Audio Engineering Society

Convention Paper

Presented at the 139th Convention
2015 October 29–November 1 New York, USA

Listener preference for height channel microphone polar patterns in three-dimensional recording

Will Howie[1,2], Richard King[1,2], Matthew Boerum[1,2], Dave Benson[1,2], and Alan Joosoo Han[1,2]

[1] Graduate Program in Sound Recording, McGill University, Montreal, Canada

[2] Centre for Interdisciplinary Research in Music Media and Technology, Montreal, Canada

correspondence: william.howie@mail.mcgill.ca

ABSTRACT

A listening experiment was conducted to determine if a preference exists amongst three microphone polar patterns when recording height channels for three-dimensional music production. Seven-channel 3D recordings of four different musical instruments were made using five-channel surround microphone arrays, augmented with two Sennheiser MKH 800 Twin microphones as height channels. In a double-blind listening test, subjects were asked to rate different mixes of the same recordings based on preference. The independent variable in these mixes was the polar pattern of the height channel microphones. Analysis of the results found that the vast majority of subjects showed no statistically significant preference for any one polar pattern.

1. INTRODUCTION

In a recent AES Engineering Brief, experimental microphone techniques for three-dimensional classical music recording were discussed. [1] Using a combination of omnidirectional and cardioid microphones, a fourteen-channel microphone array was designed to capture surround sound plus height information while recording a small baroque ensemble. Four omnidirectional microphones assigned to front and rear height channels were found to contain too much direct sound from the ensemble. This correlation of direct sound with the main layer microphones made it difficult to achieve an ideal recording balance, as increasing the level of the height channels past a certain point tended to destabilize the image of the ensemble, "smearing" the instruments upward. Based on this, and the aesthetically superior sound captured by cardioid pattern lateral (+/-90˚) microphones [1], the primary author hypothesized (after a number of listening and mixing sessions) that directional microphones would be the best choice for capturing height information in a way that yields both a strong focused ensemble image, and excellent listener envelopment.

1.1. Capturing an ideal balance of sound

Listeners of recorded classical music have become accustomed to an idealized, realistic recreation of a live performance in an acoustic space. [2] [3] The main goal of the classical music recording engineer is to capture an ideal balance of direct and diffuse sound: many microphone techniques have been developed to do just this, for both stereo and 5.1 surround sound. [4] [5] These techniques, however, fall short of capturing the fully immersive experience of listening to a live performance in a real acoustic environment. The addition of height channels allows the recording engineer to enhance the presentation by improving the depth, presence, envelopment, naturalness, and intensity of the recordings. [6] [7] [8] Most stereo and 5.1 channel microphone arrays call for the use of specific microphone polar patterns. [4] [5]

1.2. 3D audio for home listening

Japan Broadcasting Corporation (NHK) plans for 8k television broadcasts with 22.2 multichannel audio to be in common use in time for the 2020 Olympic Games in Tokyo. [9] Other three-dimensional audio formats, such as Auro 3D [10] and Dolby Atmos [11], are already available for home entertainment systems. Some record labels, such as 2L, are producing commercially available, 9.1 channel music recordings, using Pure Audio Blu-ray as a delivery format. [12] A casual review of current home theatre systems will reveal a number of products that feature seven-channel surround sound, often including the option of using two of those speakers as height channels. Given the growing availability and importance of three-dimensional audio, correspondingly few published works have discussed recording techniques for the above formats, only some of which describe actual three-dimensional music recordings in detail. [13] [14] [16] [17]

2. TEST RECORDING

In order to test the initial hypothesis stated in the introduction, a simple test recording was designed to capture height information using multiple microphone polar patterns simultaneously. The recording, of a contrabass-recorder, took place in a medium-large studio space, using a seven-channel microphone array (see below table and Fig. 1):

Main Left	Schoeps MK 21
Main Right	Schoeps MK 21
Main Centre	Neumann U87
Surround Left	Schoeps MK4
Surround Right	Schopes MK4
Height LL (+90°)	Sennheiser MKH800 Twin
Height RR (-90°)	Sennheiser MKH800 Twin

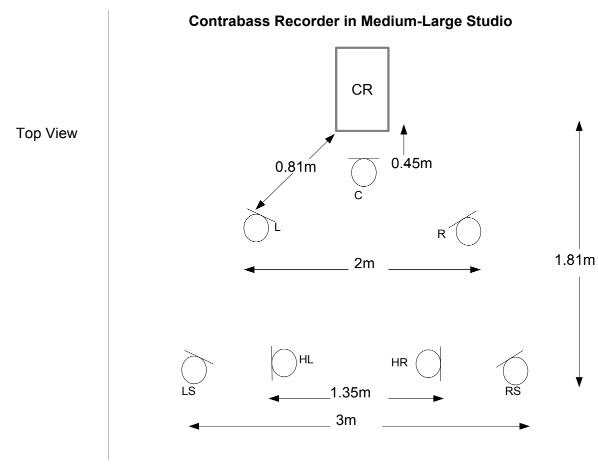

Figure 1: Overhead layout of test recording.

This small-scale array was designed to be simple to set up, and compatible with any current 3D audio system (the height microphones could, in theory, be assigned to any pair of height channels). For this recording, the decision to use the LL and RR height channels (Fig. 11) was based on a previous experimental recording [1], as well as research showing the importance of lateral sound energy and reflections for achieving listener envelopment. [18] Sennheiser MKH800 Twin microphones, which feature a back-to-back dual capsule design, were used to record height information. The microphone's dual output (one from each transducer) allows the recording engineer to create any polar pattern by adjusting the balance between the two capsules. The recordings were monitored in McGill University's Studio 22 (see section 3 for more details).

After the test recordings were completed, the primary author, performer, and composer of the piece spent time mixing and comparing the available polar patterns of the height channel microphones, focusing on cardioid, omnidirectional and bi-directional. All three listeners were surprised by how different each polar pattern sounded, and how greatly the overall sound of the recording was affected by changing the height channel polar patterns. It was observed that the cardioid height channels contributed to a strong, focused instrument image, while the omnidirectional height channels gave a less stable image, but a richer room sound. None of the listeners enjoyed the sound of the bi-directional height channels, which had a displeasing timbre.

3. LISTENING TEST

Based on the results of the recording described in section 2, a test was designed to investigate whether or not strong preferences exists among listeners for height channel microphone polar patterns.

3.1. Creating Listener Test Stimuli

Using the method from the test recording as a guide, nine more seven-channel three-dimensional music recordings were made. Seven different solo instruments were recorded in three different acoustic spaces. All three acoustic spaces are located in the Schulich School of Music's Elizabeth Wirth Music Pavilion. The large scoring stage (Music Multimedia Room) measures 24.4m x 18.3m x 17m, and has little acoustical treatment (Fig. 6, 7). The Medium-large studio, measuring 11m x 7m x 6.1m, has a combination of absorptive and diffusive acoustical panels in the lower part of the room, with the upper walls being untreated (Fig. 8). The isolation booth has similar acoustic treatments to the medium-large studio, and measures 5m x 3.2m x 6.1m (Fig. 9). The following table shows what instruments were recorded in what spaces:

Instrument	Acoustic Space
1. Harp	Large scoring stage
2. Piano (Yamaha C7)	Large scoring stage
3. Drum Kit	Large scoring stage
4. Cello	Medium-large studio
5. Trumpet	Medium-large studio
6. Drum Kit	Medium-large studio
7. Acoustic Guitar	Medium isolation booth
8. Male Vocal	Medium isolation booth
9. Drum Kit	Medium isolation booth

3.2. Microphone Choice and Placement

Aside from publications by Geluso [14], Theile and Wittek [15], Williams [16], and Hamasaki and van Baelen [17], little material exists to guide current engineers looking to record music in 3D. As such, the listening test stimulus recording sessions were viewed as an excellent opportunity to investigate a number of possible recording techniques. For all stimulus recordings, the spacing between and angle of the height channel microphones remained the same, though their height and distance from the sound source varied quite a bit depending on the instrument and room. For the main layer microphones (L, C, R, LS, RS), microphone choice and placement varied depending on the instrument, acoustic space, and repertoire being performed. For the harp and cello, fairly traditional spaced arrays were used, with a focus on achieving a strong centre image and fairly diffuse surrounds. For the drums and guitar, a more pop-based approach was taken, resulting in somewhat asymmetrical setups. (Fig. 2-9)

For all recordings, the height channel microphones were Sennheiser MKH800 Twin, facing +/-90˚ from the ensemble. All microphones were routed to a Sony SIU-100 System Interface Unit, using the internal microphone preamps and analog to digital conversion. Recordings were made to a Pro Tools HD system, at 96kHz/24bit resolution.

Instrument	Microphone
Drums Overhead L	Neumann U87
Drums Overhead R	Neumann U87
Drums Kick Spot	Audio Technica AT4047
Drums Snare Spot	Shure SM57
Drums LS	Schoeps MK4
Drums RS	Schoeps MK4
Harp L	Schoeps MK2
Harp C	Schoeps MK4
Harp R	Schoeps MK2
Harp LS	Schoeps MK4
Harp RS	Schoeps MK4
Acoustic Guitar Spot	Schoeps MK4
Acoustic Guitar L	Schoeps MK4
Acoustic Guitar R	Schoeps MK4
Acoustic Guitar LS	Schoeps MK4
Acoustic Guitar RS	Schoeps MK4
Cello L	Schoeps MK21
Cello C	Schoeps MK4
Cello R	Schoeps MK21
Cello LS	Schoeps MK4
Cello RS	Schoeps MK4

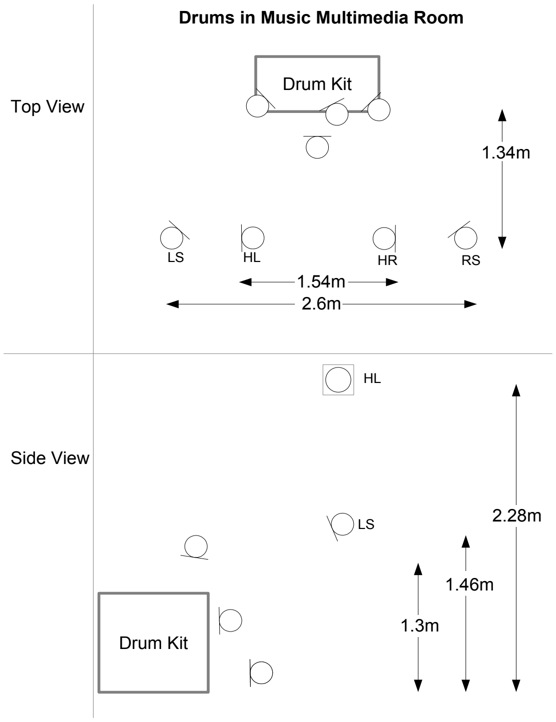

Figure 2: Drums in large scoring stage.

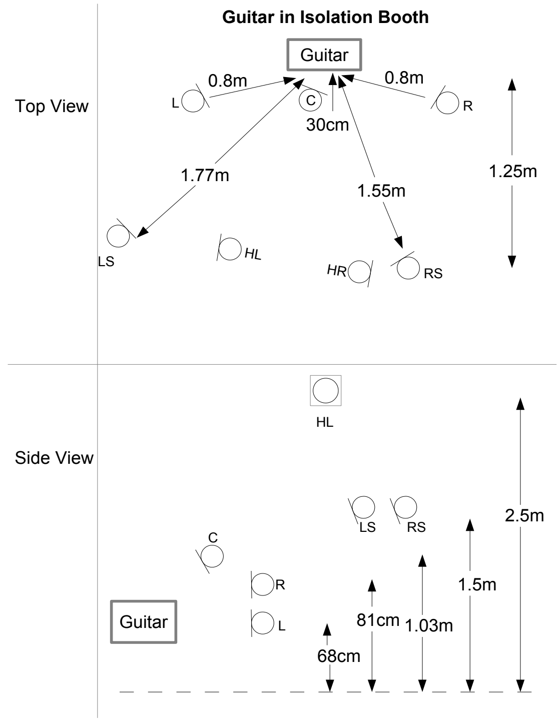

Figure 4: Guitar in isolation booth.

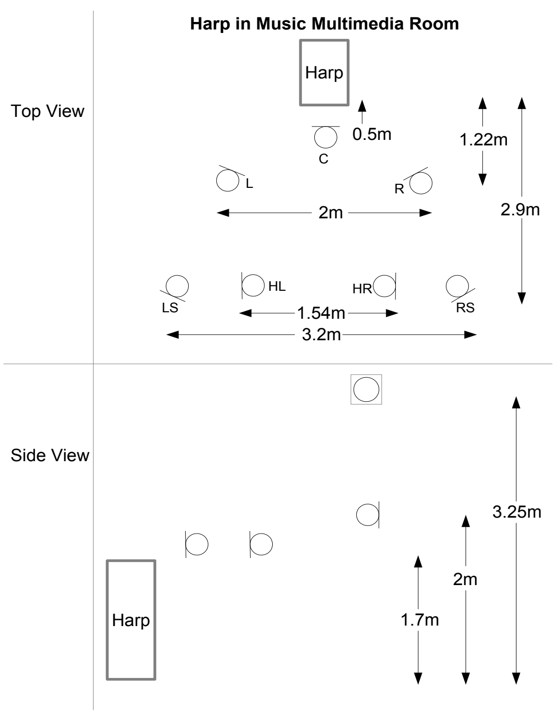

Figure 3: Harp in large scoring stage.

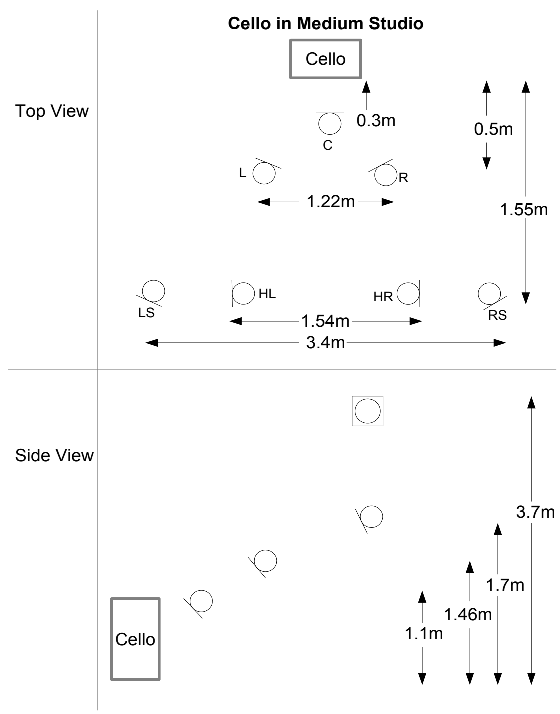

Figure 5: Cello in medium studio.

Figure 6: Harp in large scoring stage.

Figure 7: Drums in large scoring stage.

Figure 8: Cello in medium-large studio.

Figure 9: Guitar in Isolation Booth.

3.3. 3D Audio Control Room

All 3D audio playback (recording and mixing of stimuli, test administration) took place in McGill University's Studio 22 (Fig. 10), an acoustically treated listening room with 30 channels of discreet audio playback via Musikelectronic Geithain GmbH *M-25* speakers. The 30 speakers are arranged to accommodate both 22.2 multichannel sound [19] and Auro 3D [10]. Studio 22 fulfills the ITU-R BS.1116 [20] requirements.

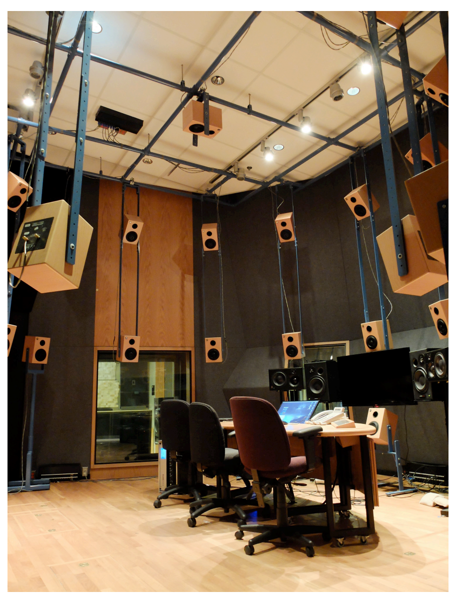

Figure 10: Studio 22.

3.4. Mixing and Level Matching Polar Patterns

For the listening test, three microphone polar patterns were chosen for the height channels: cardioid, omnidirectional, and figure 8. The authors did not want to overwhelm listeners with too many different musical stimuli, and as such, chose to focus on four recordings they considered to have the highest sound quality and greatest contrast in acoustic and musical content: harp (scoring stage), drums (scoring stage), cello (medium studio), and acoustic guitar (isolation booth). By changing the balance of the Senheisser MKH 800 Twins' dual outputs, pairs of height channels featuring each of the polar patterns under test were created.

Three audio engineers independently level-matched the different polar pattern height channel mixes for each stimulus. This was accomplished in Pro Tools: listening only to the height channels, each engineer compared the different polar pattern pairs (HLL and HRR, see Fig. 11), and balanced these pairs until they were perceived as being of equal loudness. The mix volume levels for each author were recorded and the averages of those levels were used to perform the final level matching.

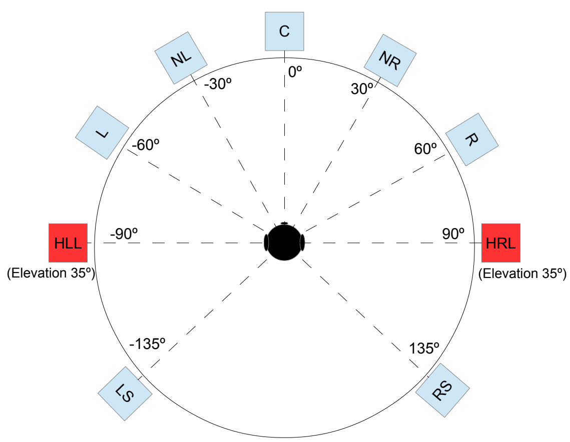

Figure 11: Speaker layout for listening test.

Each seven-channel stimulus recording was then balanced to convey a sense of depth and realism to the instruments, using a "direct sound/instrument in front, ambience to the sides, behind and above" approach. It was considered very important that mixes contain enough height channel information to be pleasant, realistic and enveloping, rather than exaggerating the differences between polar patterns. The goal was not to create an "obvious" listening test, but one that mirrored the subtle mix differences that professional engineers discriminate between on a daily basis. The use of the +/-60˚ left and right channels contributed to a greater sense of width and spaciousness in the front image. Drum overhead microphones were panned to the NL and NR speakers (Fig. 11), which gave a more realistic impression of instrument size and width. Three seven-channel mixes were then created for each stimulus, the independent variable being the polar pattern of the height channels. Each mix represented a 30 second musical excerpt.

3.5. Test Design and Implementation

A double-blind listening test was designed, using Max/MSP. Subjects were seated in Studio 22's "sweet spot", and presented with an interactive GUI (Fig. 12). For each trial, one of the four musical stimuli played on a repeating loop. Subjects were asked to select between mixes labeled as "A", "B" and "C" on the GUI. Subjects could switch between mixes at any point during stimulus playback, as many times as needed. Though subjects listened to only one mix at a time, all three stimulus mixes were synced in playback. For each trial, subjects were instructed to "rate the three mixes in order of general preference", using 100 point sliders. A comments box in the GUI allowed subjects the option to briefly explain why they made their decision.

Within the current literature, there are a number of examples of listening tests comparing different multichannel microphone techniques. Some tests have asked listeners to rate techniques based on specific attributes, such as spaciousness [21] [22], envelopment, depth, and localization [22]. Other tests have focused on general listener preference between recording techniques [23] [24]. In the present study, subjects were not given any specific subjective qualities or attributes to consider when making their preference

Each subject completed four trials of each stimulus, for a total of sixteen trials. The presentation order of the different stimuli was randomized. The order in which the different polar pattern mixes were assigned as letters "A" "B" and "C" for each trial was also randomized. A total of 29 subjects performed the listening test. The subjects ranged greatly in terms of age and listening experience:

Subject Age (in years)	Number of Subjects
18-25	14
26-32	10
33-39	3
40-50	1
51+	2
Subject Identification	**Number of Subjects**
Pro Engineer/Producer	7
Recent SR* Masters Graduate	4
Current SR Masters Student	4
SR/Music Undergrad	7
Other McGill Students	8

*SR = McGill Sound Recording

For each subject, the Max/MSP patch generated a text file showing their preferences and comments. After completing the test, subjects were asked to fill out a brief demographic survey, which included space for general comments about the test experience.

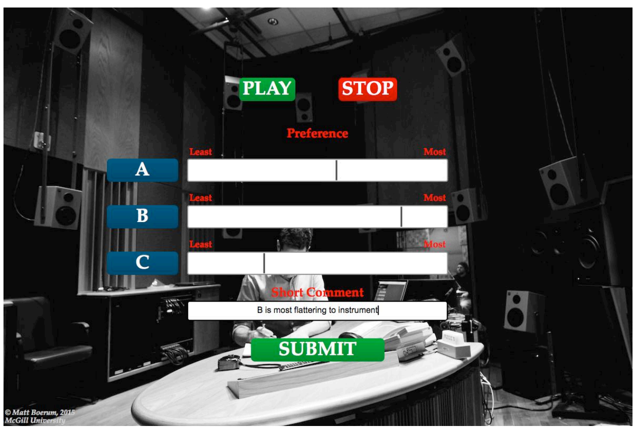

Figure 12: Screen shot of listening test GUI.

4. RESULTS

Prior to the experiment, plans were made to analyze the preference scores for the three microphone polar patterns in two ways: with all subjects pooled together, and with each subject considered separately. The first analysis would reveal general trends valid for the entire population of subjects, while the second would reveal individual differences in preference.

In all tests, the data for the four instruments were pooled together. No attempt was made to investigate interaction effects between instrument and polar pattern.

4.1. Normality Tests

4.1.1. Pooled Scores

As a first step in the analysis, the pooled preference scores were tested for normality. All three were significantly non-normal (Cardioid: $W = 0.979$, $p < .001$; Figure-8: $W = 0.984$, $p < .001$; Omnidirectional: $W = 0.983$, $p < .001$).

Histograms of the data showed two visual features that deviated from a bell-curve shape (Fig. 13).

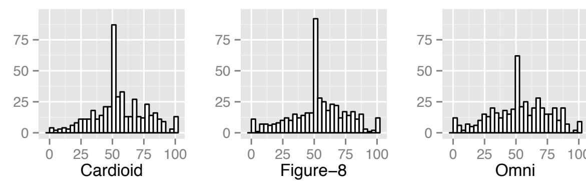

Figure 13: Histogram of preference scores by polar pattern.

First, there were a large number of responses at the centre of the scale, with a value of exactly 50. This can be attributed to the fact that the sliders were reset to this value at the start of each trial. It seems that, in many cases, subjects left the sliders at this initial value rather than moving them. Second, there were a large number of responses at the ends of the scale. This excess of extreme scores resulted from a small number of subjects who gave highly polarized ratings.

4.1.2. Individual Scores

The data from individual subjects were also tested for normality. While some subjects produced normally distributed scores, many gave responses exhibiting the features described above.

These non-normalities precluded the use of parametric tests, such as analysis of variance (ANOVA), to check for differences between groups. Instead, the Kruskal-Wallace test was used. Kruskal-Wallace is a non-parametric alternative to a one-way ANOVA that operates on ranked data.

4.2. Pooled Preferences

When the preference scores of all subjects were pooled together, no significant differences between the polar patterns were found, H(2) = 1.58, p = .45. (Fig.14)

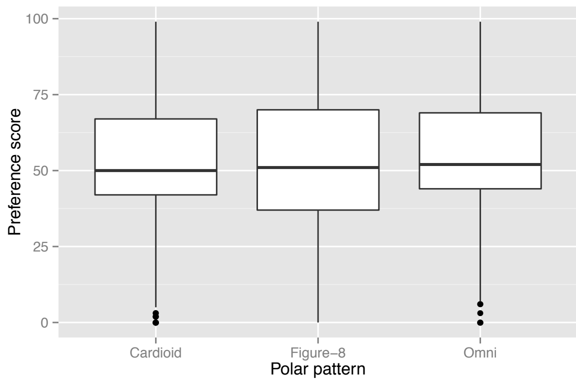

Figure 14: Polar pattern preference ratings, pooled across all subjects.

4.3. Individual Preferences

When the preferences of individual subjects were tested, differences were revealed in only two of the 29 cases: subject 2, H(2) = 15.6, p = .012; and subject 28, H(2) = 3.97, p = .037. (p-values were corrected with Holm-Bonferroni)

For subject 2, comparisons of mean ranks showed that Figure-8 had a significantly lower preference rank than Cardioid (difference = 19.3). For subject 28, Figure-8 had a lower rank than Omnidirectional (difference = 17.7). In both cases, the critical difference (α = 0.05 corrected for the number of tests) was 11.8. Raw preference scores for subjects 2 and 28 are shown in Fig 15.

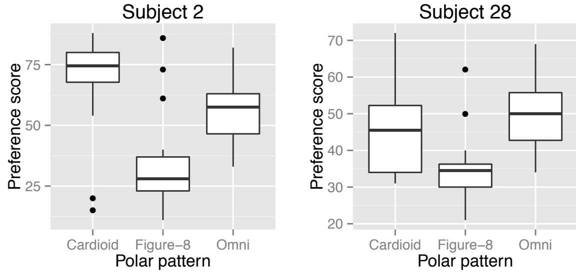

Figure 15: Preference scores for subjects 2 and 28.

For the vast majority of subjects (27 out of 29), the rankings given to the three microphone polar patterns were not significantly different (Fig 16).

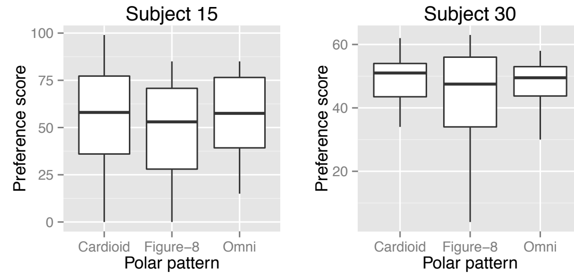

Figure 16: Preference scores for subjects 15 and 30. These subjects were typical in exhibiting no significant preference for any polar pattern.

5. ANALYSIS AND FUTURE WORK

For the vast majority of test subjects (27/29), no significant preference for any one height channel microphone polar pattern was shown.

5.1. Subtlety of Polar Pattern Differences

The sonic changes between the three different polar pattern mixes were likely too subtle for most listeners, including several professional recording engineers and producers. Either subjects did not perceive any difference between the three microphone polar patterns, or the differences were too subtle to be detected by the statistical tests. This view is fairly consistent with the results of the subject demographic surveys. Of the 29 subjects, 22 left general comments about the test. Ten of those subjects commented on the subtlety or difficulty of the test. Below are several sample comments:

Subject 001: “I found it very hard to hear any differences with the cello, harp and guitar.”

Subject 004: “In general, the differences were, for me, very subtle. In some cases, I did not not even perceive a difference.”

Subject 027: “I found the cello recordings virtually indistinguishable.”

This view would also offer some explanation as to why there were so many subject responses with a value of 50 (see section 4.1). It seems probable that had the differences in sonic quality between the three microphone polar patterns been stronger and more obvious, subjects would have felt more compelled to move the preference sliders to a correspondingly larger degree.

5.1.1. Inconsistency of Subject Preferences

When looking at the raw data for each subject, there were numerous instances where subjects were inconsistent with their preferences even with averaged responses over multiple trials. This was true for all four stimuli. This is also supported by post-test comments. For example:

Subject 010: "The subtle differences in the 3 mixes for each track had me questioning myself, especially in the middle of the test."

Subject 003: "I don't think I was consistent."

5.2. Future Work

The authors have yet to examine the comments that subjects left in the optional "short comments" text box in the listening test GUI. The majority of subjects took the time to fill in comments for at least some of the trials (many for all), and a number of subjects filled in very detailed comments for each trial. Our hope is that a thorough examination of these comments may lead to a better understanding of what perceptual attributes were important to the listeners.

6. CONCLUSION

The vast majority of test subjects showed no significant preference between three microphone polar patterns (cardioid, omnidirectional and figure-8) for height channels in three-dimensional music recordings. This was true among expert and non-expert listeners, students and professional engineers. Within the context of a music mix, the sonic differences between the height channel polar patterns are obviously very subtle, perhaps too subtle for most listeners to make strong or consistent preference ratings. The process of creating the three-dimensional test stimuli was both valuable and educational, yielding a number of ideas for microphone arrays that could be explored in future recordings.

The potentially positive result of this research is that recording engineers currently exploring three-dimensional music recording should not feel bound by the example of past microphone techniques that specify that certain polar patterns be used for certain applications.

7. ACKNOWLEDGEMENTS

This work was supported by the Social Sciences and Humanities Research Council, and the Centre for Interdisciplinary Research in Music Media and Technology. A special thanks to all the musicians who donated their time and skills to this project.

8. REFERENCES

[1] W. Howie and R. King, "Exploratory microphone techniques for three-dimensional classical music recording," in *AES 138th Convention*, Warsaw, Poland, 2015, pp. 1-4

[2] K. Hamasaki et al, "Approach and Mixing Technique for Natural Sound Recording of Multichannel Audio," in *AES 19th International Conference*, Schloss Elmau, Germany, 2001, pp. 1-6.

[3] A. Fukada, "A Challenge in multichannel music recording," in *AES AES 19th International Conference*, Schloss Elmau, Germany, 2001, pp. 1-8.

[4] M. Dickreiter, *Tonmeister Technologies*. New York, NY: Temner Enterprises Inc., 1989.

[5] "Surround Techniques," Internet: http://www.dpamicrophones.com/en/Mic-University/Surround%20Techniques.aspx [June 29, 2015]

[6] K. Hamasaki et al, "Effectiveness for height information for reproducing presence and reality in multichannel audio system," in *AES 120th Convention*, Paris, France, 2006, pp. 1-15.

[7] T. Kamekawa et al, "Evaluation of spatial impression comparing 2ch stereo, 5ch surround, and 7ch surround with height for 3D imagery," in *AES 130th Convention*, London, UK, 2011, pp. 1-7.

[8] S. Kim et al, "Subjective evaluation of multichannel Sound with Surround-Height

channels," in *AES 135th Convention*, New York, USA, 2013, pp. 1-7.

[9] I. Sawaya et al, "Dubbing Studio for 22.2 Multichannel Sound System in NHK Broadcasting Center," in *AES 138th Convention*, Warsaw, Poland, 2015, pp. 1-9.

[10] B. Van Daele and W. Van Baelen, "Productions in Auro 3D: Professional workflow and costs," Auro Technologies, Rev. 0.6, February 28, 2012, pp. 12-16.

[11] "Dolby Atmos for the Home Theatre", *Dolby Laboratories, Inc.*, San Francisco, USA, April 2015, pp. 1-18.

[12] "Use Case: Morten Lindberg, 2L – Norway," Internet: http://www.merging.com/news/use-cases/morten-linderg-2l-norway [July 12, 2015]

[13] "Multichannel sound technology in home and broadcasting applications," *Report ITU-R BS.2159-4*, May 2012, pp. 1-54.

[14] P. Geluso, "Capturing Height: The Addition of Z Microphones to Stereo and Surround Microphone Arrays," in *AES 132nd Convention*, Budapest, Hungary, 2012, pp. 1-5.

[15] G. Theile and H. Wittek, "Principals in Surround Recording with Height," in *AES 130th Convention*, London, UK, 2011, pp. 1-12.

[16] M. Williams, "The Psychoacoustic Testing of a 3D Multiformat Microphone Array Design, and the Basic Isosceles Triangle Structure of the Array and the Loudspeaker Reproduction Configuration," in *AES 134th Convention*, Rome, Italy, 2013, pp.1-14.

[17] K. Hamasaki and W. Van Baelen, "Natural Sound Recording of an Orchestra with Three-Dimensional Sound," in *AES 138th Convention*, Warsaw, Poland, 2015, pp.1-8.

[18] T. Hanyu and S. Kimura, "A new objective measure for evaluation of listener envelopment focusing on the spatial balance of reflections," *Applied Acoustics*, vol. 62, pp. 155-184, 2001.

[19] "Ultra High Definition Television Audio Characteristics and Audio Channel Mapping for Program Production", *SMPTE ST 2036-2-2008*, The Society of Motion Picture and Television Engineers: White Plains, USA, 2008.

[20] "Methods for the Subjective Assessment of Small Impairments in Audio Systems Including Multichannel Sound Systems", ITU-R Recommendation BS.1116-1, International Telecom Union: Geneva, Switzerland, 1997, pp. 1-26.

[21] L. Simon and R. Mason, "Spaciousness Rating of 8-Channel Stereophony-Based Microphone Arrays," in *AES 130th Convention*, London, UK, 2011, pp. 1-10.

[22] T. Kamekawa et al, "Corresponding Relationships between Physical Factors and Psychological Impressions for Microphone Arrays for Orchestral Recording," in *AES 123rd Convention*, New York, USA, 2007, pp. 1-13.

[23] Kassier et al, "An Informal Comparison Between Surround-Sound Microphone Techniques," in *AES 118th Convention*, Barcelona, Spain, 2005, pp. 1-17.

[24] S. Kim et al, "An Examination of the Influence of Musical Selection on Listener Preferences for Multichannel Microphone Techniques," in *AES 28th International Conference*, Piteå, Sweden, 2006, pp. 1-14.