PPVR: Teaching and Learning to Play the Piano in Virtual Reality*

Xizi (Lucy) Wang
University of Michigan
Ann Arbor, MI, USA
xiziwang@umich.edu

ABSTRACT
PPVR is a collaboration system that supports remote piano training in a virtual reality environment with the support of synchronous and asynchronous communication, various viewpoints, and detection of finger pressure. We designed two different user interfaces for the learner and the instructor, and provide different viewpoints and scaffolding tools for the learner. We also explored the usage of pressure sensing in VR to enrich the information transmission during the remote piano training. We designed a user study, but were unable to conduct it due to COVID-19.

Keywords
Remote Collaboration; Virtual Reality; Instrument Training; Wearable Computing; Pressure Sensing

INTRODUCTION
In April 2020, a news website has shown severe global impacts caused by the COVID-19, such as social distancing restrictions, public places being shut down, and even lockdown [3]. This pandemic has switched the traditional education in many countries to the remote one. There is a concern that a pandemic might happen again in the future, and remote education may become the norm in the future. Besides epidemics, there are other reasons causing the teacher or the student not to be able to be present to teach and learn an instrument. Some could be financial reasons, such as expensive in-person instrument learning or transportation. Some could be personal problems, like time conflicts. Others could be culture contractions; for example, a female is forbidden to be outside alone without a male relative in Saudi Arabia [4]. These problems have obstructed the participation of offline lessons, especially in-person piano lessons.

Although there exist many substitutes of the offline piano lesson, such as phone-based online piano lessons or YouTube videos, in-person piano lessons have shown great advantages on the following aspects: more guidance, easier correction process, adjustments of learning pace, and so on [10]. Also, people have shown great interest in remote collaboration systems for different physical task training including but not limited to instrument learning and film making [5, 7]. Besides the software, they also investigate haptic technologies that either manipulate hands or collect hand data to facilitate physical tasks learning [11, 12].

In this paper, we propose PPVR, a VR collaboration system designed to support remote piano training from different angles. We explored applications of learning theories and pressure sensors in VR, aiming to provide learners and instructors with a cognitive piano learning experience in VR. The system was designed and developed to support both synchronous and asynchronous communications, various viewpoints for the learner to observe the instructor's actions, and a wearable glove that senses the pressure of fingers, then visualizes the data across the user interface.

We want to investigate the following research questions:

1. Will the flexibility of viewpoints and the application of pressure sensing facilitate remote collaborative piano learning and teaching?

2. What are the essential functions that the learner and the instructor use to teach and learn piano playing in a VR collaboration system?

We designed a user study and purchased the equipment and planned to conduct the experiment to answer these questions.

The main contributions of this paper are:

- A design of an immersive environment to support both synchronous and asynchronous learning and teaching;
- Exploration of various interaction technologies based on learning theories to facilitate constructive learning;
- A wearable device for pressures sensing to improve the holistic piano learning experience in VR;
- A design of the user study for answering these research questions.

RELATED WORK
To better understand the existing systems and design our novel one, we conducted a literature review related to collaboration

* This work was done under the supervision of Dr. Michael Nebeling (the advisor and committee chair) and Dr. Steve Oney (committee member), who are professors at the University of Michigan School of Information.

systems, instrument learning systems with wearable devices, and learning theories.

Collaboration System

Previous studies have been exploring remote collaborative systems in various interfaces, such as AR, VR, and websites. Kumaravel et al. proposed Loki, an MR system that enables the user to be telepresent and communicate in different views synchronously or asynchronously through real-time annotations and recorded videos to support physical task training [5]. Their flexibility of switching between AR and VR views helps the learner to set up the most comfortable learning environment. Amores et al. developed a mobile collaborative system named ShowMe that allows a peer to communicate with each other through captured videos, audios, and hand gestures [2]. Nebeling et al. presented a role-based collaborative AR/VR system for designing scenes and producing firms [7]. These systems [5, 2, 7] provide various ways to enable nonverbal communication through annotations or projection of body languages (e.g., body movement and posture, and hand gestures) and diverse options of adjusting views, which has inspired us to explore the other viewpoints to leverage communication during collaborative piano training.

Instrument Learning System and Wearable Devices

Researchers also explored both the software and the hardware system to support instrument learning. For example, Rogers et al. [9] have developed an AR system named P.I.A.N.O. that projects color and notes on physical piano keys to reduce workload and support individual piano playing. Tamaki et al. [11] designed a system that uses electrical muscle stimulation (EMS) to track hand gestures and manipulate learners' finger muscles to facilitate Japanese instrument learning. Just this year, Wijaya et al. [12] presented a VR piano platform that uses two Leap Motions to capture hand gestures and pressure sensors to check if a key is pressed to enrich the piano playing experience; the result has shown that the application of pressure sensors help piano learning in VR, but still left an unknown in the piano teaching or collaborative learning.

Learning Theories

In order to better design a learning technology, we also investigated two learning theories: embodied cognition [1] and scaffolding [8]. We found that there are three different ways people learn actively and constructively: producing actions, observing the export's actions, and imaging producing actions, and these make learners more engaging and help them make connections with new knowledge during the learning process [1]. We also found that two areas of scaffolding that could help us make better connections and reduce workload: sense-making and process management. Sense-making uses representations to bridge learner's understanding of a concept or allow the learner to inspect the multiple views of the same object; process management provides guidance by using functional modes to constrain the space of activities, especially for beginning learners [8].

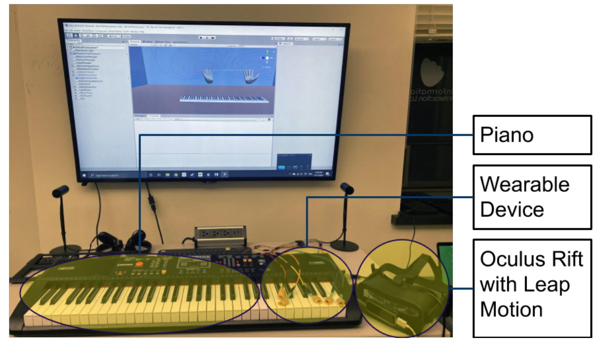

Figure 1. PPVR facility setup for one user.

PRELIMINARY STUDY

In the early stage, we conducted a preliminary study with 4 undergraduate students from the University of Michigan School of Music, Theatre & Dance who had piano skills to understand how they learned to play the piano as a beginner and what parts they think were the most difficult during the learning process. The result showed that the majority of students took in-person piano lessons, and only one learned through YouTube. All of them thought that it was difficult for beginners to remember the position of the notes. Also, two students said that learning hand gestures or how to control their muscles to play are some of the reasons they took offline lessons. All three students believe that, compared to online lessons, in-person lessons provide views from different angles so they can observe the instructor's hand more closely. Based on the results of these informal interviews and learning theories that we surveyed, we decided to design a platform that enables the visualization of hand gestures and hand pressures, provides scaffolding tools for keyboard memory and offers different viewpoints for cognitive learning.

SYSTEM DESIGN AND IMPLEMENTATION

The entire collaboration system requires 2 VR headsets, 2 leap motions, 4 hand haptics (2 haptics for each user), and skype connection, while physical pianos may be optional to leverage the piano playing experience. Our system setup is shown in Figure 1.

Hardware

The way the piano player presses the key will determine the quality of the song. To our current knowledge, most remote piano teaching and learning system do not reveal the information about how hard the player presses. To inspect the pressure the player put on the key, we designed a wearable hand device to detect the applied force. Each user wears 2 hand haptic device. Each device comes with an Arduino (Model: Arduino Uno R3) and five connected force sensitive resistors (Model: SEN-09375 ROHS) (see Figure 2). Each pressure sensor can detect 100g to 10kg force. It is wrapped inside of an elastic skin-color cot (see Figure 3), so these sensors can be fixed under the fingertip when the user is wearing the device, and the error of hand gesture recognition can be minimized. As shown in Figure 4, each wearable device tracks one hand's

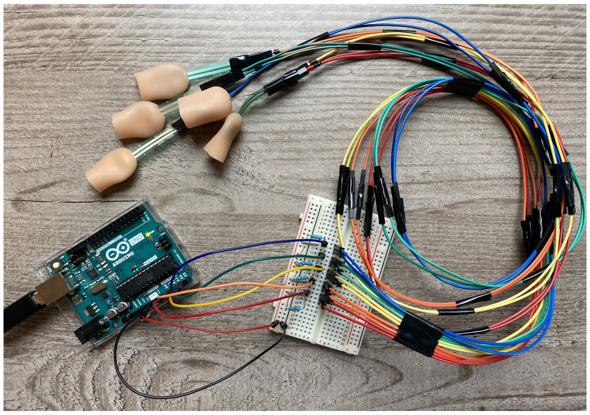

Figure 2. Components of one wearable device.

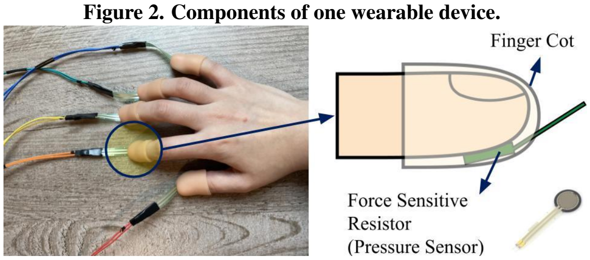

Figure 3. The structure of the fingertip part of the wearable device.

five fingertip forces, and the raw data will be packaged and transmitted to the Unity system using the serial port.

We also use a Leap Motion installed on the front of the VR headset to capture either the learner and instructor's hand gestures. Leap Motion creates both regular hand models (appearance of hands) and rigid hand models (RigidBody and BoxCollider objects) in the scene, allowing the user to see their hands and interact with the virtual environment. Then, we share the hand gesture data (both the regular hand model and the rigid hand model) via a server and cast real-time 3D hand models in the virtual environment for the local user to visualize Figure 4.

Software

The PPVR platform was developed in Unity with the support of the Mixed Reality Toolkit. The learner and instructor use two different interfaces that are alike in some aspects. The environment was designed the same, and in the default mode, the instructor and the learner share the same virtual piano, which creates a sense of presence. The system enables visual and audio communications. The student and the instructor can observe each other's real-time hand gestures from different angles and applied forces; at the same time, they also share the sound of a pressed piano key and oral voices. Besides the same visualization functions mentioned above, the instructor and the student interfaces have different functionalities. The instructor can record performance and share the recording. The students can receive the recording and replay it if they cannot catch up a complex instruction; in addition, they can switch viewpoints to match their own learning styles, and choose to use scaffolding technologies based on their piano skills.

Data Transmission

A server was set up by using Note.js on a laptop and following the Socket.IO protocol to transmit the data between two clients (student and instructor). The client-side uses a plugin named Best HTTP to send and receive the message. As the Figure 4 shown, there are three types of data that will be shared between two clients: hand gestures, finger pressures, and oral audios. Hand gestures, including positions and rotations of essential hand points, and pressures are packaged in the Unity, transmitted through the server, received by another Unity system, and then used to conduct a shared environment. The oral voices are transmitted externally through Skype.

Function 1: Synchronous and Asynchronous Teaching and Learning

PPVR allows the learner and the instructor to share the same view on the virtual learning environment and calibration system. In the default mode, they share the same virtual piano and observe each other's hand gestures in real-time, although their hand cannot physically feel touched (see Figure 5). The position and rotation of essential points (palms and fingers) of any capture hands from each client will be shared between two clients, so the other side can use this data to reproduce the hand model.

The system also enables asynchronous communication by supporting the instructor recording the performance every single frame, and the learner replaying the action. On the instructor's side, they can record continuous hand gestures and send it to the learner. On the learner's side, they would be able to receive the recording transmitted through the server and replay it by pressing the button (see Figure 6). Also, the learner would be able to stop recording and resume from where they left off; and they can drag the process bar to adjust the viewing progress (see Figure 6). The flexible replaying function was designed to help the student follow up with the instructor's fast actions.

Function 2: Flexible Viewpoints

Research has shown that providing a flexible view angle will help to learn, and the side view might be the best viewpoint [6]. Based on these findings, we designed multiple views to empower learners. The system enables the learner to visualize the instruction in three different ways: shared view, bird-eye view, and mirrored view (see Figure 7). The shared view allows the student and the instrument to share the same view on the virtual environment. The environment, including the position of the piano, on both sides is set to be the same. Since the calibration is also the same, hands from the other side can be casted in the same position, which makes the virtual environment sharing work. The bird-eye view sets the second camera fixed on the rear-right side of the instructor's side and displays what was captured on a 2D canvas of the student interface. This view enables the learner to observe the instructor's performance from the side-view angle, like how they watch a piano learning video on YouTube. The mirrored view was designed to simulate a similar learning process in the dance studio where the student can observe how the teacher dances by looking at the mirror. In PPVR, the learner can practice playing the piano in mind by watching the instructor's

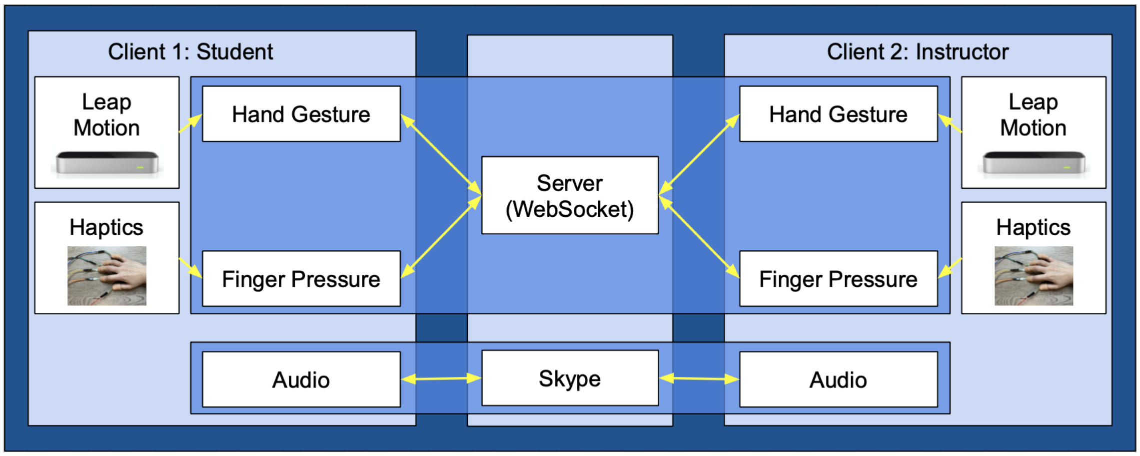

Figure 4. The overview of how data is transmitted in the PPVR system.

Figure 5. The student (left) is playing the piano, and the instructor (right) is observing.

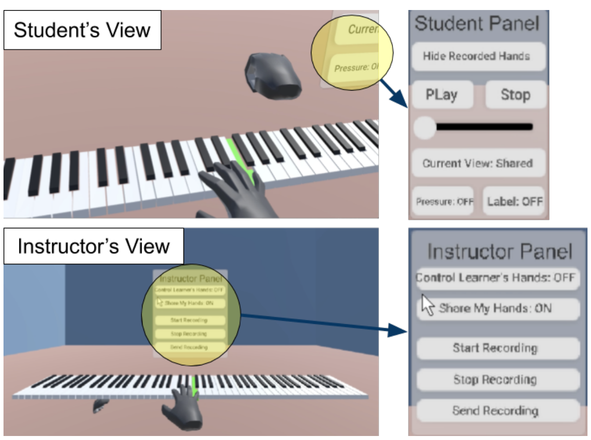

Figure 6. The views of the student replaying a recording and the instructor recording the action.

hand gestures moving in the "mirror," following the action and exercise in mind. To implement this, rotation and offset are used when the system is recasting the other user's hand.

Function 3: Scaffolding Tools

With the help of the wearable device, the system can track the user's fingertip force, and users are able to visualize the invisible force in the virtual environment. The system will display 9-level degrees of applied pressures in gradient color – the darker the color is, the greater the force is tracked (see Figure 8). In addition, the wearable device enables real piano playing experience as an option. With it, the user has to press their fingertip to make a sound, which avoids an accident virtual touching and, at the same time, indicates which finger was used to press the key.

In addition, labels and markers are designed in the system to notify the changes and reduce cognitive demands (see Figure 9 and Figure 10). A pressed piano key will be highlighted in green. To reduce the workload, the learner has the option of showing or hiding labeled note names.

USER STUDY DESIGN

In this work, we have purchased equipment, including two physical pianos, and designed the user study. Unfortunately, we were unable to execute the plan due to the COVID-19 outbreak situation. We planned to recruit 12 participants with no piano skill requirements and build piano learning experience around the physical keyboard. We will form 6 groups by randomly pairing these participate. For each pair, one person will be randomly selected to become a student, and another one will behave as an instructor. An easy song will be chosen to learn and teach, so the beginning piano player will not be stumped by the task. The same song that will be divided into two parts and used in the two experiments controls the variable (the difficulty of teaching and learning a song) among all experiments. Each pair will experience the system for 20 minutes; during the time, the instructor needs to teach the students the first part of a song, which will be prepared and casted in the virtual environment. Then, the same pair will switch the role and do the experiment again (teaching and learning the second part of the same song). The general demographic information and piano skills will be collected through a presurvey. The experiments will be recorded by setting up a camera behind each participant and collecting keys pressed and usage of functions in the system. After the two experiments, participants will be interviewed and asked questions related to playing experience and difficulty met during the task (e.g., whether they would be able to receive enough cues to learn or teach to play the piano, like how they do in person.)

Figure 7. Three different viewpoints on the student side: shared, bird-eye view, and mirrored (from left to right).

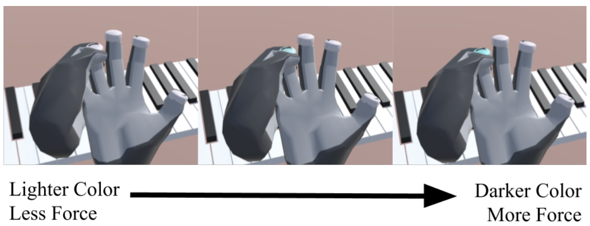

Figure 8. Visualization of pressure, having one left hand pressing the right hand's ring finger.

Figure 9. Pressed keys are be highlighted in green.

Figure 10. Keys on the virtual piano are labeled.

We define some metrics that will be used to answer the two research questions proposed at the beginning:

- The degree of completion - whether the learner has finished the task and how far the learner has achieved;
- The quality of the task, including correctness, tempo, and quality of the song played;
- The usage of each functionality - by collecting and calculate the amount of time and how often each function was used to teach or learn;
- The qualitative evaluations about the system that is collected from the interview

DISCUSSION

For the first research question, although we have not conducted the experiment, we expect to see the visualization of hand gesture and pressure helps the student and the learner to observe and evaluate the performance while they are not touching each other physical hands. Based on the previous studies [5, 6] and the internal testing we did, we also conjecture that the diversity of viewpoints will help collaborative piano training in VR. For the second question, based on our own experience, we believe that the shared viewpoint, pressure sensing, and note labels would be often used in remote piano training.

LIMITATIONS AND FUTURE WORK

Through our internal testing, we have found some limitations of the systems:

- Calibration – currently, our system does not map the height of the surface, such as the physical piano, to the virtual one. This increases the time of virtual piano setup once we move VR devices to a different place. In the future project, we believe that a marker pasted on an instrument can be used for calibration.
- Hand tracking – the complexity of the hand gesture and the color of the skin affect the correctness of the recognition. Using one Leap Motion does not always track hand position and gesture correctly sine some fingers may be covered by another, and the depth camera cannot detect them. This sometimes causes a virtual hand model to rotate or being twisted. The lack of stability may cause nausea during playing. One solution for this problem would be using two Leap Motions at the same time to recognize hand gestures.
- Pervasiveness – PPVR requires a lot of equipment and setup time. At the same time, the purchase of PC-connected VR devices that are still not uncommon to see in daily life. It would be hard to see the current system can be universally used in remote piano training.

There exists a rich space for us to make a better collaborative teaching and learning system in the future. One direction is ubiquity. We can continue to explore how to make the remote learning system in VR more ubiquitous, e.g., decreasing set-up time and size/amount of equipment, while maintaining the same functionalities. The second direction would be extending the current system to a generic one for all instrument learning and teaching. We could explore common features that occurred in different training activities and design a system that satisfied most needs. Another area would be the richness of information during remote teaching and learning. We could explore the other information – not just what we see and hear, but also what our hands feel and what our nose smell – that we can collect and transmit across systems to build a rich collaborative system that offers multimodal interactions. For example, forces could be transmitted across the system and applied on the learner's hands remotely to simulate a teacher manipulating a student's hand during hand-over-hand piano training.

CONCLUSION

In this paper, we reported on previous studies, a preliminary needfinding study, and the design and prototyping of a VR collaboration system named PPVR with wearable pressure sensors for remote piano training based on the findings. The system detects forces applied on the surface by each finger and shares the visualization of pressures applied across the learner and the instructor to make up for the lack of touching in distance education. The platform enables both synchronous and asynchronous communication, supports three different viewpoints, and provides labels and markers to scaffold learners with different piano skills. Finally, although we did not conduct a formal user study, based on previous research and internal testings, we conjecture that the pressure sensing and flexibility of viewpoints would help collaborative piano learning in VR.

ACKNOWLEDGMENTS

We would like to thank four music school students for being informally interviewed.

REFERENCES

[1] Martha W. Alibali and Mitchell J. Nathan. 2018. Embodied cognition in learning and teaching: Action, observation, and imagination. In *International Handbook of the Learning Sciences*, Frank Fischer, Cindy E. Hmelo-Silver, Susan R. Goldman, and Peter Reimann (Eds.). Routledge, New York, Chapter 8, 75–85. DOI:http://dx.doi.org/10.4324/9781315617572

[2] Judith Amores, Xavier Benavides, and Pattie Maes. 2015. ShowMe: A Remote Collaboration System that Supports Immersive Gestural Communication. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, Seoul, CHI 2015 Extended Abstracts, Republic of Korea, April 18 - 23, 2015*. 1343–1348. DOI: http://dx.doi.org/10.1145/2702613.2732927

[3] Maria Cohut. 2020. COVID-19 global impact: How the coronavirus is affecting the world. *Medical News Today* (24 4 2020). https://www.medicalnewstoday.com/articles/covid-19-global-impact-how-the-coronavirus-is-\affecting-the-world

[4] Amani Hamdan. 2004. Women and education in Saudi Arabia: Challenges and achievements. *International Education Journal* 6 (11 2004).

[5] Balasaravanan Thoravi Kumaravel, Fraser Anderson, George W. Fitzmaurice, Bjoern Hartmann, and Tovi Grossman. 2019. Loki: Facilitating Remote Instruction of Physical Tasks Using Bi-Directional Mixed-Reality Telepresence. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology, UIST 2019, New Orleans, LA, USA, October 20-23, 2019*. 161–174. DOI: http://dx.doi.org/10.1145/3332165.3347872

[6] Mamoun Nawahdah and Tomoo Inoue. 2013. Setting the best view of a virtual teacher in a mixed reality physical-task learning support system. *J. Syst. Softw.* 86, 7 (2013), 1738–1750. DOI: http://dx.doi.org/10.1016/j.jss.2012.08.060

[7] Michael Nebeling, Katy Lewis, Yu-Cheng Chang, Lihan Zhu, Michelle Chung, Piaoyang Wang, and Janet Nebeling. 2020. XRDirector: A Role-Based Collaborative Immersive Authoring System. In *CHI '20: CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, April 25-30, 2020*. 1–12. DOI:http://dx.doi.org/10.1145/3313831.3376637

[8] Chris Quintana, Brian J. Reiser, Elizabeth A. Davis, Joseph Krajcik, Eric Fretz, Ravit Golan Duncan, Eleni Kyza, Daniel Edelson, and Elliot Soloway. 2004. A Scaffolding Design Framework for Software to Support Science Inquiry. *Journal of the Learning Sciences* 13, 3 (2004), 337–386. DOI: http://dx.doi.org/10.1207/s15327809jls1303_4

[9] Katja Rogers, Amrei Röhlig, Matthias Weing, Jan Gugenheimer, Bastian Könings, Melina Klepsch, Florian Schaub, Enrico Rukzio, Tina Seufert, and Michael Weber. 2014. P.I.A.N.O.: Faster Piano Learning with Interactive Projection. In *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces, ITS 2014, Dresden, Germany, November 16 - 19, 2014*. 149–158. DOI: http://dx.doi.org/10.1145/2669485.2669514

[10] Joshua Ross. n.d. Online Piano Lessons Vs In Person. *Joshua Ross* (n.d.). https://joshuarosspiano.com/online-piano-lessons-vs-in-person/

[11] Emi Tamaki, Takashi Miyaki, and Jun Rekimoto. 2011. PossessedHand: techniques for controlling human hands using electrical muscles stimuli. In *Proceedings of the International Conference on Human Factors in Computing Systems, CHI 2011, Vancouver, BC, Canada, May 7-12, 2011*. 543–552. DOI: http://dx.doi.org/10.1145/1978942.1979018

[12] Febrina Wijaya, Ying-Chun Tseng, Wan-Lun Tsai, Tse-Yu Pan, and Min-Chun Hu. 2020. VR Piano Learning Platform with Leap Motion and Pressure Sensors. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops, VR Workshops, Atlanta, GA, USA, March 22-26, 2020*. 585–586. DOI: http://dx.doi.org/10.1109/VRW50115.2020.00143