environment and his musical-sound world (Dionysiou, 2019). Integrating active listening into activities with sounds enables children to improve their auditory perception, understanding, reflection, and creative and critical participation in the soundscape in which they live (Dionysiou, 2019; Truax, 1996; Schafer, 1977). In these activities, sound is perceived as movement, a means of creation, and an interface with the environment and culture (Etmektsoglou, 2014; Truax, 1999).

Sound pedagogy involves all students in teaching-learning processes, regardless of their musical background (see Etmektsoglou, 2019; Dionysiou, 2019), including various practices – improvisation, composition, recording, editing, mixing, listening, playing, evaluation, graphic scores, etc. – through an inquiry approach (Kokkidou, 2015). Using multimodal and multisensory perception, students can produce musical compositions and graphic scores through music-making with sounds from conventional or improvised musical instruments, even with daily objects (Tinkle, 2015). To that extent, soundscape-based activities emphasize listening and recording sounds, such as soundwalks (the exploration of the relationship of the ear with the environment (Westerkamp, 2011)), and soundmaps (rendering graphic representations of the soundscape (Schafer, 1977)), as well as sound libraries (the creation of repositories of sounds (Nicolaidou et al., 2018)).

Ubiquitous music ecologies and digital media

Nowadays, technology is integrated into everyday life, forming an inseparable unity. Within this context, pervasive and ubiquitous computing (see Weiser, 1991) provides expanded means of expression, creation, and learning (Mygdanis, 2021). We have access to music anytime, anywhere, and from anyone, at the push of a button and a click (Pimenta et al., 2014). As a result, mobile devices (smartphones, tablets, etc.), the internet, and ways of interacting with sound and the soundscape shape a new context of sound perception (Lazzarini et al., 2020). In this way, the conventional concepts of the distribution, production, recording, and reproduction of sounds acquire new meanings in the modern environment, creating conditions for pervasive or, to put it another way, ubiquitous music (Lazzarini et al., 2020).

Extending Schafer's (1977) viewpoint that the world's soundscape is changing and that the auditory environment is becoming radically different from what it used to be, this phenomenon has witnessed an augmentation in today's society whereby new digital and multimodal literacies are developing. As we move from desktop personal computers to a multi-platform environment with mobile devices that enable connection and interaction between users, a growing variety of hardware and software has been exploited in the contemporary literature on ubiquitous acoustic ecologies. The recent introduction of components from different scientific fields has enhanced the application of knowledge from other fields in an interdisciplinary approach while, simultaneously, the forms of interaction are expanding (de Lima et al., 2012). As a result, ubiquitous music perception is emerging, and requires technical knowledge from many fields, including, among others, technology, linguistics, physics, mathematics, sociology, philosophy, psychology, and music (de Lima et al., 2020), depending on the cultural and social context of these changes (Lazzarini et al., 2020).

Within this framework, we no longer refer to the use of tools but to ubiquitous music ecologies (Keller & Lazzarini, 2017) as an expanded musical-technological mindset. These technological means are divided into three major categories: a) do-it-yourself (DIY) practices; b) online platforms (browser-based platforms); and c) interconnected distributed resources (Lazzarini et al., 2020). Specifically, they include low-cost hardware (Keller et al., 2014), open-source software platforms and programming languages, electronic DIY constructions, microcontrollers and interfaces (Nikoladze, 2020), mobile computing, and the Internet of Things (Turchet et al., 2020). The integration of the above tools takes place based on the social context, which is a central factor in creative processes (Keller & Capasso, 2006). In particular, materials and resources from everyday life are ideal for collaborative artistic practices, enhancing motivation for inquiry learning and strengthening ecological consciousness (de Lima et al., 2020). Unlike conventional musical instruments strongly connected with Western music notation and requiring specific virtuoso practices, ubiquitous acoustic ecology perspective artifacts instead adapt to the social and cultural context of the individual (Keller, 2020).

STEAM approach, maker culture, and computational thinking in music education

Current trends in music education focus on the transdisciplinary STEAM model – Science, Technology, Engineering, Arts, and Mathematics – for the design of music-pedagogical activities (Gold et al., 2022). Based on the STEM framework's epistemology, STEAM incorporates the field of Arts and emphasizes the computational experiment methodology, an authentic problem-solving approach involving inquiry and experiential learning (Kalovrektis et al., 2021).

Educational scenarios designed using the STEAM approach are based on Seymour Papert's constructionism, in which learning is effective when students experiment and construct an artifact that is meaningful to them (Demetriadis, 2015). Focusing on the concept of "making," the maker culture is grounded in the mindset of STEAM (Huang, 2020), invoking the principles of inquiry, examination, iteration, designing, testing, and problem-solving in order to achieve creative, aesthetic, and self-expressive goals (Gold et al., 2022). In music education, STEAM and maker movement activities can lead to students' cultivation of computational thinking, an in-depth understanding of musical, technological, and scientific concepts, and the development of technological and musical skills, as well as a more in-depth understanding of the digital and physical world (Abrahams, 2018; Palaigeorgiou & Pouloulis, 2018).

The utilization of ubiquitous computing is intertwined with the development of computational thinking, associated with problem-solving and understanding human behaviors through various tools derived from computer science (Kalovrektis et al., 2021). Although there is no clear definition of computational thinking, researchers emphasize specific thematic areas, including abstraction, decomposition, algorithmic thinking, and pattern recognition (Psyharis et al., 2020).

In music education, computational thinking can expand existing horizons and open new doors. Although it primarily refers to computer science practices, such as coding techniques, it creates a new viewpoint where musical phenomena are perceived as computational data (Greher & Heines, 2014). The cultivation of computational-musical thinking creates conditions for new sonic, visual, audio-visual, and tactile approaches to sound, as well as hybrid forms that comprise rich, multimodal musical experiences (Mesz et al., 2012). Research activity demonstrates the added value of computer science by engaging children in authentic teaching-learning situations and connecting with the real world (Kalovrektis et al., 2021). Through maker culture activities (artifact construction), students develop the self-confidence necessary to solve complex problems with collaborative processes, acquire a positive attitude towards open-ended challenges, and cultivate their creative skills (Selby & Woollard, 2013). However, computational thinking has not been widely adopted as a strategy in music-pedagogical activities (Keller, 2020).

Rationale, participants, aim, and research questions

The rationale of the current study is based on the research gap between the incorporation of technologies and digital media in sound pedagogy activities. It focuses on designing and applying a teaching approach drawing on the current literature review about ubiquitous music ecologies following the appropriate steps and methods. The activities function as an innovative perspective and extension of existing teaching strategies in sound pedagogy, and include digital tools, block-based languages, microprocessors, and interfaces.

The educational intervention was designed and implemented in the fall of 2021 within conservatoire education in Greece after a period of quarantine brought about by the COVID-19 pandemic. Five children aged 7 to 9, who did not know each other and with no previous learning experience in a conservatory setting, participated. The primary purpose was the involvement of the children in shaping their physical and digital world through active participation in creative actions (e.g., soundscape composition) and digital storytelling with gamification elements.

The aim of the present research was to study the types of digital media integration that have been applied and implemented through the processes of ubiquitous music ecologies within the pedagogy of sound educational actions with correspondingly positive learning outcomes. At the same time, participants' perception skills, ability, and computational thinking were cultivated during their engagement in creative activities and digital storytelling. Three research questions guided the study, as follows:

1. How can ubiquitous music ecologies enrich traditional sound-based activities from a STEAM model perspective and lead to new forms of creativity?

2. How can the use of digital media contribute to the development of auditory perception and enhance students' computational thinking?
3. How can ubiquitous music ecologies form a learning environment for children to understand and shape their musical worlds?

Methodological Tools

Different methodological data collection tools were applied to increase the reliability of specific aspects of the objects being studied and lead to a more in-depth understanding of their qualitative characteristics (Denzin & Lincoln, 1994; Miles & Huberman, 1994). Data collection tools were: (a) semi-structured interviews with the participating children; (b) data from observations in diaries by the teacher-researchers; and (c) informal discussions with the children during the music creation phase. For ethical reasons, parents were required to agree to let their children participate in the research, and the participants' anonymity was ensured throughout. They were also informed about the research aims and the data collection tools.

The semi-structured interviews took place at the end of each lesson. During the interviews, the children had the opportunity to express their opinions about topics they could not mention in the activities. The diaries and forms were transcribed at the end of each lesson, minimizing the possibility of missing significant information (Denzin & Lincoln, 1994). When needed, data were recorded during the process in a coded way so as not to interrupt the whole procedure. Observation keys were used to organize the data, both to identify the areas within the patterns of behavior and to record unexpected reactions (Denzin & Lincoln, 1994), and included maintaining the children's interest, interactions with each other, self-regulation, active participation, and involvement in creative activities. These data are mentioned in this paper as field notes (FN).

The recorded material from the interviews, observations, and discussions was transcribed a second time into text, the data analysis following a triangulation perspective (Miles & Huberman, 1994). Content analysis was used, drawn from the principles of semantic condensation (Finfgeld-Connett, 2014), through a series of distinct steps that included identification, coding, and counting the frequency of the occurrence of phrases, as well as the rechecking of data (Miles & Huberman, 1994).

Designing and implementing the educational STEAM proposal based on acoustic music ecologies

The teaching intervention was implemented over a period of 14 weeks (one 60 minute lesson per week), following a STEAM approach. Teaching scenarios emphasized the involvement in creative activities (e.g., compositions of soundscapes) and digital storytelling from the perspective of ubiquitous music ecologies. The selected tools were open-sourcing online applications, such as 'SoundBlocks', 'Sampler', the block-based programming language 'Scratch', and the 'Makey-Makey' interface. Although all the steps were prepared beforehand, there was a continuous process of reflection and re-design of the content so that it could be adapted to new conditions as they arose.

Specifically, the first two lessons were an introduction to the description of the phenomenon of sound, emphasizing sound production and its characteristics – timbre, pitch, duration, and dynamics – and its transferability through the practical application of experimentations with simple materials. The actions focused on connecting the audio and the visual stimulus. Gamification elements, such as visual examples of waveforms using flashcards (when children tried to match the sound they heard or suggested their own ways of representation), also played a fundamental role.

In the following two lessons, the children's involvement with soundwalk activities allowed them to get in touch with the concept of the soundscape. In accordance with the restrictions brought about by the spread of COVID-19, the soundwalks took place inside the conservatory area before the children recorded sounds in familiar areas such as homes, neighborhoods, and playgrounds. In the second part of this unit, discussions were held about the kinds of sounds they could or could not hear in a certain environment, and then they created (in groups) artificial soundscapes with the 'SoundBlocks' app. Each group presented their composition, and the rest of the children drew an "imaginary" space for the soundscape they heard (see Figure 1). Finally, the activity was connected to the previous unit and the multiple ways of representing sounds. Discussions and concerns about various issues and the sounds of the environment were essential elements of the process and helped the children to create in a reflective manner.

Figure 1 *Drawing "imaginary" spaces based on other children's soundscapes*

The next unit focused on targeted environments from pupils' everyday lives. The children recalled sounds from their daily routines (e.g., waking up in the morning, the sound of their alarm) that could create feelings and reactions. Having recorded five to nine such sounds, they then created a digital soundmap with these sounds using the 'Sampler' app and connected conductive materials in the 'Makey-Makey' (see Figure 2). The digital map was a DIY construction using interfaces as a prototype digital musical instrument. Following this, the children exchanged sounds and collaboratively created digital maps with "imaginary" soundscapes, enriching the action with storytelling elements and utilizing digital soundmaps to accompany the narrative. In this way, the children effectively created their own sound stories by improvising and composing soundscapes. Finally, the stories were recorded, presented, and reflected upon.

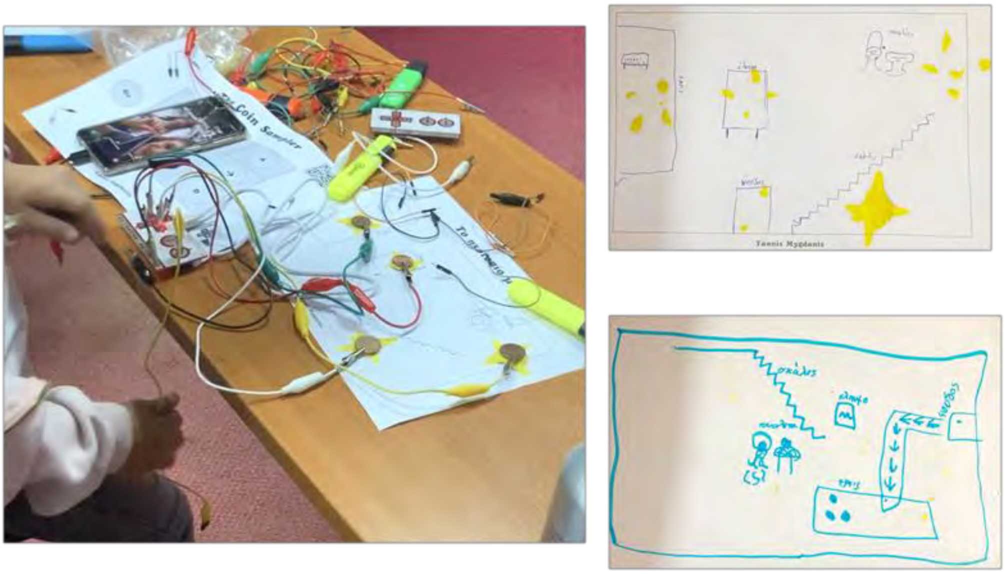


Figure 2 *Creating a digital soundmap based on prerecorded sounds*

After introducing the children to the creation of sound stories, the next unit focused on developing digital storytelling in 'Scratch'. The children invented sounds based on an environment (e.g., school, playground) and imitated them using their bodies or various sound-producing objects. These sounds were recorded in 'Scratch', after which they created the story's characters and adapted the recorded sounds to the movements. When the digital storytelling was completed, the children narrated their stories in groups while the other groups tried to draw them. The groups' role were then interchanged.

In the next stage, the pupils worked on their creations by deepening the use of technological tools based on previous activities. Using 'Scratch', they experimented further with the sound processing functions / commands (pitch, volume, stereo image) and discussed their characteristics, which led to new creative extensions. In addition, the use of ‘Scratch’ was expanded by interfacing the digital soundmap with ‘Makey-Makey’ and using sound processing and block commands to generate random selections (with parameters set by the children). A key element in this unit was the degree to which digital media was incorporated into the children’s creations. As a result, pupils made unique tactile DIY artifacts, and groups chose their favorite processing commands in ‘Scratch’. That process positively widened their sound compositions’ creative perspectives and horizons (see Figure 3).

Figure 3 *Using ‘Scratch’ to create tactile DIY artifacts and enhance their sound compositions*

At the end of the implementation phase, the pupils were introduced to augmented reality elements by drawing on the “real” objects and “virtual” soundscapes and integrating the web camera of their devices (tablets, laptops) in ‘Scratch’. Based on their previous sound stories, the children developed another story with themselves as the main characters. The designed block code was similarly based on their previous actions, forming an extension of the digital soundmap. By moving their hands, the children could "touch" the objects and cause an acoustic result based on the sounds they had previously recorded and imported.

Results and Discussion

The results of the educational intervention suggest that the planned actions opened children's minds to understanding the sound-musical world. Together with acquiring new terms related to sound and soundscape, they understood various ways and creative techniques they could apply in their sound-musical creations. Flow experiences and "aha!" moments were often observed (see Csikszentmihalyi, 2009), especially when they successfully applied a technique or discovered new creative possibilities. Through data analysis, the following thematic categories emerged: auditory perception development, creativity enhancement, computational thinking cultivation, and digital and physical world interaction.

Auditory perception development

Auditory perception development and active listening are fundamental goals in music-sound teaching-learning activities (Dionysiou, 2019). Starting from the first lessons, sound was associated by the children with kinesthetic movement and the environment (see Etmektsoglou, 2014), which contributed to improving their auditory perception, understanding, evaluation, and critical participation in the soundscape where they lived (see Truax, 1996). At first, the sound descriptions were quite general. Gradually, pupils expressed themselves more purposefully using the appropriate terminology, such as volume and pitch, which contributed to their thoughts and opinions of both their own and others' creations.

Technological means also contributed to the cultivation of the children's levels of auditory perception. From the construction of "virtual" soundscapes through online applications, as well as the creation of digital artifacts with physical objects, topic discussions emerged about sound, its sources, and methods of its production, as stated in the following children's opinion: "I can set this coin to sound whatever I want … not only this coin but also me [his body] … super!" (FN 6). At the same time, they formed a critical perspective on the soundscape, as captured in the phrase: "I chose birds [sound] because I don't hear them often. I'll put on a river, too" (FN 12). The visual representation of the waveforms, as well as the sound processing, formed the necessary conditions to allow the children to come into contact with sound in a multimodal way through experimentation and discovery. A child explained this procedure: "Now I understand why what [I recorded] doesn't sound good. It seems to be bad in 'Scratch', so I need to investigate further" (FN 25). On top of that, the audio editing capabilities embraced further creativity and participation, as stated in the phrase: "I recorded myself doing my mom … to see if I could make [the recording] sound like that" (FN 34).

Creativity enhancement

As the intervention emphasized creative activities, the children had the opportunity to experiment with sounds and make music through improvised DIY artifacts. Although they had no prior musical knowledge, they realized that creation is not limited to the sound-auditory effect but extends to the creation of ubiquitous digital musical instruments. They were willing to participate in the maker culture activities even from the first lessons. They gradually gained the confidence to create their own artifacts without the teacher's guidance. The following two phrases summarize their surprise: "Is it that easy to make a musical instrument?" (FN 5), and "All things can make a sound? Let's make music out of everything!" (FN 13).

By the final lessons, the children seemed to have acquired the appropriate musical and technological skills to decide about their creations. Decisions were taken in a collaborative framework between groups and after implementation. As they said: "We had said that we wanted to put a school bus in the story [...], but we didn't like the sounds we recorded [...], so we changed the heroes" (FN 48).

During this phase, the children came into contact with various forms of creativity, from improvised and unconventional musical instruments to ubiquitous music tools. They realized that the phenomenon of sound offers ideas for many creative actions – recording and editing – and how they could use technological means of pervasive music to expand these practices (see Lazzarini et al., 2020). It was also essential that the children utilized digital web applications and tactile interfaces with physical objects in various ways, depending on their individual and collaborative constructions and creations. All tools were explored intuitively and based on the children's musical and technological skills, satisfying complex creative tasks (see Keller, 2020) tailored to their needs and particularities. For example, when a child was exposed to the randomization commands, he adopted the new information into the artifact he was constructing:

> My code changes the sound every time without [knowing] how. When I touch the coin, the faucet [I recorded] is heard, but higher or lower [inc. pitch]. Every time it's different. Isn't it awesome? I'll do the same for the rest [sound samples] and see what happens! I might put fruit instead of coins, or maybe nothing … I don't know yet (FN 63).

Cultivation of computational thinking

Most educational activities reflected the philosophy of ubiquitous music ecologies, which constitute a transdisciplinary STEAM framework with diverse fields of knowledge, including Science, Engineering, Technology, Arts, and Mathematics (see Lazzarini et al., 2020). Within these environments, DIY practices that utilize web applications, interfaces and microcontrollers, programming languages, and mobile computing that requires technical skills and computational thinking, are integrated (see Kalovrektis et al., 2020).

The children did not have any prior knowledge of coding procedures or been involved in similar activities. However, a change in the ways that they participated in the actions was observed, especially in maker movement activities and programming practices, transforming the ways they subsequently asked for help. While at first the questions were general: "How should I start?", "What should I do first?", they gradually focused on problem-solving techniques: "After 'play sound" [a command in 'Scratch'], what else is left?" (FN 18). This focus on targeted issues contributed to the children being able to understand the stages of the problem of decomposition, actively participating in the actions and expressing willingness to expand their knowledge through abstraction: "Nice … now that we've done that, I'll change the order [of commands of the code] to see what will happen" (FN 72). This was obvious, even in the most complex actions, such as augmented reality.

More broadly, the children appeared to gain confidence and show an interest and willingness in solving complex problems through collaborative practices, leading to algorithmic thinking enhancement. This was apparent in peer-to-peer learning procedures, as it was expressed through the dialogues between the children: "Don't rush ... we have to do the steps in order, or it won't play" (FN 48), "Wait ... you haven't put in 'forever' [a command in 'Scratch'], so how do you expect it to play [the sound]?" (FN 53). Programming languages and interfaces offered new perspectives to the children's creativity (see Selby & Woollard, 2013), giving opportunities for computational thinking development. The structured steps of a process – algorithmic thinking – can also be seen from the point of view of a child who, after the construction of the digital sound map, emphasized:

> The purpose was to make an instrument [a digital sound map] unlike any other … our own unique [digital] map. First, we had to decide the things [to connect] to the board … then the shape [of the digital sound map], and then the [‘Scratch’] program. We [definitely] changed it in the end, but we still went through the same [steps] again until we got it right. (FN 37)

Digital and physical world interaction

Since digital and physical world shaping was another aim of the intervention, all educational scenarios followed a student-centered approach, promoting the conditions necessary for acquiring self-regulation through participation and encouraging the development of creative environments.

Regarding the children’s different previous musical and technological experiences in informal learning contexts (home, peers, internet, etc.), an additional goal was to observe their engagement from the very beginning. The children’s choices appeared to be meaningful (see Truax, 1996) on a personal and social level, as they obtained satisfaction and pleasure from experimentation and encouraging self-regulation. Each creation – digital in ‘Scratch’, tactile with physical objects and boards, audio with recording, editing, etc., and / or a combination of the above – varied according to the background and the expectations of the children. The choice of digital media and physical objects depended on the goals set by each team and had creative expansions: “We decided not to use a board; we’ll do it all on the computer" (FN 32).

Within a broader context, participants appeared to positively address the combination of the physical and digital worlds, perceiving them not as separate entities but as a unified whole, a philosophy reflecting that of ubiquitous musical ecologies (see Lazzarini et al., 2020). This was also observed from the interviews, where the children emphasized that: "It's perfect [the fact] that I can work with things [physical objects], change the sound [its features], this orange is now an instrument … not a fruit", "I can make a hero [in 'Scratch'] and touch it to make a sound (not just see it) … I can make it do whatever I want!". The association between the physical and digital worlds was also underlined in the pupils' statements. Although they preferred to use the computer: "We decided not to use a board; we'll do it all on the computer" (FN 32), they also stated in the interviews that "If we did it again, it would be different [...] we would put coins to change the height and water for the volume … it would be nicer [than using the keyboard]" (FN 32).

Coda

To sum up, the children's active participation in the actions of the current educational intervention meant that they were able to interact with their physical and digital worlds. The practical applications demonstrated positive outcomes from the utilization of STEAM scenarios and technological tools in ubiquitous music ecologies for the development of their auditory perception and creativity, the cultivation of their computational thinking, and self-regulation. The results of the educational intervention underlined the fundamental role of the ubiquitous music ecologies perspective in planned actions, which served to widen the understanding of the students' musical worlds. Findings demonstrated that pupils enthusiastically embraced activities and emerged in music-making as active participants. The new emerging learning environment enhanced their engagement, developed their creativity, transformed their experiences, and, in a general context, shaped and enriched their physical and digital musical worlds. The children seemed to enjoy getting acquainted with the new vocabulary related to music and soundscape characteristics. At the same time, they found various patterns and techniques that they could incorporate into their sound-musical creations.

Similar conclusions were reached by de Lima and her colleagues (2020) in their research that used different technological means and was carried out in schools in the Brazilian general education system. Through practical applications, the authors observed an enhancement of creativity through a critical look at everyday life, the development of collaboration skills, and self-regulation, as well as skills in various fields of knowledge such as technology and mathematics (de Lima et al., 2020).

The small number of participants in the present intervention does not allow for the generalization of the results and the conclusions. However, it does give an insight into the experiences and behaviors of children as they relate to new practices and strategies within the realm of ubiquitous music ecologies. In any case, the aim was to understand the process of interaction and complementarity of the educational environments so that the students could form a holistic view of their musical – physical and digital – worlds. Activities and scenarios designed for the needs of the present educational proposal seemed to raise feelings and memories and develop the imagination, which the children found meaningful. This is best illustrated in the following excerpt from a dialog between the children while making their digital map during the seventh meeting:

- I really like coming here. We have many options [apps and interfaces], and I can do whatever I want.
- It’s a game, not a lesson.
- Me too … it’s like a game, but we’re still learning!

References

Csikszentmihalyi, M. (2009). *Creativity: Flow and the psychology of discovery and invention*. New York, NY: HarperCollins.

de Lima, M. H., Flores, L. V., & de Souza, J. C. F. (2020). Ubiquitous music research in basic-education contexts. In V. Lazzarini, D. Keller, N. Otero, & L. Turchet (eds.), *Ubiquitous Music Ecologies* (pp. 109-128). New York, NY: Routledge.

de Lima, M. H., Keller, D., Pimenta, M. S., Lazzarini, V., & Miletto, E. M. (2012).

Creativity-centered design for ubiquitous musical activities: Two case studies. *Journal of Music, Technology & Education*, 5(2), 195-222.

Demetriadis, S. (2015). *Learning theories and educational software* [in Greek]. Athens: Association of Greek Academic Libraries.

Denzin, N. K., & Lincoln, Y. S. (1994). *Handbook of qualitative research*. Thousand Oaks, CA: Sage Publications, Inc.

Dionysiou, Z. (2019). Introduction to the methodology for creative and critical listening of the soundscape [in Greek]. In I. Etmektsoglou, Z. Dionysiou, & A. Mniestris (Eds.), *Towards a sound-based education: Listening, understanding, co-creating the soundscape we live in* (pp. 23-33). Kerkira: Electroacoustic Music Research and Applications Laboratory.

Etmektsoglou, I. (2014). *Basic acoustic ecology terminology for children and adults: The soundscape and the meanings of its sounds* [in Greek]. Kerkira: Hellenic Society of Acoustic Ecology. Retrieved from: http://sound.sch.gr/stable/akouseOrologia.pdf

Etmektsoglou, I. (2019). Listening to Soundscapes as Acoustic Ecologists [in Greek]. In I. Etmektsoglou, Z. Dionysiou, & A. Mniestris (Eds.), *Towards a sound-based education: Listening, understanding, co-creating the soundscape we live in* (pp. 10-22). Kerkira: Electroacoustic Music Research and Applications Laboratory.

Finfgeld-Connett, D. (2014). Use of content analysis to conduct knowledge-building and theory-generating qualitative systematic reviews. *Qualitative Research, 14*(3), 341-352.

Gold, N. E., Purves, R., & Himonides, E. (2021). Playing, Constructionism, and Music in Early-Stage Software Engineering Education. *Multidisciplinary Journal for Education, Social and Technological Sciences*, *9*(1), 14-38.

Greher, G. R., & Heines, J. M. (2014). *Computational thinking in sound: Teaching the art and science of music and technology*. New York, NY: Oxford University Press.

Huang, H. (2020). Music in STEAM: Beyond Notes. *The STEAM Journal*, *4*(2), 1-11.

Kalovrektis, K., Xenakis, A., Psycharis, S., & Stamoulis, G. (2021). *Educational Technology, Robotics, and IoT Development Platforms* [in Greek]. Athens: Tziola Editions.

Keller, D. (2020). Everyday musical creativity. In V. Lazzarini, D. Keller, N. Otero, & L. Turchet (eds.), *Ubiquitous Music Ecologies* (pp. 23-51). New York, NY: Routledge.

Keller, D., & Capasso, A. (2006). New concepts and techniques in eco-composition. *Organised Sound*, *11*(1), 55-62.

Keller, D., & Lazzarini, V. (2017). Ecologically grounded creative practices in ubiquitous music. *Organised Sound*, *22*(1), 61-72.

Keller, D., Lazzarini, V., & Pimenta, M. (2014). Prologue — ubiquitous music: A manifesto. In D. Keller, V. Lazzarini, & M. Pimenta (Eds.), *Ubiquitous Music* (pp. xi-xxii). New York, NY: Springer.

Kokkidou, M. (2015). *Teaching music: New challenges, new horizons* [in Greek]. Athens:

Fagottobooks.

Lazzarini, V., Keller, D., Otero, N., & Turchet, L. (2020). The ecologies of ubiquitous music. In V. Lazzarini, D. Keller, N. Otero, & L. Turchet (eds.), *Ubiquitous Music Ecologies* (pp. 1-22). New York, NY: Routledge.

Mesz, B., Sigman, M., & Trevisan, M. (2012). A composition algorithm based on crossmodal taste-music correspondences. *Frontiers in Human Neuroscience*, *6*(71), 1-6.

Miles, M., & Huberman, M. (1994). *Qualitative Data Analysis*. Sage Publications, Inc.

Mygdanis, Y. (2021). Virtual instruments in music teaching and learning at kindergarten-age: an educational proposal using Synth4kids web-application. In *Digital Culture & Audiovisual Challenges: Interdisciplinary Creativity in Arts and Technology* (pp. 141-148). CEUR Workshop Proceedings.

Mygdanis, Y., & Kokkidou, M. (2021). Collaborative DIY music production practices in conservatoire settings: findings from a pilot distance teaching-learning project. *ICT in Muzical Field/Tehnologii Informatice si de Comunicatie in Domeniul Muzical*, *12*(2), 7-22.

Nikoladze, K. (2020). A brief report from the land of DIY. In V. Lazzarini, D. Keller, N. Otero, & L. Turchet (eds.), *Ubiquitous Music Ecologies* (pp. 71-79). New York, NY: Routledge.

Nikolaidou, D., Tsaligopoulos, A., Economou, X., & Matsinos, I. (2018). The contribution of

Acoustic Ecology to Primary Environmental Education [in Greek]. In N. Bouparis, K. Paparrigopoulos, & G. Matsinos (eds.), *Proceedings of the 4th Conference on Acoustic Ecology "Sound, Noise, Environment"* (pp. 147-163). Hellenic Society of Acoustic Ecology.

Onwuegbuzie, A. J., & Leech, N. L. (2007). Sampling designs in qualitative research: Making the sampling process more public. *Qualitative Report*, *12*(2), 238-254.

Paynter, J., & Aston, P. (1970). *Sound and Silence*. Cambridge University Press.

Pimenta, M., Keller, D., Flores, L. V., de Lima, M. H., & Lazzarini, V. (2014). Methods in creativity-centred design for ubiquitous musical activities. In D. Keller, V. Lazzarini, & M. Pimenta (eds.), *Ubiquitous Music* (pp. 25-48). New York, NY: Springer.

Psycharis, S., Kalovrektis, K., & Xenakis, A. (2020). A Conceptual Framework for Computational Pedagogy in STEAM education: Determinants and perspectives. *Hellenic Journal of STEM Education*, *1*(1), 17-32.

Schafer, R. M. (1977). *The Soundscape: Our Sonic Environment and the Tuning of the World*. Rochester, VT: Destiny Books.

Schiavio, A., Biasutti, M., & Antonini Philippe, R. (2021). Creative pedagogies in the time of pandemic: a case study with conservatory students. *Music Education Research*, *23*(2), 167-178.

Selby, C., & Woollard, J. (2013). Computational thinking: The developing definition.

Retrieved 21 April 2022 from http://eprints.soton.ac.uk/356481

Tinkle, A. (2015). Sound Pedagogy: Teaching listening since Cage. *Organized Sound*, *20*(2), 222-230.

Truax, B. (1996). Soundscape, acoustic communication and environmental sound composition. *Contemporary Music Review*, *15*(1-2), 49-65.

Truax, B. (1999). *Handbook for acoustic ecology*. Cambridge Street Records. Retrieved 18 November 2022 from: https://electrocd.com/en/album/1881/barry-truax-ed/handbook-for-acoustic-ecology

Turchet, L., Essl, G., & Fisichione, C. (2020). Ubiquitous music and the internet of musical things. In V. Lazzarini, D. Keller, N. Otero, & L. Turchet (eds.), *Ubiquitous Music Ecologies* (pp. 154-169). New York, NY: Routledge.

Weiser, M. (1991). The computer for the 21st century. *Scientific American*, *265*(3), 94-105.

Westerkamp, H. (2011). Exploring balance & focus in acoustic ecology. *Soundscape: The Journal of Acoustic Ecology*, *11*(1), 7-14.

Asia-Pacific Journal for Arts Education

Special Issue

Unfold the Future of Music Education through Technology

Co-editors:
Prof. Bo Wah LEUNG
Dr. Cheung On TAM
Dr. Chi Hin LEUNG
Dr. Richard Guy WHITBREAD
The Education University of Hong Kong

http://www.ied.edu.hk/ccaproject/apjae/apjae.htm ISSN: 1683-6995

Volume 22 Number 6 December 2023

Pedagogical choice-making and transformation: A longitudinal study of Chinese preservice guzheng teachers in the U.K.

Dr. Xinpei Zheng

University of York

xz2250@york.ac.uk

Dr. Elizabeth Haddon

University of York

liz.haddon@york.ac.uk

Abstract

A longitudinal study over 12 months with three Chinese preservice guzheng teachers, all of whom were students on an MA Music Education: Instrumental and Vocal Teaching programme at a U.K. university, revealed their pedagogical transformation across cultures. Three rounds of semi-structured, online interviews were facilitated by a video-stimulated interview (VSI) technique: the *reflective* VSI collected the student-teachers' retrospective descriptions of their recorded guzheng lessons, while the *interpretive* VSI gained their espoused pedagogical beliefs about guzheng teaching through discovering their views concerning screen-shared videos of expert guzheng performances. Findings suggest that the student-teachers' pedagogical choices were influenced by: 1) their concurrently received U.K. teacher training, enabling an enhanced understanding of pupil-centred and process-engaged pedagogy; 2) their previously experienced guzheng tuition in China, featuring teachers' dominance and external expectations from Chinese parents / employers / audience; and 3) regional stylistic features within the guzheng repertoire.

Key words

Chinese-instrument pedagogy, cross-cultural context, pedagogical transformation and challenges, preservice instrumental teachers, video-stimulated interview (VSI)

Introduction

Developments within instrumental teaching have increased the demands on both performing expertise and the pedagogical knowledge of instrumental teachers, collectively embodied in the shift towards mentorship as a means of facilitating pupils' engagement and independence (Creech & Gaunt, 2012). Fredrickson et al. (2013) recognise that many conservatoire / college music students give private lessons in the West; in China, the situation is similar (Qin & Tao, 2021). Nonetheless, limited specialist instrumental teacher training is evidenced in the U.K. (Haddon, 2009; Mills, 2004; Norton et al., 2019); the extant research indicates a variety of influences on university music students' teaching (Haddon, 2009) and conservatoire students' perspectives on delivering Western instrumental/vocal (I/V) lessons (Mills, 2004; Perkins et al., 2015). By contrast, how Chinese music students, specifically those working with Chinese traditional instruments, establish their teaching practices, has remained largely unexplored.

In the U.K., taught Master's programmes on I/V teaching offer systematic and research-based pedagogical training, which attract Chinese students (Haddon, forthcoming, 2023). To date, however, no research has concentrated specifically on Chinese instrumental teachers on Western MA programmes who teach Chinese instruments (e.g. guzheng), exploring their pedagogical decisions for teaching Chinese instruments and their perceptions of U.K. I/V teacher training that may impact upon their pedagogical development. Through video-stimulated and semi-structured interviews, this instrument-specific longitudinal investigation of preservice trainee teachers encountering instrumental pedagogy across China and in the U.K. aimed to explore: 1) the factors that influence guzheng student-teachers' teaching decisions; and 2) manifestations of transformation reflected in both their actions and espoused beliefs in guzheng pedagogical practice during their U.K. teacher-training. The research outcomes in relation to pedagogical differences in instrumental teaching across cultures, and the effects of U.K. I/V teacher training, could serve to inform instrumental teachers worldwide who wish to develop teaching expertise cross-culturally, as well as proving of benefit to instrumental (teacher) educators in China.

Teaching Chinese instruments in China

'A long apprenticeship with a master' historically characterises the transmission mode of Chinese instrumental teaching and learning (Schippers, 2010, p. 71). Jørgensen (2000) identifies that the teacher generally represents the dominant authority in a master-apprentice model, whereby the student's role is to imitate the teacher. Yung (1987) details a student's phrase-by-phrase imitation of the teacher's guqin (seven-string zither) performance, as well as the teacher-pupil simultaneous playing of the same piece, whereby the pupil 'inherits the nuances of the music, especially its rhythm and phrasing, from the teacher' (p. 85) due to an aural transmission tradition that preserves the 'metrical, rhythmic, and phrasal aspects of the music' (p. 84). This aural approach has been recognised as a vital means of passing on Chinese instrumental music, as notation systems in Chinese music mostly record important skeletal notes (see Thrasher, 2008 for details of '*gongche*' notation and 'simple notation', Yung, 1987 for tablature notation, and Chow-Morris, 2010 for 'jiahua' (加花), the technique of adding improvised ornaments to the original music to melodically or rhythmically enrich skeletal notation). Regional stylistic differences between North and South Chinese instrumental repertoire, as they relate to musical forms, interpretations, and the application of techniques (Thrasher, 1989), may also impact Chinese instrument transmission.

Such vague notated instructions concerning the dynamics and rhythm of the music should offer spaces for performers to produce diversified yet personalised interpretations (Yung, 1987). However, the authority of Chinese master-musicians within the master-apprenticeship norm appears to challenge this, leading to questions of how Chinese instrumental teachers deploy teaching dominance. For example, in contrast to the highly-valued personal expression of individual performers within the Western musical tradition (Meissner et al., 2021), the reproduction of inherited musical interpretation is prioritised by instrumental teachers in China (Zheng & Leung, 2021a). Moreover, an exam-oriented hierarchical pedagogy, alongside the domination exerted by the teacher, has been revealed within the Chinese context (Bai, 2021; Haddon, forthcoming, 2023), in which Chinese instrumental preservice teachers consider that their teaching rarely engages pupils' autonomy and is significantly affected by inherited practices from their previous teachers

(Haddon, forthcoming, 2023). In addition to prior learning and teaching experiences, concurrently received pedagogy also shapes teachers' teaching practice (Haddon, 2009; 2019).

Teacher training and pedagogical transformation across cultures

The master-apprenticeship model is not exclusive to the Chinese context. While it historically dominates instrumental tuition (Jørgensen, 2000; Burwell, 2013), a pedagogical transformation from apprenticeship to mentoring – facilitating learner independence, confidence, and creativity – has also been identified in Western educational contexts (Creech & Gaunt, 2012). This is embodied in the U.K. I/V teacher training curricula, as acknowledged by Chinese trainee teachers in the U.K. who have gained an understanding of pupil-centredness and supporting pupils' self-efficacy through dialogic and interaction-oriented strategies (Haddon, forthcoming, 2023). Kleickmann et al. (2013) also affirm the constructive role of teacher preparation programmes in preservice teachers' ongoing development of Pedagogical Content Knowledge (PCK) through a combination of subject-specific expertise and the knowledge of how to strategically teach it (Bauer, 2020).

Pedagogical changes impact teachers' motivation and consciousness of their routinised behaviour and perceptions of teaching, awakened by new or unusual circumstances that stimulate discursive awareness to develop their practices (Burridge, 2018). This appears to align with psychological and behavioural adjustment in cross-cultural contexts. Ang and Van Dyne (2015) indicate that a new cultural setting influences self-identity and self-enhancement; adaptation actions operate cognitively and metacognitively, and are driven by motivation. Accompanied by reflective practice, trainee teachers' adaptability to changing transmission contexts that reflect societal changes should be cultivated through teacher preparation (Conway & Hibbard, 2019). Engaging with I/V teacher training in a cross-cultural context from China to the U.K. is likely to help Chinese preservice instrumental teachers become aware of their habitual, China-domiciled teaching procedures and promote the necessary degree of pedagogical transformation. The ensuing challenges, to which they are returning are, nonetheless, reflected in their actual pedagogical implementation and re-adaptation to the Chinese cultural and educational environment (Haddon, forthcoming, 2023).

Research Design

This qualitative study aligns closely with the social constructivist interpretive framework, supporting multifaceted realities as constructed by research participants' interpretations of their lived experiences and the researcher's understanding of the issue being debated (see Creswell & Poth, 2017; Ritchie et al., 2013). A longitudinal investigation (Ployhart & Vandenberg, 2010) over 12 months, comprising three rounds of semi-structured, one-to-one online interviews with the first author, was employed to explore changed and unchanged aspects in guzheng teaching identified by a sample of Chinese preservice guzheng teachers in the U.K., alongside the pedagogical impact of the University of York's one-year MA Music Education: Instrumental and Vocal Teaching programme which they studied. A technology-assisted research method – video-stimulated interview (VSI) – was adopted to facilitate the interviewees' self-revealing cognitive processes, including a *reflective* and *interpretive* process (Van Braak et al., 2018). Gazdag et al. (2019) recognise the benefits of applying this approach within teacher training, including the degree to which it aids an understanding of patterns between preservice teachers' pedagogical beliefs and actions. This research, therefore, utilised recordings of the MA student-teachers' guzheng lessons in order to explore their teaching actions via the *reflective* VSI (e.g., asking *why* they applied specific approaches at certain points in their teaching), whilst also using online videos of guzheng performances so as to gain an understanding of their espoused teaching beliefs through the *interpretive* VSI (e.g., asking *how* they would teach the music that was being performed) as Chinese instrument teachers with cross-cultural pedagogical experiences. Ethical approval was granted by the Arts and Humanities Ethics Committee at the University of York. Considering the relationship between participants' responses and the researcher's interpretation of those from an epistemological standpoint (Ritchie et al., 2013), the authors' backgrounds should be noted. The first author was a guzheng learner and teacher in China and had completed the same

MA programme; she had been studying and practising Western instrumental pedagogical approaches before and during her doctoral study at the same U.K. university, supervised by the second author.

Participants

Three participants engaged in all three rounds of interviews (i.e. nine interviews in total, averaging 63 minutes each). Table 1 details the participants' abbreviations, their guzheng learning and teaching experiences prior to and during their U.K. MA study, and provides information about their pupils and their envisaged employment as Chinese returnees.

The participants, all female, learned the guzheng one-to-one in private music studios before and during their school years. This training was arranged by their parents and was undertaken with guzheng teachers employed by Chinese conservatoires in order to support the students' preparation for the *Yishu Zhuanye Gaokao* (艺术专业高考; simplified as *Yikao* 艺考), 'The University Arts Entrance Examination' (UAEE) in China (Bai, 2021). They started teaching the guzheng part-time during and after their undergraduate studies.

Table 1 *Participant information*

Student-teachers' MA course enrolment	**Instrument; years spent learning; years spent teaching**	**Learning contexts prior to MA study**	**Teaching contexts prior to and during MA study**	**Pupils**	**Envisaged employment situation as returnees**
• **ST1** • Cohort A, September 2020 – September 2021	• Guzheng • 15 • 3	Learning 1-1 before and during conservatory study in Northern China	• Group teaching in a primary school (part-time) and 1-1/small group teaching at a private music institution in China (full-time) • Private teaching in the U.K. (for the MA course assessments, not regularly)	• Mostly Chinese (taught 1 non-Chinese pupil in China) • Mostly school students and a few adult pupils at beginner / intermediate level • Taught 2 Chinese pupils (including 1 course mate) during their MA	Private, part-time guzheng teacher (considering this as a full-time job)
• **ST2** • Cohort A, September 2020 – September 2021	• Guzheng • 10 • 5	Learning 1-1 before and during study at a comprehensive university in Southern	• Private teaching at a music studio in China (part-time) • Private teaching in the U.K. (for the MA course assessments, not regularly)	• All Chinese • Mostly school students and some adult pupils at beginner / intermediate	Full-time school music teacher and private guzheng teacher at home outside working hours

		China		level • Taught 2 Chinese pupils (including 1 course mate) during their MA	
• **ST3** • Cohort B, January 2021 – January 2022	• Guzheng • 10 • 5	Learning 1-1 before and during study at a conservatory in Southern China	• Private teaching at a music studio in China (part-time) • Private online 1-1 teaching in the U.K. (regularly)	• All Chinese • Mostly school students at beginner / intermediate level and some adult pupils • Taught 5 Chinese pupils (including 1 course mate) during their MA	Private guzheng teacher (full-time) or pursuing further study

Materials

One-to-one guzheng lessons, video-recorded by the participants before each interview, and three expert guzheng performance videos on ‘YouTube’ were used as prompts in the reflective and interpretive VSI processes. The ‘YouTube’ videos were selected by the first author for the interpretive VSI process based on the literature on regional contexts of Chinese instrumental music (e.g., Chow-Morris, 2010; Thrasher, 1989), the focus being on the inclusion and representativeness of repertoire from the Northern and Southern guzheng genres and the reputation of the performers in China: Video 1 – *Yu Zhou Chang Wan* (*Fisherman's Song at Night*; Northern genre)[1]; Video 2 – *Feng Xiang Ge* (*Song of the Flying Phoenix*; Northern genre)[2]; Video 3 – *Fen Hong Lian* (*Pink Lotus*; Southern genre)[3]. Four guzheng lesson recordings with an average duration of 31:22 were received from the participants, instead of nine as expected by the authors, due to the limited availability of the student-teachers' pupils and the practicability of simultaneous research participation and completion of the MA course assignments. ST1: no lesson recordings; ST2: 1 lesson recording (LR1: ST2); ST3: 3 lesson recordings (LR2/3/4: ST3).

Data collection procedure

Each interview with the first author included questions on the videos and a semi-structured interview (SSI), which used pre-designed, open-ended questions and related to the participants' guzheng learning and teaching experiences in China, as well as their perceptions of acquiring and applying the pedagogical knowledge obtained during their MA. Before each reflective VSI, observation notes were made by the first author by repeatedly reviewing the

---

[1] Link to Video 1: https://www.youtube.com/watch?v=xuj7kY4QCr8.

[2] Link to Video 2: https://www.youtube.com/watch?v=tBh3gghgZuo.

[3] Link to Video 3: https://www.youtube.com/watch?v=l3MonhzOfp4.

participants' lesson videos in order to generate interview questions based on timestamps of specific instructional points (see Table 2 for an example). The three 'YouTube' videos were prepared for screen-sharing in the corresponding interpretive VSI, as summarised in Table 3.

Table 2 *Sample of two of the interview questions generated with timestamps from LR2: ST3*

Timestamp	Observation	Interview questions
02:02	ST3 asked their student to practise an etude of fingerings before playing the piece.	- What do you think of the fingering practice? - Why did you choose this kind of approach?
04:19; 21:55	ST3 emphasised the student's hand shape and the skill of using their finger joints to exert force when flicking the strings.	- Why did you emphasise these techniques?

Table 3 *Interview timeline and procedure with abbreviations*

Participant	1st Round: March 2021, Spring term	2nd Round: July 2021, Summer term	3rd Round: October 2021, Post-MA
ST1	SSI	SSI + Interpretive VSI (Video 1 used)	SSI + Interpretive VSI (Videos 2, 3 used)
ST2	Reflective VSI (with ST2: LR1) + SSI	SSI + Interpretive VSI (Video 1 used)	SSI + Interpretive VSI (Videos 2, 3 used)
ST3	Reflective VSI (with ST3: LR2) + SSI	Reflective VSI (with ST3: LR3) + SSI + Interpretive VSI (Video 1 used)	Reflective VSI (with ST3: LR4) + SSI + Interpretive VSI (Videos 2, 3 used)

Data Analysis

The data included nine interviews and four videoed lessons analysed thematically based on the steps outlined by Braun and Clarke (2006) after transcription and a member check process enabling participants' verification (see Creswell & Guetterman, 2019). Adoption of those authors' '*in vivo* codes' (the actual words of the participants) and the use of researcher-assigned codes, to convey the meaning of the chosen segment , combined with observation notes, were conducted by the first author using the software MAXQDA. Whilst definitive conclusions cannot be drawn from such a small sample of instrument- and context-specific international preservice teachers learning to adjust their teaching to cross-cultural pedagogical influences, this study nevertheless offers a starting point for research on Chinese instrumental teachers and instrumental pedagogical learning across cultures.

Findings

The findings of the study are considered within the following sections: teacher-centredness and exam-focused guzheng tuition in China, teacher-training inspired pedagogical choices and beliefs, and the influence of musical features within the guzheng repertoire, together with pedagogical and performance constraints.

Teacher-centredness and exam-focused guzheng tuition in China

All the participants expressed their opinions concerning the guzheng tuition they had received in China, much of which was dominated by their teachers: 'Both teaching and teaching content are teacher-centred' (ST2); 'I feel that this kind of teaching is a matter of convention … I would learn whatever the teacher taught, because my teacher wouldn't ask me what I wanted to learn' (ST1). ST3 indicated that 'the single mode' of teacher-led delivery should be enriched in China. Specifically, 'Jiahua', the means of imparting semi-improvisational skills, exhibited their teachers' dominance, relying on demonstration and direct intervention relating to the sheet music: 'My teacher would do a demonstration of their improvised ornaments and then ask me to take notes on the score' (ST3); 'My teacher helped me to mark on the score, such as where I should add *vibrato* and where I should add a *portamento* or change the fingerings and repeat some phrases' (ST2); '[The teacher demonstrated where] I could add an extra *glissando* to this phrase and add some *vibrato* to some notes' (ST1). However, as noted by Hallam (2012), trials and 'musical doodling' are needed to cultivate students' improvisational skills, as this requires the ability to play aurally and the flexible use of mastered techniques (p. 652) unconstrained by 'effortful, conscious control' (Ayerst, 2021, p. 68). Although the semi-improvisation technique in guzheng performance is different from improvising completely new material, enabling students' autonomous and creative attempts in acquiring such capability might be a productive first step.

Specific strategies were adopted by the participants when teaching in China, learned from their received guzheng tuition. All participants expressed their reliance on demonstration, as used by their former teachers, due to its ‘direct’ (ST1) and ‘time-saving’ (ST3) effect: ‘If the student’s performance is not as ideal as expected, I give a demonstration’ (ST3). Verbal instructions concerning musical dynamics (e.g., ‘stronger’ or ‘weaker’) were simply stated by ST1 and ST3. When it came to adding ornaments, ST2 followed her teacher’s method of marking them on the pupils’ scores. Additionally, singing the melody before playing it was applied by both ST2 and ST3. As ST3 explained, ‘My teacher said you can’t play it correctly if you can’t sing it correctly’. These utilisations appear to indicate the subliminal impact of experienced, teacher-centred pedagogy on student-teacher practice.

In order to succeed in the UAEE, all three participants recalled the significant emphasis given to the technical and expressive aspects of guzheng performance by their previous teachers. ST1 thought that their technical skills had been notably improved during the preparation for the UAEE, thanks to their teacher’s demands:

> … everything was for the UAEE … how the techniques should be practised to achieve the required tempi on the scores and how fast the *allegro* should be. The metronome must be used to assist in the practice and [I] must achieve the tempi.

Similarly, ST3’s teacher ‘paid attention to the way the fingers exert strength and the emotional processing of the music’, while the teacher of ST2 ‘mainly emphasised how to deal with the expression of the music'. Nevertheless, there is room for further diversification of those dimensions that can be explored together by the teacher and student within music pedagogical contexts, including 'the relationships between the detailed honing of stylistic expression and technique on the one hand, and personal expression, spontaneity and risk-taking on the other, while also pursuing and exploiting possibilities for co-creation and experimenting with audience interactions' (Gaunt, 2017: 31).

Interestingly, after achieving the UAEE, ST1 discovered that the same teacher empowered them with more autonomy:

> I felt that my teacher's ideas and foci were completely different compared to preparing for the UAEE … [The teacher] gave us a lot of freedom to enjoy the music … and didn't rigidly stipulate how we should play … the teacher just guided us from the side or gave us some inspiration.

This situation reflects the impact of examinations on teaching. In a context with considerable consequences if the student fails the exam, deploying a directive pedagogical approach may ensure faster progress and greater success in the exam.

Owing to the importance placed on passing graded examinations by parents in China (Bai, 2021), an exam-focused guzheng tuition philosophy was common to all the participants: 'An important goal for school students learning an instrument is to take the graded exams for certificates' (ST3). This tendency affected ST2's teaching priorities: ' … for children, I paid more attention to the basics of hand shape and fingerings … because they have to take the exams' and influenced teaching content: 'I chose teaching materials according to the requirements of those institutions that organise the graded exams. They have different requirements for the repertoire and textbooks' (ST1). This exam-focused tuition, with its implications for parental influence, teaching materials and teachers' perceived value, contrasts with some of the alternative tuition priorities encountered within the U.K. MA course, e.g., teaching those students who learn for pleasure and do not wish to take exams.

Teacher-training inspired pedagogical choices and beliefs

The teaching approaches facilitated by the MA course influenced the student-teachers' pupil-centred pedagogical transformation, inasmuch as they believed that pupils should be encouraged to create their own musical interpretations. Before demonstrating a particular part of the piece in LR4, ST3 commented: 'My interpretation of the music is not the only answer. I'm just sharing my understanding of the music with you.' (9:34, ST3: LR4). During the interview, ST3 explained further:

> Many students think that what the teacher demonstrates is completely right, and they imitate exactly what the teacher does … [the pupil] doesn't need to imitate my playing because everyone has different ideas [of interpretation], and I don't want [my demonstration] to limit their performance.

Likewise, in the interpretive VSI, ST2 stated ‘I won’t tell the pupil directly how they should play [the piece] because I don’t want to dominate their learning anymore’.

Open-ended questioning was one of the learned strategies from the MA course, acknowledged by all participants as effective: ‘ … I can feel that the student actually thinks more on their own than before since I employed this method’ (ST1). For example, ST2 asked their pupil, ‘How do you feel this time when playing?’ (12:44, ST2: LR1), while ST3 asked, ‘What are you struggling with in practising this piece?’ (11:47, ST3: LR2). Those answering such questions are unlikely to respond with a single word, suggesting that using questions strategically to create dialogue in music teaching is ‘an invaluable tool’ for engaging pupils’ cognitive thinking (Allsup & Baxter, 2004: 33). Both ST2 and ST3 appeared to understand the mechanism of this approach and used the questioning technique purposefully: ‘I want the pupil to proactively think and analyse [their playing]’ (ST2); ‘[The pupil] can have a sense of participation [by pointing out challenges], which enhances their impression’ (ST3). By contrast, ST2 mentioned a lack of questioning in their teaching before the MA, ‘ …. because at that time I didn't realise that asking questions was a very important part of teaching’.

Another aspect of the transformation in the relationship between the students and the teachers was revealed in their balancing of teacher-driven progress-making and pupil-engaged learning processes. This was embedded in their attention to pupil-selected learning materials and an increased general awareness of their pupils’ wellbeing, rather than being focused solely on outcomes. They also made a conscious effort to avoid physical contact with their pupils, considering appropriateness and the importance of learners understanding how to achieve technique without physical intervention, although physical touching might be a straightforward way to make corrections in teaching and was used more freely in the China context: ‘My teacher would correct my hand shape through touching my fingers’ (ST3); ‘I didn’t realise that this [physical contact] was inappropriate … my teacher taught me the same way’ (ST1); ‘My teacher would directly grab my hands and instruct me to play’ (ST2). Additionally, all participants expressed their intention to create a mentor-friendly environment for their pupils, inspired by the MA tutors; ST2 and ST3, in particular, demonstrated an encouragement-oriented teaching technique in their recorded lessons.

As also disclosed by the participants, changes in their pedagogical practices during the MA were related to the positive effects of the teaching approaches demonstrated by their tutors, which in turn consolidated the practical role and value of these styles with their students (ST1; ST3). In addition, familiarity with the PCK learned in the training (reinforced by completing the course tasks and assessments (ST2; ST3)), and time spent on trialling strategies helped with confidence-building in pedagogical implementation (ST1; ST2). Given that change is not guaranteed within the framework of a longitudinal study (Ployhart & Vandenberg, 2010), the student-teachers’ consistent attention to background knowledge of the guzheng repertoire and consideration of sociocultural expectations in China displayed signs of some continuing practices within their subsequent guzheng tuition.

Influence of musical features within the guzheng repertoire, pedagogical and performance constraints

Based on the musical and technical features contained within the music, all the participants were able to recognise which region and guzheng genre the pieces belonged to after watching the three performance videos in the interpretative VSI ST1 mentioned the division of different guzheng genres in the North and South, leading to the development of representative musical characteristics and techniques, while ST3 identified the distinction, from a technical perspective, between the repertoire pertaining to the two regions:

> … in the Southern guzheng genres, they emphasise the variation notes using *vibrato*, but they don't press the strings as heavily as the Northern genres do. Their action of applying the *Huayin* [*portamento*] is more subtle compared to the Northern genres.

A similar regional stylistic impact on the different application of semi-improvisational techniques between the Chinese regions was acknowledged by ST2; these techniques help to enhance the musical characteristics in order to show the grandeur of the Northern and the mildness of the Southern guzheng repertoires, respectively. Based on their previous and current practices, and envisaged teaching practices of the material within all three videos, the participants consistently concentrated on introducing background knowledge (e.g., region / genre) of the pieces to their pupils to support their stylistic understanding of the music.

The pedagogical context of guzheng tuition in China, characterised and understood by the participants as an efficiency-led ideology with expectations on the part of both parents and employers, appears to challenge the prevailing notions of pupil-centredness and process-engagement. ST1 expressed her frustration of providing 'a kind of service profession' when working in a private institution in China, where only 'recognised' progress was important for the learners and parents; consequently, she would 'subconsciously speed up the pace of teacher dominance' in order to avoid pupil attrition. The ability to retain students was likewise valued by ST2's previous employer. Concerns relating to a greater reliance on pupil autonomy were raised by ST2, who found that her pupils in China were used to learning passively without being asked for active responses; similarly, pupils' lack of participation and an inactive expression of ideas resulted in a hesitation of ST3 to using a pupil-directed pedagogy. From a cultural perspective, ST1 and ST2s' perceived expectations regarding the default musical styles and advanced techniques of the guzheng repertoire that the audience was familiar with had a subliminal effect on their teaching focus. For example, ST2 pointed out that an audience's stereotypical views concerning the repertoire that 'must' be learned by guzheng players could limit the diversification of repertoire choices.

Discussion

The findings of this study indicate that a range of multidimensional factors, comprising previously and concurrently experienced educational environments in China and the U.K., combine to play a role in preservice instrumental teachers' context-specific pedagogical decision-making and transformation. By way of an example, teacher training heightened the guzheng student-teachers' awareness of pupil-centred instrumental pedagogy and their reflections on tensions between accomplishing external objectives (e.g., expectations of parents / employers and graded exams) and exploratory musical engagement within pupil-teacher interactions. That being said, teaching is not exclusively either teacher-led or student-centred. Rather, it is flexible, deliberative, and context-based (Chan, 2008). However, while the student-teachers' pedagogical implementation shows their proactivity in terms of adjusting teaching approaches, several aspects nevertheless serve to challenge pedagogical transformations, including the regional stylistic features that impact upon repertoire teaching, parents' / employers' focus on effectiveness, pupils' passive learning styles, and the Chinese audience's expectations. This aligns with Haddon's (forthcoming, 2023) findings regarding Chinese returnees' re-adaptation to pedagogical practice in the China-domiciled setting, elicited by the result-driven and hierarchical cultural norms that manifest the default control of parents and employers over pedagogical practice. Compounding the issue, professional support in China to cope with the gap in the returnees' expectations generated by the changing sociocultural environment is lacking (ibid). Further research should explore the perspectives of in-service and experienced / master-teachers who teach Chinese traditional instruments in China, focusing on the delivery of regional stylistic knowledge and dealing with the expectations generated from parents, audiences, and society.

Those musical explorations as presented by the participants' guzheng tuition in the Chinese context appear to be limited to considerations of technical issues and the expected interpretation of the stipulated material by teachers / audiences / the exam syllabus). Borel (2019) observed that conservatoire students in China are used to being trained as soloists; perfection of technique and in performance is demanded through teachers' strictness and authority, a factor that may inadvertently shape the ensuing teaching practice style of students as teachers. As revealed by the participants, guzheng graded exams remain the prime focus of school-aged pupils and their parents, a finding consistent with students taking graded piano exams in China (Bai, 2021). However, Chinese returnee piano teachers pointed out the inexplicit marking criteria and lack of aural and sight-reading tests in the Chinese graded exams (Haddon, forthcoming, 2023); this appears to be similar to graded guzheng exams, organised by varied organisations, each of which have their own syllabus. These skills appear not to have been widely included within instrumental pedagogical development in the Chinese context, in which formal training for instrumental teachers remains underdeveloped (Lee & Leung, 2022). Compared to the increased amount of research concerning creative piano pedagogy and performance in relation to Chinese settings (Zheng & Leung, 2021b), less attention has been paid to the practices and practitioners of Chinese traditional instruments, either in China or elsewhere.

The development of I/V pedagogy requires 'both evidence-based accounts of its effectiveness, and ways to improve its practice' (Carey & Grant, 2015: 19). However, empirical studies on the pedagogies and beliefs underpinning the teaching of Chinese instruments in China, including their establishment, the sociocultural context in which they are conducted, and the people who implement them, are currently lacking. Given that building collaborative, interdisciplinary communities of practice has been recognised as invaluable to facilitating connection-making and transformative learning among practitioners (Creech & Gaunt, 2012), instrumental teachers in China may benefit from further research that investigates the potential for such specialist support.

Conclusions

This study explored the impact of U.K. I/V teacher training on the pedagogical transformation of Chinese preservice guzheng teachers. The findings indicate a variety of student-teachers pedagogical changes aimed at enhancing pupils' learning engagement, including the interactive approach of asking open-ended questions, caring about their pupils' wellbeing, and creating an encouragement-oriented atmosphere. The positive effect of the research-informed strategies, familiarity with the learned PCK, and the frequency of the practical application of pedagogical practice all serve to stimulate the transformation, while the consistency exhibited in the valuing of background repertoire knowledge may be due to the student-teachers' consistent perception of the regional stylistic characteristics that influence guzheng performance. The study also identified a number of challenges relating to the practicability of implementing pupil-directed approaches in China due to the exam and efficiency-oriented teaching context experienced by the participants prior to their U.K. study; future research on simultaneously developing pupils' learning autonomy and working with external expectations (e.g., from parents / employers) in the Chinese context is suggested as being of pedagogical value. Overall, this small-scale study serves to provide a starting point for the further investigation of Chinese instrumental teachers and related pedagogies as they concerns master-apprenticeship influences, pedagogical developments, and Chinese instrumental teacher training.

References

Allsup, R. E., & Baxter, M. (2004). Talking about music: Better questions? Better discussions! *Music Educators Journal*, *91*(2), 29-33. https://doi.org/10.2307%2F3400046

Ang, S., & Van Dyne, L. (2015). *Handbook of cultural intelligence: Theory, measurement, and applications. Routledge*.

Ayerst, J. (2021). Mobilising improvisation skills in classically trained musicians. In H. Meissner, R. Timmers, & S. E. Pitts (Eds.), *Sound teaching: A research-informed approach to inspiring confidence, skill, and enjoyment in music performance* (pp. 61-72). Routledge.

Bai, B. (2021). Piano learning in the context of schooling during China's 'piano craze' and beyond: Motivations and pathways. *Music Education Research*, *23*(4), 512-526. https://doi.org/10.1080/14613808.2021.1929139

Bauer, W. (2020). *Music learning today: Digital pedagogy for creating, performing, and responding to music* (2nd ed.). Oxford University Press.

Borel, E. G. (2019). The Shanghai Conservatory of Music and its rhetoric: Building a world class musical institution with Chinese characteristics. *China Perspectives*, *2019* (2019-3), 27-35. https://doi.org/10.4000/chinaperspectives.9391

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Burridge, P. (2018). Teacher pedagogical choice. In O. B. Cavero, & N. Llevot-Calvet (Eds.), *New pedagogical challenges in the 21st century - Contributions of research in education*. IntechOpen.

Burwell, K. (2013). Apprenticeship in music: A contextual study for instrumental teaching and learning. *International Journal of Music Education*, *31*(3), 276-291. https://doi.org/10.1177/0255761411434501

Carey, G., & Grant, C. (2015). Teacher and student perspectives on one-to-one pedagogy: Practices and possibilities. *British Journal of Music Education*, *32*(1), 5-22. https://doi.org/10.1017/S0265051714000084

Chan, C. (2008). Pedagogical transformation and knowledge-building for the Chinese learner. *Evaluation & Research in Education*, *21*(3), 235-251. https://doi.org/10.1080/09500790802485245

Chow-Morris, K. (2010). Going with the flow: Embracing the "Tao" of China's "Jiangnan Sizhu". *Asian Music*, *41*(2), 59-87. https://www.jstor.org/stable/40930322

Conway, C., & Hibbard, S. (2019). Pushing the boundaries from the inside. In C. Conway, K. Pellegrino, A. M. Stanley, & C. West (Eds.), *The Oxford handbook of preservice music teacher education in the United States* (pp. 2-22). Oxford University Press.

Creech, A., & Gaunt, H. (2012). The changing face of individual instrumental tuition: Value, purpose, and potential. In G. E. McPherson & G. F. Welch (Eds.), *The Oxford handbook of music education*, *Volume 1* (pp. 694-711). Oxford University Press.

Creswell, J. W., & Guetterman, T. C. (2019). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (6th ed.). Pearson Education.

Creswell, J. W., & Poth, C. N. (2017). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). Sage Publications.

Fredrickson, W. E., Geringer, J. M., & Pope, D. A. (2013). Attitudes of string teachers and performers toward preparation for and teaching of private lessons. *Journal of Research in Music Education*, *61*(2), 217-232. https://doi.org/10.1177/0022429413485245

Gaunt, H. (2017). Apprenticeship and empowerment: The role of one-to-one lessons. In J. S. Rink, H. Gaunt, & A. Williamon (Eds.), *Musicians in the making: Pathways to creative performance*, *Vol. 1* (pp. 28-56). Oxford University Press.

Gazdag, E., Nagy, K., & Szivák, J. (2019). "I spy with my little eyes..." The use of video stimulated recall methodology in teacher training – The exploration of aims, goals and methodological characteristics of VSR methodology through systematic literature review. *International Journal of Educational Research*, *95*, 60-75. https://doi.org/10.1016/j.ijer.2019.02.015

Haddon, E. (2009). Instrumental and vocal teaching: How do music students learn to teach?. *British Journal of Music Education*, *26*(1), 57-70. https://doi.org/10.1017/S0265051708008279

Haddon, E. (2019). Perspectives of Chinese students on studying MA Music Programmes in a UK university. *ORFEU*, *4*(2), 30-58. https://doi.org/10.5965/2525530404022019030

Haddon, E. (forthcoming, 2023). Negotiating pedagogical cultures: Adaptive challenges facing Music Education graduates on their return to China. In R. Prokop & R. Reitsamer (Eds.), *Higher music education and employability in a neoliberal world*. Bloomsbury.

Hallam, S. (2012). Commentary: Instrumental music. In G. E. McPherson, and G. F. Welch (Eds.), *The Oxford handbook of music education, Volume 1* (pp. 651-657). Oxford University Press.

Jørgensen, H. (2000). Student learning in higher instrumental education: Who is responsible? *British Journal of Music Education*, *17*(1), 67-77. https://doi.org/10.1017/S0265051700000164

Kleickmann, T., Richter, D., Kunter, M., Elsner, J., Besser, M., Krauss, S., & Baumert, J. (2013). Teachers' content knowledge and pedagogical content knowledge: The role of structural differences in teacher education. *Journal of Teacher Education*, *64*(1), 90-106. https://doi.org/10.1177/0022487112460398

Lee, K. C. C., & Leung, B. W. (2022). Instrumental teaching as 'the noblest and the most under-praised job': Multiple case studies of three Hong Kong instrumental teachers. *Music Education Research*, *24*(1), 42-55. https://doi.org/10.1080/14613808.2021.2015309

Meissner, H., Timmers, R., & Pitts, S. E. (2021). 'Just notes': Young musicians' perspectives on learning expressive performance. *Research Studies in Music Education*, *43*(3), 451-464. https://doi.org/10.1177/1321103X19899171

Mills, J. (2004). Conservatoire students as instrumental teachers. *Bulletin of the Council for Research in Music Education*, 145-153. https://www.jstor.org/stable/40319248

Norton, N., Ginsborg, J., & Greasley, A. (2019). Instrumental and vocal teachers in the United Kingdom: Demographic characteristics, educational pathways, and beliefs about qualification requirements. *Music Education Research*, *21*(5), 560-581. https://doi.org/10.1080/14613808.2019.1656181

Perkins, R., Aufegger, L., & Williamon, A. (2015). Learning through teaching: Exploring what conservatoire students learn from teaching beginner older adults. *International Journal of Music Education*, *33*(1), 80-90. https://doi.org/10.1177/0255761414531544

Ployhart, R. E., & Vandenberg, R. J. (2010). Longitudinal research: The theory, design, and analysis of change. *Journal of Management*, *36*(1), 94-120. https://doi.org/10.1177%2F0149206309352110

Qin, M., & Tao, D. (2021). Understanding preservice music teachers' intention to remain in the profession: An integrated model of the theory of planned behaviour and motivation theory. *International Journal of Music Education*, *39*(4), 355-370. https://doi.org/10.1177/0255761420963149

Ritchie, J., Lewis, J., Nicholls, C. M., & Ormston, R. (Eds.). (2013). *Qualitative research practice: A guide for social science students and researchers*. Sage Publications.

Schippers, H. (2009). *Facing the music: Shaping music education from a global perspective*. Oxford University Press.

Thrasher, A. (1989). Structural continuity in Chinese Sizhu: The "Baban" model. *Asian Music*, *20*(2), 67-106. https://doi.org/10.2307/834021

Thrasher, A. (2008). *Sizhu instrumental music of south China: Ethos, theory and practice*. Brill.

Van Braak, M., de Groot, E., Veen, M., Welink, L., & Giroldi, E. (2018). Eliciting tacit knowledge: The potential of a reflective approach to video-stimulated interviewing. *Perspectives on Medical Education*, *7*(6), 386-393. https://doi.org/10.1007/s40037-018-0487-9

Yung, B. (1987). Historical interdependency of music: A case study of the Chinese seven-string zither. *Journal of the American Musicological Society*, *40*(1), 82-91. https://doi.org/10.2307/831583

Zheng, Y., & Leung, B-W. (2021a). Cultivating music students’ creativity in piano performance: A multiple-case study in China. *Music Education Research*, *23*(5), 594-608. https://doi.org/10.1080/14613808.2021.1977787

Zheng, Y., & Leung, B-W. (2021b). Perceptions of developing creativity in piano performance and pedagogy: An interview study from the Chinese perspective. *Research Studies in Music Education*. DOI: https://doi.org/10.1177/1321103X211033473

Asia-Pacific Journal for Arts Education

Special Issue

Unfold the Future of Music Education through Technology

Co-editors:
Prof. Bo Wah LEUNG
Dr. Cheung On TAM
Dr. Chi Hin LEUNG
Dr. Richard Guy WHITBREAD
The Education University of Hong Kong

http://www.ied.edu.hk/ccaproject/apjae/apjae.htm ISSN: 1683-6995

Volume 22 Number 7 December 2023

Recognition of Piano Score Difficulty Level and Application in Music Teaching

Wei Gan

Central China Normal University

musicgw@ccnu.edu.cn

Bo He

Central China Normal University

hb925822@126.com

Abstract

Scientifically and accurately grading the difficulty of music scores is an essential foundation to realizing personalized music education. At present, the classification of music score difficulty levels in piano teaching is mainly done by manual labeling, which faces the double challenges of inefficiency and tendentiousness, whereas the performance of the previous difficulty level recognition algorithm is not enough to be applied to teaching practice. In this paper, we first introduce a piano score difficulty level recognition model based on an LSTM Neural Network. We then propose a data extension method through pitch offset, which can effectively resolve the over-fitting issue that appeared in training on a small dataset. Experimental results clearly demonstrate that our model represents a significant improvement compared with existing methods (78%), as well as achieving the highest accuracy rate on the three difficulty dataset levels (88%). Finally, in a piano tutoring platform independently developed by us, we utilize this model to recommend suitable scores for adaptive practice and personalized assessment according to teaching targets and learners' mastery, both of which effectively improve learners' achievements.

Key words

difficulty recognition, classification, piano instruction, music score recommendation

Introduction

The piano is the instrument with the most learners in the world. The purpose of piano learning is to improve an individual's playing ability, which is reflected in fingering, speed and other skills indicated within the structure of the music. In the graded piano tests, assessment is based on the degree to which the examinee's performance satisfies the requirements based on the difficulty level. Music scores are likewise divided into levels of difficulty, the appropriate materials being utilized according to the abilities of the individual concerned. The aim of this near-universal teaching mode is to enable learners to gradually master various piano playing skills from easy to difficult in a hierarchical and personalized way. Stratified teaching is an effective strategy for teachers to solve the enormous differentiation of students' ability levels in practice. In order to achieve stratified teaching, however, it is not only necessary to divide learning resources according to students' basic levels, but also to specify the corresponding evaluation methods for people with different foundations, so as to "teach students according to their aptitude" and stimulate their enthusiasm for piano learning. The difficulty of the music score not only reflects learners' ability levels, but also gives an objective benchmark for the evaluation of learners. When teachers consider the difficulty of the music score itself as part of the evaluation of a student's performance, this highlights their recognition of their students' learning processes, enabling them to discover learner shortcomings and improving the effect of evaluation on learning.

Due to the lack of unified quantitative standards, the current classification of the difficulty level of music scores mainly depends on the subjective judgment of experts or professionals in the field, meaning that learners are forced to rely on existing teaching materials, resulting in a lack of diversity and autonomy in the choice of learning resources. In the process of selecting tracks for learners, teachers are inevitably influenced by their own musical preferences. Although the tracks recommended to students can determine the difficulty coefficient, it is obvious that the recommended tracks are narrow in scope and personal music preferences are obvious when they are in line with learners' musical preferences or else reflect changing musical trends. Making full use of big data and artificial intelligence technology to objectively and accurately classify and grade piano music scores, which can effectively balance teachers' or students' preferences in the difficulty coefficient of music and music categories in the teaching process, can enable learners to independently choose appropriate learning resources and contribute to personalized piano teaching.

At present, although relevant studies have made preliminary progress in the difficulty identification of piano music scores, the piano is a multi-voice instrument and the complexity of compositions, compounded by the differences in the demands of different teaching scenarios, requires a more scientific, efficient and objective intelligent analysis algorithm to optimize difficulty classification. Only by improving the accuracy and precision of "analysis in difficulty application scenarios" can we truly meet the requirements of real teaching situations. In this study, a difficulty level recognition model of piano music was established using an LSTM neural network based on the timing characteristics of the music score. At the same time, in order to solve the overfitting problem of the model for the small data set, we also propose a data optimization method based on pitch migration. The experimental results show that the recognition accuracy of our proposed model is better than those of existing methods, the recognition accuracy of the three categories of data sets reaching 88% (the existing method is 78%). Finally, in the independently developed piano teaching platform, based on the requirements of teaching objectives and the current ability levels of learners, we use this model to recommend suitable music scores for learners to practice and test, effectively improving the learning effect.

Related work

Although artificial intelligence technology is commonly used in music analysis, conversion, generation and other, similar problems, there are few studies on the automatic recognition of musical score difficulty level. Sébastien et al. (2012) regarded the recognition of musical score difficulty level as a problem of musical score classification. They took the 'MusicXML' score as the analysis object and defined seven characteristics of musical score difficulty, such as the rhythm, fingering, and hand displacement required for performance, chord, rhythm and polyphony. Since these features interact with each other in a complex way with regard to the difficulty level of the score, they first obtained the key features by principal component analysis (PCA), and then realized the difficulty classifier by using an unsupervised clustering algorithm. The score difficulty classifier, based on the quintessential features, achieved a classification accuracy of 0.66 on the data set of 50 samples.

Chiu & Chen (2012) regarded the recognition of musical score difficulty level as a regression problem. Based on statistical features such as pitch entropy, range, average pitch and performance speed were added. The ‘RreliefF’ algorithm was used to filter the feature set, and other algorithms, such as linear regression (Robnik-Šikonja & Kononenko, 2003), multilinear regression, step-up regression and support vector regression (Smola & Schölkopf, 2004), were used to construct the music difficulty prediction model on the optimal special solicitation. The results of the model evaluation experiment showed that the performance of the support vector regression model was the best, and the regression coefficient (R2) of the model was close to 0.4 on the data sets of the 4 difficulty and the 9 difficulty categories.

The study by Nakamura et al. (2014) showed that scores with more unusual fingerings were very difficult to play. Consequently, Ramoneda et al. (2022) used fingering characteristics in piano playing techniques to analyze the difficulty of the music score; by using two kinds of finger-extracting systems, namely knowledge-driven and data-driven, four kinds of finger-extracting sequences were obtained from the music score, with ‘XGBoost’ and ‘DeepGRU’ deep learning models being utilized to construct music score difficulty classifiers on different finger-extracting sequences. At the same time, in order to make the study more relevant to the educational setting, they constructed a three-difficulty level data set that utilized the musical score used in piano teaching. In this dataset, the performance of the 'DeepGRU' model was the best, the classification accuracy reaching 0.78.

Ghatas et al. (2022) believed that the features based on manual definition had limitations and were not applicable to the end-to-end deep learning model. They therefore divided the music score into equal-length fragments and used the convolutional network to extract the difficulty features, and then combined the features obtained by the neural network and the manually defined features to form the feature input of music score difficulty. Finally, a multiple perceptron classifier was used to classify the music. The highest accuracy of this method was 0.80 and the highest F1 score 0.76 on the data set of three difficulty-level categories.

The evaluation of music score difficulty itself has a high degree of subjectivity, and the factors that need to be considered in different subdivision themes are uniform, so the construction of a music score difficulty evaluation model requires a certain complexity. There are subjective deviations in the artificially defined features; at the same time, the normalization, filtering, and other numerical calculation processes will also produce errors. More importantly, some statistical characteristics (such as vocal range, rhythm, and intervals) cannot describe the dynamic changes in the time sequence of the music itself. The difficulty of analyzing musical scores from the performance process is primarily reflected in the changes in distinct nodes, such as those to unit beat, post-bar rhythm, scale, and rotation. The model built on a time series has a natural structural advantage in capturing such information.

Approach

The construction of a time series classification model usually involves two main tasks: firstly, it is necessary to find a time series representation that can fully describe the research object, specifically one that solves the problem of time dimension division and the feature representation of a single time step; secondly, the time series is used as the input to construct the classification model. In order to more comprehensively represent the changes of musical notation in time sequence and make the data extensible, this chapter proposes improvements to the existing musical notation time sequence representation method, together with designing an automatic recognition model of musical notation difficulty features based on the improved representation method.

Sequential representation of piano scores

MIDI (Musical Instrument Digital Interface) (David, 2019), as the international standard of digital music electronic synthetic instruments, is the most extensive music standard format in the field of music programming. MIDI records music through the digital control signals of notes, and uses ‘MIDI Message’ to describe the information of the music to be played, such as at the specific moment, what type of instrument will be used, which notes will start to sound, which notes will sound at the end, whether the tone of the beat changes, etc. For any keyboard instrument, including the piano, each note message contains a key number and force. The key number corresponds to the pitch, indicating the frequency at which the note is made, and ranges from 0 to 127; the force represents the weight of the “down” key, and the “stop” key is represented by a message with a force of 0. [9] In addition, MIDI files contain other metadata that are used to control playback.

Even though the “MIDI message” stream accurately captures the temporal characteristics of the score, algorithmic models cannot use it as an input directly. After being divided by time steps in related studies, it is typically encoded as a two-dimensional Piano Roll matrix (Dervakos et al., 2021). Taking the matrix D_{k*t} ($k \in \{0,1,2...127\}$; $t \in N$ is the time step) to express the piano roll, each column of the matrix corresponds to a time step (e.g., quarter note duration), the row number corresponds to the key number in MIDI, and the element $d_{ij} = 1$ of the matrix indicates that the key number i is pressed at the jth time step, while $d_{ij} = 0$ indicates that it has not been pressed.

There are two segmentation techniques for the time dimension that are currently used: fixed step and fixed duration. Fixed step represents the duration of each unit beat by the same number of steps. Fixed duration divides the duration of notes by the base time. The challenge with fixed duration lies in determining the base duration, whereas the division approach of fixed steps frequently results in excessively lengthy coding sequences that hinder the algorithm’s ability to converge. In order to achieve convergence, we created a relative reference coding, selecting 1/n of a unit beat as the time step for any score (n is typically 16, 32, or 64).

Considering that splitting by time step will truncate the notes with large durations into multiple small fragments, and in order to distinguish whether the keys with the same number and in the pressed state in adjacent time steps are the same notes in the corresponding score, a matrix S_{k*t} (k ∈ {0,1,2...127; t ∈ N is the time step length}) was defined to record this state. The element s_{ij}=1 of the matrix indicated whether the pressed state of the ith key in the j_{th} time step was a continuation of the j-1 step; s_{ij} = 0 if it was not a continuation of the pressed state. The score was finally represented as a piano roll by aligning the matrices D and S in time steps and then splicing them together as P = (DS), where P is a two-dimensional matrix of m*n, m is the double of the number of keys, and n is the overall number of time steps after splitting. M = 88*2 if the MIDI representation is a piano score (generally, a piano score has only 88 keys, corresponding to the MIDI key numbers 21 to 108). Figure 1 shows a score fragment of the Piano Roll sequence.

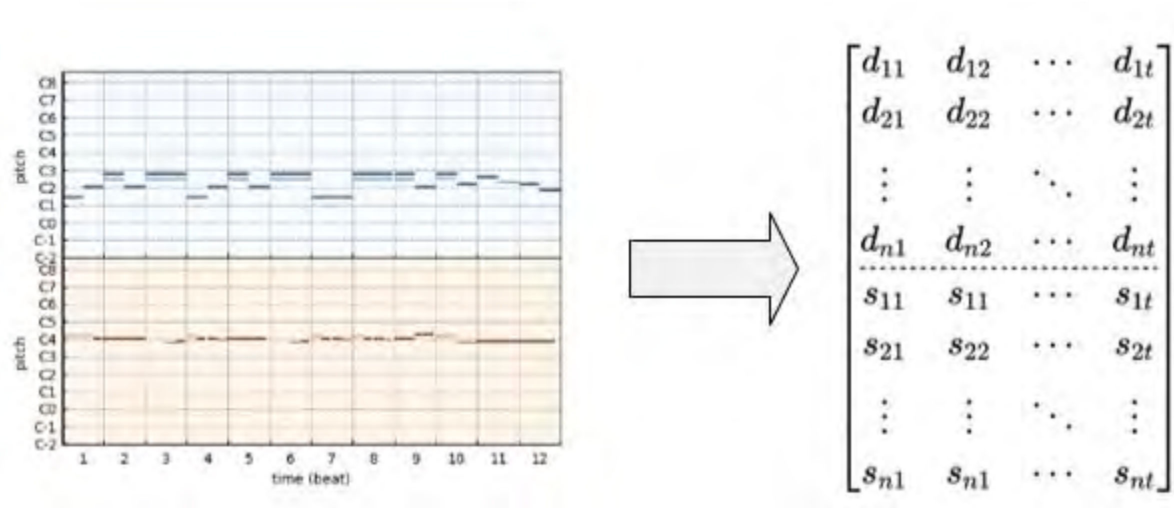


Figure 1 *Music sheet encode example by proposed method*

Classification method

An LSTM network is a recursive network designed to optimize the gradient disappearance problem of a recurrent neural network (RNN) (Pascanu et al., 2013). A logic gate module is utilized in neurons in order to replace activation function in RNN, so that it can store and transmit context information between, and realize the long-term dependence of, time series (Cheng et al., 2016). In LSTM-based sequence prediction and classification problems, LSTM output state sequence $\{h_1,h_2... h_t\}$ is often regarded as the characteristic information of the original sequence, which is seen as the difficulty related information implied in the musical score sequence in this study. Since the original LSTM only considers the sequential dependence of the time series, and the reverse characteristics of adjacent notes in the score and the difficulty of the combined sequence may also be related to the score, the experiment of this research also uses the 'Bi-LSTM' and 'Attention-LSTM' network models improved from LSTM.

The state sequence $\{h_1, h_2, \ldots h_t\}$ containing music difficulty information is obtained through the LSTM network. After h_t, the classification task can be expected to be completed only after subsequent processing. Figure 2 describes the general structure of the LSTM-based classifier, which is composed of the fully connected LSTM network layer, 'SoftMax' function and 'Argmax' function, where 'hlast' represents the output state of the last neuron in the LSTM network. In other words, as it pertains to the difficulty feature vector of the score, the full connection layer is the high-dimensional state vector 'hlast' transformed into a vector of the same dimensions as the category, where the 'SoftMax' function is used to calculate the probability of each category and the 'Argmax' function selects the maximum probability category from the probability distribution.

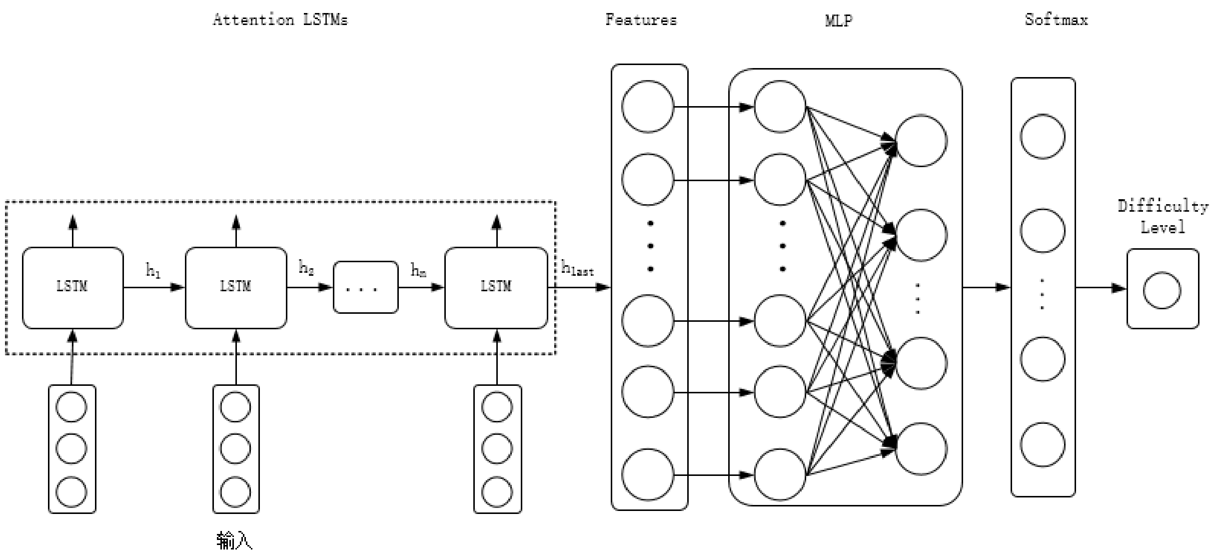


Figure 2 *Framework of proposed music sheet classification model*

Since the LSTM network in our model is used to extract the difficulty features from the score sequence, and there is no Encoder-Decoder structure, we use an attention mechanism that was initially proposed in the Encoder-Decoder structure for obtaining the contextual information of the input sequence (Sundermeyer et al., 2012) and which can be propagated in one direction (Elliot, 2016). As seen in Figure 3, when determining the output of the current time step, the output state of the previous n steps is taken into full consideration, with the attention part being calculated as shown in Equations 1-3:

$$u_i^t = v^T \tanh(W_h h_i + W_c c_t) \tag{1}$$

$$a_i^t = \mathrm{softmax}(u_i^t) \tag{2}$$

$$h_t' = \sum_{i=t-n}^{t-1} a_i^t h_i \tag{3}$$

where vectors v, W_h, and W_c are the parameters to be learned in the model, c_i is the state of the LSTM cell at the current time step, h_i is the output of the LSTM cell at the previous n time steps (h_{t-n},... ,h_{t-1}); the weights of each time step are calculated using the state of the current time step and the outputs of the previous n time steps.

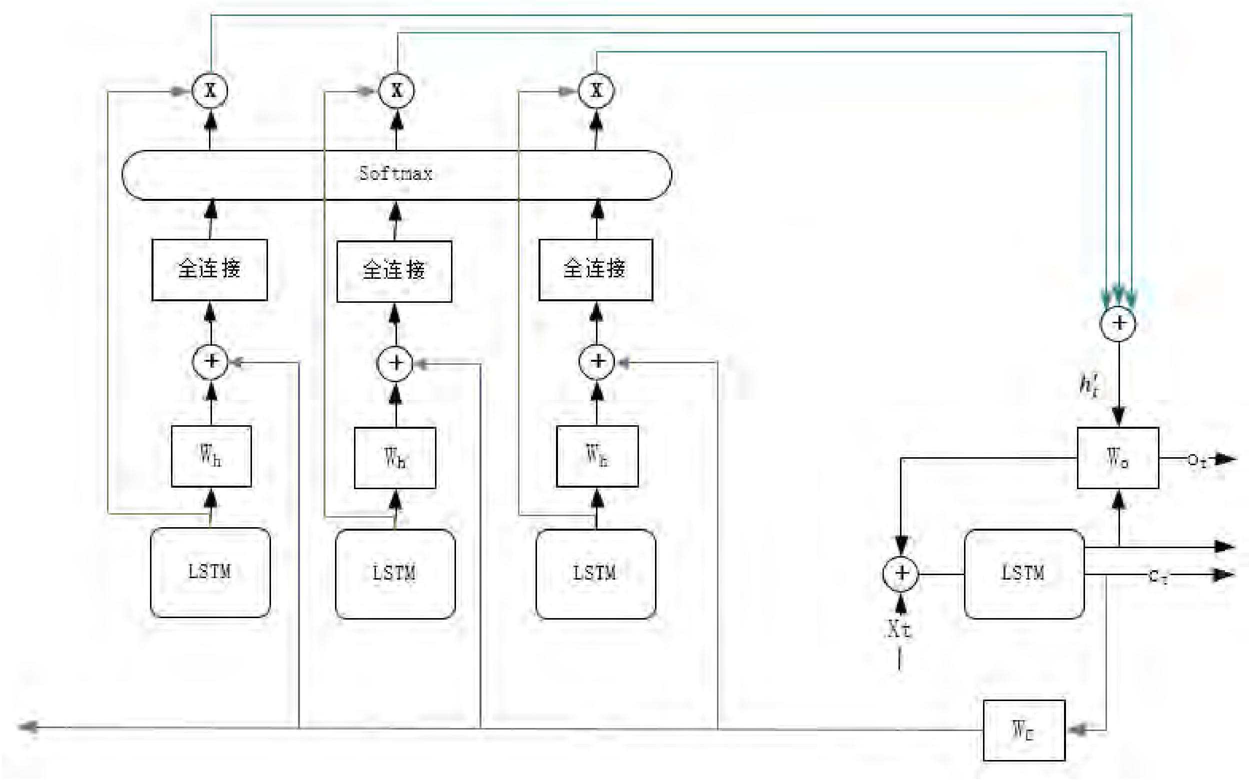


Figure 3 *The inner structure of attention LSTM cell unit*

The attention value corresponding to each time step is obtained after 'Softmax' normalization, following which the attention value is weighted and summed with the output to obtain the attention representation h_t' for the n time steps. The vector obtained by concatenating the attention h_t' with the hidden state $\boldsymbol{h}_t^*$ of the last time step is the difficulty feature value of the score sequence used in Equation 4:

$$\boldsymbol{h}_t = [\boldsymbol{h}_t', \boldsymbol{h}_t^*] \tag{4}$$

In the experimental section we will use the basic LSTM and 'Bi-LSTM' models as benchmarks to verify the performance of the attention mechanism in the score difficulty classification problem.

Experiment

Dataset

We selected two sheet music datasets, '80notes' (8notes.com, 2020) and 'Mikrokosmos' (Pedro, 2021/2022), which have been used in previous literature as experimental data. The '80notes' dataset contains four difficulty levels (Beginner, Easy, Intermediate, Advanced) and 'Mikrokosmos' contains three difficulty levels (Beginner, Moderate, Professional). Due to the large length of the segmented scores in the Advanced category of '80notes' and the performance problems of the LSTM network in dealing with very long sequences, only the first three levels are used in this study.

Preprocess

The number of scores and the average length of each difficulty level in the two datasets are shown in Table 1, where the average length of scores is counted in the number of sixteenth notes.

Table 1 *Dataset statistical analysis*

Dataset Name	Difficulty Level	Count	Average steps($1/16_{th}$ note per step)
'80notes'	Beginner	91	245.54
	Easy	624	449.7

Dataset Name	Difficulty Level	Count	Average steps(1/16th note per step)
	Intermediate	1144	773.97
'Mikrokosmos'	Beginner	52	299.12
	Moderate	53	399.62
	Professional	30	731.73

Data enhancement.

Data enhancement is a data preprocessing technique used in deep learning to improve the performance of models on small data sets and is commonly used in research related to computer vision, such as flipping, rotating, and the offsetting of images. From the features selected in the literature (Chiu & Chen, 2012), it can be seen that the difficulty of a musical score is related to features such as the length of the note itself and the change in the pitch of adjacent notes. Therefore, expansion of the sample size can be achieved by shifting the pitches in the whole score up and down in order to form a new score with the same difficulty level. Figure 4 shows a score fragment before and after pitch shifting, whereby throughout the process the duration of all the notes remains the same.

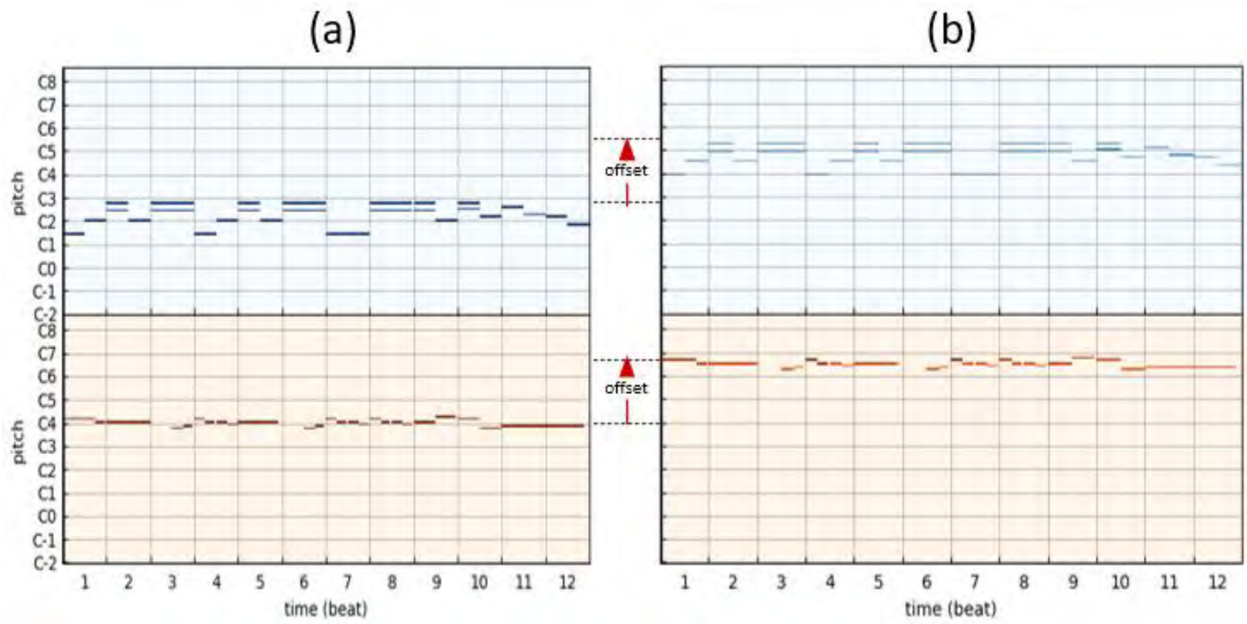


Figure 4 *Pitch offset example of one fragment, where (a) is the original score and (b) is the offset one*

According to this method, the dataset was expanded with different offset units and the pitches were guaranteed to be between 22 and 109, so the maximum offset unit was set as a threshold with the lowest note of the offset score not lower than 22 and the highest note not exceeding 109. After data selection and expansion, the number of scores in the final dataset was 701 for each level in ‘note80’, totaling 2804 scores; there were 292 scores for each difficulty level in the ‘Mikrokosmos’ dataset, totaling 876 scores. In each set of experiments, the original data was used as the test set and the extended data as the training set.

Experimental setup

The number of LSTM, ‘bi_LSTM’ and ‘att_LSTM’ networks in the experiment was 2, and the number of hidden state units in each layer 128 and 64, respectively; the attention window size was 64 in 'att_LSTM'; to prevent model overfitting, each layer used the 'DropOut' mechanism with a pass rate of 0.5, and added L2 regularization with a weight of 0.0001 for all trainable parameters to the model loss. The batch gradient descent learning method was used in the training process, using the 'AdamOptimizer' gradient optimizer with a learning rate of 0.001, a batch size of 256 (i.e., 256 samples were randomly selected for each iteration), and a gradient cropping threshold of 5.

Since the difficulty level of the score is a multi-category problem, we used four metrics, Accuracy, Recall, Precision, and F-value (F1), for comparative analysis (Equation 5~6). Among them, Recall, Accuracy, and F-value, calculated separately by different categories, were used to find the average value. For each difficulty level, TP was the number of samples correctly identified as level I; TN was the number of other difficulty levels correctly identified; FP was the number of difficulty level i identified as other difficulty levels, and FN was the number of other levels identified as level i (i = the difficulty level):

$$\text{Accuracy} = \frac{\text{Correct Predictions}}{\text{Total Number of Samples}} \tag{5}$$

$$\text{Recall} = \sum_{i=1}^{n} \frac{\text{TP}_i}{\text{TP}_i + \text{FN}_i} (\text{n is the total difficult level number}) \tag{6}$$

$$\text{Precision} = \sum_{i=1}^{n} \frac{\text{TP}_i}{\text{TP}_i + \text{FP}_i} \quad (\text{n is the total difficult level number}) \tag{7}$$

$$F_1 = \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{8}$$

The experimental procedure consisted of two sets of independent experiments for model validation and data enhancement method validation. In the model comparison experiments, the recognition accuracy of different models was compared with the enhanced 'Mikrokosmos' dataset; in the data enhancement validation experiments, the 'att_LSTM' model was selected to test the effectiveness of the proposed data enhancement method by comparing its effect on the extended dataset and the original dataset.

Results and Discussion

The results of the model validation are shown in Table 2.

Table 2 *Enhanced Dataset classification results*

Model Name	Accuracy	Precision	Recall	F1
LSTM	0.85	0.85	0.85	0.85
'att_LSTM'	0.90	0.88	0.89	0.88
'bi_LSTM'	0.82	0.82	0.82	0.82

Precision, recall, accuracy, and F-value of the 'att_LSTM' model were the best among the three models (0.90, 0.88, 0.89, and 0.88, respectively), that of precision exceeding other methods in the literature on the same data set and indicating the potential of 'att_LSTM' for extracting the difficulty features of musical scores.

In order to compare the differentiation of the model on varying difficulty levels, we selected the recognition of ‘att_LSTM’ model in different categories, as shown in Tables 3 and 4.

Table 3 *Confusion matrix of LSTM model on ‘Mikrokosmos’ dataset*

	Beginner	Moderate	Professional
Beginner	0.85	0.15	0
Moderate	0.18	0.61	0.21
Professional	0	0.36	0.67

Table 4 *Confusion matrix of an ‘att_LSTM’ model on ‘80notes’ dataset*

	Beginner	Easy	Intermediate
Beginner	0.86	0.14	0
Easy	0.16	0.64	0.20
Intermediate	0	0.34	0.66

The results show that there is a higher probability of recognition error for adjacent difficulty levels, in which 0.15 of difficulty level 1 is incorrectly recognized as difficulty 2; 0.57 of difficulty level 3 is incorrectly recognized as difficulty 2. Also, since difficulty level 2 is an intermediate difficulty level, there is a possibility that the samples are incorrectly identified as difficulty level 1 and difficulty level 3, and that the incorrect identification rates are 0.18 and 0.21, respectively. At the same time, ‘att_LSTM’ model’s ability to distinguish adjacent difficulty levels still needs to be improved.

Considering that the enhanced dataset was used in the model validation experiment in order to verify the effectiveness of the proposed data enhancement method, we chose the 'att_LSTM' model with the highest recognition accuracy; these evaluation indexes on the original dataset and the enhanced dataset are shown in Table 5.

Table 5 *Classification results of an 'att_LSTM' model on original and enhanced dataset*

Dataset Name	Type	Accuracy	Precision	Recall	F1
'Mikrokosmos'	original	0.63	0.65	0.67	0.67
	enhanced	0.83	0.82	0.84	0.82
'80notes'	original	0.65	0.62	0.67	0.67
	enhanced	0.79	0.81	0.80	0.77

Recognition of the 'att_LSTM' model's accuracy on the extended dataset was 0.90, which is much higher than that on the original dataset. Therefore, it can be concluded that the data enhancement method proposed in this paper can effectively improve the accuracy of the difficulty recognition model of the score, as well as confirming that the change of minimum and maximum pitches in the score does not affect the difficulty level of the score itself within the same pitch range.

The data in the tables shows that although the difficulty features extracted by the LSTM network lack interpretability compared with the manually defined ones, the resulting difficulty features have good representability and can more accurately distinguish between scores of different difficulty levels. Our proposed data augmentation method effectively solves the problem of small data sets within the larger issue of sheet music difficulty recognition, and provides a reference for other deep learning models that can be applied to such problems.

Educational applications

When learning the piano, the learning space is constrained by the learning material's rigidity content, which also erodes students' confidence and interest when they repeatedly play the same compositions with little variation. This study integrates musical difficulty evaluation into the teaching process and uses a suggestion system based on musical difficulty level to prevent thems from becoming frustrated or worn out. Therefore, when students begin to feel discouraged or exhausted, it is advised that they practice a score with a comparable level of difficulty.

The final output h_{last} of the LSTM network in the sheet music difficulty level recognition model can be considered as the feature value of the sheet music difficulty level. When retrieving teaching resources, making individualized recommendations, and conducting personalized assessments, the similarity of characteristics can be utilized to differentiate between the varying degrees of difficulty of different pieces of sheet music. Figure 5 displays the recommendation process based on the difficulty similarity of the score.

Algorithm 1 Recommendation based on difficulty similarity

Input: Last practised sheet music $\mathbf{P}_{pre}$,

Sheet music collection $\mathbf{P}\{p_1, p_2, p_3 \cdots, p_n\}$.

Output: The score closest to $\mathbf{P}_{pre}$ in difficulty $\mathbf{P}_{next}$.

Random choose a candidate set $\mathbf{P}_{candi}\{p_1, p_2, p_3 \cdots, p_m\}$ from $\mathbf{P}$;

Extract difficulty features of $\mathbf{P}_{pre}$,$\mathbf{P}_{candi} \to h_{pre}; h_1, h_2, h_3, \cdots, h_m$

Caculate similarity$\mathbf{D}_{candi} = \{d_1, d_2, d_3 \cdots, d_m\}$,where $d_i = |h_{pre} - h_i|_2^2$ $(i = 1, 2, 3 \cdots m)$;

Index of the best item $j = argmin(\mathbf{D}_{candi})$.

return $\mathbf{P}_{candi}[j]\mathbf{P}_{candi}[j]$

Figure 5 *Algorithm flow of score recommendation based on difficulty similarity*

Since there are only a finite number of scores in the resource library, the actual application recommends multiple scores of varying degrees of difficulty as candidate sets for learners to choose from on their own after setting a threshold in order to avoid making duplicate recommendations. As seen in Figure 6, when students practice a piece in the textbook using our self-created piano teaching APP, multiple scores of comparable difficulty are suggested for them so that they can practice on various scores depending on their learning circumstances.

Figure 6 *Recommended algorithm example*

Conclusions and Future Work

This study examined the utilization of a classifier based on an LSTM neural network in order to categorize the degree of difficulty of musical scores and validated it using real data. The experimental findings indicate that given enough experimental data, the classifier proposed in this paper can effectively extract difficulty-related information from score sequences, and its classification accuracy of 88% is superior to that of the best study to date (78.67%). Furthermore, the comparison of various coding techniques in terms of classification effect establishes from one side that the level of difficulty of a score should be assessed by factoring both the complexity of the performance and the score structure, both of which are essential requirements for students' score recognition and performance ability in piano learning. Finally, we explored how score difficulty could be applied practically in piano instruction and designed and implemented a system for recommending valuable functions based on the score difficulty recognition model.

We will continue to refine the classification model for sheet music and increase the accuracy with which difficulty level is recognized in the follow-up study. Further research will be done on the application of sheet music difficulty in individualized learning and instruction sessions, such as adaptive evaluation and automated testing based on our piano teaching platform. This will enhance the platform's individualized educational experience and help the system become more intelligent and personalized.

References

8notes.com. (2020). Retrieved December 2, 2022, from Free sheet music on 8notes.com website: https://www.8notes.com/

Cheng, J., Dong, L., & Lapata, M. (2016, September 20). *Long Short-Term Memory-Networks for Machine Reading*. arXiv. Retrieved from http://arxiv.org/abs/1601.06733

Chiu, S.-C., & Chen, M.-S. (2012). A Study on Difficulty Level Recognition of Piano Sheet Music. *2012 IEEE International Symposium on Multimedia*, 17-23. Irvine, CA, USA: IEEE. https://doi.org/10.1109/ISM.2012.11

David, B. (2019). Standard MIDI file format, updated. Retrieved December 2, 2022, from http://www.music.mcgill.ca/~ich/classes/mumt306/StandardMIDIfileformat.html

Dervakos, E., Kotsani, N., & Stamou, G. (2021). Genre Recognition from Symbolic Music with CNNs. *Artificial Intelligence in Music, Sound, Art and Design*, 98-114. Springer, Cham. https://doi.org/10.1007/978-3-030-72914-1_7

Elliot, W. (2016, July 15). Generating Long-Term Structure in Songs and Stories. Retrieved December 2, 2022, from Magenta website: https://magenta.tensorflow.org/2016/07/15/lookback-rnn-attention-rnn

Ghatas, Y., Fayek, M., & Hadhoud, M. (2022). A hybrid deep learning approach for musical difficulty estimation of piano symbolic music. *Alexandria Engineering Journal*, *61*(12), 10183-10196. https://doi.org/10.1016/j.aej.2022.03.060

Nakamura, E., Ono, N., & Sagayama, S. (2014). *MERGED-OUTPUT HMM FOR PIANO FINGERING OF BOTH HANDS*. 6.

Pascanu, R., Mikolov, T., & Bengio, Y. (2013). On the difficulty of training recurrent neural networks. In S. Dasgupta & D. McAllester (Eds.), *Proceedings of the 30th International Conference on Machine Learning* (Vol. 28, pp. 1310-1318). Proceedings of Machine Learning Research: PMLR. Retrieved from https://proceedings.mlr.press/v28/pascanu13.html

Pedro. (2022). *Mikrokosmos-difficulty dataset*. Retrieved from https://github.com/PRamoneda/Mikrokosmos-difficulty (Original work published 2021)

Ramoneda, P., Tamer, N. C., Eremenko, V., Serra, X., & Miron, M. (2022). Score Difficulty Analysis for Piano Performance Education based on Fingering. *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 201–205. https://doi.org/10.1109/ICASSP43922.2022.9747223

Robnik-Šikonja, M., & Kononenko, I. (2003). Theoretical and Empirical Analysis of ReliefF and RReliefF. *Machine Learning*, *53*(1), 23-69. https://doi.org/10.1023/A:1025667309714

Sébastien, V., Ralambondrainy, H., Sébastien, O., & Conruyt, N. (2012). *SCORE ANALYZER: AUTOMATICALLY DETERMINING SCORES DIFFICULTY LEVEL FOR INSTRUMENTAL E-LEARNING*. 6.

Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, *14*(3), 199-222. https://doi.org/10.1023/B:STCO.0000035301.49549.88

Sundermeyer, M., Schlüter, R., & Ney, H. (2012). LSTM neural networks for language modeling. *Thirteenth Annual Conference of the International Speech Communication Association*.

Asia-Pacific Journal for Arts Education

Special Issue

Unfold the Future of Music Education through Technology

Co-editors:
Prof. Bo Wah LEUNG
Dr. Cheung On TAM
Dr. Chi Hin LEUNG
Dr. Richard Guy WHITBREAD
The Education University of Hong Kong

http://www.ied.edu.hk/ccaproject/apjae/apjae.htm ISSN: 1683-6995

Volume 22 Number 8 December 2023

A preliminary study of an online training mode for choirs in colleges and universities during an epidemic: an exemplar from the Choir of Capital Normal University

Xia Qing

Capital Normal University, Beijing

qingxia5411@163.com

Abstract

The developmental history of music education in Chinese colleges and universities over the course of the past 150 years can be said to have been characterized by its unevenness. In recent years, through the introduction of a number of policies, the state has attached greater importance to the cultivation of aesthetic education, a move that has yielded, and continues to yield, positive results within school music tuition. Nevertheless, 2020 witnessed the global COVID-19 pandemic, which resulted in severe and sustained alterations to people's lives and the transference of education to online teaching. This situation made it necessary to explore new modes of instruction for the training of a vocal ensemble, which had previously almost always taken the form of on-site rehearsals. In 2022, the Choir of Capital Normal University in Beijing, China, successfully explored an online training mode, in the process achieving positive teaching results. This paper discusses the action research and investigation research involved in the hope of providing a body of reference for the healthy development of other art ensembles in similar future situations.

Key words

online training mode, music education in colleges and universities, choir, epidemic

Background

Music education in Chinese colleges and universities is little more than 150 years old. With the founding of New China, especially since the reform and opening up period, the state has attached greater importance to the development of school music education through the introduction of a number of aesthetic education policies. This has enabled schools of all levels and types to achieve progress in music education, curriculum planning, community activities, scientific research results, and teacher quality, achieving remarkable results in the process.

In 2020, the outbreak of the COVID-19 pandemic worldwide severely hindered economic and social development, along with people's daily lives. Within the field of school music education, the lack of face-to-face contact, whether it be rehearsals, performances or lessons, had serious repercussions in terms of adversely affecting students' participation in art activities; some even went so far as to quit their ensembles. The outcome was that in order to maintain enthusiasm and commitment to music education, many schools had to make full and creative use of the Internet.

The Choir of Capital Normal University in Beijing, China, is composed of teachers and students who are naturally drawn to the art of choral singing. Full of youthful vitality, the chorus members not only possess musical ability, but also demonstrate a high degree of unity and cooperation. During the epidemic, the conductor, the voice training teacher, and the members jointly explored an online training mode in order to maintain the ongoing development of the choir, in the process achieving measurable learning results.

With an eye to sharing the findings, this paper summarizes the online training experiences of the school choir in the spring semester of 2021-2022. It is hoped that this preliminary model can provide a reference for the survival and development of other art ensembles during similar times of duress.

Development of the choir

Founded in 2015, the Choir of Capital Normal University in Beijing is an off-shoot of the Youth League Committee of the University. Under the direction of artistic director and permanent conductor Professor Xiao-Yong Shao, the members consist of non-music major undergraduates from different departments. In addition to winning the gold medal in the choral performance category of Beijing University's Students' Art Festival on numerous occasions and frequently holding special concerts, the artistic director / conductor was invited to cooperate with the National Symphony Orchestra in December 2015 in order to successfully premiere Wang Xilin's *Requiem for the War of Resistance Against Japanese Aggression* at the National Center for the Performing Arts. In July 2019, the choir won first prize in the National Adult mixed Chorus category at the sixth National Nie Er Music (Chorus) Week Chorus Performance. The ensemble also participated as part of a large-scale chorus performance in Tiananmen Square in 2019 to mark the 70th anniversary of the founding of the People's Republic of China and the 100th anniversary of the founding of the Communist Party in 2021.

Before the epidemic, the choir used a traditional training mode whereby each of the four vocal parts trained separately before working together. Collectively forming the teaching assistant team, each vocal section had its own teacher (the author was in charge of the mezzo-sopranos, while each of the other three parts were taught by Professor Shao's graduate students). Each vocal section trained for three hours on Sunday afternoons; this was followed, on Sunday evenings, by a two-hour full rehearsal of the choir by Professor Shao.

As the epidemic grew more severe, the choir switched to online training starting in November 2021. In the following section, the author introduces and explains the exploration process of the online training model.

Online training mode

Due to the limitations of the network technology available, the author only successfully trained one student during the initial online period (November 2021). In the spring semester of 2022, due to the ongoing epidemic, the online training method continued to be adopted. The resultant online singing group often sounded chaotic, and whilst one-on-one training proved workable, efficiency was low. Following a discussion, it was decided to divide the members of each vocal section into several groups, with about five members per group. One advanced-level student was selected as the leader, with responsibility for the training of each group. Between March and May 2022, the choir went through three stages of self-exploration: preliminary, development, and maturity, followed by the production of a summary document outlining a set of scientific and practical precepts for use as a model for online teaching and development.

1. Preliminary stage (March 6, 13; two training sessions)

(a) Main modes

In the preliminary stage, the members submitted the recording homework, after which the group leaders made comments to each of the groups. In the first instance, the training teacher assigned several works to the group leaders and members. After self-learning, the group leaders submitted singing recordings to the training teacher, who listened to them and gave feedback in the form of suggestions. The group members submitted their recordings to the group leaders, who similarly listened before giving suggestions according to the requirements of the training teacher.

Given that there were 11 choral works to be studied in the space of two weeks, the group leaders needed to ensure that a balance was struck between the amount of homework that was allocated and the quality of responses received in return. The training teacher selected and sent everyone high quality reference models of the chosen works, additionally collecting, sorting and sharing the lyrics (together with their translations), and providing the participants with background information of difficult works such as Orff's *O Fortuna*.

(b) Teacher training

In order to improve efficiency, Professor Shao suggested that the focus should be on teacher training by enabling the group leaders to become 'qualified' music teachers. The logic behind this was that it would result in a higher degree of accuracy in terms of the comments given, in the process serving to convince the members of the leader's efficacy. The training teacher taught the group leaders simple music theory and effective teaching methods in a bid to gradually improve the quality of the group leaders as music teachers. Their comments were to be targeted, specific and in-depth, as well as being explanatory in terms of the reasons provided in relation to the correct performance of the work. Far from being afraid of offending team members, it was stressed that the group leaders should make bold and accurate suggestions for improvement in a non-accusatory way.

(c) Establish a strict evaluation system

Although the choir takes part in competitions every year, there is a limit to the number of participants that can be accommodated at any one time. Given that sopranos and mezzo-sopranos are the voices in greatest demand, it was decided that in order to promote fair competition and enable participants to succeed based on their own efforts, the establishment of a strict evaluation system was necessary. This would require the group leaders to provide scores or language evaluations as a basis for selection based on certain criteria, such as singing quality, attendance rate, and homework completion.

(d) Problem summary and new findings

After listening to the recordings, the group leaders collected and discussed the group members' questions and attempted to find solutions to the singing problems being encountered, as shown in Table 1. These suggestions were then relayed by the group leaders to the members.

Table 1 *Group members' identified singing problems and proposed solutions*

Identified problems	Proposed solutions
1. Some group leaders are not good at singing and do not give detailed comments to group members.	Conduct professional training for the group leaders, improve the level of the group leaders, and require the group leaders to give more strict and detailed comments to the group members.
2. Some group members cannot submit the recording homework on time.	The group leaders should be required to pay attention to the daily performance of the group members as a basis for evaluation.

3. Some group members had more problems with their recording homework than others.	Group leaders should give suggestions for improvement and ask the group members to submit their recordings again.
4. Most of the singing problems are due to a lack of music theory knowledge.	Share theoretical knowledge of music expression, the accuracy of music style, the use of breathing, beating time, reading music, etc.
5. There is a lack of aesthetic feeling in singing. First, the members appreciate the quality works less; second, the singing standard stays at the basic stage of intonation and rhythm.	Members are advised to listen to the singing of first-class choirs in order to improve their appreciation levels. After completion of the study of intonation and rhythm, the requirements of the pursuit of musical beauty should be put forward with a view to improving the beauty of singing through online training.
6. Singing generally lacks emotional expression.	Training members to fully understand the expression marks, while recitation should be used to promote singing emotions and the bold expression of emotions.

At the same time, it was discovered that individuals brought their own unique skills, such as comprehensive musical accomplishments, outstanding singing ability, or strong teaching and management styles, to the table, all of which could be drawn upon as supportive forces for training purposes.

2. Development stage (March 20, 27; two training sessions)

During the development stage, the members submitted their recording homework and then trained in groups, each group member being instructed one-on-one by the training teacher.

Over the course of the previous two weeks, the members had learned 11 pieces by themselves. However, the lack of an established training process meant that singing problems could not be fully solved. It was therefore decided to add some online group training to the process, whereby the team members could choose the problematic passages to sing. Three group sessions took place on the first week (20 March) and two on the second week (27 March), with each group being allocated 40-45 minutes. The details are shown in Table 2 below.

Table 2 *Results of the online group training sessions*

Identified problems	Proposed solutions
1. Most group members sing at a low volume.	Group members are required to warm up their voices before rehearsing and breathe before singing.

2. Most of the singing problems are due to a lack of understanding regarding what constitutes a high standard, and they do not know how to solve the problems.	Strengthen the popularization of vocal music theory, enhance group members' cognition of the theoretical issues such as the relationship between music style, speed and singing, and give corresponding sound training.
3. Too much attention to intonation and rhythm, singing standards are not high, and the music lacks beauty.	Strengthen emotional expression training, require group members to sing on the basis of good intonation and rhythm, encourage a bold performance of the music, and pursue beautiful singing.

One of the advantages of online one-on-one training, compared to group singing training, was that an individual's specific problems could be addressed in a targeted way. It was found that many of the members' problems were cognitive in nature. For example, one of the students thought that their singing was flat because of an inability to modulate the sound properly, whereas it was actually a fault of her singing technique. After advice-based adjustment, the right approach was found and there was a marked improvement in intonation after the two weeks of online training. On the negative side, the limited amount of training time available meant that each person could only sing for about ten minutes. In order to improve efficiency and try and solve everyone's individual problems, it was decided that the group leader should train his / her group simultaneously, while the author attended each group session for spot check training.

3. Mature stage (April 10 - May 29; six training sessions)

During the mature stage, in addition to the spot check training that was carried out by the author while delivery of the group instruction took place, some high-quality learning materials were selected in April in order to help the members study by themselves.

Having previously asked a music major student to collate the playing and singing recordings of the new works as the exemplar material, the author firstly recorded an analysis and explanation of the difficult works for the students in order to better help them learn by themselves. For example, each of the paragraphs in 'Give Everything To The Party' encompass three key changes and contrasts in the paragraph styles, while the intonation in the middle of 'Bless The Motherland' needs careful practice. The group leader then listened to the recorded homework before the third, group training stage, the aim being to ensure that each member received sufficient training and significantly improve learning efficiency. Finally, after weekly training, organized internal meetings and discussions of each voice part took place; the training teachers and group leaders also all met for discussions following the monthly training sessions.

In May, consolidation of the difficult works began, the members learned to sing a new song, and all the works studied in the semester were reviewed. During the course of the last training session, a spot check on each member was carried out in order to ascertain the effect of the learning model.

Consequently, after three months' exploratory work, a set of relatively perfect online training models had been created and a group of "little music teachers" with enhanced musical abilities duly cultivated. With the help of the (online) learning materials, the members were able to study new works independently after class, seeking out the author only when difficult problems arose.

Three summary meetings

The exploration process of online training mode during the semester in question cannot be separated from the guidance of Professor Shao, whose valuable suggestions ensured the direction and quality of the online training.

After the rehearsals in March, the four voice training teachers reported their progress to Professor Shao, who responded with the following suggestions:

- The training teachers should strengthen the promotion of students with weak majors. In doing so, care should be taken to ensure that the students gained something while simultaneously ensuring that their interest in music did not diminish.
- Strengthen the professional training of the group leaders.
- Keep the momentum going by regularly taking part in competitions.
- Pay attention during the rehearsals to the identification of potential members who could serve as the next choir leaders. (Leadership elections take place every year.)

The author organized two meetings of the group leaders on April 4th and 24th in order to convey the suggestions of Professor Shao, to exchange the experiences of the online training sessions among the various voice parts, and to provide answers to the questions raised by the choir leaders, as shown in Table 3.

Table 3 *Author's replies to the questions raised by the group leaders*

Questions raised by the group leaders	**Author's replies**
1. The recording homework of some group members was completed by splicing, and the singing was not complete, so one-to-one training was necessary.	Strengthen the professional evaluation of the recording work, and insist on high standards and requirements; at present, a combination of recording and online rehearsals is the best way.

2. A special time could be set aside every week to popularize music theory knowledge.	The popularization of music theory knowledge can be added in the training process, but theoretical learning cannot be isolated, nor can it replace training, which is essential.
3. A group leader is responsible for 4 group members, which can fully realize one-to-one training. If there are too many group members, it may affect the quality of training.	Each group leader should improve his / her own levels and strictly require the same of the members. If online time is limited, the individual concerned can make an appointment in private to communicate with each other instead of making comments.
4. Sometimes the group leader’s comments are not professional or deep enough, or for fear of offending people, dare not speak boldly.	In order to guide group members as effectively as possible, teacher training for group leaders should be strengthened with the aim of improving their level of music theory, singing and teaching, and to teach them how best to use their professional knowledge and appropriate language.
5. If training is always carried out online without the experience of singing with the other members, it may affect the enthusiasm of the choir.	Members should be actively guided to understand the advantages of online rehearsals, solve the specific problems of each member in a one-to-one

	way, try to eliminate the negative emotions of the members, and give clear learning objectives and enough guidance so that the members can really improve.
6. The learning effect of submitting recorded works is not as good as that of online training.	Since recording and online training have their own advantages, it is suggested to combine them. Each voice situation is different, so members should be allowed to decide for themselves what works best.

On the evening of May 8th, after the rehearsal, the third meeting was held online and Professor Shao, along with 21 important members of the choir, were invited to attend. Among other things, a series of specific requirements relating to the next online training session were put forward.

Over the course of the three sessions, details of experiences were actively exchanged, problems and countermeasures proposed and summarized, and specific requirements tabled, all of which were intended to improve the efficiency and quality of the online training sessions.

4. The effects of online training

Once the online training for the semester had ended, a questionnaire was conducted on the effects of the training model among the participants (N = 50). Twelve questions, as shown in Table 4, were based on a 5-point Likert scale and related to the improvement of singing ability.

Table 4 *Questionnaire on the effects of the training model*

Investigation on the effect of online choral training

in the spring semester of the 2021-2022 academic year

(Four voice parts of a female high school, female middle school, a male high school and a male middle school were investigated, respectively.)

Please give your own answer to the effect and feeling of the online training of the choir in the spring semester of 2021-2022 academic year. (The degree is represented by numbers: 1 is low; 5 is high)

1. I have a more solid grasp of basic music knowledge: 1, 2, 3, 4, 5

2. The music feels more delicate and smooth: 1, 2, 3, 4, 5

3. I have a more accurate grasp of the relationship between melody and lyrics: 1, 2, 3, 4, 5

4. I have a deeper understanding of the relationship between solo and choral singing: 1, 2, 3, 4, 5

5. My self-learning ability in terms of music has been significantly enhanced: 1, 2, 3, 4, 5

6. I am more sensitive to the quality of music: 1, 2, 3, 4, 5

7. I am capable of leading group training tasks: 1, 2, 3, 4, 5

8. I prefer singing Chinese choral works: 1, 2, 3, 4, 5

9. My singing skills have improved: 1, 2, 3, 4, 5

10. Recording the homework can improve my singing ability: 1, 2, 3, 4, 5

11. Online training can improve my singing ability: 1, 2, 3, 4, 5

12. I am satisfied with the effect of online choral training: 1, 2, 3, 4, 5

13. Other suggestions or feelings:

The results, which were presented to the members as percentages, were as follows:

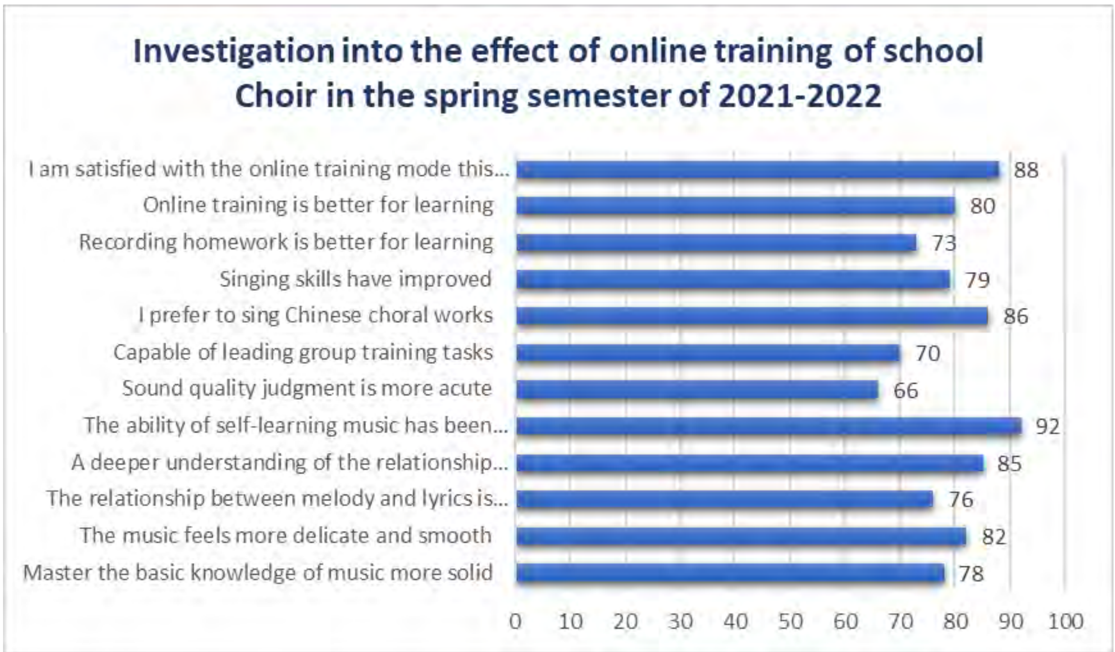


Figure 1 *The effects of the online choral training model in the 2021-22 spring semester*

Analysis of the results of the questionnaire survey reveals that the members generally recognized the positive effects of online training, and that their musical abilities had improved in all aspects. At the same time, it emerged that the more actively the students participated in the training, the higher the recognition of the benefits of the process, and vice versa. Among the four voice parts of the choir, the mezzo-sopranos, of whom the author was in charge, demonstrated a relatively high attendance rate, learning effect and recognition of online training.

5. Results of online training

Teaching practice results

(a) An efficient online training mode

After three months of online training, the choir undertook the preliminary, development and mature stages of exploration, discovering in the process that working with recordings had the ability to accurately detect problems. Online training could then be targeted to solve those issues as part of an efficient choral tuition model.

(b) Music learning harvest

Nineteen pieces of music were studied during the course of the semester, including seventeen Chinese pieces and two foreign pieces. The members' intonation, rhythm, musical feeling, musical style, and singing skills significantly improved as a result, with a number of singing-related problems being solved online.

(c) Good communication helps music teaching

Through timely communication with the group leader, the author developed a better understanding of the mentality and ability of the members and was able to undertake a reasonable analysis of the learning situation. This increased the accuracy of the teaching plans and methodologies, leading in turn to greater teaching efficiency.

(d) Popularize musical knowledge and expertise

Consistent adherence to high standards is the successful key to any arts education program. (Gardner, 2012) Sometimes, however, over-specialized requirements are not suitable for non-music majors. Combining the suggestions of teachers and students, the author summed up the most suitable teaching methods for them as: passing on professional knowledge of music in the most concise and popular language within a limited time, along with accurately revealing the laws which relate to music as an art, and; combining the training of vocal music training with an explanation of music theory, so that members can understand the relationship between theory and practice and improve in both aspects.

(e) Cultivate an independent learning ability

The survey revealed that after three months' study, the members had developed the ability to independently learn new works, make full use of the learning materials provided, and significantly improve their ability to read music, analyze and process music, and judge the quality of music.

Management and teacher training

(a) Establish a strict evaluation method

Based on scores and written records, a comprehensive evaluation was made of the singing ability and attendance of each member involved in the online training. Process and evaluation, based on the test results of the stage concerned, were combined to form the basis of selecting personnel to participate in the choral competition. However, this evaluation mechanism was not fully implemented, meaning that a high-quality evaluation mechanism needs to be explored further.

(b) Train groups of qualified "music teachers"

As the individuals responsible for undertaking the training tasks for each group, the aim of the group leaders' work was to improve the level of music theory and teaching, a process collectively termed "teaching and learning". Different members took turns as group leaders, so that their music teaching abilities could be exercised. By the end of the semester, a large number of members had duly "qualified" for the task of leading group rehearsals.

Theoretical research

(a) Clarify the relationship between solo and choral singing

In order that the members could better understand the relationship between the parts and the whole, in terms of the value of individuals to the sound of the overall choir, online, one-on-one training was delivered to the same high standard as it would have been for solo training. At the same time, it was hoped that this approach would better prepare them for completing the choral task.

(b) Clarify the relationship between learning music theory and singing practice

Many of the problems to do with members' singing skills were caused by a lack of understanding of the relevant theoretical issues. It was felt that timely explanations, aimed at resolving these misunderstandings, could help guide the students towards a better unification of theory and practice.

6. The rationale behind the online training mode

Improve the musical instruction of the teaching assistant team

Although the survey results showed that the teaching had been quite effective overall, a number of shortcomings on the part of the teaching assistant team were nevertheless identified. For example, the standard of members had a tendency to stay at the basic level, and an inordinate amount of energy was spent on improving the accuracy of intonation and rhythm, all of which indicated that their reserves of musical knowledge were limited, that their musical skills were insufficient, and that their pursuit or understanding of music learning was still relatively rudimentary. Regardless of the identity of the individual directing the proceedings (offline rehearsals, consisting of mixed choral training under the guidance of Professor Shao, took place for two hours a week), the ultimate aim for the choir is to feel a spiritual connotation to the work. Professor Ciz-Zao Wang, a famous music aesthetician and former President of the Central Conservatory of Music, talks about “spiritual connotation being the deep content of musical beauty”, further stating that “There are indeed heroic, tragic, civic, ordinary, noble and other factors embedded in music. These factors are neither external emotional types, nor internal meanings that reflect the formal characteristics of music, but a deeper spiritual connotation than those basic emotional and stylistic characteristics mentioned above. The spiritual connotation of music is a reflection of the composer’s innermost ideology, and often reflects the characteristics of the composer’s time.” (Wang, 2021) Bearing these words in mind, since the process of learning music is the aspiration of taking part in a dialog with the composer, the purpose of joining any choir is not simply to exercise cooperation, or else to relax for a few hours, but to engage musically at a deep and fundamental level. In order to more fully reflect the rationale of the choir as a whole, with the help of a systemic plan participants can be provided with basic music theory courses and the strengthening of solfeggio ear training in order to improve their musical performance / actualize the spiritual aspect of the music as effectively as possible.

Create an atmosphere for choral learning

Through this study, the author attempted to guide the members to pay more attention to the advantages of online training, together with reminding them to keep in mind their reasons for originally joining the ensemble. Although the aforementioned set of preliminary online training methods were successfully explored, the limitations of network technology meant that the model was unable to achieve online group singing training. Nor did it prove possible to conduct four-part choral singing online, with the result that the members were unable to experience the effects of choral harmonies. More regrettably still, some members chose to quit, undoubtedly a brutal test for any choir.

To assist in counteracting these issues, the epidemic needs to be viewed as a temporary setback, after which face-to-face rehearsals can resume. In the meantime, organizers should strive to create a united, mutually helpful, positive, and enterprising atmosphere for choral learning, as well as providing sufficient learning resources and psychological care so that members can eliminate negative emotions, enjoy learning, and gain pleasure from the choral experience. Fortunately, in this instance, most of the members were engaged and persistent in their pursuit of the art of choral singing. They actively cooperated with the leaders and completed the online semester training, demonstrating a high level of competence, good teamwork and a collective spirit. Similarly, the online training results indicated that the students had reached a level of artistry, pointing as evidence to the success of the adopted model.

Develop an online training mode

The mode outlined above has since been updated and developed. In the case of those music majors taking the choral class within the School of Music, the combined video homework and online training model has been adopted, the video serving to more intuitively visualize the level of students' singing. At the same time, the evaluation mechanism has been better implemented: every week, the leaders of each group give detailed evaluations of students' singing performances in the form of scores and written notes. Together, these constitute the process evaluation element of the choral online training course, which can be combined with the final result evaluation to form a more reasonable assessment mechanism. This mechanism can serve to stimulate students' enthusiasm, allowing them to pay more attention to the learning process, and giving rise to higher quality work.

Concluding Remarks

After a semester of exploration, the school choir's online training mode has now become normal practice. Members now generally recognize this mode, have abandoned whatever negative emotions they may have harbored, and regularly participate in the weekly online training sessions.

The successful introduction of this mode cannot be separated from the strict requirements, close attention to detail, and the ardent expectations of the participants on the part of the conductor. Success is also due, in no small part, to the responsibility shown by the voice training teachers and the league leaders, and the persistence and adherence of the members who love choral singing. Training quality and an efficient management system are likewise inseparable. Only with each member and league leaders' full cooperation in completing the training tasks can progress emerge. In this instance, every member of the league has seen their singing skills, teamwork ability, and other aspects of communal music-making, improve, while every leader has been trained in music learning, teaching, management, communication, and other choral requirements.

When these students graduate and find their places in the workforce of the future, it is hoped that they will have opportunities to make use of their musical talents. Music not only serves to relieve the pressures of heavy work, but can also create a good artistic atmosphere in collective activities and lead to more people living creative lives. In addition to contributing to the development of music education in society, some of the non-music majors, as a result of their choral exposure, may even choose to embark on a professional music development path themselves, become professional musicians or music educators, or alternatively assist in the development of another individual's music career.

Likewise, it is to be hoped that the achievements of exploring online training modes as reported in this study will provide some ideas for emulation that can lead to the development and survival of other art ensembles or school art educational areas in the current era of epidemics. The authors strive to maintain their persistent pursuit of art with the greatest passion, stay true to their original intentions, and strive to overcome whatever difficulties they may encounter. In the spirit of jointly embracing a bright future for arts education, they look forward to more offline exchanges and cooperation with art colleagues in the future, both within China and further afield.

References

Gardner, H., translated by Shen, Z. L. (2012) Two approaches to art education, *New Vision of Multiple Intelligences* (p. 164). China Renmin University Press.

Official public account of the Choir of Capital Normal University.

Wang, C. Z. (July 2021). *The Beauty of Music and Its Appreciation* (p. 150). People's Music Publishing House.

Asia-Pacific Journal for Arts Education

Special Issue

Unfold the Future of Music Education through Technology

Co-editors:
Prof. Bo Wah LEUNG
Dr. Cheung On TAM
Dr. Chi Hin LEUNG
Dr. Richard Guy WHITBREAD
The Education University of Hong Kong

http://www.ied.edu.hk/ccaproject/apjae/apjae.htm ISSN: 1683-6995

Volume 22 Number 9 December 2023

Creativity in three one-to-one videoed piano lessons in Chinese higher education institutions

Dr. Mengyao Zhao

University of York

mengyao.zhao@york.ac.uk

Dr. Elizabeth Haddon

University of York

liz.haddon@york.ac.uk

Abstract

This research focuses on contemporary piano pedagogy in the one-to-one lesson in higher music education institutions in mainland China. This qualitative study focuses on three video-recorded lessons in which creativity was evaluated through analysing teaching approach, teaching strategy, teaching content and student behaviour. The master-apprentice teaching approach was demonstrated in all lessons; these were teacher-directed, showing unequal relationship status between teacher and student, and minimal dialogue involving the students' own ideas. The findings also show that the students were learning through imitation and obeying directives. Although some creative performances occurred, these were first created by the teachers and then conveyed to their students; this might imply a reproductive creativity rather than the development of students' individual creativity. This may have implications for the development of student independence in relation to expressive and interpretative creativity.

Key words

creative piano pedagogy, higher education, video analysis technology, Jian Ying, MAXQDA

Introduction

Studies have been conducted on teachers' and students' perceptions of creativity and creative teaching (de Souza Fleith, 2000) and creativity in music education and higher education (Burnard & Haddon, 2015; Haddon & Burnard, 2016; MacDonald, Byrne & Carlton, 2006). Scholars in mainland China have likewise begun to realise the importance of cultivating students' creativity (Yi, Plucker & Guo, 2015), examining how Confucian ideology influences creativity (Niu, 2012) and conducting comparative studies on creativity (Niu & Kaufman, 2013). Running parallel to this, the Chinese government has announced new educational initiatives aimed at fostering creativity, while academics in Hong Kong (Cheung, 2012, 2013, 2016) and Taiwan (Horng et al., 2005; Lin, 2009, 2011) have conducted research on the characteristics of creative instructors and creative teaching practices.

Research undertaken by Niu and Sternberg (2006) indicates that the Chinese tend to regard creativity as having 'social and moral values' (p. 18), a finding that may be linked to collectivism in society and culture. Doing something creative and unusual may run the risk of alienating others, suggesting that a society with a collectivist culture might be detrimental to the growth of creativity. In addition, Chinese creativity is not always associated with novelty, instead 'making a connection between the new and the old' (Niu & Sternberg, 2006, p. 18). This is supported by Fung (2017), who discusses how Confucius – who believed that examining the known is necessary for discovering new perspectives [*wengu er zhixin, 温故而知新*], implying that creativity is founded on past knowledge and experience, and thus 'an established frame of reference' (Fung, 2017, p. 148) – may have influenced music education. In Zheng and Leung's (2021a) empirical study, knowledge and experience also emerge as two major characteristics associated with creativity. As Fung (2017) explains, without a frame of reference a piece of music cannot be composed; likewise, a new teaching approach cannot be implemented without the necessary theory. However, while knowledge and experience effectively combine to constitute the concept of expertise (Amabile, 1998), the other two components of creativity discussed by Amabile, motivation and creative thinking skills, have received considerably less attention in China.

Yan (2014) believes that creativity promotes students' individual growth, and that creative teaching is a vital component of music education that should be encouraged. However, the Chinese educational system is deeply influenced by examination-oriented structures (Mullen, 2017). The traditional style of teaching music in China is teacher-centred: teachers deliver knowledge to their pupils, and pupils imitate their teachers; this appears to apply not only in school classes, but also in one-to-one instrumental tuition in China (Guo & Xu, 2015). This teaching method does not emphasise learners' emotions in music (Burwell, 2016), along with providing little room for developing students' creativity (Guo & Xu, 2015). Research on fostering students' creativity and strengthening creative teaching practices in music education, particularly within piano pedagogy, is limited and elementary (Zheng & Leung, 2021b). The current research addresses this situation, analysing three videoed piano lessons that took place at three tertiary level music departments.

Methodology

A paradigm is defined as ‘a whole system of thinking’ (Neuman, 2007, p. 96). Two research paradigms are frequently applied in the field of social science research: the social constructivism paradigm, and the interpretive paradigm. Social constructivism is concerned with the historical, cultural, and contextual relevance of the environments in which individuals work and live, understanding what happens in society and constructing knowledge through this information (Creswell & Creswell, 2009). In other words, what individuals perceive and experience in the social environment is created socially. Creswell & Creswell indicated that social constructivists believe that ‘individuals seek understanding of the world in which they live and work, and individuals develop subjective meanings of their experiences’ (p. 8). Furthermore, there is a ‘complexity of views rather than narrowing meanings into a few categories or ideas’ (ibid.). In other words, all the participants’ views or beliefs are seen as valid by the researcher. Interpretivism emphasises that knowledge is constructed through people’s experiences and perceptions (Thanh & Thanh, 2015), which the researcher seeks to understand. The interpretivist paradigm tends to be inclusive, implying the potential for numerous realities instead of a single reality.

Research methods

With the aim of investigating creativity in one-to-one piano lessons in the Chinese higher education context, this paper reports selected data and findings from a case study carried out by the first author. Case study research is defined as 'an in-depth examination of an extensive amount of information about very few units or cases for one period or across multiple periods of time' (Neuman, 2007, p. 42). This may involve various research methods; Thomas (2013) indicates that observation, in which the researcher obtains access to a group, and then watches and listens to what those in the group, such as teachers and students, say and do, is a significant method for gathering data in social science. In the present study, the first author used video to aid the observation process. Video recording is being used increasingly in educational research (Pirie, 1996) as a significant means for recording human interaction, since it enables the researcher to gain a detailed understanding of the context and to carry out detailed analysis (Roschelle & Goldman, 1991; Suchman & Trigg, 1992). The advantage of using video recording is that it can 'capture a social scene far more quickly than taking field notes' (Thomas, 2013, p. 224). The researcher can also gather various perspectives of an activity, such as complex behavioural data (Roschelle & Goldman, 1991. Daniel (2006) has demonstrated that in instrumental teaching and learning, detailed analysis of student-teacher interaction and diverse teaching strategies can be accomplished via video analysis.

With regard to the present study, the first author recorded three piano lessons and analysed them from several perspectives in ways that demonstrated 'interactions between teachers and students, teaching strategies, teaching roles, learning opportunities presented to students and learning experiences provided for students' (Daniel, 2006, p. 192), as well as pupil behaviours. These aspects can help understand how the teacher-student relationship operates in one-to-one teaching within the Chinese higher music education context, what creative teaching strategies may apply, and how the teacher-student relationship might affect creative pedagogies.

Ethical considerations

Ethical approval was granted by the Arts and Humanities Ethics Committee at the University of York, UK. Information and consent forms were distributed to potential participants; these were presented in Chinese and stated the aims and objectives of the research, the methods of data collection, the issue of anonymity, how the data would be stored and used, and the risks or benefits of taking part in the research. Participants were informed that involvement would have no detrimental consequences or potential benefits, and that there would be no consequences if they decided to withdraw at any point.

To limit the possibility of disrupting teacher-student interaction, the first author was invited to sit in the corner of the teaching room by each of the three teachers and to operate the recording equipment (the first author's iPad). After each lesson, the first author asked the teacher and the student to again confirm whether they were happy for the recording to be used in the study. Each person indicated their willingness for the first author to use and analyse the video and confirmed that the lesson had followed their normal routine.

Data Analysis

The software Jian Ying [*剪映*] was used by the first author to analyse the videoed lessons. After being uploaded, captions were manually added in both Chinese and English for each lesson; the second author reviewed them to ensure that the captions were understandable to a native English speaker before transferring the captions to 'Word' documents as transcripts in order to calculate word counts. Additionally, MAXQDA was employed to produce codes for teaching methods and teaching foci by uploading lesson transcripts, quantifying the number of times the teacher concentrated on technique, expression, elements of the music score, or other areas. The teachers and students who took part in the lessons will be referred to using the following numbering: Lesson A took place with T1 and S1; Lesson B with T2 and S2, and Lesson C with T3 and S3. The main goal of the study was to analyse the process of teaching and learning in order to improve understanding of how one-to-one piano teaching is delivered in Chinese higher education institutions. Each of the three lessons will be examined separately before being discussed together in the 'Conclusions' section.

Findings

Lesson A

Lesson A took place inside a practice room in a Music Conservatory in China. S1 was studying piano performance; at the time of the research, he was a Year 3 student preparing for his first term final piano exam.

Teaching approach

The lesson began with S1's performance of the work that had been practised, the teacher providing feedback on various elements. Comments were frequently made in response to errors in S1's performance. T1 started the dialogue with an exploratory question, 'When you practise on your own, what do you think about?' While this opening suggests a mentor-friend rather than master-apprentice relationship, T1 later asked numerous rhetorical questions, such as: 'I told you before that you need to pedal first, didn't I?' and 'This is where you can immerse yourself, isn't it?' Previous research has advocated for the use of exploratory questions as the best way to stimulate students' active participation (Burwell, 2005). By contrast, rhetorical questions serve as a model for the learner (ibid). In this case, the frequent use of 'isn't it?' at the end of the teacher's statements seems to require the learner to agree rather than disagree. As a result, although both exploratory and rhetorical questions are used, the emphasis on the rhetorical serves to reiterate the authority of the master.

Within instrumental lessons, the proportion of student talk to teacher talk, as well as the quantity and types of questions being asked, can significantly indicate the type of teaching approach favoured (Young, Burwell & Pickup, 2003). In Lesson A, the student's contribution was minimal; they learned through the teacher's instructions and comments. Consequently, the teaching approach in this lesson was considered to be consistent with the master-apprentice format, whereby the teacher is viewed as a model of imitation and source of knowledge, and the student learns through the teacher's instruction (Jørgensen, 2000).

Teaching content

The focus of Lesson A was on reproducing a Chopin Nocturne from the written score. Most of the time was devoted to technique, followed by reproduction of the musical score and musical expression. Technique-oriented issues appeared 55 times, the most frequent aspects being related to dynamics (17 times) and speed (11 times), followed by pedalling, articulation, phrasing, and rhythm. Moreover, although T1 was very strict in terms of requiring S1 to play according to the marks on the score, there were only five instances where musical expression was emphasised. The specific teaching strategies that were applied are discussed in the next section.

Teaching strategy

Several strategies, such as imagery, metaphor, demonstration, and directive teaching, in which the student followed the teacher's instructions, were demonstrated in Lesson A. For example, T1 used imagery to illustrate how a specific section should be expressed by stating: 'It seems like the door will open, so I must go out and find something that I long for; that's how this [harmonic] tendency, [it] needs to be expressed'. Metaphors were frequently used to support musical expression. For example, one of the teacher's metaphors for a section was mournfulness: ‘It’s starting to get rather sad here, isn’t it?’ However, the most frequently-used teaching strategy was demonstration, typically accompanied by explanation. Sometimes, explanations were followed by demonstrations; at other times, demonstrations were followed by explanations. Demonstrations were of two types: those on the piano and those presented through gestures. Directive teaching, where T1 gave simple and short instructions and the student followed what the teacher indicated, occurred multiple times during the video. S1 mostly responded by playing according to T1’s directives, rather than offering verbal reactions.

The evaluation of student performance is an important aspect of instrumental music teaching and is linked to effective teaching (Mills & Smith, 2003; Zhukov, 2012). Both positive and negative feedback can be divided into general and specific comments. For example, ‘yes’, ‘right’, and ‘much better’ can be viewed as generally positive feedback, particularly if followed by a specific reason or explanation (Zhukov, 2012). In Lesson A, S1’s performance appeared to elicit positive comments from T1 in the form of remarks such as ‘It’s not bad’ (5:43), ‘It was good just now’ (15:30) and ‘Yes, good, good, good’ (17:08). However, what exactly was considered to be good was notable by its omission.

Pupil behaviour

S1’s behaviour showed several characteristics, with a primary emphasis on positively engaging throughout the lesson, as well as listening closely to T1’s instructions and demonstrations. In addition, S1 often nodded to indicate agreement with T1, made eye contact, and looked at T1's hands when she demonstrated, together with following the score and asking questions or responding to T1's questions. The pace of the lesson was moderate; while T1 spoke quickly, S1 was given the opportunity to reflect, answer, and ask questions.

Lesson B

Lesson B was conducted at a piano inside a practice room of a Normal University[1]. At the time of the research, S2 was a first-year student preparing for the final exam of the first term. T2 is a young piano teacher who has been teaching at this particular university for several years.

Teaching approach

On five occasions during the course of the lesson, S2 asked T2 for clarification about an aspect of the latter's instructions. However, T2 did not follow up with any questions to inspire S2 to think and help S2 take responsibility for her own learning. Instead, it appeared that T2 was used to being viewed as the authority figure, while S2 appeared to receive knowledge passively, suggesting a strong master-apprentice relationship.

Teaching content

The three pieces played by S2 included a contemporary Chinese piece, '*Colourful clouds*

---

[1] A Normal University is one which focuses on teacher-training

chasing the moon' *[Caiyun zhui yue,* *彩云追月]* by Jianzhong Wang, a Chopin Ballade and a Haydn Sonata in F Major. The observed instruction focused on the Chinese repertoire and the Haydn sonata, covering fingering, pedalling, phrasing, dynamics and musical expression. Frequent topics of instruction in Lesson B included those relating to technique, which occurred 64 times, and those covering notation and expression, which occurred nine and three times, respectively. Therefore, it can be seen that the teaching content was strongly technique related.

Teaching strategy

Directive teaching was frequently employed in Lesson B. Verbal directives were aimed at developing S2's performance according to T2's understanding of the music, using shushing to encourage a quieter dynamic and finger clicking to emphasise pulse / speed. Singing was also employed by T2 while S2 was playing in order to indicate aspects such as speed, dynamics and musical expression. In addition to demonstrating finger technique and dynamics on the piano, T2 indicated, through performing, that the performance should be more passionate, saying: 'Get a little riled up'. However, since this was played at a speed that S2 could not yet master, this was perhaps intended to be more in the way of motivation. Either way, the demonstration seemed to have little influence on S2's performance during the lesson. There were a few brief instances of praise from T2, but they did not seem to be precise. For instance, T2 did not state what was good and how it was good when the student had finished playing, stating instead: ‘Overall, much better than the last lesson. Good, good, let’s start from the beginning’. After S2 finished her first performance, T2 said: ‘OK, good. Let’s start with these. Let’s start again from the beginning’. As a result, T2’s compliments seemed to be only general, perhaps motivational, but neither full nor informative.

Pupil behaviour

Overall, S2 participated positively throughout the lesson. The student spent the majority of the time attending to T2’s talk, directives and demonstrations, contributing and responding by nodding, establishing eye contact with T2, and generally demonstrating awareness and active engagement. The general pace of Lesson B was rapid; T2 spoke immediately after S2 made errors and expected S2 to correct the errors after the lesson on her own if she was unable to do so during their time together. As with the previous example, the lesson appeared to be strongly embedded in a master-apprentice teaching model.

Lesson C

Lesson C took place in a university teacher’s practice room with T3 and S3. At the time of the research, S3 was a Year 3 student studying piano as her main instrument; T3 had many years’ teaching experience at this particular institution.

Teaching approach

T3’s speech accounted for 1,861 words, or 99.8 percent, of the total word count, based on the transcription of the lesson. As further evidence of a highly dominant master-apprentice relationship, this teacher only used questions to confirm whether the student had understood or not before going straight on to the following point. Since the student was not afforded enough time to respond, this might indicate that T3 was indifferent to whether the student understood or not, or else that they had become used to silence from the student. Either way, no form of dialogue was evident.

Teaching content

S3 played Etude Op. 299, no. 5 by Czerny in the lesson, a work set for the student's Year 3 piano exam. The teaching content tended to be technique-oriented. As a result, technique-related themes featured 16 times, whereas practice-related themes appeared eight times.

Teaching strategy

In Lesson C, T3's most common teaching strategy was modelling. The reason might be that T3 had a piano of her own (there were two upright pianos in the practice room), or it could simply have been as a result of T3 believing that modelling was the most efficient way for S3 to learn. There were also five instances of student-teacher collaboration, which consisted of S3 and T3 playing together on separate pianos, rather than S3 playing one hand and T3 supplying the other part. In addition, directive teaching was employed at several points, and there were also a few instances of the use of the metaphor strategy.

In Lesson C, T3 used more specific positive feedback, and there were more instances where compliments were given than in either of the previous two lessons, e.g., 'Um, good. The completeness is good' and 'That's what I just told you about doing *crescendo*, you did well'. Overall, T3 was relatively positive and encouraging in her evaluation of S3's performance.

Pupil behaviour

S3 listened attentively to T3's speaking and playing, establishing eye contact and looking at T3's hands. In addition, the student added markings to the score and silently imitated T3's playing several times. Overall, the lesson was conducted at a moderate tempo; T3's pacing enabled S3 to attempt individual sections more than once and (to) play at a speed which was the most comfortable for her.

Discussion

The analysis of the piano lessons in this study explored teaching approaches, teaching strategy, teaching content, and pupil behaviour. The master-apprentice model was demonstrated in all three lessons, which were teacher-directed; as such, they revealed an unequal relationship status between teacher and student and minimal dialogue involving the students' own ideas. As indicated by Luo (2018), piano education at the higher education level in China is mostly didactic in style. Carey et al. (2013) define this as transfer pedagogy. Since teachers place the focus on the outcome rather than the process, this type of teaching has a tendency to make students more reliant on their tutors and take less ownership of their own learning. With the exception of a few teacher-student interactions in Lesson A, there was very little verbal communication between students and teachers, and relatively few opportunities for the students to speak, other than to answer questions. The students appeared to be in a relatively passive position, learning through imitation and instruction. Dialogue between the teachers and the students was infrequent.

Questioning techniques seemed to be largely overlooked by the teachers in all three lessons. Even where they were in evidence, these inquiries tended to be closed questions, and the dialogue did not appear to be an equal discussion that respected the students' ideas. Allsup and Baxter (2004) emphasise the significance of asking more open-ended questions and the role of dialogue in music lessons, since students can gain additional knowledge. Kassner (1998) believed that skilled questioning could stimulate students' higher-level thinking, promoting their own evaluative capabilities.

It can be seen that the teacher in Lesson A (conservatory) used slightly more varied teaching strategies than the others. There appeared to be more verbal communication between T1 and S1, although this might be due to the longer duration of this lesson (33 minutes, compared to 27 minutes and 19 minutes for the others, respectively). A common feature of all three lessons was the instructive approach of pointing out immediately where the students had made mistakes. This is in line with the findings of previous research; for example, Yeh (2018) indicated that piano teachers in their study mostly focused on analysing students' playing mistakes in one-to-one lessons.

The data analysis revealed the use of a variety of teaching methods, including imagery, metaphor, demonstration, directive teaching, and student-teacher cooperation. However, where there were two pianos in the classroom, the teacher (in Lesson C) chose modelling as her primary teaching strategy. Constant directive teaching was also seen, particularly by T2 in Lesson B. According to Zhukov (2012), while modelling, directives and praise are frequently used for teaching instrumental music in higher education, this research discovered a predominance of directives and demonstration as teaching strategies, resulting in an underdeveloped level of praise.

Although the students were learning through the processes of imitation and obeying directives, this approach also perhaps deprives them of the ability to play according to their own preferences. Laukka (2004) interviewed teachers from UK and Swedish conservatoires and found that they emphasised verbal inspiration while developing students' independence. Teachers in Laukka's study felt that if the students' primary learning strategy was imitation, they might fail to learn independently, resulting in a lack of their own ideas and an inability to really express the music.

When comparing the most significant aspects of creative teaching revealed by Cremin and Chappell's (2021) comprehensive study of the relevant literature (idea generation and exploration, co-construction and collaboration, supporting autonomy and agency, and problem solving) with these piano lessons, very few – if any – of these features are evident in the three videoed lessons. However, T1's use of imagery in Lesson A, in addition to relating more to musical expression than technique, also seemed to have a positive influence on S1's performance; more *rubato* was displayed, and both tone quality and musical expression were enhanced. Furthermore, T3 developed several types of exercises to play the same piece in different rhythms. This not only made the exercises more enjoyable, but also engaged S3's attention, enabling him to generate a more creative performance. Given that the performance was first created by the teachers and then conveyed to their students in each of these three lessons, this might imply a reproductive creativity rather than the development of students' individual creativity.

Additionally, the students learned repertoire that included at least one, if not two, pieces of Western classical music. As indicated by Wang (2018), Western piano music has dominated piano education in higher education in China, in part a reflection of the fact that many piano teachers have grown up following the Western pedagogical system. Lesson B did include a contemporary Chinese piece, though it is a somewhat dated work ('*Colourful clouds chasing the moon*' [*Caiyun zhui yue, 彩云追月*] was first composed in 1935 for orchestra and later arranged for piano in 1975 by Jianzhong Wang). However, all of the material in these lessons was taught with the same approach, the primary focus being on accuracy and a realisation of the teacher's interpretation.

Overall, the three lessons were found to favour detailed and informative teaching, the focus being mainly on technique-related content; issues related to emotion and expression were mentioned infrequently. This might be due to the students' levels of learning in relation to each piece, as well as their year of study. This particularly applies to Student B; this filmed lesson may have involved more extensive and informative teaching related to technique than other lessons participated in during the academic year. The students in Lessons A and C were in Year 3 at the time of filming. While Lesson A appeared to contain slightly more discussion and reflection about musical expression compared with the other two piano lessons, it seems as though the student was still not given much freedom to develop their own independence and creativity.

Conclusions

The teaching style and teaching strategies employed in these one-to-one lessons were largely consistent across all three types of institutions. Teachers seemed to have complete control within the one-to-one teaching and learning context, a hierarchical relationship that did not obviously appear to help students to become independent learners. Furthermore, technique-oriented teaching, demonstration and directives as the predominant teaching strategies limited encouragement, and the choice of repertoire all served to limit students' motivation, and thus support them to take responsibility for their own learning, as well as failing to foster the development of creative approaches to instrumental learning. Future research might examine more teaching contexts, including Year 4 students and lessons for Master's degree students, while simultaneously exploring these phenomena over a longer period of teaching and learning. Findings from this study could also be used to advocate for wider strategies and the facilitation of increased student communication in one-to-one piano lessons, along with considering the implications for how creativity can be further fostered within the context of piano teaching.

References

Allsup, R. E., & Baxter, M. (2004). Talking about music: Better questions? Better discussions!. *Music Educators Journal*, *91*(2), 29-33. doi.org/10.2307/3400046

Amabile, T. M. (1998). How to kill creativity. *Harvard Business Review*, *87*, 77-87.

Burnard, P., & Haddon, E. (Eds.). (2015). *Activating diverse musical creativities: Teaching and learning in higher music education*. Bloomsbury.

Burwell, K. (2005). A degree of independence: teachers' approaches to instrumental tuition in a university college. *British Journal of Music Education*, *22*(3), 199-215. doi.org/10.1017/s0265051705006601

Burwell, K. (2016). *Studio-based instrumental learning*. Routledge.

Carey, G. M., Bridgstock, R., Taylor, P., McWilliam, E., & Grant, C. (2013). Characterising one-to-one conservatoire teaching: Some implications of a quantitative analysis. *Music Education Research*, *15*(3), 357-368. doi.org/10.1080/14613808.2013.824954

Cheung, R. H. P. (2012). Teaching for creativity: Examining the beliefs of early childhood teachers and their influence on teaching practices. *Australasian Journal of Early Childhood*, *37*(3), 43-52. doi.org/10.1177/183693911203700307

Cheung, R. H. P. (2013). Exploring the use of the pedagogical framework for creative practice in preschool settings: A phenomenological approach. *Thinking Skills and Creativity*, *10*, 133-142. doi.org/10.1016/j.tsc.2013.08.004

Cheung, R. H. P. (2016). The challenge of developing creativity in a Chinese context: The effectiveness of adapting Western creative pedagogy to inform creative practice. *Pedagogy, Culture & Society*, *24*(1), 141-160. doi.org/10.1080/14681366.2015.1087419

Cremin, T., & Chappell, K. (2021). Creative pedagogies: A systematic review. *Research Papers in Education*, *36*(3), 299-331. doi.org/10.1080/02671522.2019.1677757

Creswell, J. W., & Creswell, J. D. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage Publications.

Daniel, R. (2006). Exploring music instrument teaching and learning environments: Video analysis as a means of elucidating process and learning outcomes. *Music Education Research*, *8*(2), 191-215. doi.org/10.1080/14613800600779519 de Souza Fleith, D. (2000). Teacher and student perceptions of creativity in the classroom environment. *Roeper Review*, *22*(3), 148-153. doi.org/10.1080/02783190009554022

Fung, C. V. (2017). *A way of music education: Classic Chinese wisdoms*. Oxford University Press.

Guo, B., & Xu, J. (2015). Analysis of the necessity of carrying out the educational training program for creative music talents. *Beijing Education Journal*. Retrieved from: http://www.jyb.cn/theory/rcpy/201507/t20150719_630630.html [assessed 30 November 2022]

Haddon, E., & Burnard, P. (Eds.). (2016). *Creative teaching for creative learning in higher music education*. Routledge.

Horng, J.-S., Hong, J.-C., ChanLin, L.-J., Chang, S.-H., & Chu, H.-C. (2005). Creative teachers and creative teaching strategies. *International Journal of Consumer Studies*, *29*(4), 352-358. doi.org/10.1111/j.1470-6431.2005.00445.x

Jørgensen, H. (2000). Student learning in higher instrumental education: Who is responsible? *British Journal of Music Education*, *17*(1), 67-77. doi.org/10.1017/s0265051700000164

Kassner, K. (1998). Would better questions enhance music learning? Teachers who ask effective and carefully crafted questions may help students improve their learning and performance. *Music Educators Journal*, *84*(4), 29-36. doi.org/10.2307/3399113

Laukka, P. (2004) Instrumental music teachers' views on expressivity: A report from music conservatoires. *Music Education Research*, *6*, 45-56. doi.org/10.1080/1461380032000182821

Lin, Y. S. (2009). *Teacher and pupil responses to a creative pedagogy - Case studies of two primary sixth-grade classes in Taiwan*. (Doctoral dissertation), University of Exeter, UK.

Lin, Y. S. (2011). Fostering creativity through education: A conceptual framework of creative pedagogy. *Creative Education*, *2*(3), 149. doi.org/10.4236/ce.2011.23021

Luo, J. (2018). Discussion on the inheritance, integration and promotion of Chinese traditional folk music culture in college piano education. *Advances in Social Science, Education and Humanities Research*, *180*, 122-124. doi.org/10.2991/essd-18.2018.35

MacDonald, R., Byrne, C., & Carlton, L. (2006). Creativity and flow in musical composition: An empirical investigation. *Psychology of Music*, *34*(3), 292-306. doi.org/10.1177/0305735606064838

Mills, J., & Smith, J. (2003). Teachers' beliefs about effective instrumental teaching in schools and higher education. *British Journal of Music Education*, *20*(1), 5-27. doi.org/10.1017/s0265051702005260

Mullen, C. A. (2017). *Creativity and education in China: Paradox and possibilities for an era of accountability*. Routledge.

Neuman, L. W. (2007). *Social research methods.* Pearson.

Niu, W. (2012). Confucian ideology and creativity. *The Journal of Creative Behaviour*, *46*(4), 274-284. doi.org/10.1002/jocb.18

Niu, W., & Kaufman, J. C. (2013). Creativity of Chinese and American cultures: A synthetic analysis. *The Journal of Creative Behaviour*, *47*(1), 77-87. doi.org/10.1002/jocb.25

Niu, W., & Sternberg, R. J. (2006). The philosophical roots of Western and Eastern conceptions of creativity. *Journal of Theoretical and Philosophical Psychology*, *26*(1-2), 18. doi.org/10.1037/h0091265

Pirie, S. E. (1996, October). Classroom video-recording: When, why and how does it offer a valuable data source for qualitative research? Paper presented at the Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, Panama City, FL.

Roschelle, J., & Goldman, S. (1991). VideoNoter: A productivity tool for video data analysis. *Behavior Research Methods, Instruments, & Computers*, *23*(2), 219-224. doi.org/10.3758/bf03203368

Suchman, L. A., & Trigg, R. H. (1992). Understanding practice: Video as a medium for reflection and design. In J. Greenbaum & M. Kyng (Eds.) *Design at work: Cooperative design of computer systems* (pp. 65-90). Routledge.

Thanh, N. C., & Thanh, T. T. (2015). The interconnection between interpretivist paradigm and qualitative methods in education. *American Journal of Educational Science*, *1*(2), 24-27.

Thomas, G. (2013). *How to do your research project: A guide for students*. Sage.

Wang, P. (2018). Rational thinking on the teaching of Chinese piano music in colleges and universities. *Advances in Social Science, Education and Humanities Research*, *220*, 39-41. doi.org/10.2991/iceemt-18.2018.8

Yan, F. L. (2014). The function of creative teaching in music curriculum. [音乐课堂创新教学的意义作用]. Retrieved from: https://wenku.baidu.com/view/28ad690db84ae45c3a358c18.html [accessed 30 November 2022].

Yeh, Y. L. (2018). An investigation of Taiwanese piano teachers' reflection on teaching challenges and pupils' learning difficulties. *Music Education Research*, *20*(1), 32-43. doi.org/10.1080/14613808.2016.1249359

Yi, X., Plucker, J. A., & Guo, J. (2015). Modelling influences on divergent thinking and artistic creativity. *Thinking Skills and Creativity*, *16*, 62-68. doi.org/10.1016/j.tsc.2015.02.002

Young, V., Burwell, K., & Pickup, D. (2003). Areas of study and teaching strategies instrumental teaching: A case study research project. *Music Education Research*, *5*(2), 139-155. doi.org/10.1080/1461380032000085522

Zheng, Y., & Leung, B.-W. (2021a). Perceptions of developing creativity in piano performance and pedagogy: An interview study from the Chinese perspective. *Research Studies in Music Education*. doi.org/10.1177/1321103x211033473

Zheng, Y., & Leung, B.-W. (2021b). Cultivating music students' creativity in piano performance: a multiple-case study in China. *Music Education Research*, *23*(5), 594-608. doi.org/10.1080/14613808.2021.1977787

Zhukov, K. (2012). Teaching strategies and gender in higher education instrumental studios. *International Journal of Music Education*, *30*(1), 32-45. doi.org/10.1177/0255761411431392