THE CRITICAL EVALUATION OF IMMERSIVE BINAURAL AUDIO

by

Nathan Thomas McCubbin

A Dissertation
Submitted in Partial Fulfillment of the Requirements for the
BSc Hons in Audio and Music Technology

Confetti Institute of Creative Technologies
Nottingham Trent University
May 2020

Supervisor: Ken Liston

Contents

Abstract

This dissertation is asking the question of whether Binaural Audio is ready for more commercial use. As Binaural is the next step in immersive audio and opens up a lot of methods of further immersion for video, music and games but does not see much use currently as it can be complicated to produce.

Current research says that binaural audio is effective at creating an immersive experience and can even make listeners confused as to whether it is real or synthesised, and there are many methods to correct issues in Binaural audio like 3D printed Pinnas using silicone, correcting Room Divergence with visual stimulation and recording individual HRTFs to create a more accurate representation of how an individual hears but unfortunately it just does have a footing in the commercial world as not many popular artists use it and it is more of a niche market.

A dummy head recording was made at the same time as a stereo recording and some music and a soundwalk were made into binaural via a binaural panner as well as a stereo version, after that the testers were asked to listen to the audio and answer some questions via a survey. The results in the Data Analysis section and Conclusion show that the Dummy heads are ready for use to record binaural music for commercial use whereas artificial binaural panners are not effective enough to make commercial binaural music as they cannot be used to create a binaural track on their own due to aliasing but are effective when used more subtly and on less complicated audio.

Introduction

The research question for this Dissertation is, The Critical evaluation of Current Immersive Binaural Audio, which will give the answer to whether the current state of binaural audio is ready for consumer consumption. The research for this dissertation is primarily about binaural audio, what is binaural audio, what common issues are associated with it and what technologies and techniques are available to produce binaural audio.

The reasons behind this question are to see whether binaural audio compares well against traditional methods of playback, to see where binaural audio fits in the consumer world and to gauge consumer reactions and opinions on binaural audio. The aims of the testing for this dissertation are to create four different binaural audio experiences, a listening piece that is an acoustic song recorded with a dummy head, a rock track using artificial binaural panners, an electronic atmospheric track and a soundwalk, both electronic track and soundwalk are also using the artificial binaural panners.

This research topic was chosen as binaural audio is the next step in immersive audio, but it is complicated to create effectively which has made it difficult to work with and currently binaural audio is not being fully utilised; so the goal is to discover whether it is ready to be used more commercially.

Binaural is the process of compiling two signals to replicate the human hearing by retaining all the acoustical and spatial information that comes with pinnas, a head and occasionally a torso (Rumsey, 2014) (Moller, 1992) (Zhang, et al, 2017). Binaural audio is recorded via a few different methods, with equipment like the dummy head (Figure 1), head worm microphone (Figure 2) and a baffled microphone. (Figure 3) (Hoose, 2015)

Figure 1, Dummy Head (Hoose, 2015)

Figure 2, Head Worn Microphones (Sennheiser, 2019)

Figure 3, Baffled Microphone (Hoose, 2015)

1. Main Body

1.1 Literature Review

The dummy head is designed to replicate a head and is based on the average size of a human head, pinnas, nose and ear canals to help make the binaural recording more realistic (Hoose, 2015). The microphones are placed in, or outside, the ears of the dummy head depending on what microphones are used. This versatility of the dummy head makes it excellent for binaural recording as head worm microphones can be used, as well as regular omnidirectional microphones (Figure 1) and placement is not restrictive as Moller (1992, p.2) states that anywhere within the ear can be used for recording. This also goes for a few millimetres outside the ear as the sound pressure still has all the spatial information needed, which helps to make the dummy head more effective for binaural recording in comparison with the baffled microphone technique, as despite it being similar to a dummy head with the use of an object between the omnidirectional microphones, it cannot compete because it does not have the filtering effect of the pinnas and is only useful for producing the shadowing effect of the head and the dummy head has both the filtering and shadowing. (Hoose, 2015)

A localization issue that occurs with binaural audio is that humans use small head movements to understand front and back spatial information as a right head turn will make the frontal sound source reach the left ear first and the right ear last, the opposite occurs when the sound source comes from behind the listener and what makes this an issue is that binaural playback does not follow head movement (Moller, 1992), however, tests have shown that with broadband sound sources, listeners can tell the difference between the front and back with their head kept still. (Moller, 1992) (Rumsey, 2011)

There are some main complications that can make binaural audio not completely immersive and realistic for the listener (Rumsey, 2011), one of the main elements that negatively affects binaural playback is the listener themselves, as everyone has different shaped; heads, pinnas, ear canals and torsos thus influencing the signals for each individual listener differently. (Rumsey, 2011) (Moller, 1992) (Spagnol, Tavazzi and Avanzini, 2017)

In order to overcome the listener's individual spatial influence, the head-related transfer function (HRTF) was created, which is a left and right filter for each ear that simulates the acoustical information that the listener's ear drums, pinnas, head and torso create when interacting with the sound source. (Spagnol, Tavazzi and Avanzini, 2017) (Hoose, 2015) (Zhang, et al, 2017) (Moller, 1992)

Interaural time difference (ITD), interaural level difference (ILD), colouration and interaural phase difference (IPD) are responsible for directional hearing in each of their domains (Frequency range), in the horizontal plane low frequencies are assessed by interaural phase differences, mid frequencies by interaural time differences and high frequencies by interaural level differences. Colouration is responsible for where there are no interaural differences, in the elevation plane. (Zhang, et al, 2017) (Moller, 1992)

Creating individual HRTFs is done by sitting the subject in an anechoic room or an acoustically treated room with head worn microphones in their ears and a single loudspeaker to measure at different directions around the subject (Figure 4); More expensive HRTF set ups use multiple loudspeakers surrounding the subject. (Figure 5) (Spagnol, Tavazzi and Avanzini, 2017) (Rumsey, 2014) (Rumsey, 2011)

Figure 4 Single Loudspeaker HRTF (Ausim3d, 2011)

Figure 5 Multiple Loudspeakers HRTF (Future Reality Lab, 2018)

Individual HRTFs are very time consuming and complicated to properly measure (Rumsey, 2011), so the commonly used alternative is general HRTF which is measured with a dummy head or head and torso simulator (HATS) that is based on the average dimensions of a human head and torso (Zhang, et al, 2017) (Rumsey, 2011) (Spagnol, Tavazzi and Avanzini, 2017). However, despite the general HRTF being averaged on the human body there will be a degrade in localization thus resulting in a poorer experience. Fortunately, this localization error will gradually decrease as the listener gets used to the general HRTF, given that it does not change, but this takes time and is infrequent as some can learn it faster than others (Rumsey, 2011) .

In Rumsey's (2014) article he has a section called plausibility of binaural synthesis. In this he explains and discusses a test done by Chris Pike where they tried to figure out if listeners could tell the difference between binaural synthesis and their own expectations of spatial realism and see if a simulation can fit within a listener's experience and knowledge of real audio stimulations.

They set up the testing by creating one hundred audio samples each from a real loudspeaker or synthesized through headphones, they created all the synthesized sounds with a general HRTF by measuring a Neumann KU100 dummy head (Figure 6), they placed open-back headphones on the dummy head while measuring because the testers were also going to have those open-back headphones on while the loudspeakers signal was tested; head tracking was imployed for the testers to monitor their head movements and to adjust the binaural audio with their movements.

Figure 6 Neumann KU100 dummy head (Neumann, 2018)

When testing began the testers were asked whether the audio is real or simulated. The results showed that the sensory differences between the real and simulated signals was minimal, as the testers found it difficult to tell the difference between the two, they used head movement to try and determine what was being played. Testers found it easier to tell the differences with elevated signals, one reason was due to the signal going beyond the loudspeaker. An issue with the testing that the Chris Pike mentioned was that the testers had open-back headphones over their ears which affected the signal from the loudspeaker. (Rumsey, 2014)

Reddy and Hegde (2016) also has an interesting article based around creating their own dummy head with 3D printed bionic ears for rendering binaural audio. They found that their bionic ears showed to have an effective perceptual experience especially in the elevation plane compared to HRTF databases. A main contributing factor in this was the development of the bionic ears and the material they used, as they decided on silicone because silicone has mechanical properties very close to the cartilage based human ears. (Reddy, Hedge, 2016)

It is interesting to note that if Reddy and Hegde's (2016) bionic ears approach was adapted to the testing from Rumsey's (2014) article then the elevation being easier to identify would be reduced since Reddy and Hegde's (2016) improved elevation localization, it would also be useful to hide the loudspeakers for future testing. (Rumsey, 2014) (Reddy, Hedge, 2016)

Gutierrez-Parera and Lopez (2016) has an interesting and influential article based on discovering the affects low quality headphones have on binaural audio. They used the same headphones for testing and developed virtual headphones to simulate different types of high and low quality headphones in order to remove visual and tactile biases, as it has been found that testers can be influenced by these factors, and it has also been found that blind folding the testers to switch headphones can still result in influenced testing due to the feel of the headphones. As such, Gutierrez-Parera and Lopez decided to simulate the headphones to remove these influences. (Gutierrez-Parera and Lopez, 2016)

In conclusion, Reddy and Hegde's (2016) article is about improving HRTF, by using individual HRTF measurements from 3D printed bionic ears on a dummy head instead of databases. Whereas Rumsey (2014) argues for using general HRTFs instead of individualized HRTFs due to the impracticalness of individual HRTFs as well as supporting his claim by discussing a test's results that showed that listeners could not tell the difference between real or synthesized sounds.

With the knowledge presented, database HRTFs are great at immersing the listener into the binaural audio and are a lot easier to use than measuring individualized HRTFs. They get the job done but it is worth noting that Reddy and Hegde's testing resulted in a better result than general HRTF as the elevation plane had better accuracy. However, without a 3D printer and the technology and knowledge to measure HRTF and to implement the HRTF into a computer it would unreliable to pursue this bionic ears method. (Reddy, Hegde, 2016) (Rumsey, 2011)

This dissertation will use general HRTF due to it being practical and less time consuming than individual HRTFs (Rumsey, 2014), it would also be unreasonable to expect every tester to let themselves be measured up for their HRTFs, as this has been known to cause fatigue due to the length of measuring HRTF and the testers not being allowed to move (Reddy, Hegde, 2016). This dissertation is about whether or not binaural audio is currently immersive, it is not about creating new methods but rather seeing if the most common and practical binaural playback is worthy for consumer use, so using general HRTF will give better results due to it being commonplace.

Moller (1992) and Rumsey's (2017) articles about binaural audio will be a vital tool in the creation of binaural audio, as they help establish frameworks on how to record binaural audio as well as issues that accompany binaural recordings. As stated earlier by Moller (1992), there is a certain amount of space that can be recorded inside or outside the ear, as to provide the spatial information needed.

Another thing to consider for binaural recording is Room Divergence (Rumsey, 2017). This is the issue that comes with mismatched room acoustics and reverberation, an example being a tester listening to a binaural recording of their bathroom while they are sat in their living room. This will cause the tester to experience cognitive mismatch as they will assume that the recording is a lot more exaggerated and more reverberant. This is excellent information to know for the methodology and creation of the binaural recordings as they will be steps taken to avoid room divergence by making sure the binaural audio's acoustic and reverberation matches the images/ videos shown in the test (Rumsey, 2017).

Hoose (2015) and Zhang, et al (2017) both go in depth about methods of recording binaural audio, mainly discussing microphone techniques, such as the dummy head, baffled, head worm and also discuss sound field microphones.

The binaural microphone techniques that will be used for this dissertation are the Dummy head technique and the head worm technique as they are the most practical. Although, the Baffled microphone technique is useful and works similarly to the dummy head technique, it does not fulfil the requirements of binaural recording as well as head worm and dummy head because they both work better for capturing the acoustical presence of the head (Hoose, 2015)

The soundfield microphones will also not be used as they are designed primarily for surround sound recording, its design shows this as all the capsules are very close together and designed to be facing each angle. What sets this apart from a binaural microphone technique is the fact that there is no head or pinnas, all the sound energy goes directly to the microphone whereas a dummy head will have the sound be influenced by the head and pinnas. (Hoose, 2015) (Zhang, et al, 2017) So using a soundfield microphone would work but would result in poor localization and confusion for the testers. (Hoose, 2015)

1.2 Methodology

Due to the Coronavirus pandemic, any human interaction that was planned for this methodology could not be done. This resulted in the focus group and class room plan not being achievable, also meaning the same headphones strategy was not possible. This has affected the results as there will now be visual and tactile biases in the testing. Fortunately, the goal was to also test the testers own headphones and that is still possible. In addition to this, the surveys can still be done. As such, the methodology for this dissertation was not possible at the time.

It is worth noting that using particular groups of people for testing has changed but not necessarily for the worst as instead of audiophiles and standard music consumers being tested separately with one half doing same headphones and the other using user headphones, the testers were combined for testing. The pilot test consisted of Music Technology students whereas, the main test consisted of Music Technology students and standard music consumers. This meant that the original goal of having separate results for each type of group is not possible but testing both types of people will still result in more accurate discovery of whether binaural audio is ready for more commercial use. (Kruger, 2018).

The other two main parts of the dissertation that were not possible were the visual soundwalk, as time outside was limited to a one-per-day exercise and imperative reasons such as medical appointments. Also, the equipment to be used for the visual soundwalk could not be booked anymore meaning even if there was no restrictions as to going outside, it still was not possible to record the walk. Brown's (2017) article is still useable as instead of creating a visual soundwalk, a normal artificial soundwalk can be created, by gathering samples, foley and creating sound design to build a soundwalk from scratch. Unfortunately, it did not accommodate any visuals. This makes the dissertation's question more narrow as to what answer it can achieve, this is due to both tests being similar, alongside only being audio, as well as the issue known as Room Divergence can occur due to the room/ space not being visible. Regardless, it can still isolate whether people prefer binaural in use for soundwalks over music or vice versa. Additionally, the soundwalk being artificial will not suffer from room divergence as it will not try to replicate a room/ space. (Brown, 2017) (Rumsey, 2017)

The other part of the dissertation that was not possible was the recording a band in binaural and stereo, as the studios was closed including access to the equipment and due to the coronavirus, there was restrictions on socialising in groups, being limited to 3 persons from a single house hold only. Due to this, the Cambridge-mt website was used to gather stems from a track, the track was then placed into Pro tools for mixing.

To adapt Bayley's (2008) dissertation, the mix was done in stereo then the session was reopened and then the plugin AMBEO Orbit by Sennheiser was applied to all the tracks and set to match the stereo panning. This is done to keep the testing fair and not make obvious which one is which. Unfortunately, binaural panners are known for their spatial aliasing which causes a trough in the frequencies at 8kHz which causes the track to sound dull but due to the situation there is no other option to solve this. (Bayley, 2008) (Rumsey, 2014) (Otani, et al, 2016).

Fortunately, before the coronavirus pandemic a band was recorded in binaural and stereo, but the recordings are not great and are not similar enough but still can be used and can offer different variety for testing and a comparison from artificial binaural and dummy head binaural. (Hoose, 2015) (Bayley, 2008) (Rumsey, 2014). (Otani, et al, 2016)

The testing for the dissertation consisted of one listening test, one listening/visual test, a survey for the listening test and another survey for the listening/visual test. The listening test consisted of an acoustical set being recorded in binaural and stereo and the listening/visual test was a visual soundwalk that was also recorded in binaural and stereo (Brown, 2017).

The decision to make a visual/listening test came from Brown's (2017) article as he created a binaural soundwalk to discover whether a binaural soundwalk can be a form of art and believes it can be used to tell a story of an event at the soundwalks location. As such, his work and methods were adapted to see whether soundwalks can be immersive and tell a story but with the extra benefit of making it a visual experience as this dissertation's goals was to discover whether binaural audio can fit multiple forms of media, video, music and gaming. (Brown, 2017). The listening test was derived from Brown (2017) and Rumsey's (2014) articles as Brown, as mentioned above, offered an idea to create a binaural experience with intention of being a form of art and to tell a story and Rumsey stated that binaural audio is not fully utilized, despite the majority of listening done with headphones. Rumsey also stated that there is no popular binaural music available so the listening test was created to see whether music consumers will enjoy binaural music and to evaluate if binaural music is even necessary to make music more immersive. (Brown, 2017) (Rumsey, 2014).

These extra tests were decided upon due to the pilot tests and due to see how testers react to different styles of audio with binaural, especially how the dummy head recording compares to the artificial binaural panner, as well as the testers recommended further examples for the main testing. (Appendix 2.) (Hoose, 2015)

The listening test was recorded with the Soundman Dummy Head, Roland head-worm microphones and the Zoom H4n Pro Handy Recorder, this allowed for the music to naturally fit in the dummy head's azimuth instead of artificially panning the musical elements to fit in the head space (Hoose, 2015).

The stereo part of the listening test was being recorded at the same time as the binaural part by using two C404 XLII microphones set up above the dummy head and placed in the X/Y stereo technique, which was in the crossed cardioid pattern (B) (Figure 7). This was done because this supplied a stereo image as it captured each individual instrument (Figure 8) (Huber, 2017). The C414 microphones were set to the same settings, cardioid polar pattern, no pass filter, and no pad; this was done so there were no different microphone characteristics influencing the sound. Since the Roland head-worm microphones were recording every piece of the acoustic set, using the C414s to record the instruments will result in everything sharing the same microphone characteristics, much like the Roland (Huber, 2017) (Rumsey, 2014).

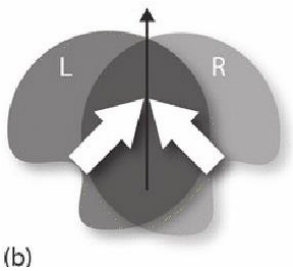

Figure 7 X/Y Stereo technique, crossed cardioid pattern (Huber, 2017)

The acoustic instruments used for the recording were vocals, two rhythm guitars, one lead guitar, one guitar playing the bass parts and one cajon, they were setup in a manner to surround the front of the dummy head and X/Y setup, shown in (Figure 8), this was done to make each stereo and binaural test have similar panning information to give a fair comparison (Rumsey, 2014). The listening/visual test was a soundwalk going through Forest Park in Nottingham and the audio was recorded with the Roland Head-worm microphones placed inside a human head, Zoom H4n Pro Handy Recorder, and the GoPro Hero 7 Silver head-worm camera to capture the visuals of the walk. This test being a listening/visual experience was done to overcome the issue of room divergence, to gauge the interest in binaural audio visuals and to resolve the issue of not being able to create a video game binaural experience (Brown, 2017) (Rumsey, 2017).

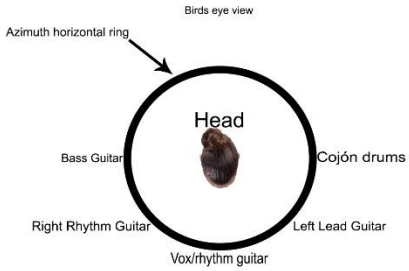

Figure 8 Binaural recording panning plan

For the stereo version of the visual soundwalk, the equipment used to record it was the Zoom H4n Pro Handy Recorder and a Rode NTG-2 Microphone which was attached to a Rode Boom Pole and was held above the person who performed the walk. The Rode NTG-2 was chosen primarily its polar pattern, which is a supercardioid as this will allow for the sound image to be wide as it covers the front and sides (Brown, 2017). It is worth noting that the binaural and stereo recordings of the visual soundwalk happened simultaneously to keep both binaural and stereo tests as close as possible (Rumsey, 2014).

This test was interesting as this used an individual HRTF due to the recording being done with the head-worm microphones placed inside a human head which resulted in better localization for this listening/visual test than the listening test, which was effective as this listening/visual test was designed to be primarily immersive to gauge the tester's immersion compared to the stereo listening/visual test, whereas the musical test was designed to be immersive but also creatively diverse to see if the testers would find it engaging (Spagnol, Tavazzi and Avanzini, 2017).

The questions asked in the surveys were based on the combination of the quantitative and qualitative approach as it captured the emotional reaction while also maintaining a numerical data approach. (Brannen, 2016). Two surveys were created (Figure 9), one for the listening test and the other for the listening/visual test and they were designed to evaluate immersion, noticeable difference, necessary, quality and accuracy (Figure 10).

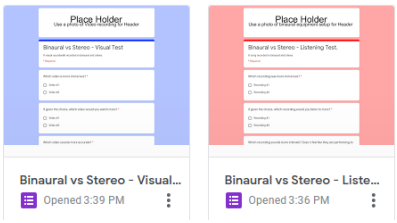

Figure 9 Both surveys for both tests

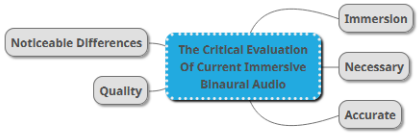

Figure 10 Subcategories for Dissertation question

It is worth explaining what necessary, quality and accuracy were trying to discover; necessary was to determine if the testers believed binaural audio was worthy of use over stereo audio for these tests and whether stereo is just good enough to be used, quality was to determine whether the audio matched up to the quality of stereo as stereo recording has more options and better microphones than binural recording microphones and accuracy was to determine if the localization was good despite being a general HRTF and someone else's individual HRTF. (Hoose, 2015)

Gutierrez-Parera and Lopez (2016) have influenced the testing methods used in this dissertation due to them using the same headphones for testing to remove tactile and visual bias, while also using filters and processing to emulate consumer headphones to get a more consumer accurate result. I adapted their methods by separating them into two types of tests, the first test being with the same headphones for each tester, which achieved a much more uninfluenced result by having the same playback, which meant that the only difference in the playback was the testers themselves. The second test had the testers using their own headphones which achieved the consumer accurate result like in Gutierrez-Parera and Lopez's testing. The reason for splitting the test methods into two tests were due to not having enough time to emulate consumer headphones nor the knowledge to do so.

The testing was done at one of Confetti's classrooms as they allowed for enough space for the testers, and also give access to the laptops in the room as they were using the laptops to listen to the music test and watching the visual test. Four focus groups were made, each group having five participants for the testing, Groups A and B were audiophiles as they are a community of music consumers who desire the highest quality music and playback systems, they are keen on new musical experiences. Groups C and D were standard music consumers as they supplied the information for whether the general public may want more binaural audio productions. It is worth mentioning that groups A and C were given the same headphones, while groups B and D were asked to bring their own headphones. (Gutierrez-Parera and Lopez, 2016) (Kruger, 2018). These groups comprised of Confetti's university students, the audiophiles being second year audio technology students and the standard music consumers being second years from the visual effects production course.

Each of the focus groups were sat down with every tester having their own laptops ready to go, then asked to listen to the binaural and stereo listening tests and then answering the listening survey and afterwards asked to watched the binaural and stereo visual/listening tests and asked to answer the visual/listening survey. After testing was completed the results were compiled into an Excel spreadsheet to be analysed.

1.3 Results

Pilot Test.

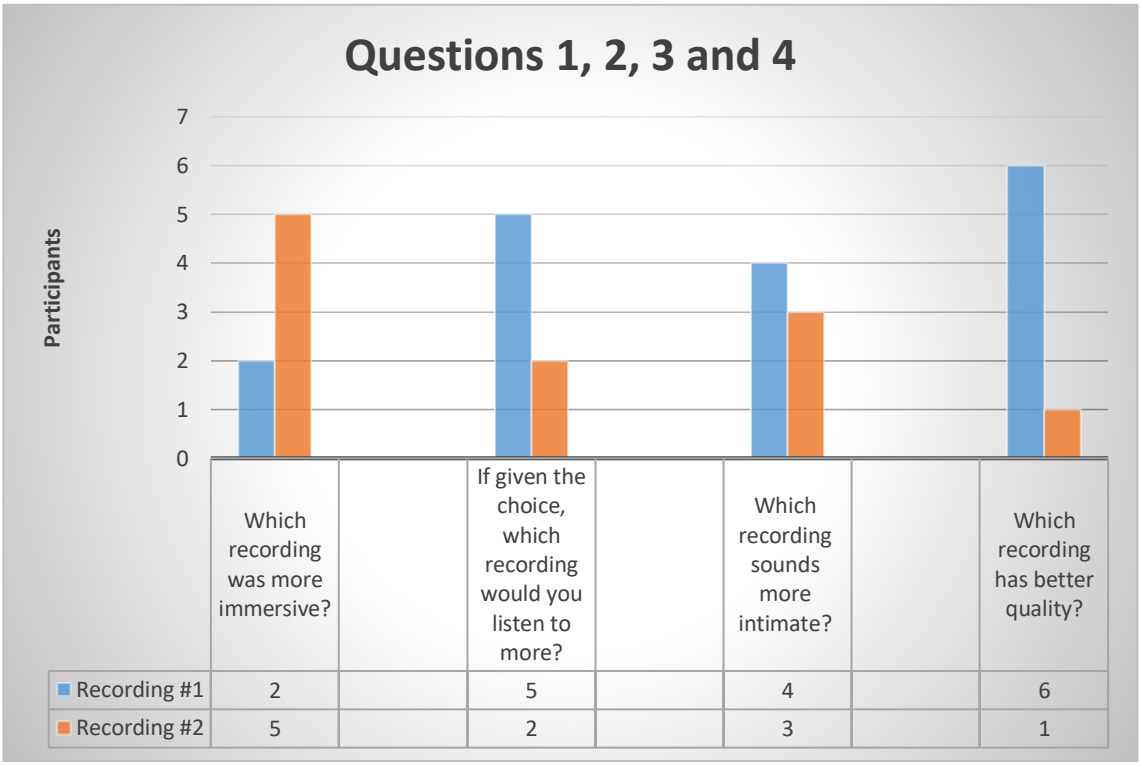

Figure 11 Question 1, 2, 3, and 4

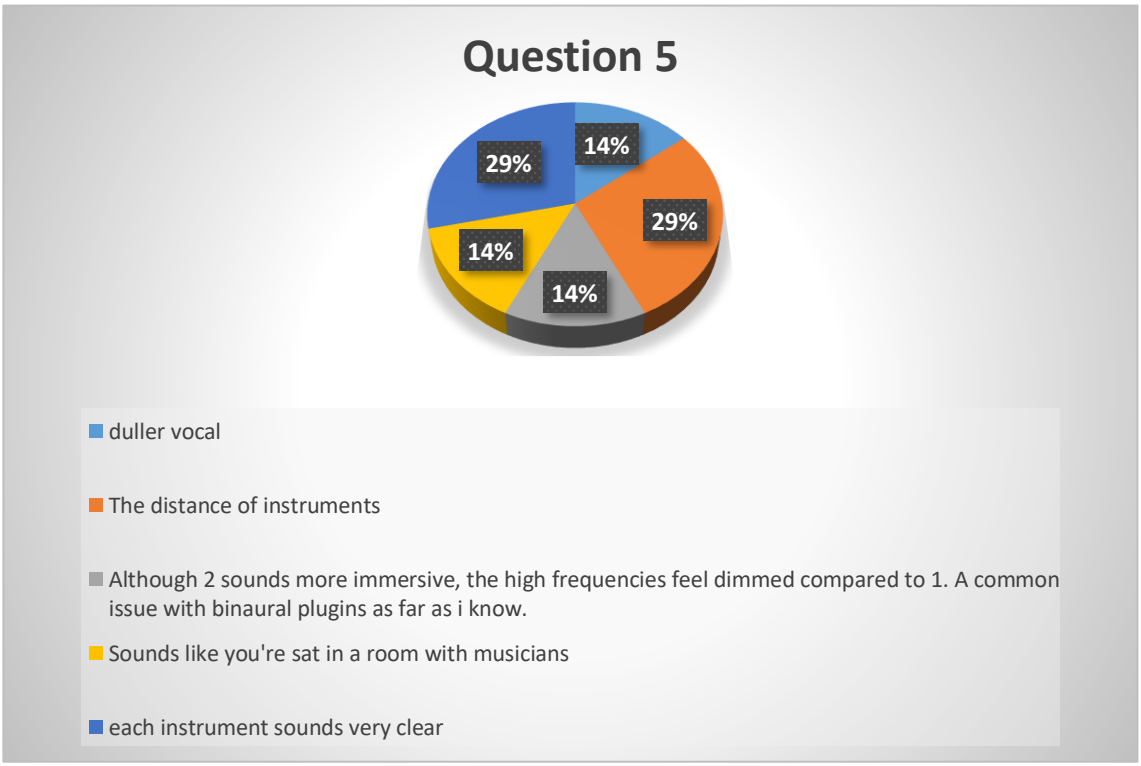

Figure 12 Question 5

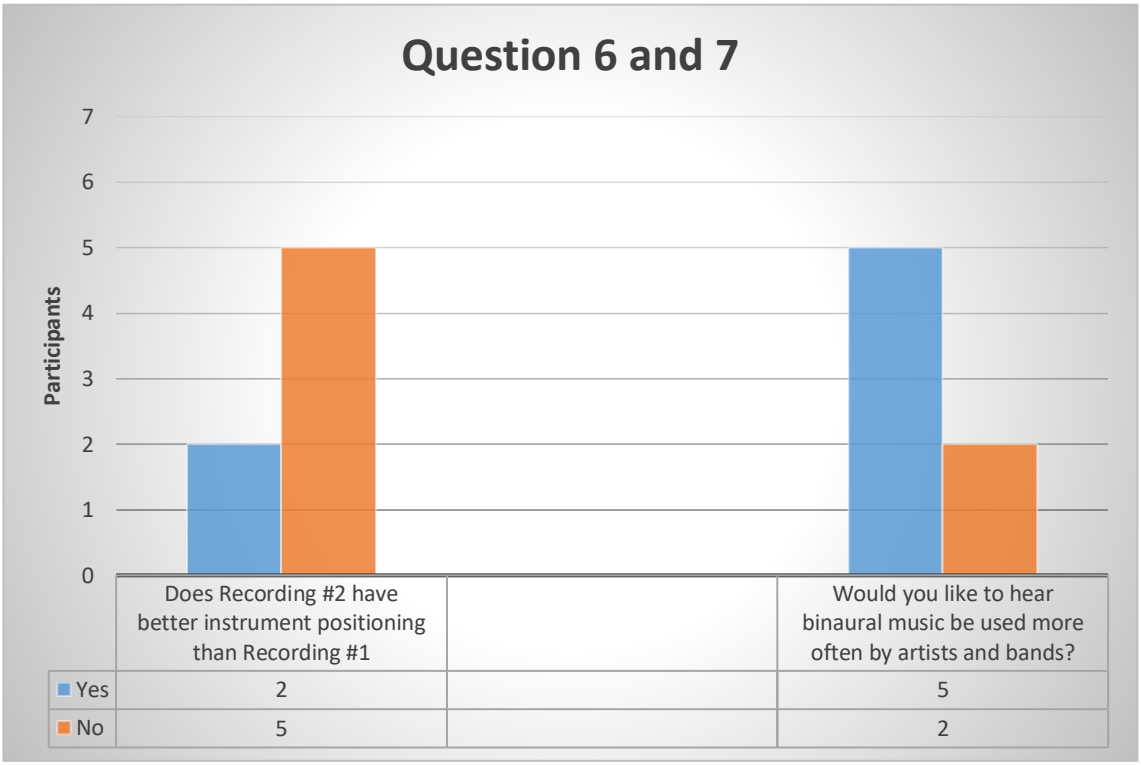

Figure 13 Question 6 and 7

Main Test.

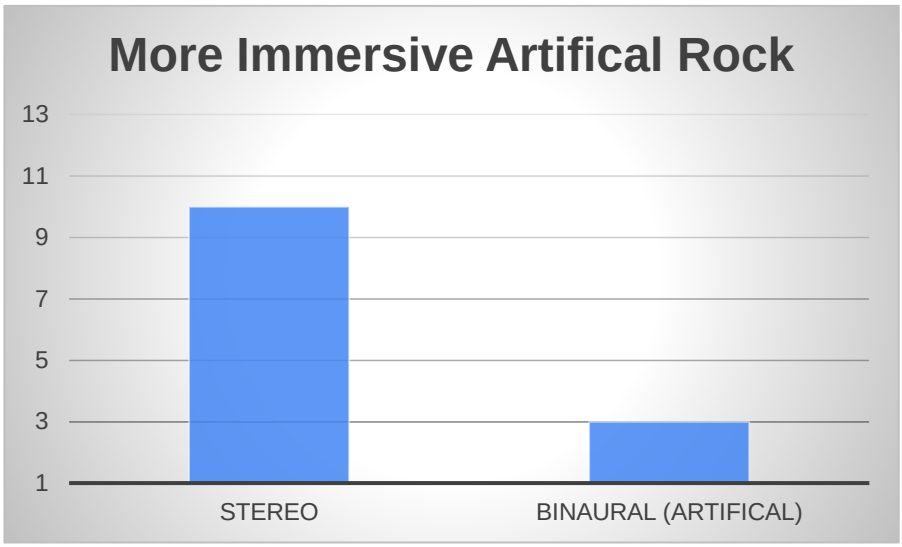

Figure 14 Immersion results Test A

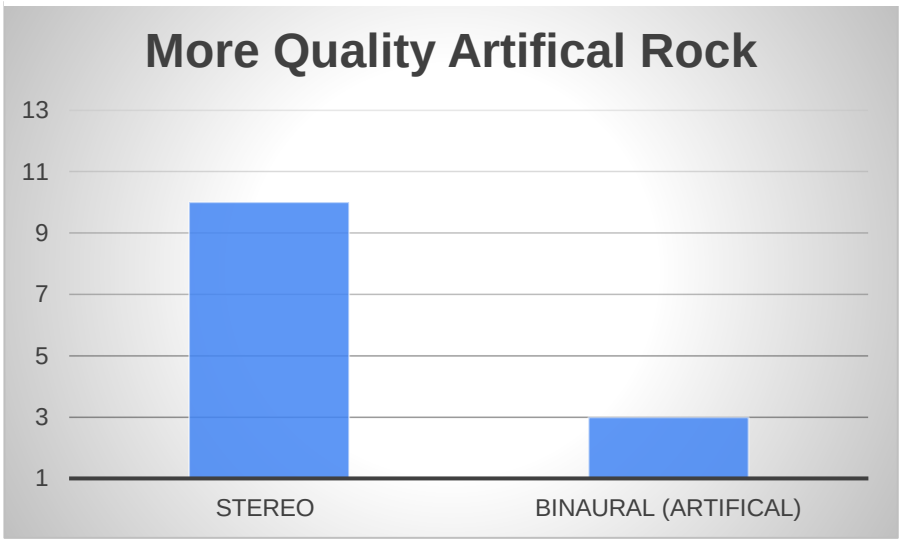

Figure 15 Quality results Test A

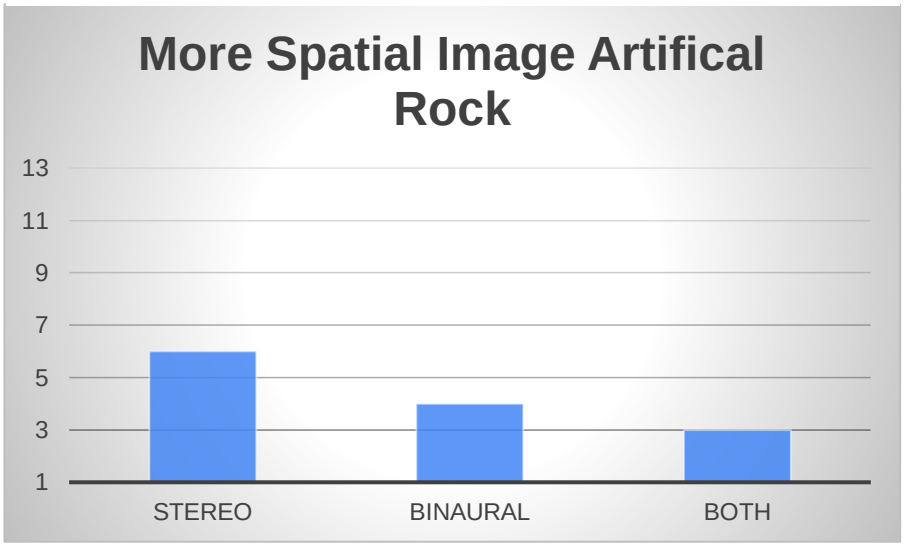

Figure 16 Spatial Image results Test A

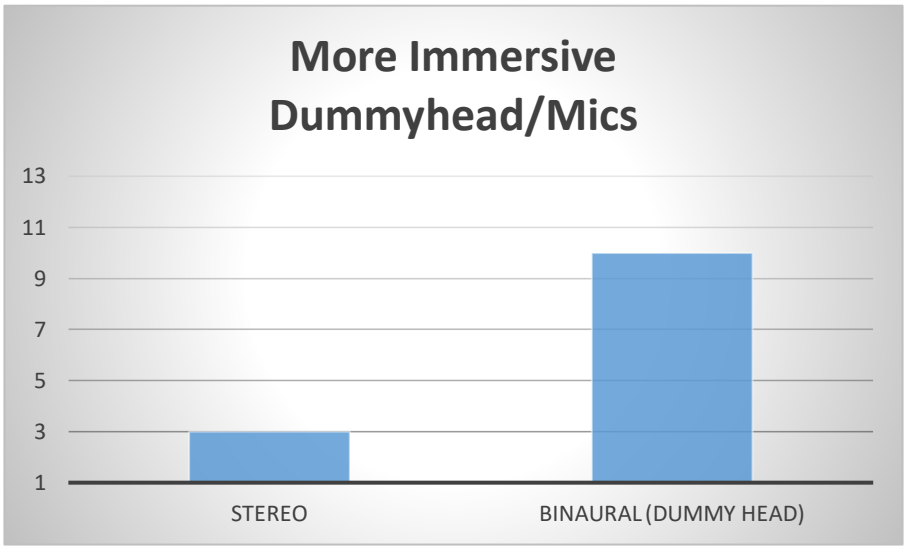

Figure 17 Immersion results Test B

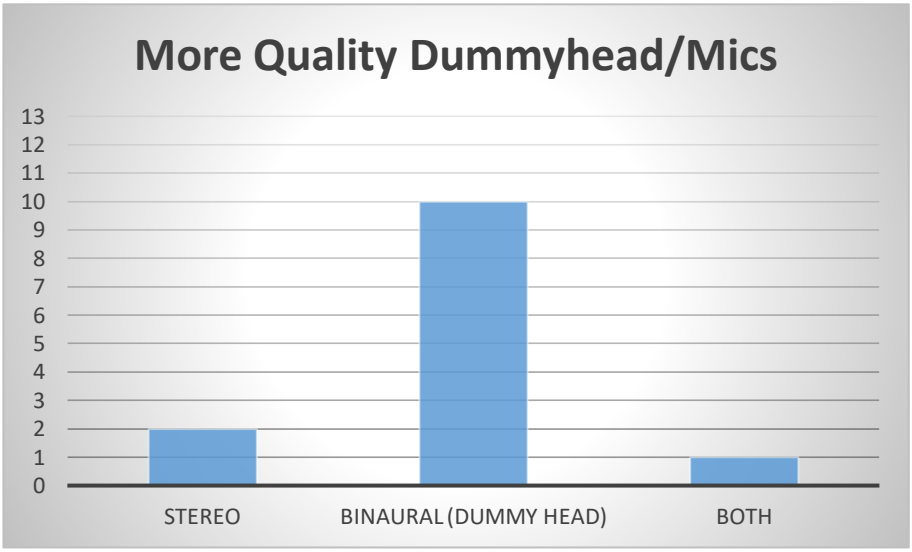

Figure 18 Quality results Test B

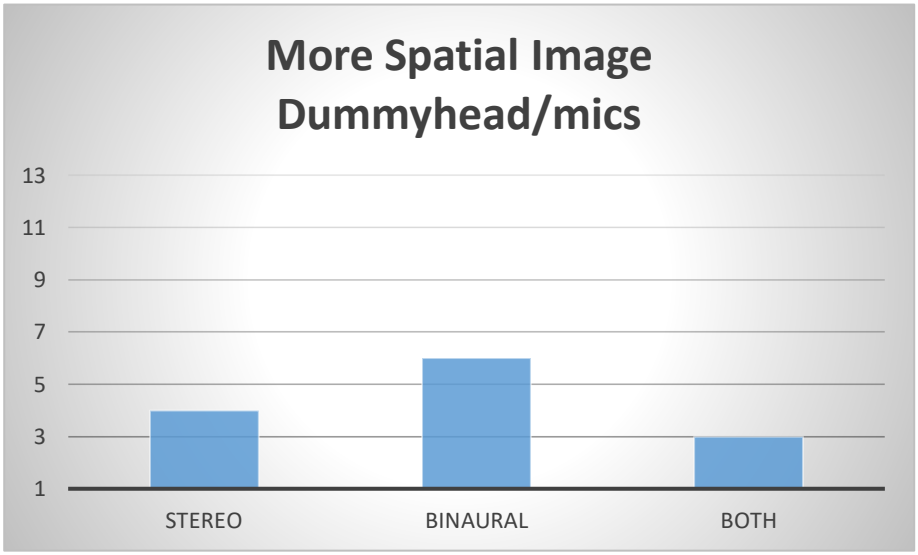

Figure 19 Spatial Image results Test B

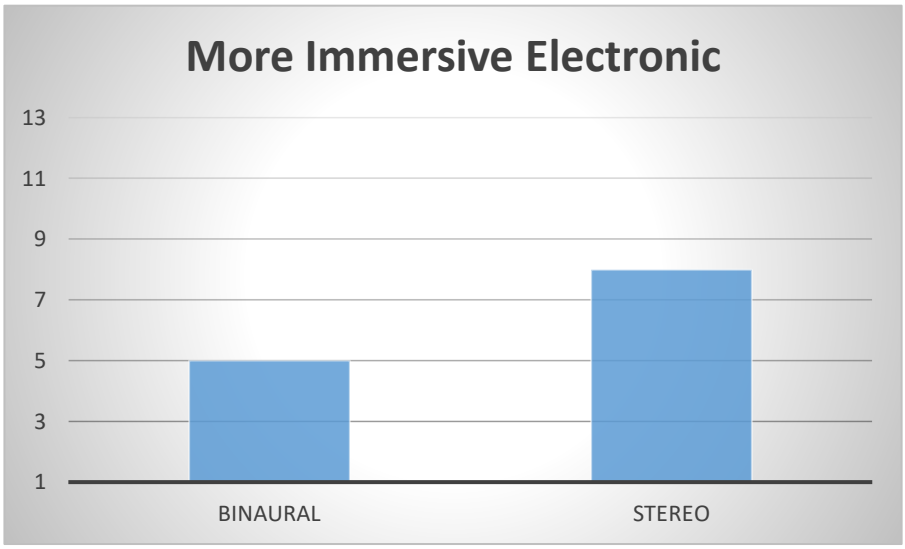

Figure 20 Immersion results Test C

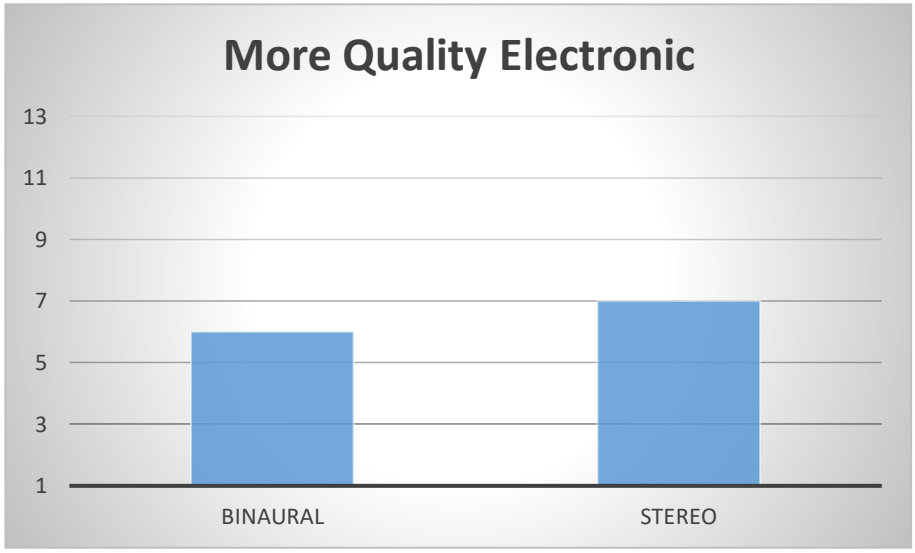

Figure 21 Quality results Test C

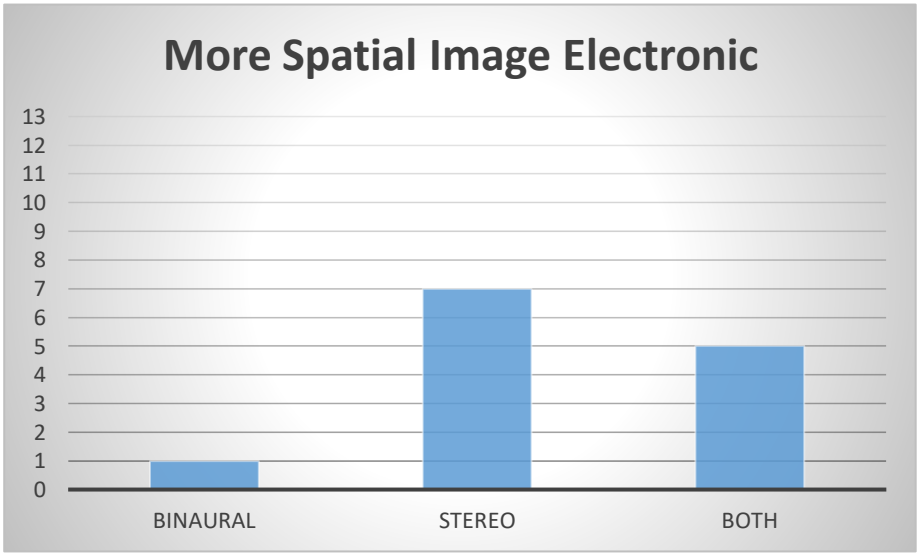

Figure 22 Spatial Image results Test C

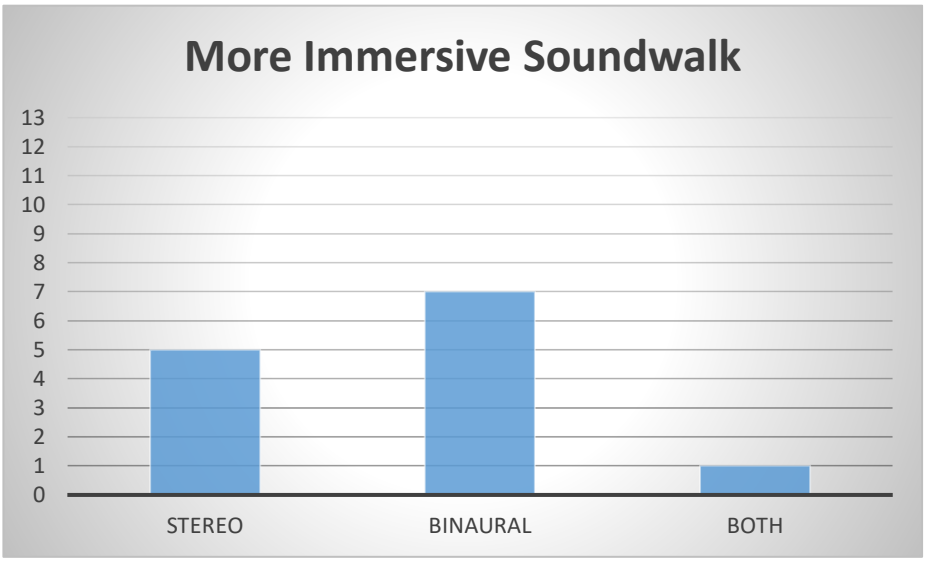

Figure 23 Immersion results Test D

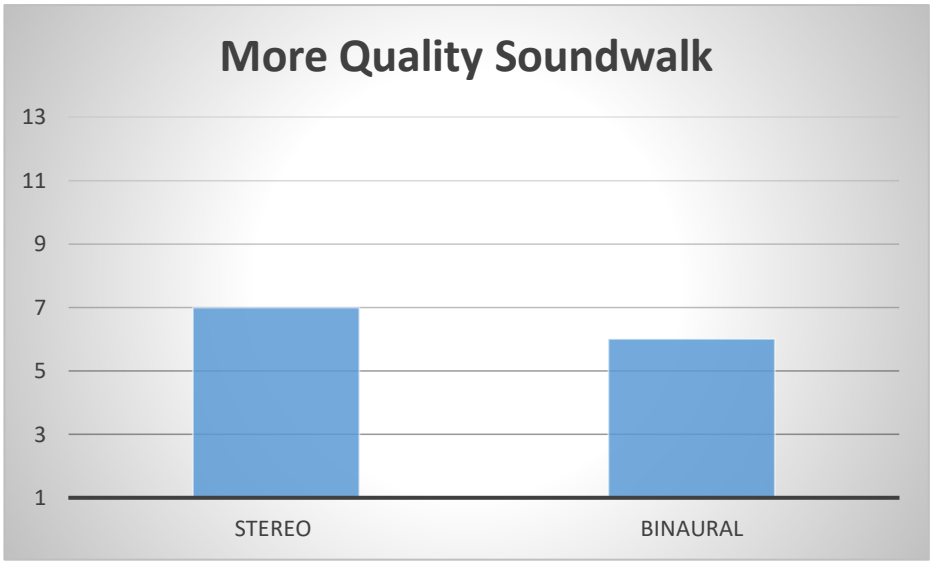

Figure 24 Quality results Test D

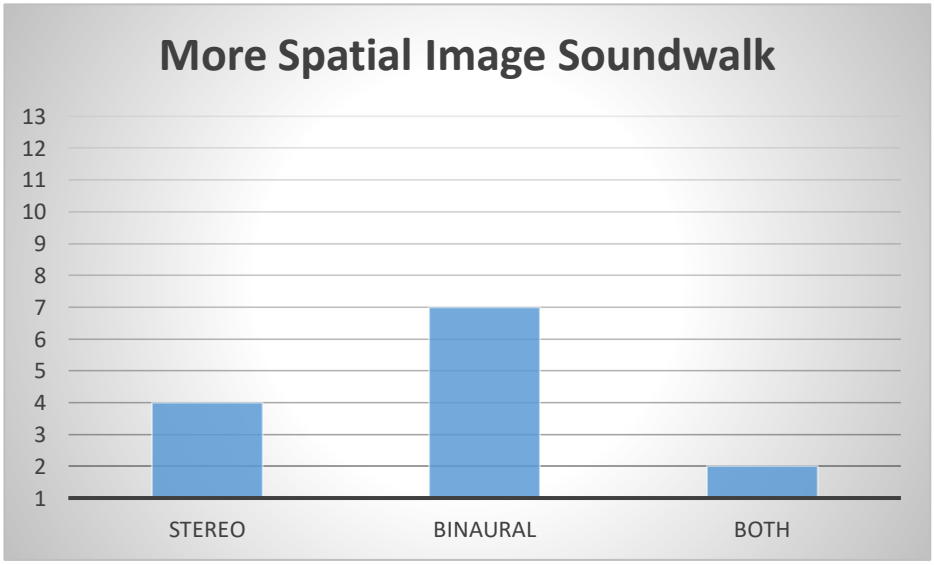

Figure 25 Spatial Image results Test D

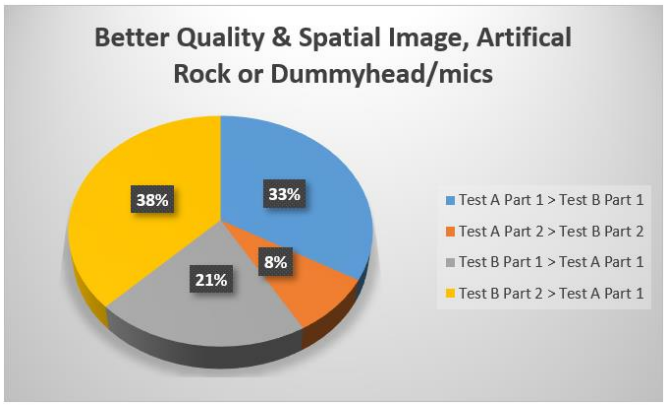

Figure 26 Pie Chart results for comparison between Test A and Test B

Figure 27 Tell a Difference results

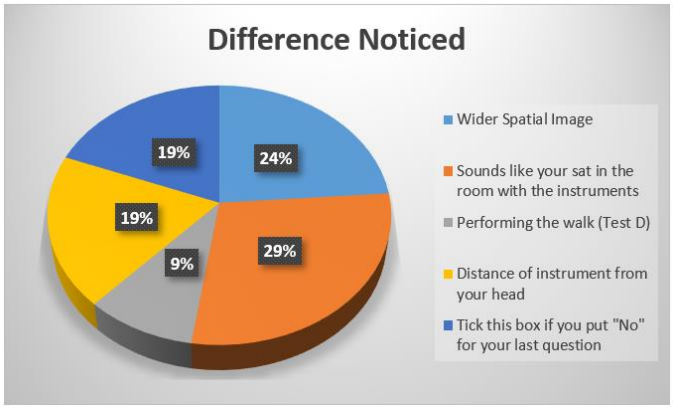

Figure 28 Difference Noticed results

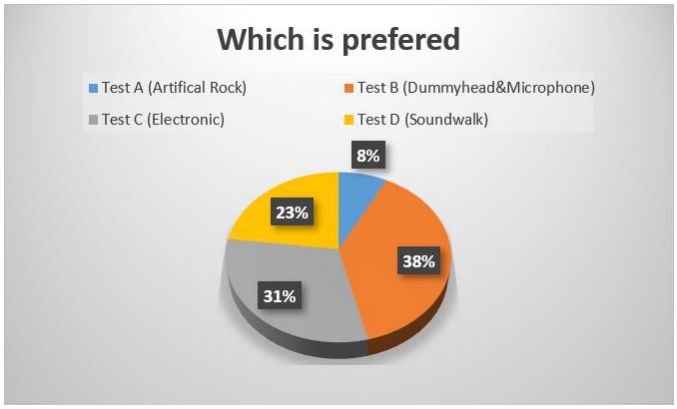

Figure 29 Which test did the testers prefer results

1.4 Data Analysis

For the pilot test, the Test A track was used as it was the only one ready for use at that time. And due to being more straight forward, being that it is a rock track and was only using Artificial binaural panners. It is worth noting that during this pilot test the Test A track was referred to recording, and the full tracks were used for testing instead of exerts, the binaural version is Recording 2, and the stereo version is Recording 1.

(Figure 11) Shows the first four questions of the test, and by looking at the data it is interesting to note that question 1 has five out of seven votes for Recording 2 which asked which one is more immersive. This shows that the testers believe that the binaural version of the track is much more engaging and immersive, which is interesting since question 4 asks which one has better quality and 6 out of the 7 testers voted for Recording 1, this is due to the aliasing associated with artificial binaural panners (Otani, et al, 2016). But despite this error in the binaural panners, the testers still mostly voted for Recording 2 when it comes to the immersion.

Question 3 also is an interesting question, as it asks which is more intimate out of Recording 1 and 2,. The results show 4 votes for Recording 1 and 3 votes for Recording 2. This shows that there is some confliction between the two recordings, the quality and immersion are definitely defining factors on this question, this is evident from question 4 and question 1, as half of the testers cannot get past the lack in quality thus making their choice the stereo version whereas 3 of them can look past it and can enjoy the enhanced spatial image and immersion. But Recording 1 still had an extra vote showing that quality is still important and higher quality leads to better intimacy, perhaps due to not being distracted by the quality unlike Recording 2. (Rumsey, 2011)

Another factor into question 3's confliction could be due to the HRTF of the AMBEO Sennheiser Binaural plugin, as it is known that general HRTFs can sound confusing for the listeners and confuse them (Rumsey, 2011) (Moller, 1992) (Spagnol, Tavazzi and Avanzini, 2017) but the fact that question 1 has the majority vote for Recording 2, and question 4 has the majority vote for Recording 1, it is evident that the quality is definitely a major factor into the binaural tracks success. (Rumsey, 2011)

Question 2 asks which track they would prefer to listen to, 5 out of 7 testers choose Recording 1, this simply backs up question 4 and the overall importance of quality of the recordings being a major factor into preference. (Rumsey, 2011)

(Figure 12) shows the question 5 pie chart, this pie chart shows the responses from the question Does Recording 2 sound different to Recording 1 due to?. This question had some responses for the testers to choose from but also had the option to add their own response, it is interesting to note that two of the testers added their own response and both responses address the quality and dullness of Recording 2. The yellow response goes a bit in depth with it stating that Recording 2 is more immersive but high end dip was the main difference noticed, the tester also mention that this is a common issue with binaural panners which is very negative in regards to Recording 2 as this implies they only recognised it being binaural due to this dip, which is certainly not a quality binaural music wants to be recognised for as despite the tester knowing this issue and most music listener do not, if binaural music became more common place, this issue would stand out. (Otani, et al, 2016).

The main trend from this pie chart is the response about the instruments, as both the red response and orange response have the most votes and the green response despite only having one vote still contributes the instrument trend of this pie chart. This shows that the main element the testers are noticing is the instruments themselves and how they are positioned and what spatial image they create. This gives more weight to question 1 as it shows what exactly is making Recording 2 so immersive for the listeners and it is interesting to note that the instruments were positioned to imitate the stereo version meaning that if the binaural track made use of the entire 3D space, it could prove to be even more immersive. This is evident with question 6 (Figure 13) as the question asked was does recording 2 have better instrument positioning than recording 1 and 5 out of 7 testers voted for No, which shows that the testers could tell a different between the two versions but it is not necessarily better than the stereo version's instrument placement. (Rumsey, 2014) (Moller, 1992) (Zhang, et al, 2017) (Spagnol, Tavazzi and Avanzini, 2017)

Question 7 asks whether the testers would like to hear more binaural music from artists and bands (Figure 13), and 5 out of the 7 testers voted Yes, proving that there is some interest in binaural music, this is interesting as this example used within the pilot test is flawed due to the aliasing issue associated with binaural panners (Otani, et al, 2016) but despite this, the testers still heard a difference and found it more immersive but even with these positive factors for this binaural track, it is still clear that the decreased quality compared to the stereo version is a dominating factor as evident with question 2, 4 and the blue and yellow responses from question 5 (Figure 11) (Figure 12)

The main test consisted of four different pieces of audio, labelled Test A (Artificial Rock), Test B (Dummyhead&Microphones), Test C (Electronic) and Test D (Soundwalk). Artificial Rock refers to the use of binaural panners and dummyhead&microphones refers to one recording using a dummy head and Roland head-worm microphone and the other recording using the C414 microphones mentioned in the Methodology.

(Figure 14, 15, 16) shows the results for Test A. The pilot test already shows that Test A's stereo version is preferred due to the clarity and quality compared to the binaural version which suffers from binaural panner aliasing (Otani, et al, 2016) and shown that quality is a defining factor on the tracks other qualities but is interesting is how in the pilot test the immersion question had the most votes for the binaural version whereas within this main test the testers have mostly voted for the stereo version.

It is worth noting that some of the testers within the main test were standard music consumers which explains why the results differ from the pilot test, it further emphasizes the importance of quality that the recordings need to have in order to be successful and this is also further backed up by (Figure 15, 16) as both of these results show how the stereo version is preferred when considering these points of quality and spatial image. But it is worth noting that the spatial image results (Figure 16) show that the tester were more conflicted and despite the stereo version still being preferred, it only had two more votes than the binaural one and three testers voted for both being as good as each other.

(Figure 17, 18, 19) shows the immersion, quality and spatial image results for Test B, this audio test is particular interesting compared to the others due to being recorded in binaural with the use of the Dummy head instead of being made to be binaural via artificial binaural panners.

When it comes to the immersion results (Figure 17), ten of the testers voted for the binaural version of the track which shows that the binaural version is definitely considered to be more immersive which falls in line with binaural being made to imitate how humans hear so when listening to anything in binaural it will sound like you are in that space, of course this is not always completely accurate with HRTFs and such but it generally more immersive than stereo audio and this test results show this. (Hoose, 2015) (Rumsey, 2011) (Moller, 1992) (Spagnol, Tavazzi and Avanzini, 2017)

(Figure 18) shows that the testers believed the quality of the binaural version to be superior compared to the stereo version, except for one testers who believed them to be equal. This is interesting and falls in line with what has been discovered currently in this test, but it is also interesting as the microphones used on the stereo version were C414s microphones which are high quality microphones and there are no negative effects like Test As binaural version having aliasing. (Otani, et al, 2016)

Looking at the results from (Figure 17, 19) it is safe to conclude that the reason the binaural version of Test B is believed to be higher in quality is due to the factors like better immersion and better spatial image as these factors enhance the track and make it sound better due to the ability to feel within the room, which is interesting as all these factors like quality, immersion and spatial image all interlinked and have effects on each other depending on their own individual effectiveness. This shows there is no defining factor that determines the overall quality of the track. (Moller, 1992)

It is also interesting to the note that Test B is the only binaural test that is create from a dummy head and is also the only test that is generally more one sided for the binaural recording, and this shows that the dummy head and head-worm microphone way to create binaural audio is much more effective than artificial binaural panners as they are able to capture the instruments and space using an general HRTF head shape in real time and does not need to convert something that was in stereo to binaural and not suffer any negative effects whereas despite binaural panners using generalised HRTF to also create binaural audio, it also comes with complications like aliasing which reduces the higher frequencies of the audio resulting in a duller sound and this effects the quality which effects the overall tracks effectiveness. (Hoose, 2015) (Moller, 1992) (Otani, et al, 2016)

(Figure 20, 21, 22) show the results from Test C, the electronic track, and it backs up the previous tests as stereo is the most favoured version when it comes to immersion and quality and this will be due to the binaural panner aliasing (Otani, et al, 2016) and the fact that the binaural version of the electronic track is duller due to the aliasing. But what is interesting is the results from the spatial image question (Figure 22) as only 1 tester voted for the binaural version, 7 voted for stereo and 5 voted for both which shows that the testers were very conflicted on this, perhaps due to the wider spatial imaging of the binaural version but higher quality and more clarity of the stereo version leaving the testers confused, this may also be due to the instruments positioning being the same between both versions, or perhaps due to Room Divergence as track emits a very spacey feeling and tonality so perhaps that there is no visual aid to this, it has confused them (Rumsey, 2017)

(Figure 23, 24, 25) shows the results from the Test D, Soundwalk. It is interesting how the immersion results have binaural being considered better but the quality results have stereo being better, which is interesting as every other test up to this point have shown that if something like the quality is favoured on the stereo version then the same goes for the binaural version and this test is conflicting with that trend.

Expect (Figure 25) shows that binaural is majority preferred in terms of spatial imaging than the stereo version which means that this test does not conflict and still falls in line with what has been discovered up this point and also shows that even if the quality is not favoured in the binaural version, the immersion and spatial image factors, if effective enough, can still allow for the binaural version to be more preferred and still be an effective piece of binaural audio. (Rumsey, 2011) (Moller, 1992) (Spagnol, Tavazzi and Avanzini, 2017)

It is worth noting that the only question within the Test D section that is one sided is the spatial image question, this is because both the immersion and quality questions are neck-to-neck which also further proves the point of all three qualities being interlinked as if the spatial image was not effective enough neither would be the immersion and thus also the quality but due to having a very effective spatial image it is able to make the other factors better. (Moller, 1992)

(Figure 26) shows the comparison question between Test A and Test B, this was created to discover whether the testers prefer Artificial Binaural or Dummy head (Hoose, 2015) (Otani, et al, 2016) It is worth noting that there was a typo within this question that was noticed after all of the testers answered it, and that being that the Yellow answer as it supposed to say Test A Part 2 instead of Part 1. It is safe to assume that the testers understood what it was meant to be regardless, but it is still unfair to use it as it is wrong. Fortunately, the Orange answer is still valid and still gives the same results as the Yellow answer.

For the Blue and Gray answers (Figure 26), they are the stereo versions of each test and it is interesting to note that Test A is preferred, this is perhaps due to the quality of each test as Test A is a multitrack mix that was taken from the Cambridge-mt website while Test B is a L&R X/Y Stereo Pair Acoustic recording and to keep each version of Test B as close to each other as possible, the C414s were set up above the dummy head and had to be set far enough to it could fairly capture each of the elements (Huber, 2017) (Rumsey, 2014) so the instruments and vocals sound slightly distant and roomy whereas Test A is all close miced and sounded great before any mixing was done.

The Orange answer shows that most of the testers did not believe the binaural version of Test A was better than the binaural version of Test B, this shows that the testers think the dummy head binaural audio is better than the artificial binaural audio and this will be due to the dummy head version having a wider spatial image and better clarity, as the dummy head does not suffer from aliasing like the binaural panners do and the dummy head supplies a real space that the binaural microphones record whereas the binaural panners only synthesis a binaural space with the instruments. (Hoose, 2015) (Otani, et al, 2016)

(Figure 27, 28) show results from a two-part question, (Figure 27) is the question asking whether they notice a difference between each parts and (Figure 28) is the question of what they noticed to be different. 10 out of 13 testers heard a difference when listening to both parts and the most voted for noticeable differences are Wider Spatial Image (Cyan) and Sat in the room (Orange) which means that generally the noticeable difference is a positive one and something that enhances the audio creating more immersion and from the other responses on the other questions, it is clear that these differences are welcomed ones. (Rumsey, 2011)

(Figure 29) is the last question of the test and it asks which is preferred out of all tests, and the results show that Test C and Test B are the most preferred. Which shows that the testers liked the way the binaural compliments these two styles of music as the electronic track being very spacey suits and the dummy head track being an acoustic session in a real environment, what both of these tracks have in common is their small amounts of instruments as the electronic track has 4 instruments and the dummy head track has 6 which means that binaural is perhaps best suited for more simpler compositions as both of the other tests has much more instruments and are sonically more complex. Another reason for Test B being the most preferred will be due to the use of the dummy head as this allowed to create a real space instead of a synthesised one as well as not suffering from side effects like aliasing. (Hoose, 2015) (Rumsey, 2011) (Otani, et al, 2016) (Moller, 1992)

1.5 Conclusion

(Figure 14, 15, 16, 17, 18, 19, 26, 29) show the results from Test A and Test B, as well as the comparison question and which test the consumers preferred. As analysed in the data analysis section, it is clear that the testers prefer the Dummy Head binaural track over its stereo counterpart and the artificial binaural track. This is due to a couple factors; The first being the immersion, quality, and spatial image effectiveness of each Test. It is clear by the results that the dummy head has superior overall quality, the artificial binaural is not completely ineffective compared to the dummy head track but due to the aliasing from the binaural panners it negatively effects all of the other factors of the track and as discovered, these factors are interlinked and effect each other and the overall quality. (Otani, et al, 2016) (Rumsey, 2011) (Moller, 1992) (Spagnol, Tavazzi and Avanzini, 2017) (Hoose, 2015)

The other factor is the method of creating the binaural track, Test B is superior due to utilization of the dummy head and head worn microphones as this allows for a more natural, realistic recreation of how humans hear, whereas with the binaural panners it has no contextual structure to rely on instead relying on mathematical equalisation, coding and a general HRTF, the dummy head is also crafted with the general HRTF in mind but due to the more natural and real approach it lends to better sounding binaural audio as well as the dummy head suffering no negative side effects. The results from this test show that dummy head binaural reproduction is much more successful at making quality, immersive binaural audio. (Hoose, 2015) (Rumsey, 2011) (Otani, et al, 2016) (Moller, 1992) (Spagnol, Tavazzi and Avanzini, 2017)

(Figure 20, 21, 22, 23, 24, 25, 29) show the results from Test C and Test D back up the main trend of the factors, immersion, quality and spatial image being interlinked and all effect each other's perception on the overall effectiveness of the track as well as also show the binaural panners can work in dependence of context of which tracks they are being used on, as Test A is a rock track with vocals and many instrumentation whereas Test C and D are not standard mainstream style music and Test D is not music at all, and the elements within Test C and D lend themselves to the binaural panner as they are not significantly affected by the aliasing and can still be enjoyed despite being technically slower quality because of the aliasing. (Rumsey, 2011) (Moller, 1992) (Spagnol, Tavazzi and Avanzini, 2017) (Otani, et al, 2016)

It is clear from all that has been discovered that the dummy head method of creating binaural audio is far superior to binaural panners, but is also worth noting that binaural panners are not a lost cause and can be used in particular context, that being either used more subtly or used on music that are not mainstream style genres like Rock music, and more utilized on atmospheric electronic music. It is also worth noting that the dummy head would not have such limitations on musical context and can be used on anything due to having no negative side effects and a more realistic binaural image. (Rumsey, 2011) (Moller, 1992) (Spagnol, Tavazzi and Avanzini, 2017) (Otani, et al, 2016) (Hoose, 2015)

The answer the question, as shown from the results binaural music is ready for commercial use as the dummy head track test shows that the binaural version is completely preferred over the stereo counterpart and it is also preferred over the other binaural tracks which show that the dummy head method is ready to be used to record binaural music and thus making it more useable for commercial use whereas the artificial binaural panners are not ready and need to be -+improved upon to remove the aliasing or be used more subtly rather than be used to on every element of the track. (Rumsey, 2011) (Moller, 1992) (Spagnol, Tavazzi and Avanzini, 2017) (Otani, et al, 2016) (Hoose, 2015) (Rumsey, 2014)

If this question were ever to be asked again in the future, post coronavirus, it is worth noting that the test results would be more accurate as some of the issues listed in the methodology could have been avoided. Further research in to this question would be to try to discover new methods of recording music in binaural using dummy heads and other binaural audio recording techniques to discover what the best methods of binaural recording are and what they are more suited for recording. For example maybe using head worm microphones in the drummers ears to record the drums.

Reference List

1. Bayley, 2008. Surround Sound For The DAW Owner. Masters. University of Colorado Denver
2. Brown, A. 2017. Soundwalking: Deep Listening and Spatio-Temporal Montage. [e-journal] 69. Available at: <https://www.mdpi.com/2076-0787/6/3/69/htm> [Accessed 18 January 2020]
3. Hoose, S. 2015. Creating Immersive Listening Experiences with Binaural Techniques. [e-journal] 55. Available at: <https://www.jstor.org/stable/26574394?read-now=1&seq=3#page_scan_tab_contents> [Accessed 23 December 2019]
4. Huber, M, D and Runstein, R. 2017. Modern Recording Techniques. 711 Third Avenue, New York. Taylor and Francis Group.
5. Moller, H. 1992. Fundamentals of Binaural Technology. [e-journal] Available at: < > [Accessed 23 December 2019]
6. Otani, et al. 2016. Effects of spatial aliasing in sound field reproduction: reproducibility of binaural signals. [e-journal] Available at: <https://www.researchgate.net/publication/316604156_Effects_of_spatial_aliasing_in_sound_field_reproduction_Reproducibility_of_binaural_signals>
7. Reddy, S, C and Hegde, R. 2016. Design and development of bionic ears for rendering binaural audio. [e-journal] Available at: <https://ieeexplore.ieee.org/document/7746678/authors#authors> [Accessed 23 December 2019]
8. Rumsey, F. 2014. Spatial Audio: Binaural Challenges. [e-journal] 62/11 Available at: <http://www.aes.org.ntu.idm.oclc.org/e-lib/browse.cfm?elib=17553> [Accessed 23 December 2019]
9. Rumsey, F. 2011. Whose Head is it Anyways? [e-journal] 59/9 Available at: <http://www.aes.org.ntu.idm.oclc.org/e-lib/browse.cfm?elib=15982> [Accessed 24 December 2019]
10. Spagnol, S, Tavazzi, E and Avanzini, F. 2017. Distance rendering and perception of nearby virtual sound sources with a near-field filter model. [e-journal] Available at: <https://www.sciencedirect.com/science/article/pii/S0003682X16302407> [Accessed 24 December 2019]
11. Zhang, W, et al. 2017. Surround by Sound: A review of Spatial Audio Recording and Reproduction. [e-journal] 7/5 Available at: <https://www.mdpi.com/2076-3417/7/5/532/htm> [Accessed 23 December 2019]

Bibliography

1. Brannen, J. 2016. Mixing Methods: Qualitative and Quantitative Research. 2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN 711 Third Avenue. Taylor and Francis Group.
2. Durrington, M. 2008. Record Store. [Video] Berkeley, CA, United States. Berkeley Media
3. Hollebeek, L. 2016. Sounds of music: Exploring consumers' musical engagement. [e-journal] 33/6 Available at: <https://search.proquest.com/docview/1828657713/fulltextPDF/95A50CB6DE794655PQ/1?accountid=14693> [Accessed 21 January 21 2020]
4. Kruger, L, et al. 2018. Individual interviews or focus groups? Interview format and women's self-disclosure. [e-journal] 245/255 Available at: <https://www.tandfonline.com/doi/abs/10.1080/13645579.2018.1518857?journalCode=tsrm20> [Accessed 21 January 2020]

Appendix

1. Ausim3d, 2011. HeadZap Sustem. [image online] Available at: <http://www.ausim3d.com/products/HeadZap/index.html> [Accessed 27 December 2019]
2. Dissertation Main Test, 2020. Binaural vs Stereo – Main Test (Responses). Available at: <https://drive.google.com/open?id=1SWCPMxKIS3mUvbr1w3OmpcJhav0JZd_6>
3. Dissertation Pilot Test, 2020. Binaural vs Stereo – Pilot Test (Responses). Available at: <https://drive.google.com/open?id=1SWCPMxKIS3mUvbr1w3OmpcJhav0JZd_6>
4. Marta, 2018. Znalezione obrazy dla zapytania hrtf measurement. [image online] Available at: <https://frl.nyu.edu/mysterious-hrtf/> [Accessed December 2019]
5. Neumann, 2018. KU 100. [image online] Available at: <https://en-de.neumann.com/ku-100> [Accessed 28 December 2019]
6. Rode, 2020. NTG2. [online] Available at: <http://www.rode.com/microphones/ntg-2> [Accessed 21 January 2020]
7. Sennheiser, 2019. Ambeo Smart Headset. [image online] Available at: <https://en-uk.sennheiser.com/in-ear-headphones-3d-audio-ambeo-smart-headset> [Accessed 27 December 2019]