How Audio Versus Visual Cues Affect the Ability to Learn Rhythm

DYLAN NIERMAN, Colorado State University, The United States of America
DYLAN ARMSTRONG, Colorado State University, The United States of America
BENJAMIN SAATHOFF, Colorado State University, The United States of America
SHAMSHER SINGH SODHI, Colorado State University, The United States of America

In this paper, a virtual environment was designed to collect data. The virtual environment is a two-drum 3D space, the drums replicate the functionality of a traditional, physical drum-set. The experiment aims to collect data that will be used to analyze the differences in success rates of attempting to play rhythms using alternative ways; those being: visual images corresponding to rhythm along with audio, and without audio. This experiment is conducted using an Oculus Quest 2 headset in an environment created with Unreal Engine 5. The data that is found from this experiment could result in new ways for music educators to teach young students how to read and play music.

Additional Key Words and Phrases: music, VR, rhythm, learning

ACM Reference Format:
Dylan Nierman, Dylan Armstrong, Benjamin Saathoff, and Shamsher Singh Sodhi. 2023. How Audio Versus Visual Cues Affect the Ability to Learn Rhythm. 1, 1 (May 2023), 11 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

In music education, two primary modes of learning are used: Auditory and Visual learning. Some people find visual aids more helpful than audio, and vice versa. But there are diverse ways that musical rhythms can be expressed visually, other than traditional sheet music such as a step sequencer. This experiment aims to find the success rates of learning rhythms with different visuals and audio playback to see which way of learning the best is using a drum set created in virtual reality. The ability to learn and replicate rhythms is a fundamental skill in many musical traditions and is an essential component of musical performance. Studying the success rate of these different visual aids for learning rhythms could have practical implications for music educators and may change the way students could learn music. While sheet music will still be the standard way music is learned, rhythm wheels and step sequences may help students struggle with the complications of sheet music.

Our experiment goes through 10 different rhythms and participants are required to go through them, attempting to replicate them as accurately as possible. Playing a rhythm perfectly on the beat is very challenging, so there is lenience with the accuracy. We discuss the results and what the data might mean, along with what should be changed about the experiment if it were to be done again.

Authors' addresses: Dylan Nierman, dylan14@colostate.edu, Colorado State University, Fort Collins, Colorado, The United States of America; Dylan Armstrong, dylan2@colostate.edu, Colorado State University, Fort Collins, Colorado, The United States of America; Benjamin Saathoff, bensaat@colostate.edu, Colorado State University, Fort Collins, Colorado, The United States of America; Shamsher Singh Sodhi, thesham@colostate.edu, Colorado State University, Fort Collins, Colorado, The United States of America.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
© 2023 Association for Computing Machinery.
Manuscript submitted to ACM

2 RELEVANT WORKS

Learning styles can vary widely and some people consider the ability to retain information to be subjective to the individual. There are two big styles of learning music, visual and auditory, though there may be extra factors that apply to the learning process as well. Current research tends to show that different forms of sensory input may increase the ability to learn rhythms [5, 11, 14, 22, 27]. From this research, we can see an increased ability to learn solely on auditory and visual representations of environments and music. For example, research shows that the eyes or visual cues are the most effective, as that is how the brain interprets an environment. However, additional sensory inputs trend towards an additive relationship when increasing the capacity to remember an environment [11]. Additionally, the ability to perceive a visual and audio relationship has been applied to AI to create a robot accompanist that responds to changes in a musician's behavior and sound, adapting to match their rhythm and sound [27]. There are some caveats, however, because some research has shown the ability to learn short sentences or short melodies is tied to the preferred learning style of the individual who is participating [22]. Though the research seems more general, there are some cases where there is a larger benefit from the preferred learning style. The research states that formally trained musicians may have a distinct advantage in the auditory field of learning due to their extra experience [5]. On the other side, we can also see an increase in the ability to learn rhythms through a visual system and less experienced musicians [14]. The "SeeGroove" system shows how inexperienced musicians have an easier time learning specific rhythms through visual input, rather than auditory [14]. The research shows how specific styles of learning may be a form of the individual's adaptation to their own experience. From this we can derive that additional sources of input may help with increasing the ability to learn, but it will come down to the individual's preferred learning style.

Even without specifically investigating the difference between the specific learning styles, current research shows how VR environments and music can help with memory improvement and improve general psychological conditions [1, 9, 10, 12, 34, 35]. Different forms of visual stimuli can help improve mood, productivity, and ability to enhance general well-being [1]. Additionally, the visual stimuli can further be improved to advance further in the field of well-being and memory improvement. Research indicates that VR is effective for memory management, however, it's difficult to tell whether it is more effective or less effective than current environments [9]. While difficult to tell, VR can provide statistics automatically that otherwise may not be available, showing the effectiveness of measuring learning and memory management [9]. Research also trends toward a better user experience leading to a more effective and enjoyable learning environment. Increasing the user experience promotes creativity and exploration of the tools at hand, despite any frustrations that happen along the way [10]. The studies at hand demonstrate another congruence between music and a reduction in general anxiety in VR [35]. Different studies also show different mentalities that arise when immersed in VR. These are the action mentality that prioritizes physical movements and interaction, the game mentality that focuses on accuracy and speed, and the musical mentality the focuses on responses to rhythm [12]. Through this, we can determine that not only is the combination of music, learning, and VR effective in memory management, there are further fields of study, such as psychology, where the concepts apply the same. Additional data is also demonstrating that there are improvements in hand-eye coordination and reaction time in musicians learning in VR [34].

Current experimental research is showing how VR itself can be an effective tool in reforming how music education is done today [3, 13, 17, 36]. Music is a difficult skill to learn, however, current VR technology and its advancements create an alternative learning environment that can be effective in learning music for the first time [36]. There are a few benefits that are included in having a VR environment to learn music. A few examples include teaching rhythmic skills to young children, teaching music production, as well as providing an environment to practice with others, without needing to be together [36]. Though these benefits are good, some more advancements in the design of such programs in VR may need to be further ahead to provide the most effective environment for music education. Additionally, VR is trending toward being a more efficient tool for music education [3]. The extent of the potential of VR education is unknown, but the 3 programmabilities and customizability of VR programs in teaching can make it a far more efficient tool when learning music, teaching core concepts such as proper breathing, resonance, and others [3]. VR can also be used as a tool for music education, without needing to get a physical instrument itself [17]. Through the use of haptic feedback, it is possible to create a real feeling of the instrument, while there is no physical instrument. This is demonstrated through the "Air Piano", which was developed using haptic feedback to create a real feel of the piano [17]. Additional research also shows how VR could potentially be useful for ear training to recognize musical pitches [13]. These tools have been demonstrated to be successful at a core level, showing how they may be effective at reforming musical and audio education. All of these tools combined show the potential effects of VR on music education, and its potential to replace or be used alongside traditional music education.

In a more general sense, current research is also investigating the effectiveness of using VR in many fields and seeing success [2, 24, 31–33, 37]. VR has some decent potential in the atmosphere of learning, specifically when it comes to presentation. For example, it can help students see 3D models of objects and manipulate them within VR to understand different concepts such as the rotational force of the Earth. [2] It's also possible for VR to increase the number of senses used while learning, allowing the material and knowledge to be more meaningful and clear [2]. From this, it's clear to see some of the potential benefits of creating and using VR environments for education. Additionally, VR can be used to create a more personal learning experience when learning a second language [24]. Due to the different learning styles of students, one method of teaching may not be effective. However, through personalizing the content to the user in VR, a far more effective learning environment can be created [24]. In theory, this may help with learning the second language faster, but also in a calmer and potentially more effective environment. VR has also been demonstrated to increase the ability to learn concepts when compared to power-point slides. In a study done between VR and PowerPoint slides learning, students were seen to answer questions much more effectively when they learned in the virtual environment. They were able to understand the core concepts much easier and answer the ensuing questions more accurately, while more specific questions were more difficult to answer [37]. Additionally, in some circumstances, 360 video VR has also been shown to be effective for educational purposes [32]. It has been demonstrated that these VR environments are good for knowledge retention of core concepts within the scope of VR [32, 37]. It is possible that including VR in younger generations of students, can increase their capability to learn in STEM and learning problem solving early [31, 33]. All of these different fields combined show the potential of VR as a learning tool, not only in a musical sense, but in a more general sense as well.

In recent years, numerous studies have investigated rhythm perception and processing in relation to learning styles and sensory inputs, particularly auditory and visual modalities. Patel and Iversen (2014) and Hove et al. (2014) explored the neuroscience behind musical beat perception and the role of bass-ranged instruments in establishing rhythms, respectively [16, 30]. These studies align with the notion that auditory and visual sensory inputs, as well as individual learning styles, play crucial roles in musical learning [5, 11, 14, 22, 27]. Su and PÃ¶ppel (2012), Iversen and Balasubramaniam (2016), and Petrini and Carneiro (2016) examined body movement's role in auditory perception, synchronization with external rhythms, and the role of vision in processing rhythmic information, respectively [7, 19, 38]. Additionally, Koelsch (2010) delved into the neural basis of music-evoked emotions, highlighting another important aspect of music perception and learning [21]. These findings resonate with the concept of using additional sensory inputs to enhance learning capabilities while recognizing individual preferences in learning styles.

As technology advances, virtual reality (VR) has emerged as a potential tool for music education, with studies demonstrating its potential for memory improvement and psychological well-being [1, 9, 10, 12, 34, 35]. The integration of VR with music learning has shown promise in various applications, such as teaching rhythmic skills, music production, and facilitating practice sessions in virtual environments [36]. Moreover, VR's customizability and the incorporation of haptic feedback create realistic instrument simulations that contribute to the efficiency and effectiveness of music education [3, 13, 17]. Consequently, the collective knowledge from rhythm perception and processing studies, learning styles research, and VR advancements has the potential to revolutionize the landscape of music education.

Though VR can seem like a very promising technology when it comes to education, there are always trade-offs for using different forms in almost anything. Current research investigates some of the potentials and challenges that are faced when considering VR in education [4, 8, 15, 25, 28, 40]. Since VR is an emerging technology is is essential to see what sort of challenges arise when considering its use in education. For example, VR is heavily associated with gaming, which could potentially lead to it being less useful while learning [40]. It also requires some stronger computers to handle the intense graphics, as well as individual headsets to actually even use VR in the first place [4, 40]. Additionally, depending on implementation, VR can have an immense effect on the cognitive load that people use to learn with [25]. Since the brain doesn't have unlimited resources and VR introduces a lot of different variables that may alter what is actually being learned versus what is targeted to be learned [25]. Because there are also so many levels to designing VR education, it can be difficult to actually consider usability in classrooms [4]. This is mostly due to the multi-sensory inputs that must be considered at each level, while in the real world, these properties are more intuitive.

After analyzing some of the challenges in VR, considerations for the potential of the technology can be made as well. While it is currently expensive as of now, prices for technology tend to go down, making it more affordable as time goes on [4]. It is possible that due to advancements made, the cost of both the computers and headsets required may be lowered to decrease the barrier of entry required. Additionally, due to the environments created, VR education can increase the personalization of environments and scenarios. For example, doctors and dentists can practice a wide variety of situations, some of which may be rare, without needing an actual patient to practice on [40]. VR is also highly exploratory, which means that students can follow their own paths and interests while learning, but still manage to learn necessary information [15, 28]. This could potentially lead to more engagement from students, hypothetically meaning they could learn more, faster. With the potentials and challenges in mind, it seems that education is best completed in a more hybrid structure of classical curricula and VR-guided curricula [8, 15]. These potentials and challenges show what should be kept in mind when considering reforming education. While it may not be the best idea to swap to a VR-only system, there are potential benefits to having a balance between classical and VR education.

The impact of music on human emotions and performance has long been an area of interest in psychological research. A growing body of evidence suggests that music has the potential to significantly influence mood and productivity in various contexts [20, 39]. Thompson, Schellenberg, and Husain (2001) found that specific musical pieces can induce changes in arousal and mood, demonstrating the powerful effect music can have on our emotional state [39]. Similarly, Juslin and Västfjäll (2008) highlighted the importance of considering the underlying mechanisms that drive emotional responses to music, shedding light on the complex relationship between music and emotion [20].

Music has also been shown to have a direct impact on work performance and productivity [18, 23, 26]. Lesiuk's (2005) study demonstrated that listening to music while working could improve performance, with the optimal level of music engagement varying depending on the individual's personality and the task at hand [26]. In the realm of academia, Ziv and Hoftman (2016) discussed the association between music and academic performance, emphasizing the potential benefits of incorporating music into learning environments [18]. Furthermore, music has been found to play a significant role in sports and exercise, with athletes frequently using music to regulate their emotions and enhance motivation [23].

Despite the growing recognition of music's positive effects on mood and performance, it is essential to consider the individual differences and preferences that may influence these outcomes [6, 29]. For instance, Chamorro-Premuzic and Furnham (2007) discovered that individuals' personality traits could explain their music preferences and the ways in which they use music in everyday life [6]. Additionally, the so-called "Mozart effect" has been challenged by Nantais and Schellenberg (1999), who suggested that the observed improvements in cognitive performance might be an artifact of preference rather than a direct result of listening to Mozart's music [29]. This highlights the importance of understanding the personal and contextual factors that determine the extent to which music can influence mood and performance across different situations.

3 METHODOLOGY

For our experiment, we will be gathering 26 people to participate. This will be a within-subjects experiment. Everyone will be playing 10 different rhythms with 2 different visuals and audio playback (2 step sequences with audio, 2 sheet music with audio, 2 audio only, 2 step sequences only, 2 sheet music only). These separate ways of learning the rhythm are our independent variable.

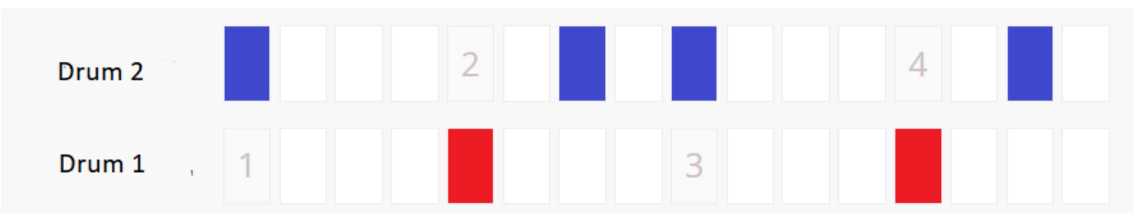

Fig. 1. Example rhythm shown as a step sequence, via Musicca (https://www.musicca.com/).

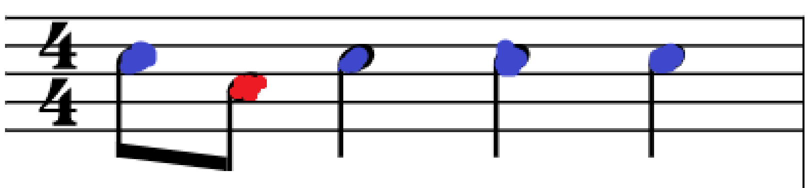

Fig. 2. Example of a sheet music rhythm. (Created by Dylan Nierman)

The participants will all be using the same VR headset (Oculus Quest 2) and the same drum set created in Unreal Engine 5. The drum set the user will be shown as two different drums, one is red and one is blue. The user can hit these drums just like they would hit a drum in real life since we are using a VR headset. Moving the controller in their hand into the drum will cause the drum to make a sound. The color of these drums corresponds to the color of the notes that we can see in Figure 1 and Figure 2. A red note in the rhythm will mean that the user should play the red drum. In front

Fig. 3. Colored drums and first rhythm shown to participants in Unreal Engine

of the user is a screen that shows the different visuals. There will be two different types of visuals; step sequences and sheet music, as shown in Figure 1 and Figure 2. For the audio playback parts, the user will be able to open a menu that has multiple buttons. One reads "Replay Audio", and pressing this button will replay the rhythm back for them. They can press this button as many times as they like.

When the experiment is started for a user, we decided that the visuals should be shown along with the audio playback of the function. Since our participants are all going to be people with very little to no music experience, we need to show them how these visuals work with support from the audio.

While presented with the rhythm, the user is required to replicate the rhythm using the drums and playing at the required tempo. Each rhythm has a metronome playing along with it at 120 beats per minute. This tempo is slow enough that the user will not be overwhelmed by the speed of the rhythm, but also quick enough that playing the rhythm on the correct beats will not be difficult. We are able to measure the accuracy of the users playing by keeping track of the time at which they hit the drum. If the times of these hits are close enough to the actual beat, it will count as successful. The user must hit all of these beats at the correct time. With this measurement, we have implemented a timing forgiveness gap. This is because hitting a drum at the exact time a beat happens is nearly impossible, especially having to do it for each note in the rhythm. For example, the program might have a note on beat 2. The user will not have to hit it exactly on beat 2, they have a range between 0.1 seconds before and 0.1 after the beat that they can hit. This forgiveness time may be adjusted for each rhythm, depending on the difficulty.

Our dependent variable for this experiment is the time it takes the individual to successfully play the rhythm shown. The rhythms shown to the participants will be beginner-level rhythms that should take no more than a couple of minutes to play accurately. In the case where the participant cannot play the rhythm at all, we will cap the time for the given rhythm at 4 minutes. After we conduct the experiment, we may be able to see that one of the different visuals has a lower average time. It is possible that we will also notice that every individual has separate times for every rhythm, meaning the visual does not matter at all, and there is no real trend in the data.

4 RESULTS

After conducting our experiment, the first thing we wanted to observe were the averages of each different rhythm delivery. Looking at the averages of each different type will be the first indicator as to which method is the best for learning rhythms.

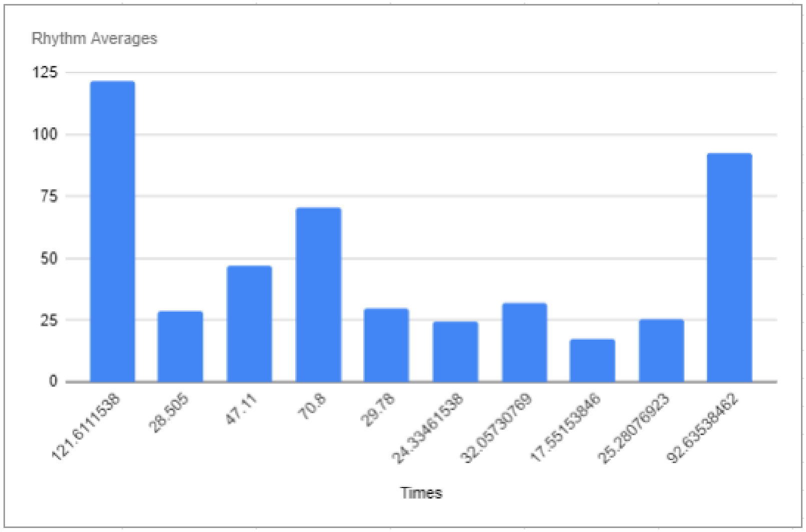

Fig. 4. Averages of each rhythm 1 through 10 respectively

The first method we had was the step sequence with audio. The average of the first rhythm was 121.61 seconds and the average of the second was 28.51 seconds. The first rhythm having a much higher average time to complete than the second one could be due to the fact that the participants were learning how to do the experiment and accurately play the drums. After completing the first rhythm, participants had a better idea of how to complete these rhythms, which is why we see a shorter time.

For the second method, the participants were shown sheet music along with audio. The averages for these rhythms were 47.11 and 70.8 respectively. While both of these are lower than the average for the first rhythm, they are both higher than the second step sequence rhythm, which could be a more accurate reading of how people perform with the step sequence.

The next method was audio only, which has two very low averages with 29.78 for the first and 24.33 for the second rhythm. By this point, participants had gone through 4 rhythms, so they all had learned how to approach the drums. Having no visual present could have helped since new visuals may still be confusing for them.

The last two methods were the visuals only. The first two rhythms were the sheet music only, and the averages for these were 32.06 seconds and 17.55 seconds. They ended up being lower than the other sheet music rhythms, but these are also later rhythms, so participants have learned how to do the experiment at this point, as stated above. The last two rhythms were the step sequence with no audio, and the averages for these were 25.28 seconds and 92.64 seconds. We see that the first rhythm has a similarly low average compared with audio-only and sheet music-only, but the second rhythm has an incredibly high average. This is the second-highest average, the highest being the first rhythm. The last rhythm was exactly the same as a previous rhythm they were shown as a step sequence, but it was missing two hits between beats 1 to 2 and 3 to 4 (The and of 1, and the and of 3). During the experiment, we noticed a lot of people struggle with this space, as they usually waited too long, treating the space like a whole quarter note instead of an eighth note.

We ended up doing two different ANOVA tests on our data. This was done because we wanted one that took into account every rhythm the participant played, and another one for the audio-only rhythms, and the visual-only rhythms. We think calculating this second p-value might get rid of some of the noise caused by participants learning how to do the experiment, which could have been a large reason the earlier rhythms have higher averages. We found that the p-value for all 10 rhythms across all 26 participants was <.001, meaning that our data was statistically significant. The p-value for our second test was also <.001. While both of these values show that our data is significant, there are a few factors that could be affecting our data.

5 DISCUSSION

After observing the results of our experiment, the data gives some interesting insights as to what methods of visualization people prefer. As explained in the results, the lowest average times with visuals are the sheet music rhythms. This may be because sheet music is the standard for reading music so even people with no musical experience might have some idea of how to read sheet music. One thing to take into account is that because all of these rhythms are different, some people might find different rhythms to be challenging and others very easy. All of these rhythms are based on just eighth notes and quarter notes, so they are all very similar, but the rests (spaces between beats) could have caused confusion with some participants. Although these rhythms were all very similar, there was one rhythm that people seemed to struggle with, which was the first one, as shown below in figure 5. Since this rhythm is constant eighth notes at 120 beats per minute, many people struggled to play this one fast enough for many attempts. This factor is a reason we chose to do a second ANOVA with the last 6 rhythms and leave out the first 4 rhythms.

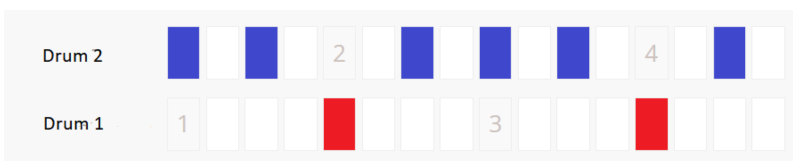

Fig. 5. First rhythm shown to participants

A factor that might have affected the data as well would be the program itself. Testing rhythmic accuracy through virtual reality is not the best way to approach the research topic, but it did allow us to see a realistic musical experience compared to pressing buttons on a keyboard or controller and calculating accuracy that way. In our program, we did not have 100% accuracy for the correct rhythms being played. This means that when a rhythm was played that was close enough to being perfect, it should have given a checkmark to go to the next level and be marked as successful, but this did not always happen. This meant that the participant had to keep replaying the rhythm until they got it right again. While this was not common in our experiment, it was something that did happen a few times which means it could have affected our data.

Although there were some slight problems with our implementation of VR drums, our data still gives us some insight as to what is best for people to learn rhythms. Our goal was to see if the step sequence visual was a better way for new musicians to learn rhythms, but our data is showing that that is not the case. Two of the highest average times to complete the rhythm were step sequence rhythms, which were the first one and the last one. Something that could have contributed to this is the lack of understanding of where the metronome beat lands in the boxes. In figure 5, we are able to see a 1, 2, 3, and 4 every 4 boxes. This means that each metronome note would happen at those intervals exactly. Our goal with playing audio along with the visual was to show participants where the beat is and how to interpret the rhythm, but due to their lack of musical understanding, this was not clear.

If we were to attempt to do this experiment again, there are a few things that should be changed. Although creating drums in virtual reality allowed for a unique experiment, this prevented us from truly gauging rhythmic accuracy in a short time frame. We could simulate the same experiment with a controller or other type of input device that would be more accurate. Another good thing to change would be to use a latin square to change the order of the rhythms, so each person is getting slightly different rhythms for each different visual/audio method. Since the averages tend to lower as the rhythms go on (except for the last one) we know that the later rhythms might not be an accurate assessment because participants had the earlier rhythms to learn. One last change that should be made is an introduction. In our experiment, participants were given some time to play around with the drums, but were not shown any rhythms and could not hear the metronome. An introduction level should be added that gives a rough overview of what the metronome means and how the step sequence and sheet music can be interpreted.

6 CONCLUSION

The goal of our experiment was to see if there was a better alternative to sheet music for learning rhythms. We conducted our experiment through a program created in Unreal Engine with an Oculus Quest 2. Through testing 26 participants, we weren't able to find a definitive conclusion that there is a better alternative to learning music than sheet music. The data suggested the participants struggled the most with the step sequence method due to much higher averages in learning the rhythms, indicating it likely isn't a good replacement for sheet music. The data also didn't suggest that audio in the rhythms increases the participant's ability to learn, as the average for visuals only ended up being lower than both audio and visual. One factor that could have affected this though is the participants learning how to do the experiment as they go. They are more likely to do worse earlier in the first rhythms since this is a new experience. Something that could have been done to fix this would be to change the rhythms for each person so no one gets the same order of rhythms and to also provide clear instructions on how to read the visuals so the order of method deliveries could be changed. Having visual only first and then audio and visual after could have given us different insights into the data. After the conclusion of the experiment, we weren't able to find significant data that suggests the step sequence method was better than sheet music for learning regardless of audio inclusion. This may be because of noise that might have been generated in our experiment. Some factors that might have affected the data would be the difficulty of the very first rhythm, minor program glitches, and unclear instructions on how to understand the metronome completely. Although with our current research, we didn't find significant enough data to determine a better method than sheet music, the inclusion of more advanced methods of research may be able to yield different results.

REFERENCES

[1] Stephen Alexanian, James DeSimone, Kai Koelsch, Soojin Kim, Kyogu Kim, and Marc Leman. 2022. Virtual Reality and Music's Impact on Psychological Well-Being. *Frontiers* (Jul 2022). https://doi.org/10.3389/fresc.2022.864990

[2] Hadi Ardiny and Esmaeel Khanmirza. 2018. The Role of AR and VR Technologies in Education Developments: Opportunities and Challenges. In *2018 6th RSI International Conference on Robotics and Mechatronics (IcRoM)*. 482–487. https://doi.org/10.1109/ICRoM.2018.8657615

[3] Hong-xuan Bian. 2016. Application of virtual reality in music teaching system. *International Journal of Emerging Technologies in Learning (Online)* 11, 11 (2016), 21. https://www.proquest.com/openview/8ca350f07858d22b90b3d5d6ca0338ad/1?pq-origsite=gscholar&cbl=5452619%7D,doi={10.

3991/ijet.v11i11.6039}

[4] Meredith Bricken. 1991. Virtual Reality Learning Environments: Potentials and Challenges. *SIGGRAPH Comput. Graph.* 25, 3 (jul 1991), 178–184. https://doi.org/10.1145/126640.126657

[5] Kirsten E. Smayda Casey L. Roark and BBharath Chandrasekaran. 2022. Auditory and visual category learning in musicians and nonmusicians. *Journal of Experimental Psychology* 151, 3. https://doi.org/10.1037/xge0001088

[6] Tomas Chamorro-Premuzic and Adrian Furnham. 2007. Personality and music: Can traits explain how people use music in everyday life? *British journal of psychology* 98, 2 (2007), 175–185.

[7] Antoni B. Chan and Janet H. Hsiao. 2016. Information Distribution Within Musical Segments. *Music Perception* 34, 2 (12 2016), 218–242. https://doi.org/10.1525/mp.2016.34.2.218

[8] Matt Cook, Zack Lischer-Katz, Nathan Hall, Juliet Hardesty, Jennifer Johnson, Robert McDonald, and Tara Carlisle. 2019. Challenges and strategies for educational virtual reality. *Information Technology and Libraries* 38, 4 (2019), 25–48.

[9] 'Ad'am B Csap'o, Istv'an Horv'ath, P'eter Galambos, and P'eter Baranyi. 2018. VR as a Medium of Communication: from Memory Palaces to Comprehensive Memory Management. In *2018 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, 000389–000394. https://doi.org/10.1109/CogInfoCom.2018.8639896

[10] Thomas Deacon, Graham Wakefield, Wayne Piekarski, Zoran Sterjovski, and Ian Livingston. 2017. User Experience in an Interactive Music Virtual Reality System: An Exploratory Study. *Researchgate* (2017). https://www.researchgate.net/publication/319861197_User_Experience_in_an_Interactive_Music_Virtual_Reality_System_An_Exploratory_Study%7D,doi={10.13140/RG.2.2.23776.46083}

[11] H.Q. Dinh, N. Walker, L.F. Hodges, Chang Song, and A. Kobayashi. 1999. Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. In *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)*. 222–228. https://doi.org/10.1109/VR.1999.756955

[12] R. Dongas and K. Grace. 2023. Designing to Leverage Presence in VR Rhythm Games. *Multimodal Technologies and Interaction* 7, 2 (2023), 18. https://doi.org/10.3390/mti7020018

[13] Connor Fletcher, Vedad Hulusic, and Panos Amelidis. 2019. Virtual Reality Ear Training System: A study on Spatialised Audio in Interval Recognition. In *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*. 1–4. https://doi.org/10.1109/VS-Games.2019.8864592

[14] Issei Fujishiro, Naoki Haga, and Masanori Nakayama. 2015. SeeGroove: Supporting Groove Learning through Visualization. In *2015 International Conference on Cyberworlds (CW)*. 189–192. https://doi.org/10.1109/CW.2015.65

[15] Caroline Graeske and Sofia Aspling Sjöberg. 2021. VR-Technology in Teaching: Opportunities and Challenges. *International Education Studies* 14, 8 (2021), 76–83.

[16] Michael J. Hove, Céline Marie, Ian C. Bruce, and Laurel J. Trainor. 2014. Superior time perception for lower musical pitch explains why bass-ranged instruments lay down musical rhythms. *Proceedings of the National Academy of Sciences* 111, 28 (2014), 10383–10388. https://doi.org/10.1073/pnas.1402039111

[17] Inwook Hwang, Hyungki Son, and Jin Ryong Kim. 2017. AirPiano: Enhancing music playing experience in virtual reality with mid-air haptic feedback. In *2017 IEEE World Haptics Conference (WHC)*. 213–218. https://doi.org/10.1109/WHC.2017.7989903

[18] Jonathan Impett. 2009. Making a mark. In *The Oxford handbook of music psychology*. Oxford University Press, USA, 403.

[19] John Rehner Iversen and Ramesh Balasubramaniam. 2016. Synchronization and temporal processing. *Current Opinion in Behavioral Sciences* 8 (2016), 175–180. https://doi.org/10.1016/j.cobeha.2016.02.027 Time in perception and action.

[20] Patrik N. Juslin and Daniel Västfjäll. 2008. Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences* 31, 5 (2008), 559–575. https://doi.org/10.1017/S0140525X08005293

[21] Stefan Koelsch. 2010. Towards a neural basis of music-evoked emotions. *Trends in cognitive sciences* 14, 3 (2010), 131–137.

[22] Lisa M. Korenman and Zehra F. Peynircioglu. 2007. Individual Differences in Learning and Remembering Music: Auditory versus Visual Presentation. *Journal of Research in Music Education* 55, 1 (2007), 48–64. https://doi.org/10.1177/002242940705500105

[23] Petri Laukka and Lina Quick. 2013. Emotional and motivational uses of music in sports and exercise: A questionnaire study among athletes. *Psychology of Music* 41, 2 (2013), 198–215.

[24] Michael Jongseon Lee, Adam Spryszynski, and Eric Nersesian. 2019. Personalizing VR Educational Tools for English Language Learners.. In *IUI Workshops*, Vol. 19.

[25] Ryan Lege and Euan Bonner. 2020. Virtual reality in education: The promise, progress, and challenge. *JALT CALL Journal* 16, 3 (2020), 167–180.

[26] Teresa Lesiuk. 2005. The effect of music listening on work performance. *Psychology of Music* 33, 2 (2005), 173–191. https://doi.org/10.1177/0305735605050650

[27] Angelica Lim, Takeshi Mizumoto, Louis-Kenzo Cahier, Takuma Otsuka, Toru Takahashi, Kazunori Komatani, Tetsuya Ogata, and Hiroshi G. Okuno. 2010. Robot musical accompaniment: integrating audio and visual cues for real-time synchronization with a human flutist. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 1964–1969. https://doi.org/10.1109/IROS.2010.5650427

[28] Fabrizia Mantovani et al. 2001. 12 VR learning: Potential and challenges for the use of 3d environments in education and training. *Towards cyberpsychology: mind, cognition, and society in the Internet age* 2, 207 (2001).

[29] Kristin M Nantais and E Glenn Schellenberg. 1999. The Mozart effect: An artifact of preference. *Psychological Science* 10, 4 (1999), 370–373.

[30] Aniruddh D. Patel and John R. Iversen. 2014. The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in Systems Neuroscience* 8 (2014). https://doi.org/10.3389/fnsys.2014.00057

[31] Nikolaos Pellas, Andreas Dengel, and Athanasios Christopoulos. 2020. A Scoping Review of Immersive Virtual Reality in STEM Education. *IEEE Transactions on Learning Technologies* 13, 4 (2020), 748–761. https://doi.org/10.1109/TLT.2020.3019405

[32] Johanna Pirker and Andreas Dengel. 2021. The Potential of 360° Virtual Reality Videos and Real VR for Education—A Literature Review. *IEEE Computer Graphics and Applications* 41, 4 (2021), 76–89. https://doi.org/10.1109/MCG.2021.3067999

[33] Maria Roussou. 2009. A VR Playground for Learning Abstract Mathematics Concepts. *IEEE Computer Graphics and Applications* 29, 1 (2009), 82–85. https://doi.org/10.1109/MCG.2009.1

[34] Sebastian Rutkowski, Mateusz Adamczyk, Agnieszka Pastuła, Edyta Gos, Carlos Luque-Moreno, and Anna Rutkowska. 2021. Training Using a Commercial Immersive Virtual Reality System on Hand–Eye Coordination and Reaction Time in Young Musicians: A Pilot Study. *International Journal of Environmental Research and Public Health* 18, 3 (Feb 2021), 1297. https://doi.org/10.3390/ijerph18031297

[35] Sofia Seinfeld, Ilias Bergstrom, Ausias Pomes, Jorge Arroyo-Palacios, Francisco Vico, Mel Slater, and Maria V. Sanchez-Vives. 2016. Influence of Music on Anxiety Induced by Fear of Heights in Virtual Reality. *Frontiers in Psychology* 6 (2016). https://doi.org/10.3389/fpsyg.2015.01969

[36] Stefania Serafin, Ali Adjorlu, Niels Nilsson, Lui Thomsen, and Rolf Nordahl. 2017. 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)*. 1–4. https://doi.org/10.1109/KELVAR.2017.7961562

[37] Yoana Slavova and Mu Mu. 2018. A Comparative Study of the Learning Outcomes and Experience of VR in Education. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 685–686. https://doi.org/10.1109/VR.2018.8446486

[38] Yi-Huang Su and Ernst Pöppel. 2012. Body movement enhances the extraction of temporal structures in auditory sequences. *Psychological research* 76 (2012), 373–382.

[39] William Forde Thompson, E. Glenn Schellenberg, and Gabriela Husain. 2001. Arousal, Mood, and The Mozart Effect. *Psychological Science* 12, 3 (2001), 248–251. https://doi.org/10.1111/1467-9280.00345 PMID: 11437309.

[40] Dimiter Velev and Plamena Zlateva. 2017. Virtual reality challenges in education and training. *International Journal of Learning and Teaching* 3, 1 (2017), 33–37.