MusicLesson: An Exploratory Augmented Reality Learning Solution for Music Educators

by

Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

Bachelor of Science

in the

Department of Mathematical Sciences
Faculty of Natural and Applied Sciences

TRINITY WESTERN UNIVERSITY
Spring 2021

Acknowledgements

The author would like to thank Trinity Western University as well as its Computing Science faculty, especially Dr. Herbert H. Tsang for his support, advice, and help in securing funds for this thesis. In addition to this, the author would like to thank both Mitacs and Credo Interactive for funding and support. Finally, thanks to Nathan Hampson for his sage networking advice, as well as Cory Murguly for helping to test the software.

Abstract

Due to the COVID-19 pandemic, the demand for virtual learning environments has skyrocketed. While not perfect, online group video calls can work for many types of classes. Music classes, however, are not well suited to these environments for a number of reasons, including difficulties observing form and lack of notation capabilities for scores. Both teachers and students suffer because of this. The author proposes a new application to solve some of these problems, *MusicLesson*. The software will include capabilities for score annotation, as well as an interface to allow for multiple camera angles. The software will also be integrated with an AR system for increased immersion and utility. As setup is a big barrier in remote learning, the software should be easy to use for both students and teachers, and not be dependent on external tools. The software will be developed as a web based application, allowing any device that can browse the internet to use it.

Keywords: Augmented Reality; Remote Learning; Music Education;

Contents

List of Tables

List of Figures

Chapter 1

INTRODUCTION

Due to the COVID-19 pandemic, educators and students everywhere face the challenge of virtual learning. While current online video solutions such as *Zoom* [7] and *Skype* [5] can provide an alternative to in-person learning, they have several weaknesses that music educators are particularly vulnerable to. The primary objective of this research is to improve the state of remote learning for music teachers, as well as their students. That will be achieved by creating an online virtual learning application with considerations for music education. The software will use Augmented Reality (AR) technology, utilizing EPSON Moverio BT-300 Smart Glasses [2].

Chapter 2

BACKGROUND

2.1 Problems with current remote learning

This section will explore some of the issues with current online video chat applications, as well as proposed solutions. While these challenges aren't exclusive to music teachers, they disproportionately impact them.

2.1.1 Document sharing

There is also a lack of document sharing and annotation functionality in current popular remote learning applications. While applications that can already perform such services may already exist, it is important to consider the increased complication in an already foreign environment for both the student and teacher. By integrating those capabilities into a virtual learning application, we can avoid the necessity of doing so with a 3rd party program that both the student and the teacher would have to download, configure, and learn, leading to potential problems and frustrations.

2.1.2 Multi camera angle interface

When a music teacher is giving a lesson in person, they have freedom to observe their student play the instrument from any angle they need to. In an online environment, this is not possible. It can be important for the teachers to observe things like technique, and in order to do this in an online environment, they may need to be able to see multiple angles. For instance, a violin teacher could require a close-up camera angle of the student's fingers, as well as their bow hand in order to properly instruct them. The application could support this by allowing participants to stream video from multiple cameras. On the teacher's side, there would be an interface for viewing the multiple angles. While many students may not own a second webcam, smartphones are now widely used, meaning most students would likely have a secondary internet enabled device with a camera in it. For this reason, the multi-camera interface would have to include affordances for multi-device usage.

An alternative to solution to this problem is Virtual Reality (VR). Theoretically, a virtual reality environment could be fabricated in which a teacher can observe the student in 3D. However, research suggests that implementing VR in an education setting can be a challenge. One major barrier is hardware, as it can be expensive. VR also contradicts this research's prerogative to create a simple and easy to use application.

A more helpful middle ground, in the authors opinion, is Augmented Reality (AR). The specifics on this implementation will be covered in a later section.

2.1.3 Latency

Internet connection is not always stable or reliable. While there is nothing that can be done to guarantee internet outages don't occur, latency, especially in the context of time-sensitive activities (such as playing along with a reference track), has potential to pose an issue. Scenario A in Figure 2.1 exemplifies this issue. If the teacher starts the reference track on their end of the video call, there will be a significant delay between the reference track and the students performance. With that methodology, the synchronization of the backing track and the student's response will be offset by double the latency of the connection. This could be improved by allowing the teacher to trigger an audio file to play on their student's computer, synchronizing the reference track with the student's accompaniment. Scenario B in Figure 2.1 demonstrates this principal. Note that the offset between the student's performance and the play-along track is zero. It should be noted that while this was identified as an important issue,

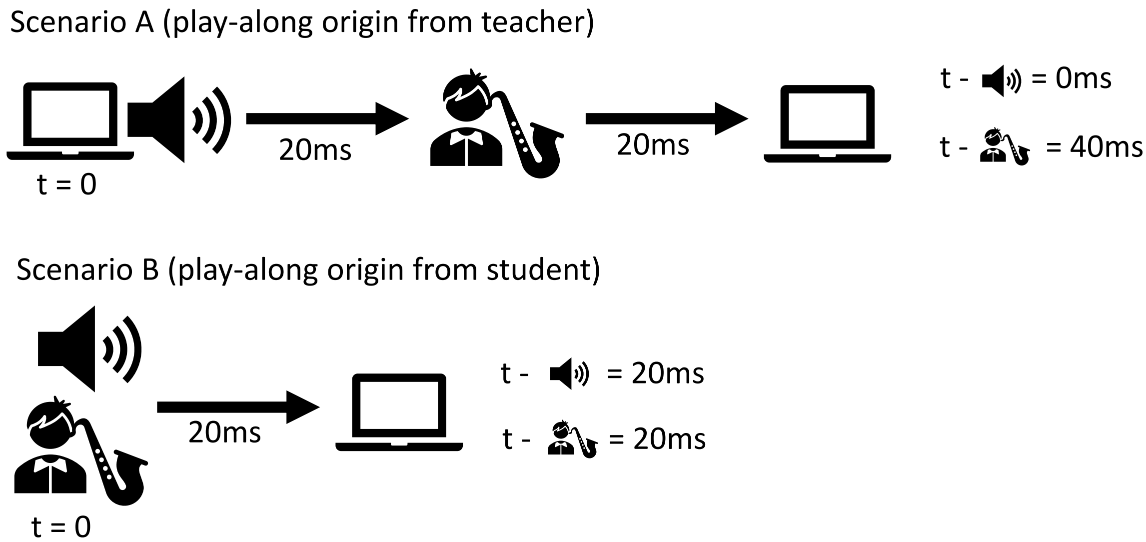

Figure 2.1: Visualization of latency problem discussed in Section 2.1.3. Scenario B uses novel methodology described.

the suggested solution to it is not currently implemented in *MusicLesson*. This section is included to be thorough in an exploration of relevant issues.

2.2 Literature review

2.2.1 Terminology

Much research has been done on distance learning. As far as etymology goes, in 2011, Moore et al. wrote a paper entitled *e-Learning, online learning, and distance learning environments: Are they the same?* [21]. It was found that use of the terms was inconsistent among research papers. Keeping this in mind, the terminologies that are used in this literature review may vary based on the research that is being reviewed.

2.2.2 Online Education and Teaching

There has been ample research on the effects of distance learning from a teaching perspective. While lacking in proposed solutions, the Cambridge University Press published an editorial on the importance of proper music education during a pandemic [11]. The research largely focused on the under-preparedness of teachers going into the pandemic. Additionally, there has been research affirming the value of tailored virtual learning environments in other disciplines, such as medical education [27], and Orthopaedic education [19]. It is clear that best-practice virtual education differs depending on the subject that is being taught.

2.2.3 Online Education and Performance

Research has also been done on student performance in online learning environments. In 2013, Pike et al. did research on the effect of distance learning on the acquisition of piano sight-reading skills [22]. They used *Skype* [5] for their testing, but it should be noted that they also had secondary internet MIDI software as well as MIDI enabled keyboards. MIDI technology is not applicable to many instruments. It was found that there was no significant differences between remote and in-person learning in regards to piano sight-reading development. In 2017, Rassaei did a case study of Persian EFL (English as a Foreign Language) students, and found that there was no significant difference between online and in person learning [23]. They used *Skype* [5] as their virtual learning application. These papers show that optimistically it is possible to implement a solution that works as well as in-person teaching.

2.2.4 Integration of Existing Technology

Research has also been done on using standalone existing video chat technology for music education. Kruse et al. used *Skype* [5] for a musical education case study [20]. It was found that the lessons had a "natural feel". They cite technical complications, as well as disconnectedness as issues. Another music education case study was conducted by Kruse et al. in

Jay Dorfman did research on one-to-one computer based remote learning for music teachers in 2016 [14]. One concern that was found was the ability of the teachers to integrate the technology. He used a concerns-based adoption model. It was found that many teachers shared concerns about technical support, pedagogical support, and authenticity of integration.

2.2.5 Use of Virtual Reality

González-Zamar and Abad-Segura argue that a Virtual Reality (VR) approach to online arts education improves creativity and immersion [17]. They also point to the growing relevance of this topic, observing that 48.69% of all articles on the implications of VR in Arts Education were published in the last 5 years. Callaghan and Garner also take a VR approach. They created an online game to simulate a virtual classroom environment. Their goal was to try to create a sense of community among students who had to learn remotely, as well as provide a way for the physically present students to interact with the virtual ones [8] [16]. It was found that while the software had some issues, in general, it "had a minimal impact on normal patterns of teaching, and the teachers perceptions of the learning occurring in their teaching environment" [16]. However, it was found that ease of use was an issue. In 2017, Shin did research on the impact of motivational affordances in virtual reality [26]. He defined affordances as features that allowed users to track and record goals. It was found that affordances are an effective tool to enhance the user experience in remote learning scenarios. Serafin et al. did research on the use of Virtual and Augmented Reality (AR) technology for learning to play an instrument [25]. They used Virtual Reality Musical Instruments (VRMIs). While VRMIs as music education tools were found to have many potential benefits, such as assisting with stage fright and training rhythmic skills, the researchers acknowledge that there is a lack of research and development on the topic.

2.2.6 Document Annotation

Pertaining to document annotation, in 2009, Vallence et al. did research on the conditions for successful online document collaboration [29], providing a useful set of rules. One of the points that they emphasize is that the presence of collaborative tools alone does not guarantee their effective use. They argue that proper implementation and use case is important. They would likely agree that an integrated document sharing service would be more effective than a external one. In 2012, Dodd did research on signalling, which refers to cues that emphasize important information in materials and documents [13]. In his research, he notes that there is a lack of existing literature pertaining to signaling in remote learning.

2.2.7 Latency

Research has also been done on achieving low latency conditions for online performance. Riley et al. developed a low latency online audio performance system (LOLA), and research participants found that low latency significantly improved their experience [24]. Chafe et al. conducted research on the impact of short time delays on musical synchronization. They found that small offsets (3ms) between performers tended to make the performance speed up, while large offsets (greater than 15ms) tended to slow the performance down. Driessen et al. also did research on the impact of latency on collaborative performance [15], and discovered the same negative correlation between latency and performance speed.

Chapter 3

METHODOLOGY

3.1 MusicLesson

After researching the viability of different solutions to the problems posed above, development on *MusicLesson* began. The application runs over a local network, supported by a Node.js server [4]. Websockets are the primary technology that the application makes use of, powering both the annotation subsystem as well as handling the video and audio streaming. Two external resources in particular, from github user wesbos [6] and tech blogger Jirka Hladis [3] were instrumental in understanding and

Figure 3.1: EPSON Moverio BT-300 Smart Glasses. The wearer can see as normal, with a display superimposed over their vision. The front facing camera captures a first-person view.

implementing websocket technology. Source code can be found in a publicly available repository at *https://github.com/davidhampson/Thesis*.

3.1.1 Augmented Reality

The first step in the concrete implementation was to decide on an AR device to use. As mentioned above, the EPSON Moverio BT-300 Smart Glasses [2] were chosen. This was because the price was better than a lot of competitors, and the device had all the features that were required (front facing camera, AR). The firmware of the device is a modified version of Andriod, so it was not required to learn any unfamiliar development environments.

The ideal use case for the glasses is as follows: the student wears them while playing and streams a first-person view to their teacher. Using the display on the glasses, the student can see a score being annotated in real time by their teacher.

Configuring the glasses turned out to be more of a challenge than initially anticipated. Because the port of Andriod that the glasses are running forces the use of an outdated browser, the web-socket technology that was being used to stream video was not working.

Thankfully, a package called Epson WebView Installer [1] conveniently updated the browser to allow for the technology to work. Once this was configured, the glasses were ready to go.

3.1.2 Camera Angles

The biggest camera angle related challenge was making sure that the application was compatible with as many devices as possible. Thanks to websockets being so widely supported, this was not a big issue. The only exception was iPhones, and unfortunately, the application is not compatible with them. However, this issue would likely not be very difficult to solve, but it was not deemed important enough to delay getting other features working on time. Adding a new camera angle to the application is as simple as loading a browser page on any modern internet enabled device with

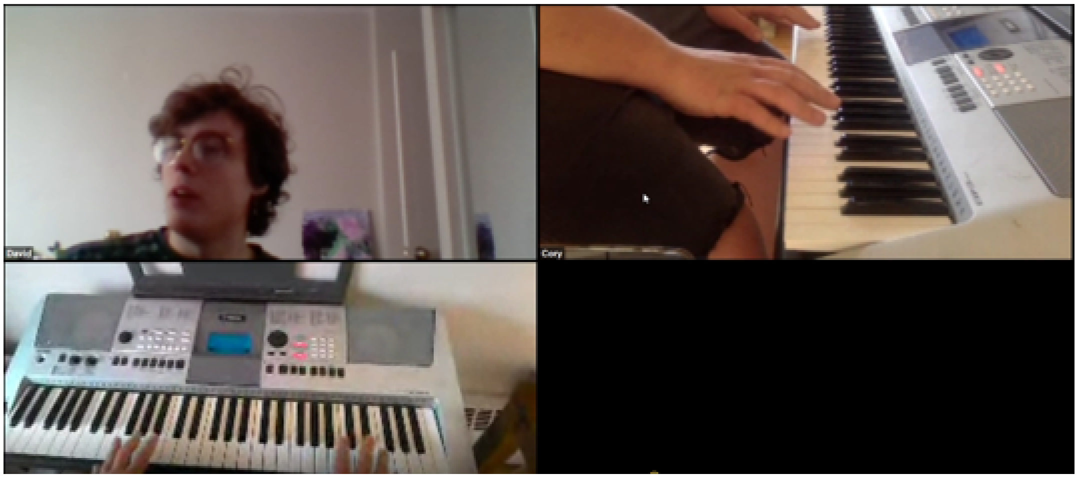

Figure 3.2: Screenshot from MusicLesson. You can see multiple camera angles of the student playing the piano, while the author watches.

a camera. There are not many more technical or implementation specific details to share. Figure 3.2 shows a screenshot from the application demonstrating a student that is showing two camera angles to an observer.

3.1.3 Ease of Use

As made clear in the literature review, ease of use is a very important consideration in this type of application. This was made a priority in development. No more than one click is required for a user to connect to a session. There are also no configurable settings. While it could be argued that this is a downside, the clear upside is that it is impossible for an end-user to "break" the application. When testing the application with an individual who had no prior experience with it, ease of use was not found to be an issue. While this is only a case study, the result was encouraging. Finally, since the application is hosted in a browser, there is no need for users to download anything. This is useful in scenarios where end users don't have download privileges on their computers (common in school settings).

3.1.4 Document Annotation

Thanks to the existing websocket infrastructure from the video chat component of the application, implementing shared annotation was quite trivial. An HTML canvas watched by Javascript events simply echos anything the teacher draws on the score in

Figure 3.3: Screenshot from MusicLesson. The red marking represents an annotation.

front of them to all connected sessions via the existing websocket server. The teacher can draw on the canvas as if they are drawing with a pen on a page. Any marking that they make will leave a red line behind. This allows for real-time annotation that is viewable by all parties involved in the lesson. Figure 4.1 shows a screenshot of an annotation inside the application.

In the current application, the document that is being annotated is set in the back end. This was mostly done because the software was intended as more of a proof of concept than an actual end-user application. The code to allow an end user to upload custom scores (and not have to set them in the back end) would be quite simple, and would not have any impact on the collaborative annotation part of the software.

Chapter 4

RESULTS/FUTURE WORK

4.1 Testing

The software was tested using the help of the author's roommate, Cory. A session was established, and connected to on multiple devices, including the AR glasses. No actual musical instruction was done, but testing of the annotation subsystem was successful. Cory also had no issues connecting to the session on either device (having had no previous experience with the system). A demonstration video can be found in the repository for the project (*https://github.com/davidhampson/Thesis*, file is called *Demo.mp4*).

4.2 Future Work

4.2.1 User Feedback

Due to time constraints and various delays, user testing was not done with this software. The author plans to pursue this with ongoing funding from Mitacs/Credo Interactive. The target audience will be music teachers that already have some experience with remote learning. An alternative (and perhaps easier to achieve in spite of current restrictions) approach could be to simply share a demonstration video of the software with them.

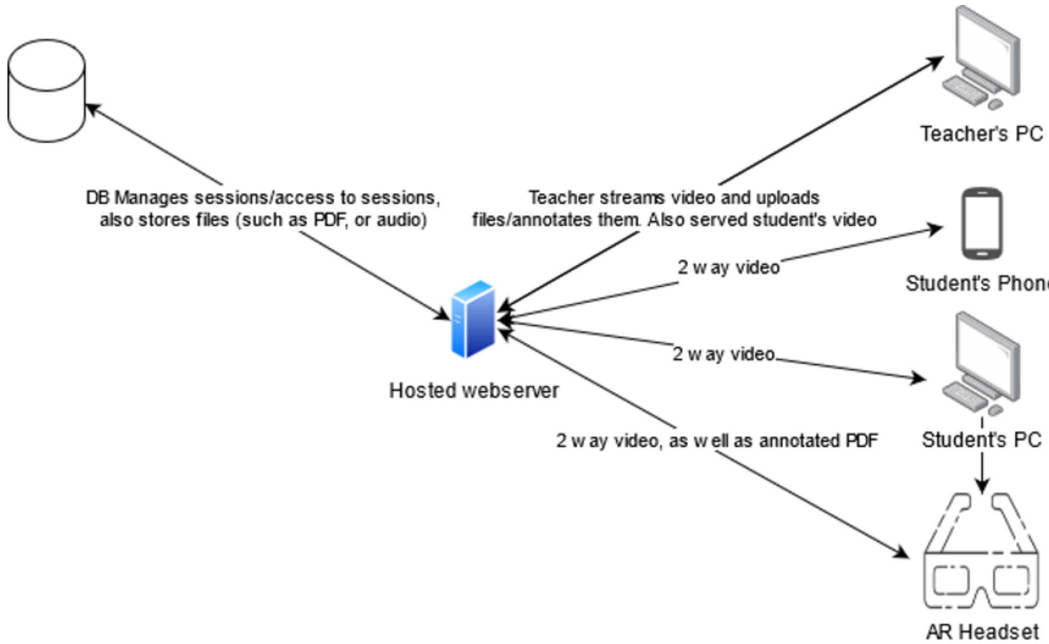

Figure 4.1: Desired network configuration for future build of MusicLesson, as opposed to running on a local network.

4.2.2 Remote Hosting

In the early stages of development, *MusicLesson* was running on a remote server. However, in light of several networking issues, the software was moved to run on the author's local network. These issues are known (to host this service remotely, a TURN server needs to be hosted and configured), but the other features of the software were more exploratory and higher priority to implement. Moving the system back onto a remote server is an important step if user testing is to be done easily. Figure **??** shows the desired network configuration.

4.2.3 Latency Solution

Section 2.1.3 discusses a feature that was planned for this software. Although it is not currently implemented due to time constraints and higher priorities of other features, the author plans to include it in the next build of *MusicLesson*.

Chapter 5

CONCLUSIONS

Overall, the results of this research are encouraging. The application is very simple, and works consistently. The features feel novel and intuitive. No critical issues were encountered in testing. Although the pandemic may be nearing an end, the novelty of this technology allows for easier remote music education, which has many use cases. Perhaps the reason that remote learning has historically not been a commonly implemented solution is not because it is inferior to in person learning, but that not enough work has been put into optimizing it. Hopefully as future work on this project proceeds, real music teachers and students will get to enjoy an increased quality of life during lessons.

Bibliography

[1] Epson webview. https://moverio.epson.com/jsp/pc/pc_application_detail.jsp?pack=com.epson.moverio.bt3.webviewinstaller&page=0&key=&cat=000&tab=recommend&device=3. Accessed: 2021-04-25. 9

[2] Moverio bt-300 smart glasses. https://epson.com/For-Work/Wearables/Smart-Glasses/Moverio-BT-300-Smart-Glasses-(AR-Developer-Edition)-/p/V11H756020. Accessed: 2021-04-25. 1, 9

[3] Multi user video chat with webrtc. https://www.dmcinfo.com/latest-thinking/blog/id/9852/multi-user-video-chat-with-webrtc. Accessed: 2021. 8

[4] Nodejs. https://nodejs.org/en/. Accessed: 2021. 8

[5] Skype website. https://www.skype.com/en/. Accessed: 2020-12-16. 1, 5

[6] Websocket canvas draw. https://github.com/wesbos/websocket-canvas-draw/. Accessed: 2021. 8

[7] Zoom website. https://zoom.us/. Accessed: 2020-12-15. 1

[8] Victor Callaghan, Michael Gardner, Bernard Horan, John Scott, Liping Shen, and Minjuan Wang. A mixed reality teaching and learning environment. In Joseph Fong, Reggie Kwan, and Fu Lee Wang, editors, *Hybrid Learning and Education*, pages 54–65, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg. 6

[9] Chris Chafe, Juan-Pablo Cáceres, and Michael Gurevich. Effect of temporal separation on synchronization in rhythmic performance. *Perception*, 39(7):982–992, 2010.

[10] Mervyn Cook. Augmented reality: Examining its value in a music technology classroom. practice and potential. *Waikato Journal of Education*, 24:23–38, 11 2019.

[11] Alison Daubney and Martin Fautley. Editorial research: Music education in a time of pandemic. *British Journal of Music Education*, 37(2):107–114, 2020. 5

[12] Beatrice A Digolo, EA Andang'o, and J Katuli. E-learning as a strategy for enhancing access to music education. *International Journal of Business and Social Science*, 2(11):135–139, 2011.

[13] Bucky J. Dodd and Pavlo D. Antonenko. Use of signaling to integrate desktop virtual reality and online learning management systems. *Computers & Education*, 59(4):1099 – 1108, 2012. 7

[14] Jay Dorfman. Music teachers' experiences in one-to-one computing environments. *Journal of Research in Music Education*, 64(2):159–178, 2016. 5

[15] Peter Driessen, Thomas Darcie, and Bipin Pillay. The effects of network delay on tempo in musical performance. *Computer Music Journal*, 35:76–89, 03 2011. 7

[16] M. Gardner and L. O'driscoll. Mirtle (mixed-reality teaching and learning environment): from prototype to production and implementation. 2011. 6

[17] Mariana-Daniela González-Zamar and Emilio Abad-Segura. Implications of virtual reality in arts education: Research analysis in the context of higher education. *Education Sciences*, 10(9):225, 2020. 6

[18] S. Jana, A. Pande, A. Chan, and P. Mohapatra. Mobile video chat: issues and challenges. *IEEE Communications Magazine*, 51(6):144–151, 2013.

[19] Monica Kogan, Sandra E. Klein, Charles P. Hannon, and Michael T. Nolte. Orthopaedic education during the covid-19 pandemic. *Journal of the American Academy of Orthopaedic Surgeons*, 28(11), 2020. 5

[20] Nathan B. Kruse, Steven C. Harlos, Russell M. Callahan, and Michelle L. Herring. Skype music lessons in the academy: Intersections of music education, applied music and technology, Apr 2013. 5

[21] Joi L. Moore, Camille Dickson-Deane, and Krista Galyen. e-learning, online learning, and distance learning environments: Are they the same? *The Internet and Higher Education*, 14(2):129 – 135, 2011. Web mining and higher education: Introduction to the special issue. 4

[22] Pamela D. Pike and Kristin Shoemaker. The effect of distance learning on acquisition of piano sight-reading skills. *Journal of Music, Technology and Education*, 6(2):147–162, 2013. 5

[23] Ehsan Rassaei. Video chat vs. face-to-face recasts, learners' interpretations and l2 development: a case of persian efl learners. *Computer Assisted Language Learning*, 30(1-2):133–148, 2017. 5

[24] Holly Riley, Rebecca B. MacLeod, and Matthew Libera. Low latency audio video: Potentials for collaborative music making through distance learning. *Update: Applications of Research in Music Education*, 34(3):15–23, 2016. 7

[25] S. Serafin, A. Adjorlu, N. Nilsson, L. Thomsen, and R. Nordahl. Considerations on the use of virtual and augmented reality technologies in music education. In *2017 IEEE Virtual Reality Workshop on K-12 Embodied Learning through Virtual Augmented Reality (KELVAR)*, pages 1–4, 2017. 6

[26] Dong-Hee Shin. The role of affordance in the experience of virtual reality learning: Technological and affective affordances in virtual reality. *Telematics and Informatics*, 34(8):1826 – 1836, 2017. 6

[27] Kuldeep Singh, Shival Srivastav, Abhishek Bhardwaj, Abhinav Dixit, and Sanjeev Misra. Medical education during the covid-19 pandemic: A single institution experience. *Indian Pediatrics*, 57(7):678–679, 2020. 5

[28] Robert Ubell. *Virtual Teamwork Mastering the Art and Practice of Online Learning and Corporate Collaboration*. Wiley, 2011.

[29] Michael Vallance, Phillip A. Towndrow, and Charles Wiz. Conditions for successful online document collaboration. *TechTrends*, 54(1):20–24, 2009. 6