5-4-2019

Mid-Side to X/Y Equivalence: Turning Microphone Theory into Practice

Brent T. Hauer
brenthauer@gmail.com

Mid-Side to X/Y Equivalence: Turning Microphone Theory into Practice

Master's thesis presented to the faculty of the
Audio Engineering Graduate Program
of
The Mike Curb College *of* Entertainment *&* Music Business
Belmont University, Nashville TN

In partial fulfillment of the requirements for the degree

Master of Science
with a major in
Audio Engineering

Brent T. Hauer
May 4, 2019

Advisors

Wesley A. Bulla
Doyuen Ko
Eric Tarr

ABSTRACT

The mathematical transformation of the Mid-Side microphone array (M-S) to stereophonic left and right outputs results in an equivalent coincident microphone configuration (X/Y). The output pattern of the combined mid (M) and side (S) signals can be altered by manipulating the polar pattern of the mid (M) microphone and by adjusting the ratio between the M and side (S) microphones. This study investigated the technical and perceptual attributes of equivalent M-S and X/Y patterns in order to determine the consistency of the theoretical with the practical application. A simulated jazz sextet was recorded using a selection of M-S, and their equivalent X/Y, techniques. Each pair of recordings was measured to compare the timbral and spatial attributes of the equivalent configurations. Listening tests provided a perceptual metric of this theoretical equivalence. Comparisons and subjective ratings provided a more complete picture to determine whether the mathematical transformation was supported by perception. A significant preference for M-S techniques suggested the mathematical equivalence transformation model does not reliably predict an equivalent stereophonic perception of X/Y and M-S recording configurations.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

DEFINITIONS OF TERMS

Stereophonic: The processes of recording and reproduction that utilize two loudspeakers to emulate sound on the horizontal plane in front of the listener [1].

Omnidirectional Microphone: A microphone that is equally sensitive to sound from all directions [2].

Cardioid Microphone: A microphone that is most sensitive to sounds on the frontal axis and relatively less sensitive sounds from the sides and rear of the microphone [3].

Bidirectional Microphone: A microphone that is most sensitive to sounds coming from the front and rear of the microphone and relatively less sensitive to sounds from the sides [3].

hypercardioid microphone: A cardioid microphone with a sensitivity similar to, but less than, the bidirectional microphone [3].

X/Y: A stereophonic microphone technique where the directional microphones are vertically aligned on a common axis (i.e., coincident) and set at an angle to each other in the horizontal plane [4].

Mid-Side (M-S): A stereophonic microphone pair that places one microphone (M) aimed directly at the centerline of the sound source with a bidirectional microphone (S) oriented perpendicular on the centerline and coincident with the frontal microphone. Outputs of the two microphones must be processed through a sum and difference procedure to resolve the signals into conventional stereophonic signals [4].

1. INTRODUCTION

The equivalence between Mid-Side (M-S) and coincident stereophonic microphone (X/Y) configurations is well supported by theory; however, this equivalence lacks empirical evidence. The mathematics used to convert M-S recordings for stereophonic playback was well documented by Wesley Dooley and Ronald Streicher in their extensive review of the M-S stereophonic technique [5]. By utilizing the mathematical concepts necessary to matrix M-S for stereophonic playback, a series of tables were presented to display the transformation of various M-S configurations and their equivalent X/Y configuration (see Figure 1). While the mathematics used to derive these tables, described in Section 2.3, are well-documented, several practical concerns justify the use of M-S over X/Y techniques. The frequency response of M-S configurations benefits from discrete on-axis pickup of the center of the sound image. Coloration of the reverberant field is further reduced by the on-axis pickup from the S microphone. Additionally, M-S configurations exhibit absolute monophonic compatibility. Lastly, as a practical convenience, M-S provides the ability for an audio engineer to adjust the stereophonic signal either by varying the pattern of the middle (M) microphone, or by altering the difference in level between the M and side (S) microphone channels while recording or during post-production signal processing.

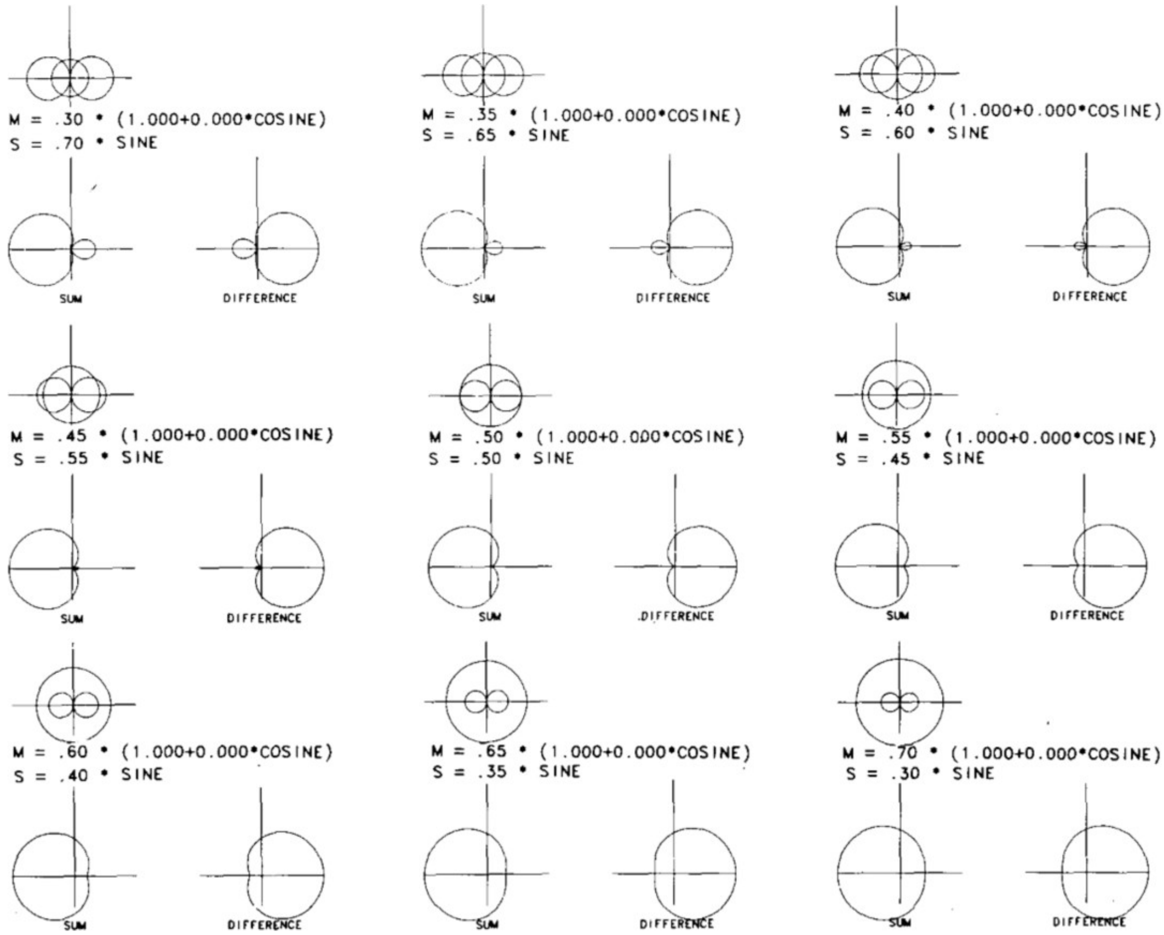


Figure 1. M-S to equivalent X/Y transformation table for M-S with an omnidirectional M microphone. Adapted from [5].

By applying the equations presented in [5], a script was created in MATLAB [6] to generate plots for the five M-S patterns under investigation. The patterns varied in the configuration of the M microphone utilizing omnidirectional, wide cardioid, cardioid, hypercardioid, and bidirectional patterns. The M-S patterns with their equivalent X/Y configurations are shown in Figure 2 through 6. Each transformation from M-S to X/Y was based on equal gain for the M and S microphones (see Appendix A.1 for MATLAB® code).

This research investigated the effects that the M-S transformation process had on the resulting stereophonic output. Using physical testing and a series of listening tests, this study compared a variety of M-S and equivalent X/Y techniques. These comparisons sought to answer the following research question:

- Is the mathematical equivalency of the M-S to X/Y transformation process supported by perception?

Listening tests were executed to determine if:

- Listeners could identify M-S from equivalent X/Y produced recordings;
- Comparative spatial image ratings for sound image width and depth perspective, when reproduced through a stereophonic system, were equivalent between M-S and X/Y pairs;
- Comparative spectral ratings for timbral balance and sound source definition, when reproduced through a stereophonic system, were equivalent between M-S and X/Y pairs.

Answering these questions may inform the design and implementation of stereophonic microphones. The implications of this study could prove useful for audio engineers when choosing which stereophonic microphone technique to use as a main stereophonic pair in live, as well as studio recording scenarios. If M-S can capture a stereophonic sound that is equal to or better than that produced by X/Y, audio engineers could take full advantage of the flexibility provided by the adjustable nature of M-S configurations without compromising the overall quality of the recordings.

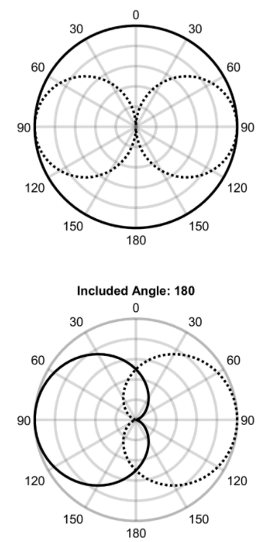


Figure 2. M-S with omnidirectional M (top), equivalent X/Y configuration (bottom). Included angle is 180 degrees.

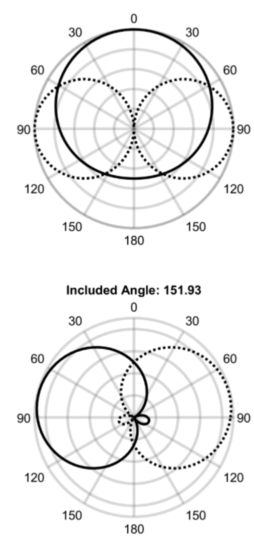


Figure 3. M-S with wide-cardioid M (top), equivalent X/Y configuration (bottom). Included angle is 151.93 degrees.

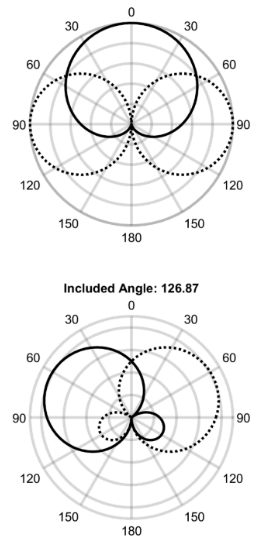


Figure 4. M-S with cardioid M (top), equivalent X/Y configuration (bottom). Included angle is 126.87 degrees.

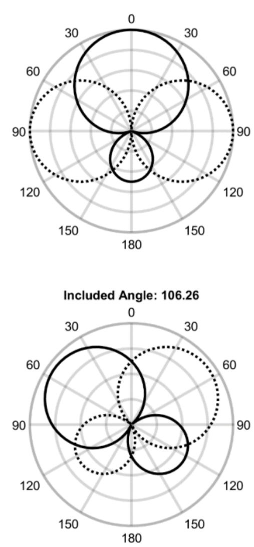


Figure 5. M-S with hypercardioid M (top), equivalent X/Y configuration (bottom). Included angle is 106.26 degrees.

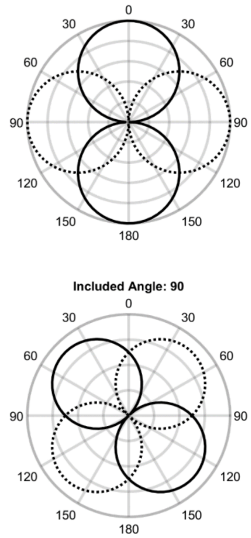


Figure 6. M-S with bidirectional M (top), equivalent X/Y configuration (bottom). Included angle is 90 degrees.

2. PRIOR ART

Stereophonic microphone techniques can be used to convey a sense of the recording environment by providing acoustical cues for width and depth [7]. Commonly used techniques can be organized into three categories: spaced, near-coincident, and coincident pairs [4], [7]–[9]. The first use of stereophonic microphones at the 1881 Paris Electrical Exhibit made use of spaced microphones [10]. Spaced techniques were utilized by A.C. Keller in his telephony experiments conducted at Bell Laboratories in the 1930's. These experiments resulted in two albums of stereophonic recordings of the Philadelphia Orchestra that were considered the finest recordings that conductor Leopold Stokowski had ever heard [11], [12]. Coincident techniques, which utilize a pair of microphones placed as close together as possible, include configurations such as "X/Y", "Mid-Side", and "Blumlein" [4], [7] – [9], [13]. These techniques, also referred to as "intensity stereo" techniques, rely on capturing differences in level while minimizing differences in time between the microphones [4], [7], [8]. Coincident techniques provide stereophonic imaging which is unaffected by the distance from the sound source. It has been claimed that the absence of inter-channel time differences (ICTD) may cause these techniques to lack a sense of space [4], [14].

2.1 Psychoacoustics of Stereophonic Listening

In a stereophonic loudspeaker system, which utilizes two sound producing sources, the ears receive signals from both loudspeakers [15]. While working at Bell Laboratories in 1934, JC Steinberg and W.B. Snow experimented with inter-channel level difference (ICLD) as the primary driver of stereophonic localization [16], [17]. However, natural free-field localization is accomplished through a combination of time and level difference cues [18]. Their explanation failed to consider the inter-channel time difference (ICTD), which has been shown to be the primary localization method in the presence of lower-frequency sounds [19] – [22]. Although stereophonic loudspeakers are capable of reproducing both ICLD and ICTD, stereophonic microphone techniques vary in their utilization of time-based cues [4].

2.2 X/Y and M-S Recording Techniques

Coincident microphone techniques are those in which the microphones are aligned vertically, and set at an angle to each other horizontally referred to as the included angle [4], [7]. These types of configurations are often referred to as "X/Y" and can utilize several types of microphone patterns [4], [7], [8]. Cardioid microphones set with an included angle between 90 and 120-degrees are commonly used in an X/Y configuration, with the use of hypercardioid microphones allowing for more distant placement [4]. Alternatively, bidirectional microphones can be used in what is known as the "Blumlein" technique [4], [7] – [9], [13]. This method is used when the capture of ambience or room reverberation is desired [4], [23]. The Blumlein recording technique is unique because it allows the listener to judge distance from the sound source in a similar manner as one could while listening to a live performance [23].

Mid-Side recording is accomplished using a bidirectional S microphone combined with a M microphone that can vary in its pattern. These microphones are placed in a coincident configuration with the M microphone pointed toward the center of the sound source. The S microphone is set at an angle of 180-degrees to the M microphone. The outputs of these two microphones must be processed to form left and right signals. With the positive lobe of the S microphone facing to the left, equation 1 shows the calculations to convert M-S to left and right [4], [7] – [9].

$$L = M + S$$

$$R = M - S \tag{1}$$

M-S recording techniques benefit from the ability to alter the width of the output by varying the ratio between the M and S channels, even after the source has been recorded [4], [7], [8]. Additionally, M-S recording is purely compatible when summed for monaural playback. Equation 1 shows that a monoaural signal (or $L + R$) is the same as $(M + S) + (M - S)$, or more simply, $2M$. As a result, when summed for monaural playback, the S channel disappears completely, leaving only the monaural M channel [4], [7] – [9].

2.3 M-S to Stereophonic Equivalence

As noted by [4], the conversion of M-S to left/right results in conventional X/Y signals. In a more in-depth analysis, Streicher and Dooley parsed out the equations that allow M-S signals to be directly related to X/Y configurations [5]. Equation 2 shows the process used to generate polar diagrams for the M and S channels.

$$M = |A(B + (1 - B) \cos \theta|$$

$$S = |(1 - A) \sin \theta|$$

$$for\ 0 \leq \theta \leq 2\pi \tag{2}$$

For these equations, "A" represents the decimal fraction for the M microphone's contribution to the M-S matrix. "B" represents the polar pattern of the M microphone using a decimal fraction of its omnidirectional to bidirectional components [5]. Using the values found for M and S, the application of equation 1 produces polar diagrams for the equivalent X/Y configuration. Equation 3 shows the calculation used to find the included angle of the X/Y configuration [5].

$$\theta = \arctan\left(\frac{1 - A}{A(1 - B)}\right) \tag{3}$$

Using these equations, [24] concludes that the M-S configuration is more versatile than X/Y due to three characteristics of M-S, where:

1. M-S can utilize an omnidirectional microphone, whereas X/Y is dependent on directional microphones.
2. The reproduction angle of M-S recordings can be easily altered.
3. The recording angle of M-S systems can be altered between 0-degrees and 120-degrees, whereas X/Y techniques can only be altered between 90-degrees and 120-degrees.

Additionally, where X/Y microphone configurations capture the center of the sound source off-axis, M-S configurations capture the center on-axis. This reduces the effects of off-axis capture when the recording is summed for monaural playback [24].

2.4 Mathematical and Perceptual Comparison Tests

Microphone techniques can be measured both mathematically and perceptually. Mathematically studies of microphones can often lead to the discovery of new techniques. Perceptual testing allows a researcher to determine if differences between techniques can be heard by a listener. These two approaches are mutually beneficial, as mathematical testing can help explain the results of perceptual testing, while perceptual tests help determine if mathematically derived differences are significant enough to be heard.

2.4.1 Mathematical Approaches

To compare spatial images created with common microphone techniques, [25] designed a binaural model to predict the imaging of a given microphone technique. This model relied on calculating head-related transfer functions (HRTFs) as well as estimations of hair-cell stimulation and the central nervous system's decision-making process based on aural cues. The study resulted in an improved model that can be used to analyze microphone techniques with less difficulty accounting for multiple peak phenomena than previous models [25].

Calculating inter-channel time and level differences for a microphone technique is one way to derive a localization curve for that technique. Listening tests can also be used to directly measure the accuracy of localization for recordings produced with various microphone techniques. By comparing localization curves derived mathematically with those captured with listening tests, [26] found consistency between the calculations and real-world performance for several common stereophonic microphone techniques. The study concludes that while testing is time-consuming and work-intensive, modelling can provide accurate results for stereophonic techniques. Additionally, the study recommends simple listening tests for the study of surround microphone techniques as the calculations for such methods are highly advanced and expensive to implement [26].

By accounting for the frequency dependent nature of correlation functions, [27] developed a "Diffuse-Field Image Predictor" to predict perceived width. While this model was developed mathematically, listening tests were used to verify its effectiveness. The model uses a weighting function, along with frequency-dependent correlation equations derived for each type of microphone technique studied, to assign a single value to describe the coherency of a given microphone technique. A listening test based on MUSHRA was used to show that the DFI Predictor is an effective tool for predicting the perceived spatial width of any stereophonic microphone technique [27].

2.4.2 Perceptual Approaches

Listening tests are a common tool for the audio researcher, however, many studies fail to properly account for psychological and acoustical variables. Attempts have been made to standardize methods for listening tests using several key principles: testing that is reproducible, reflects only the audible characteristics of the system under examination, and display the significance of audible differences appropriately [28]. Given the importance of sound quality to any work in the field of audio engineering, it is a term that is often poorly defined in audio research. Researchers have developed a more defined set of perceptual parameters, producing three distinct classes of listening assessments. Class A assesses loudness, pitch, and/or duration, while class B assesses timbral and/or spaciousness, and class C assesses a combination of any or all sensations from classes A and B. While class A parameters are easily defined, class B and C require more detailed descriptions. Definitions for several parameters, including timbral and sound quality, as well as a mural of practical terms that can be used to describe specific aspects of sound quality can be found in [29].

These attempts to standardize listening tests have not been entirely successful and the search for improved methods persists. A scene-based paradigm for spatial quality assessment, discussed by [30], has resulted in unidimensional descriptions of spatial attributes which separate descriptions of sources, environments, and scene-based parameters. Rumsey notes that this paradigm is far from definitive and simply adds to the debate over the standardization of listening test parameters [30]. More recently, several standards produced by the International Telecommunication Union (ITU) and the European Broadcasting Union (EBU) have been found to provide adequate global parameters for spatial quality that appropriately describe stereophonic sound quality and can be used in surround spatial quality tests as well [31]. These methods include evaluations of small and intermediate impairments in audio samples, as well as assessments specific to "classical" music [32], [33].

Comparisons of stereophonic microphone techniques have often employed listening tests. Orchestral recordings using several simultaneously implemented microphone techniques made it possible to develop listening tests that utilized the same stimulus for each technique being tested. This ensured that listeners rated only the aspects of the techniques and not the differences between multiple performances by live musicians.

Several stereophonic microphone techniques were rated for parameters such as liveness, intimacy, continuity, warmth, and brilliance [34]. These vague parameters would likely not satisfy the scrutiny of more modern audio researchers. More recently, [35] used a digital piano to produce stimuli for a similar listening test. Using an instrument that can recreate the same performance multiple times allowed for corrections to a glaring issue with Ceoen's study by utilizing consistent placement of microphone techniques. The results of this study suggest that listeners prefer clarity over spatial width, and that classifying microphone techniques by the distance between microphones may not adequately describe the attributes of each technique studied [35].

3. METHODS

Research was conducted in two separate phases. The first phase focused on measuring the physical parameters of the microphones and techniques used throughout the study. The second phase utilized two listening tests to determine perceptual differences between equivalent techniques and capture listener preference ratings for several criteria.

3.1 Objective Microphone Measurements

The AKG C-426 B microphone is a dual-diaphragm stereophonic microphone that features a rotating capsule placed directly above a stationary one. This design allows for a wide variety of coincident configurations. Additionally, the microphone features a remote power unit that allows the user to adjust the polar pattern of each capsule independently. The polar patterns available include omnidirectional, cardioid, bidirectional, and six additional intermediate patterns [36]. Polar and frequency response diagrams are provided by [35] for the cardioid, omnidirectional, and bidirectional polar patterns.

In order to determine the real-world characteristics of the AKG C426-B microphone, measurements were taken and compared to the published specifications. Additionally, averaged frequency response and inter-channel correlation measurements for each technique were used to ensure that microphone placement was consistent and that the microphone techniques were performing as expected in both the M-S and X/Y configurations.

3.1.1 Microphone Specifications

The published data from AKG was verified, and polar diagrams for the additional polar patterns were generated in an anechoic environment at Belmont University. To create the polar diagrams, the microphone was rotated by an electronic turntable multiple times while several frequencies of sine waves were produced through a time-aligned, coaxial loudspeaker, shown in photograph 1. Measurements were made at octave intervals from 125 Hz through 16 kHz. These recordings were captured at a sampling rate of 96 kHz and converted into polar diagrams [6]. The code used for this process divided the signals into 500 segments and calculated the RMS value of each segment. These RMS values were then plotted on a polar diagram to generate continuous patterns for each frequency band.

Photograph 1. AKG C426-B microphone set up in anechoic environment to generate polar response diagrams.

Frequency response diagrams for each capsule of the microphone were also produced in an anechoic environment at Belmont University. The same time-aligned, coaxial loudspeaker, shown in Photograph 2, was used in conjunction with a 2700 Series Audio Analyzer from Audio Precision [37]. Measurements were conducted over the bandwidth from 40 Hz to 20 kHz.

Photograph 2. AKG C426-B microphone set up in anechoic environment to generate frequency response diagrams.

3.1.2 Stereophonic Configurations

Objective measurements of the stereophonic microphone configurations included averaged frequency response as well as inter-channel correlation. These measurements relied on multitrack recordings and sine sweeps (20 Hz – 20 kHz) that were captured inside McAfee Concert Hall at Belmont University. Processing was performed using [6].

Sine sweep recordings were processed using the Fast Fourier Transform (FFT) to extract frequency-domain amplitude. Each of the five M-S and X/Y techniques were processed by performing the FFT on the left and right channels and averaging the results. Inter-channel correlation was calculated for each of the five M-S and X/Y configurations using equation 4 [38].

$$r = \frac{\overline{S_{xy}}}{\overline{S_x} * \overline{S_y}} \tag{4}$$

Where: r is the correlation coefficient; $\overline{S_{xy}}$ is the covariance of the left and right $\overline{S_x}$ signals; is the standard deviation of the left $\overline{S_y}$ signal; and is the standard deviation of the right signal. The MATLAB code for these measurements appears in Appendix A.2.

Comparisons of stereophonic image were conducted using a goniometer script [39]. Stereophonic audio recorded with each of the ten microphone configurations (see Section 3.2.2) was processed to display the phase differences between the left and right channels.

3.2 Listening Tests

The second phase of research for this study was conducted using two listening tests. The first, an ABX comparison, asked listeners to identify differences between matched pairs of recordings. The second, a preference test, compared matched pairs of recordings based on several spatial and timbral parameters.

3.2.1 Subjects

Nine subjects participated in the ABX listening test. All subjects were graduate students in audio engineering technology at Belmont University and reported normal hearing at the time of the listening test. For the preference listening test, 13 subjects participated. The subject pool for this portion of the listening tests came largely from the same pool of participants as the ABX listening test. However, additional participants were either graduate audio engineering students or professors in the audio engineering technology program. Each of these subjects self-reported normal hearing at the time of the listening test.

3.2.2 Stimuli Creation

Stimuli for the listening tests were generated inside McAfee Concert Hall at Belmont University. A series of loudspeakers, amplifiers, and subwoofers was placed on the stage of the concert hall to simulate a live performance of a jazz sextet, shown in Photographs 3 through 7. The instrumentation for the jazz sextet was drum set (rear, right), bass (rear, center), piano (rear, left), trombone (front, right), trumpet (front, center), and tenor saxophone (front, left). A multitrack recording was played through this array and was recorded using each of the five M-S configurations as well as the five equivalent X/Y configurations, as shown in Table 1. The AKG C426-B microphone was connected to a two-channel preamp which passed line-level signals to a USB powered interface. Laser measurements were taken for each configuration to ensure consistent placement of the microphone twelve feet above the center of the front edge of the stage. The recorded tracks were edited to create five separate stimuli, ranging from 25 to 35 seconds, that were each peak-normalized to reduce any differences in level. No additional post-processing was performed on any of the recorded stimuli.

Table 1. Mid-Side and X/Y microphone techniques used to generate stimuli.

Mid-Side Mid Component Polar Pattern (B)	X/Y Polar Pattern (B)	X/Y Included Angle (deg.)
Bidirectional (0)	Bidirectional (0)	90.00
hypercardioid (0.25)	hypercardioid/Bidirectional In-between (0.125)	106.26
Cardioid (0.5)	hypercardioid (0.25)	126.87
Wide Cardioid (0.75)	Cardioid/hypercardioid In-between (0.375)	151.93
Omnidirectional (1)	Cardioid (0.5)	180.00

The track used was titled “All the Gin is Gone” and was performed by the Maurizio Pagnutti Sextet [40]. This recording was recorded at Artesuono Recording Studios and was made available for educational use by Cambridge Music Technology. The drum set was recorded using monaural microphone configurations for the kick drum, snare drum, toms, trombone, trumpet, and tenor saxophone. The bass was recorded via direct box while the drum set overheads and piano were recorded using stereophonic microphone techniques. Stereophonic room microphones were included with the original multitrack recording but were not utilized to generate stimuli. While the use of stereophonic microphone techniques in the original recording will introduce ICTD unique to the recording space, the playback system used to create stimuli was identical for each M-S and X/Y configuration.

Photograph 3. Jazz sextet setup used for stimuli creation.

Photograph 4. Loudspeaker configuration used for drum set.

Photograph 5. Amplifier used for bass.

Photograph 6. Loudspeaker configuration used for piano.

Photograph 7. Loudspeaker configuration used for tenor saxophone (left), trumpet (center), and trombone (right).

3.2.3 ABX Testing

The ABX test was used to determine if trained listeners could correctly identify pairs of M-S and X/Y recordings. This test was required prior to preference testing in order to assure that trained listeners could provide reliable preference data. Listeners were presented with recordings using equivalent microphone techniques. In each trial the “A” and “B” stimuli were randomly assigned to the M-S technique or its equivalent X/Y technique. Listeners were asked to identify “X” by comparing it to “A” and “B” and deciding which stimulus matched “X”. Nine listeners each performed 100 trials, which included 20 trials from each of the five M-S and equivalent X/Y pairs. Stimuli varied in length from 25 to 35 seconds. This testing was performed using loudspeakers and an ABX application developed using [6].

3.2.4 Preference Testing

Testing for listener preference was used to compare M-S and equivalent X/Y recordings based on spatial and timbral qualities. For this test, listeners rated pairs of stimuli based on their preference for the following parameters, as defined by [41]:

- Spatial Qualities
  - Sound image width: The subjective impression of an appropriate width of the sound stage in the stereo sound field.
  - Depth perspective: The subjective impression that the sound image has an appropriate front to back depth.
- Timbral Qualities
  - Timbral balance: The subjective impression of the accurate portrayal of the different sound characteristics of the sound source.
  - Sound source definition: The subjective impression that different instruments or voices sounding simultaneously can be identified and distinguished.

The ratings provided by listeners were entered using an electronic survey with a touchscreen interface. Each parameter was represented by a slider that started in the middle with the label "No Preference." Listeners were able to drag the slider towards either the "A" or "B" stimulus to indicate preference for that recording. The slider had three stopping points for "A Little Better," "Somewhat Better," and "Much Better." Listeners were provided with a physical switch to compare the "A" and "B" stimuli while using the same pair of loudspeakers. The test was comprised of sixty trials. Each of the five techniques investigated by this study was presented using five different sections of music and were presented twice for each listener. These fifty trials were used to generate data for the experiment while the additional ten trials acted as a control data. For these trials the "A" and "B" stimuli were identical. The stimuli varied in length from 25 to 35 seconds and repeated until the listener stopped playback.

The preference testing was performed at Robert E. Mulloy (REM) Studio D at Belmont University. Objective measurements were performed using Room EQ Wizard [42] to ensure that this studio meets or exceeds the recommendations related to listening rooms found in [43] (see Appendix B). REM Studio D has been shown to meet or exceed most of the recommendations presented in [43]. Where the room fails to meet the recommendations (total area and low-frequency reverberation) exceptions already built into the recommendation can be applied. For these reasons, REM Studio D was chosen for preference testing throughout this study.

Loudspeakers in REM Studio D were calibrated so that each loudspeaker produced 78 dBC, slow-weighted when -20 dBFS pink noise was produced through a signal generator plugin. Listeners were given the option to control the output using a single fader. Additional user-operated hardware included an AB switchbox, a touchscreen tablet computer, and transport controls. The setup of the listening room is shown in Photographs 8 and 9. The AB switchbox contained two stereo inputs, which were connected to a USB audio interface, and a single stereo output, which was connected to the loudspeaker amplifiers. This allowed listeners the ability to switch seamlessly between stimuli. Preference was indicted on the touchscreen tablet using a series of sliders. Transport controls were used to stop playback, indicating that the listener had completed their ratings and was ready for the next trial.

Photograph 8. REM Studio D set up for preference testing.

Photograph 9. User-operated hardware for preference testing. AB switchbox (left), touchscreen tablet (center), and volume/transport control (right).

4. RESULTS

4.1 Objective Microphone Measurements

4.1.1 Microphone Specifications

Polar patterns generated using the script shown in Appendix A.1 were compared to the published specifications for the AKG C426-B microphone. Results show that the microphone performed as expected for frequencies up to 4 kHz. Above 4 kHz, the directionality of the microphone begins to deteriorate. While this is shown in the published specification, the measured results show that this effect was more pronounced than expected. Figures 7 and 8 show the generated polar patterns, while Figure 9 shows the published specification sheet.

Frequency response curves were compared to the published specifications for the AKG C426-B microphone. Results show that the microphone performed as expected. Figures 10 and 11 show the generated frequency response curves and Figure 12 shows the published specification sheet.

Generated plots for both polar and frequency response show that the AKG C-426 B microphone used in this study was operating as expected and within the published specifications. Anomalies in the directionality of the microphone at frequencies above 4 kHz can be attributed to imperfections in the test environment. Variances between the generated frequency response curves and those in the published specification sheet can be attributed to the frequency response, particularly in low frequency bands, of the loudspeaker used to test the microphone.

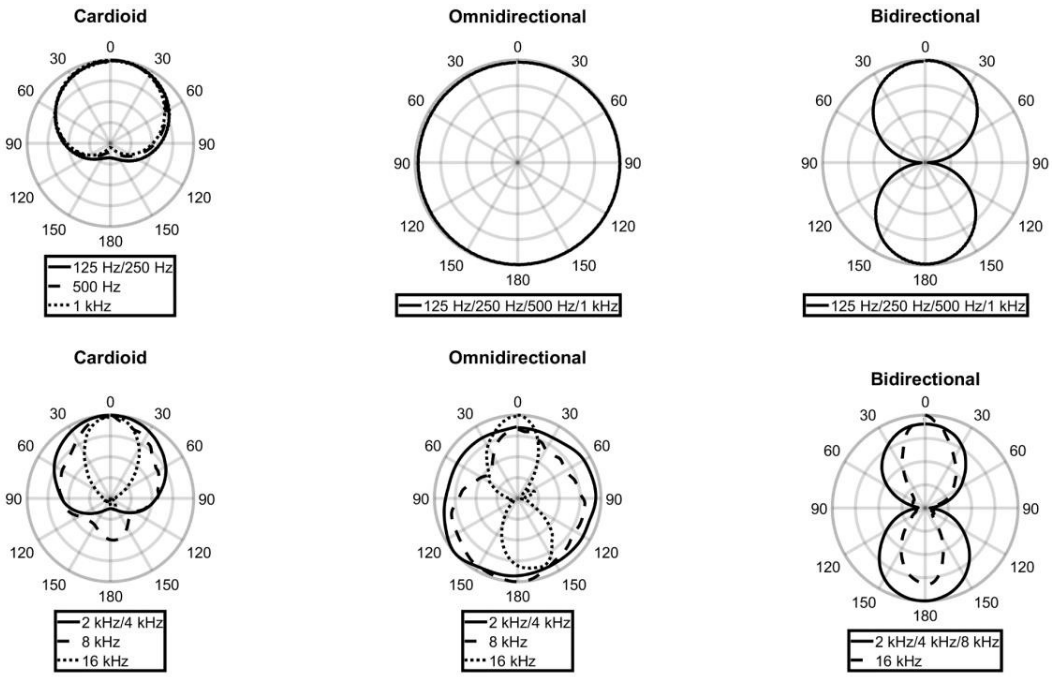


Figure 7. Polar patterns for channel 1 of the AKG C426-B microphone.

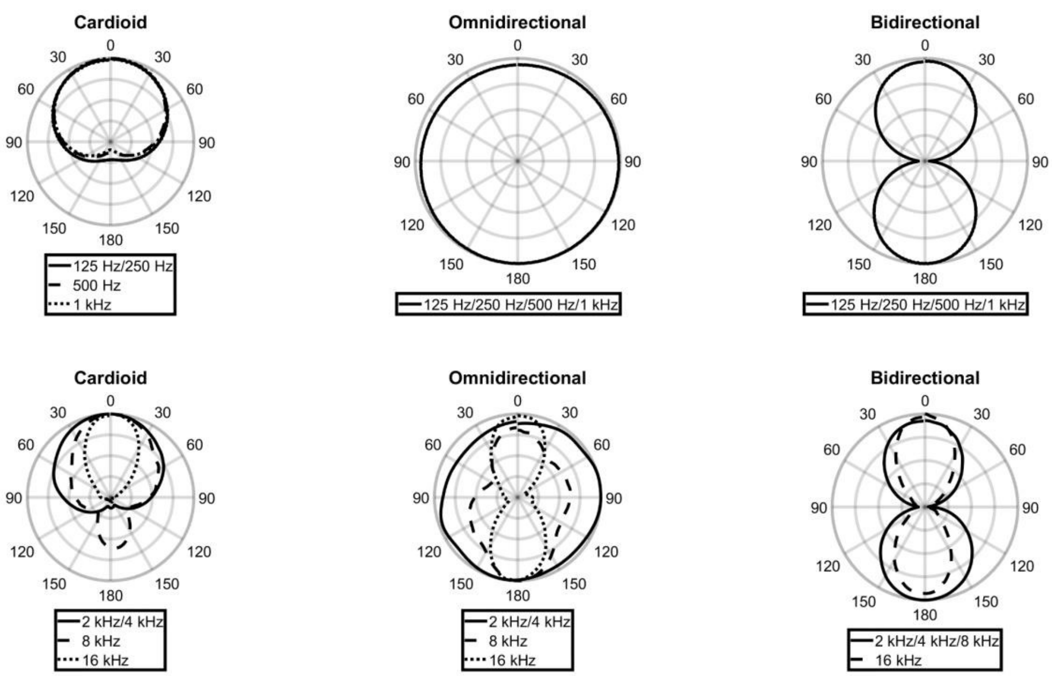


Figure 8. Polar patterns for channel 2 of the AKG C426-B microphone.

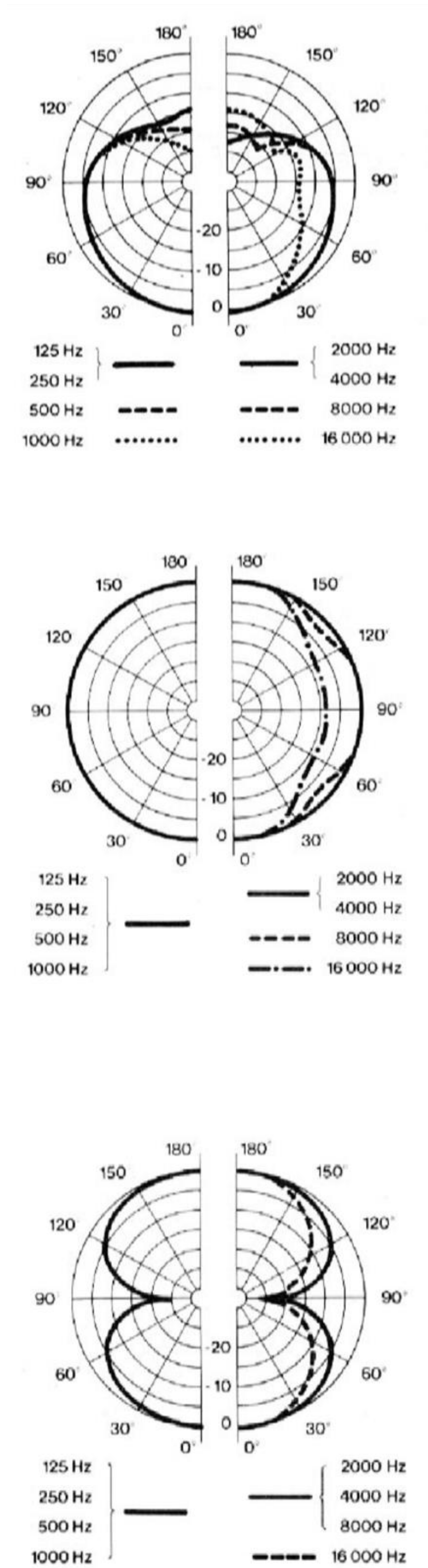


Figure 9. Published polar diagrams for the AKG C426-B microphone from [36].

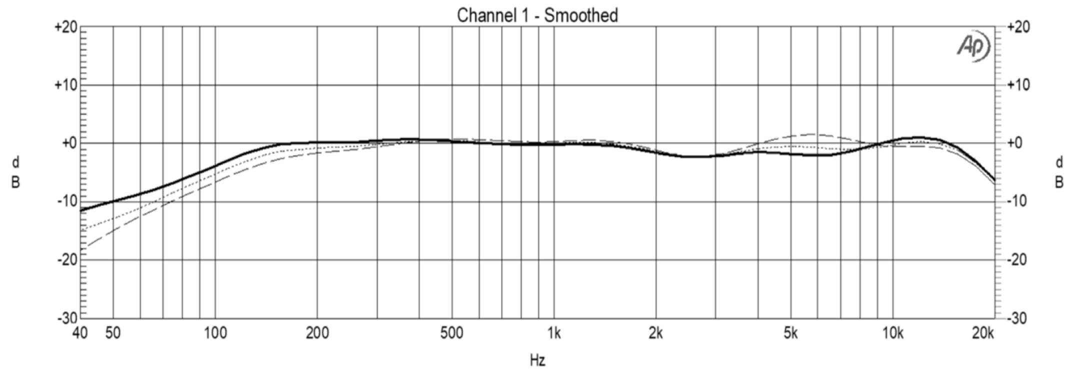


Figure 10. Frequency response curves for Channel 1 of the AKG C426-B microphone. omnidirectional (solid), cardioid (dotted), and bidirectional (dashed)

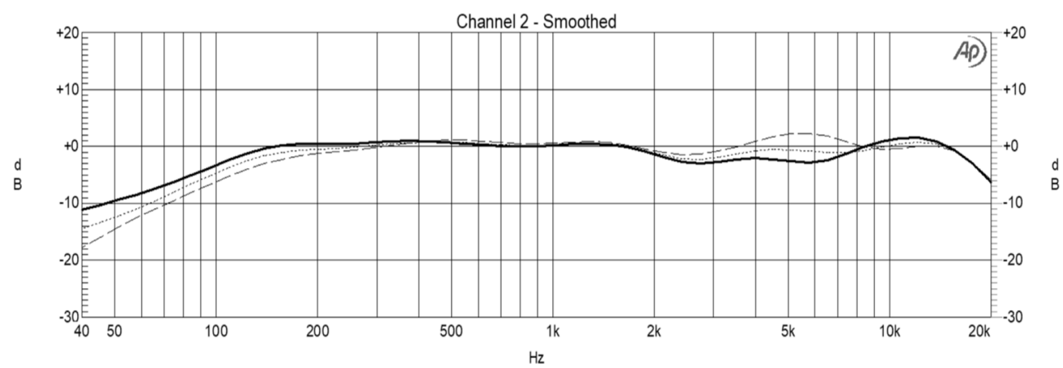
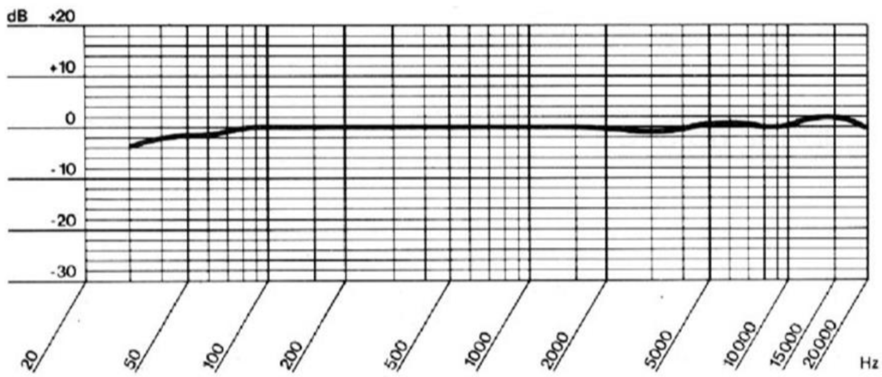
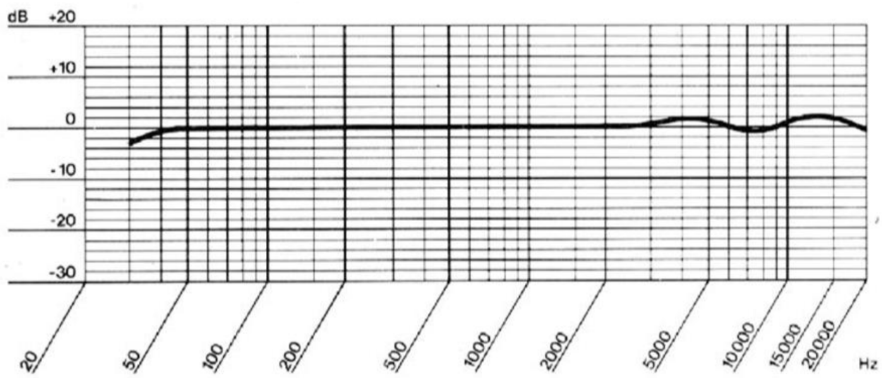


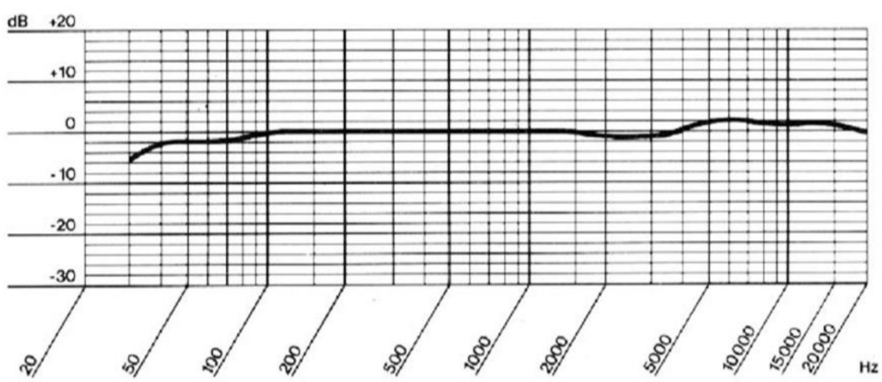
Figure 11. Frequency response curves for channel 2 of the AKG C426-B microphone. omnidirectional (solid), cardioid (dotted), and bidirectional (dashed).



Cardioid



Omnidirectional



Bidirectional ("figure-eight")

Figure 12. Published frequency response curves for the AKG C426-B microphone from [36].

4.1.2 Stereophonic Configurations

Results for averaged frequency response indicate strong agreement between M-S and X/Y pairs up to 8 kHz. In each of the comparisons, differences of up to 20 dB can be found above 8 kHz, with the X/Y configuration producing a more drastic reduction of high frequency content. Low frequency attenuation mirrors the published, and verified specifications show in Figures 10 through 12. Figures 13 through 17 show the averaged frequency response, 20 Hz to 20 kHz, for each M-S and equivalent X/Y pair. These results are ordered according to the "B" value (see equation 2) of the M-S mid component.

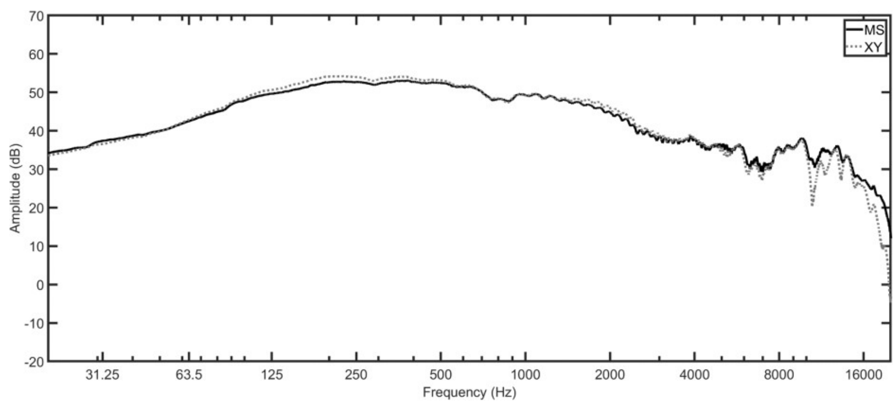


Figure 13. Frequency response comparison for M-S with bidirectional (B = 0), mid component (solid), and equivalent X/Y (dashed) configurations.

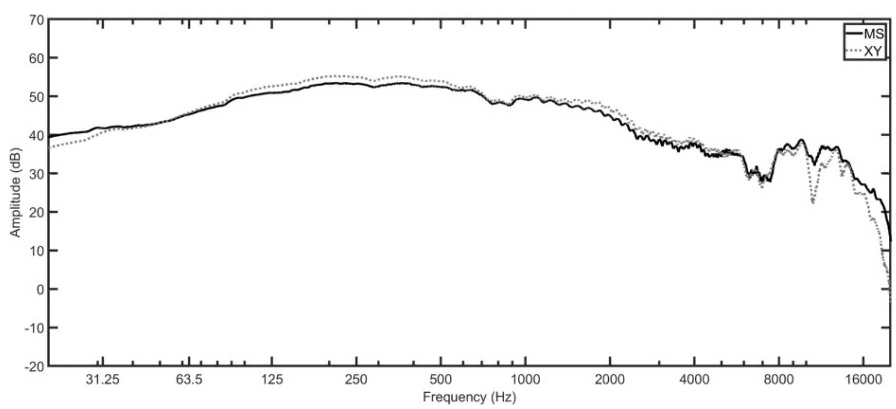


Figure 14. Frequency response comparison for M-S with hypercardioid (B = 0.25), mid component (solid), and equivalent X/Y (dashed) configurations.

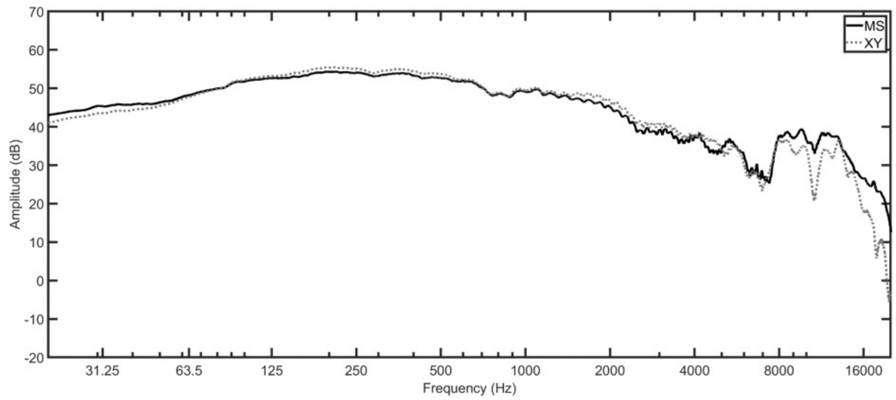


Figure 15. Frequency response comparison for M-S with cardioid (B = 0.5), mid component (solid), and equivalent X/Y (dashed) configurations.

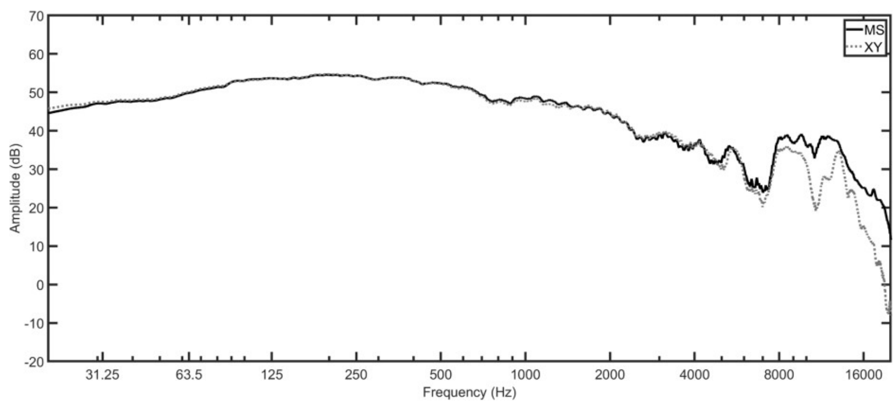


Figure 16. Frequency response comparison for M-S with wide-cardioid (B = 0.75), mid component (solid), and equivalent X/Y (dashed) configurations.

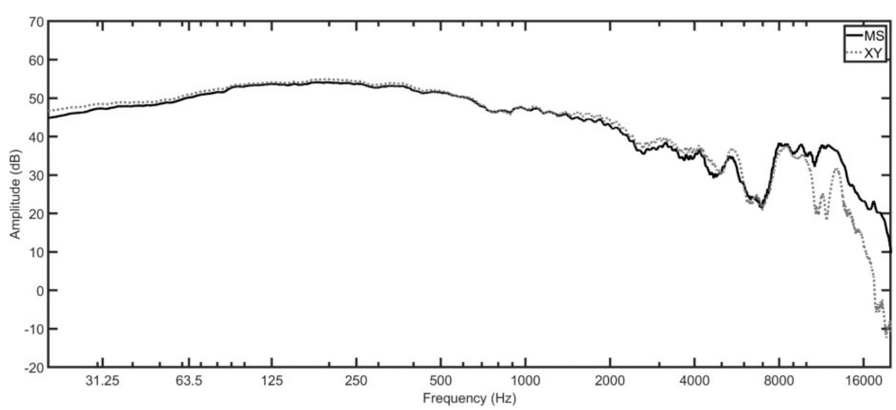


Figure 17. Frequency response comparison for M-S with omnidirectional (B = 1), mid component (solid), and equivalent X/Y (dashed) configurations.

Results for inter-channel correlation, shown in Figure 18, indicate that the configurations display fairly high correlation, above 0.7 for all measurements. Additionally, the slight trend towards higher correlation as the mid component becomes less directional, is mirrored by the equivalent X/Y configurations. The arrangement of the microphone patterns along the x-axis reflects the value of “B” in equation 2 where the omnidirectional microphone has a B value of 1 and the bidirectional microphone has a B value of 0. The agreement between M-S and X/Y configurations indicates that the X/Y configurations were placed in the correct position and utilized the correct polar pattern and included angle for M-S equivalence. The widest discrepancy, between the M-S with wide cardioid mid-component and the equivalent X/Y configuration, is 0.04 and may indicate unintentional deviation from the prescribed included angle.

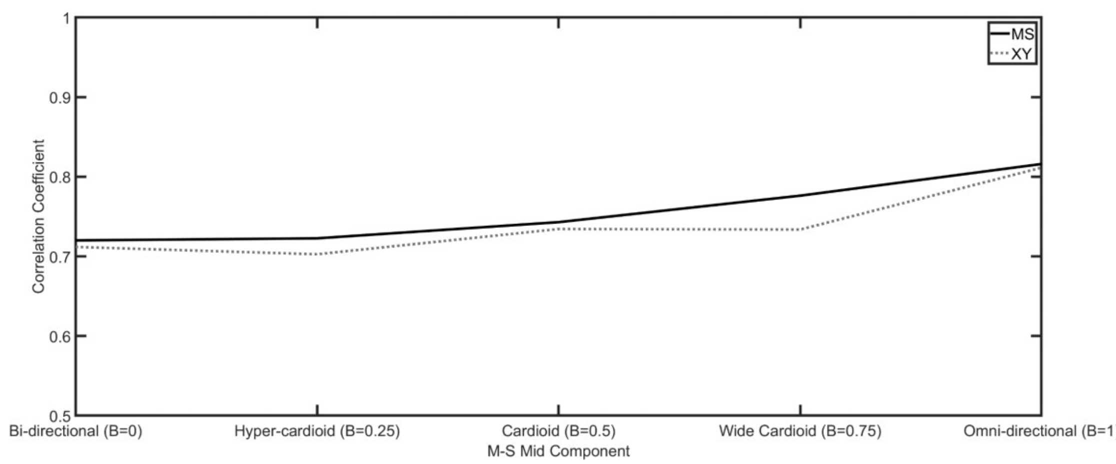


Figure 18. Inter-channel correlation coefficients for each M-S and equivalent X/Y technique.

Results of stereo image comparisons using a goniometer show increased width for M-S configurations compared to equivalent X/Y configurations, particularly as the directionality of the mid component decreases (B value increases). The outputs of the goniometer, being a measure of correlation between parts of a stereophonic signal, further verify the upward trend shown in the calculations of inter-channel correlation. Figures 19 through 23 show goniometer outputs for M-S and equivalent X/Y configurations.

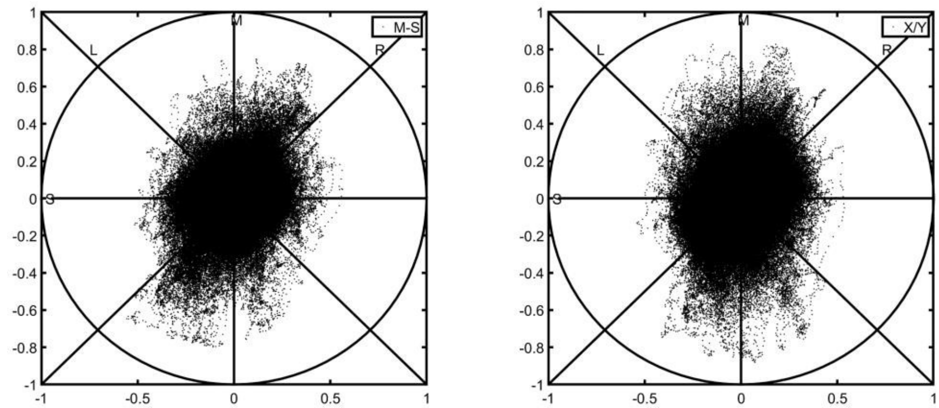


Figure 19. Stereo image comparison for M-S with bidirectional (B = 0), mid component (left), and equivalent X/Y (right) configurations.

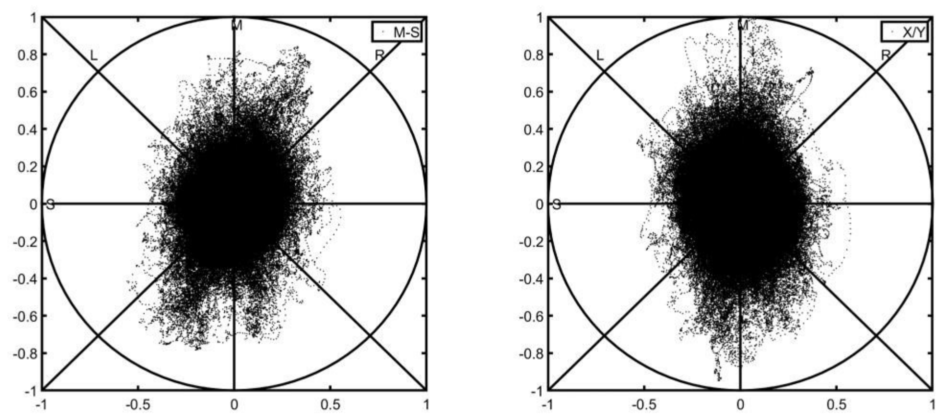


Figure 20. Stereo image comparison for M-S with hypercardioid (B = 0.25), mid component (left), and equivalent X/Y (right) configurations.

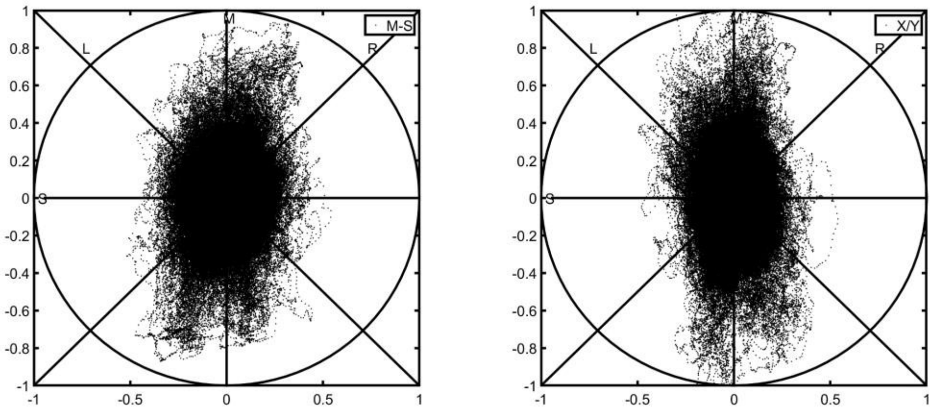


Figure 21. Stereo image comparison for M-S with cardioid (B = 0.5), mid component (left), and equivalent X/Y (right) configurations.

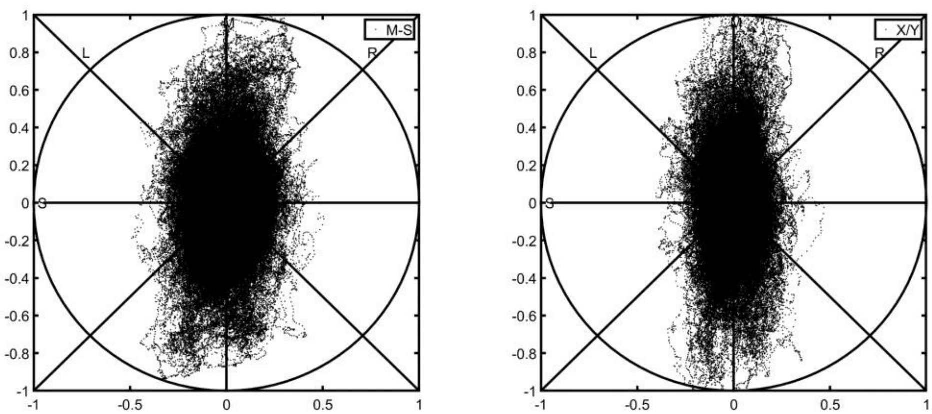


Figure 22. Stereo image comparison for M-S with bidirectional (B = 0.75), mid component (left), and equivalent X/Y (right) configurations.

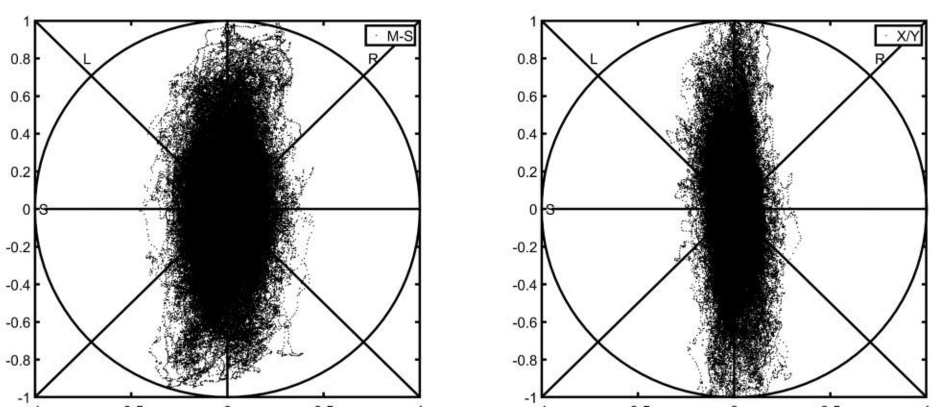


Figure 23. Stereo image comparison for M-S with omnidirectional (B = 1), mid component (left), and equivalent X/Y (right) configurations.

4.2 Listening Tests

4.2.1 ABX Testing: Descriptive Data

Of the nine subjects that participated in ABX testing, four were able to successfully differentiate between M-S and X/Y stimuli on each of the 100 trials. Only two subjects scored below 95% threshold with one subject scoring 89% and the other 80%. Using binomial distribution, shown in Tables 2 through 7, results for each subject and each technique display highly reliable scores ($15/20 = p < .05$) that indicates perceptual differences between M-S and equivalent X/Y recording techniques can be identified by trained listeners.

Table 2. Binomial distribution for ABX testing (M-S with bidirectional-mid component).

Subject	1	2	3	4	5	6	7	8	9
Correct	19	20	20	19	20	20	19	15	20

Table 3. Binomial distribution for ABX testing (M-S with hypercardioid-mid component).

Subject	1	2	3	4	5	6	7	8	9
Correct	19	20	20	16	20	20	20	16	20

Table 4. Binomial distribution for ABX Testing (M-S with cardioid-mid component).

Subject	1	2	3	4	5	6	7	8	9
Correct	19	20	20	18	18	20	18	18	20

Table 5. Binomial distribution for ABX Testing (M-S with wide-cardioid-mid component).

Subject	1	2	3	4	5	6	7	8	9
Correct	18	20	20	20	19	20	19	16	20

Table 6. Binomial distribution for ABX testing (M-S with omnidirectional-mid component).

Subject	1	2	3	4	5	6	7	8	9
Correct	20	20	20	16	20	20	20	15	20

Table 7. Cumulative results of binomial distribution for ABX testing

Subject	1	2	3	4	5	6	7	8	9
Correct	95	100	100	89	97	100	96	80	100

4.2.4 Preference Testing: Descriptive Data

Preference ratings were sorted so that values below zero indicated a preference for the M-S technique and values above zero indicated a preference for the X/Y technique. The comparisons were categorized based on the polar pattern of the M-S mid component using the "B" value found in equation 2.

Ratings for width show mean results that indicate preference for M-S in four out of five techniques. When classified by the mid component of the M-S technique, the bidirectional technique averaged 0.06 (SD = 1.29). The hypercardioid technique averaged -0.19 (SD = 1.34). The cardioid technique averaged -0.71 (SD = 1.25). The wide cardioid technique averaged -0.91 (SD = 1.17). The omnidirectional technique averaged -1.42 (SD = 1.34). The trendline for width preference shows a coefficient of determination ($R2$) of 0.983. The average score with standard deviation above and below for each technique is shown in Figure 24.

Ratings for depth show mean results that indicate preference for M-S in 3 out of 5 techniques. When classified by the mid component of the M-S technique, the bidirectional technique averaged 0.27 (SD = 1.15). The hypercardioid technique averaged 0.03 (SD = 1.10). The cardioid technique averaged -0.51 (SD = 1.14). The wide cardioid technique averaged -0.78 (SD = 1.15). The omnidirectional technique averaged -1.05 (SD = 1.29). The trendline for depth preference shows a coefficient of determination (R^2) of .982. The average score with standard deviation above and below for each technique is shown in Figure 25.

Ratings for timbral balance show mean results that indicate preference for M-S in 3 out of 5 techniques. When classified by the mid component of the M-S technique, the bidirectional technique averaged 0.46 (*SD* = 1.03). The hypercardioid technique averaged 0.07 (SD = 0.99). The cardioid technique averaged -0.50 (*SD* = 1.11). The wide cardioid technique averaged -0.98 (*SD* = 0.89). The omnidirectional technique averaged -1.43 (*SD* = 0.98). The trendline for timbral balance preference shows a coefficient of determination (R^2) of .998. The average score with standard deviation above and below for each technique is shown in Figure 26.

Ratings for sound source definition show mean results that indicate preference for M-S in three out of five techniques. When classified by the mid component of the M-S technique, the bidirectional technique averaged 0.46 (*SD* = 1.22). The hypercardioid technique averaged 0.05 (*SD* = 1.18). The cardioid technique averaged -0.53 (*SD* = 1.28). The wide cardioid technique averaged -1.19 (*SD* = 1.09). The omnidirectional technique averaged -1.57 (*SD* = 1.16). The trendline for sound source definition preference shows a coefficient of determination (R^2) of .992. The average score with standard deviation above and below for each technique is shown in Figure 27.

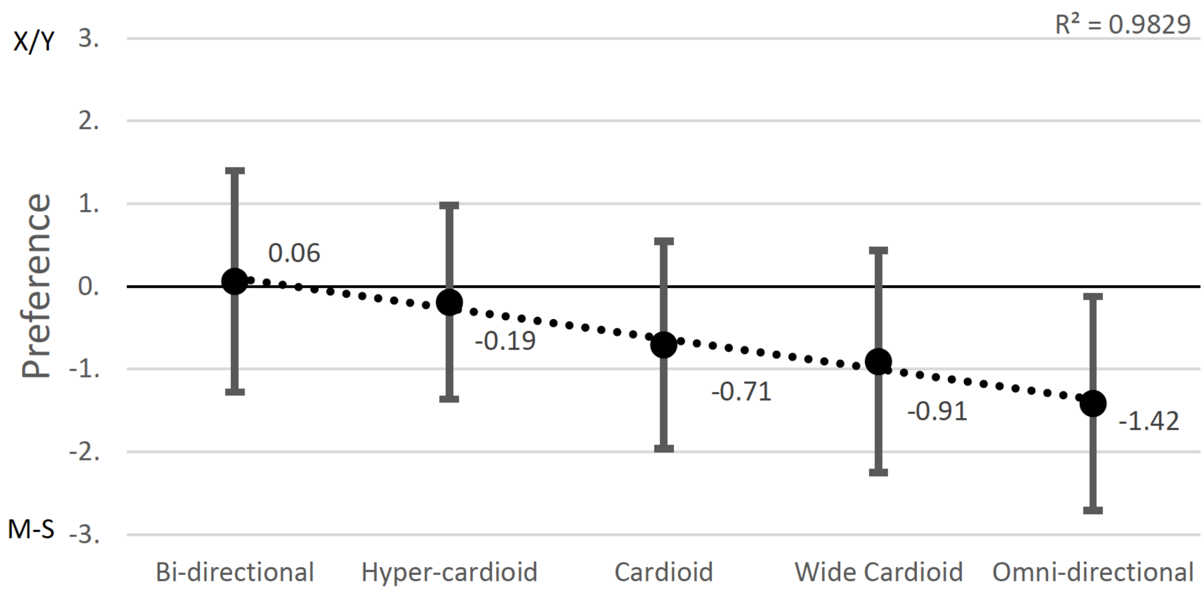


Figure 24. Mean with standard deviation for width ratings.

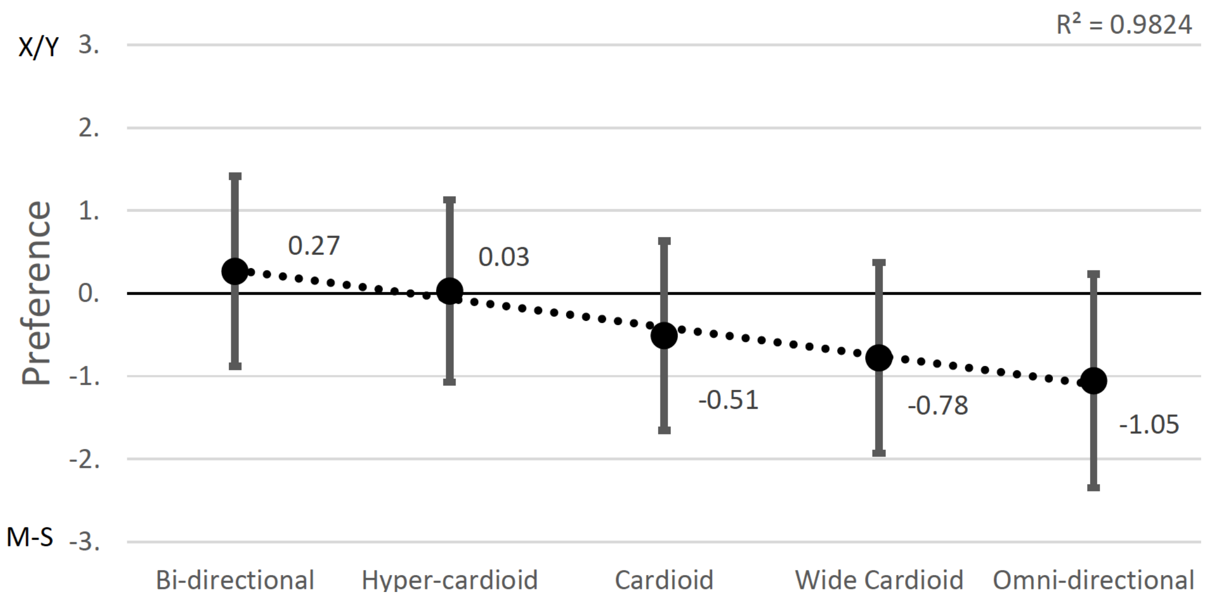


Figure 25. Mean with standard deviation for depth ratings.

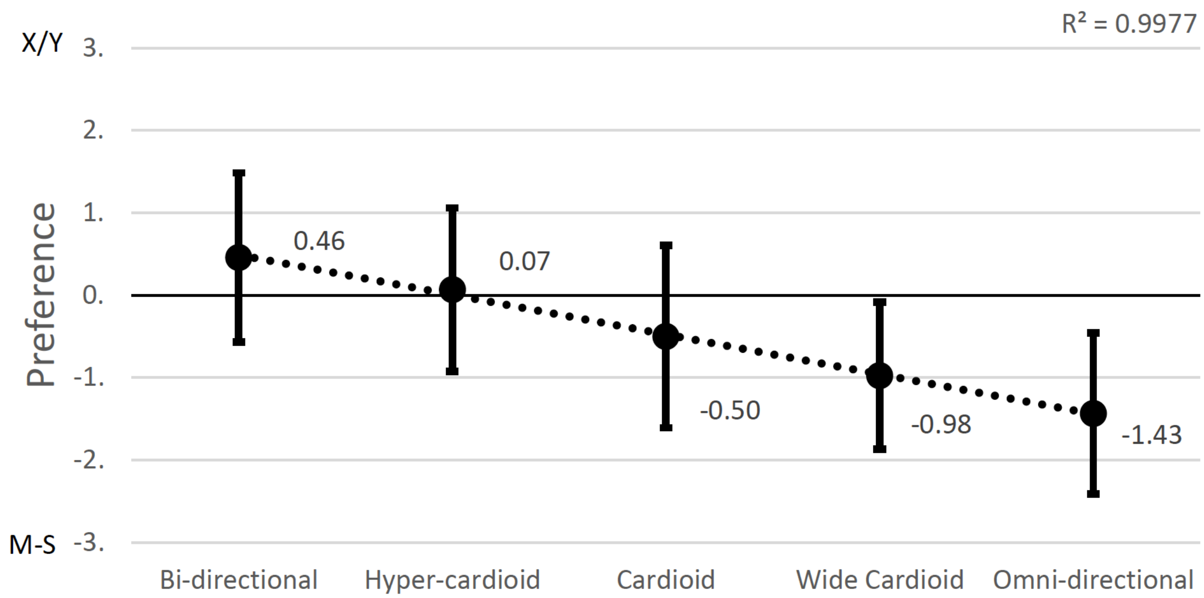


Figure 26. Mean with standard deviation for timbral balance ratings.

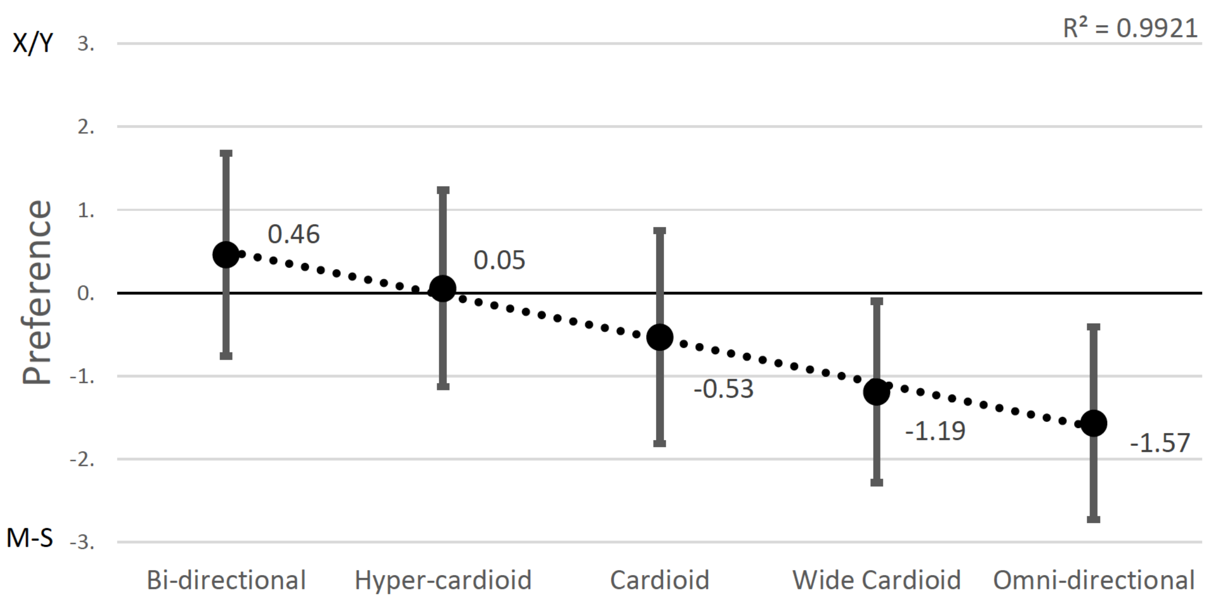


Figure 27. Mean with standard deviation for sound source definition ratings.

4.2.3. Preference Testing: Results of Statistical Testing

Analysis of variance indicated a main effect of technique on the preference for width ($F(4, 48) = 10.56$, $p < .001$) and no effect of subject ($F(12, 48) = 1.13$, $p = .361$). *Post-hoc* analysis using

Fisher LSD indicated significant differences when the techniques being compared were not adjacent (i.e. omnidirectional and wide-cardioid). Two-way ANOVA data for width is displayed in Appendix C.

Analysis of variance indicated a main effect of technique on the preference for depth ($F(4, 48) = 9.59$, $p < .001$) and no effect of subject ($F(12, 48) = 1.55$, $p = .141$). Post-hoc analysis using Fisher LSD indicated significant differences between omnidirectional and hypercardioid, omnidirectional and bidirectional, wide-cardioid and hypercardioid, wide-cardioid and bidirectional, as well as cardioid and bidirectional. Two-way ANOVA data for depth is displayed in Appendix C.

Analysis of variance for timbral balance was conducted twice due to statistical anomalies present in the first ANOVA that showed a main effect of subject ($F(12, 48) = 1.99$, $p = .047$). *Post-hoc* analysis using *Fisher* LSD indicated that Subject 2 varied significantly from five out of 12 subjects. Normality analysis indicated that Subject 2 displayed bias towards a preference for "A" when rating stimuli that were identical to each other ($M = -0.275$, $SD = 0.72$). This condition appeared ten times within the 60 trials. Subject 2 indicated a preference for 14 out of 40 ratings, with 11 of these 14 ratings indicating preference for the "A" stimulus. When analyzing only the timbral balance ratings, Subject 2 indicated preference for three out of ten ratings, with each of the three indicating preference for the "A" stimulus. For this reason, timbral balance ratings for Subject 2 were removed and a second ANOVA was performed. The uncorrected two-way ANOVA for timbral balance is displayed in Appendix C.

The second analysis of variance indicated a main effect of technique on the preference for balance ($F(4, 44) = 20.93$, $p < .001$) and no effect of subject ($F(11, 44) = 1.47$, $p = .177$). *Post-hoc* analysis using *Fisher* LSD indicated significant differences when the techniques being compared were not adjacent. The corrected two-way ANOVA data for timbral balance is displayed in Appendix C.

Analysis of variance indicated a main effect of technique on the preference for sound source definition ($F(4, 48) = 27.76$, $p < .001$) and no effect of subject ($F(12, 48) = 1.92$, $p = .056$). *Post-hoc* analysis using *Fisher* LSD indicated significant differences between all techniques except for omnidirectional and wide-cardioid as well as hypercardioid and bidirectional. Two-way ANOVA data for sound source definition is displayed in Appendix C.

4.2.4 Preference Testing: Interpretation of Statistical Results

Each of the four attributes that were rated in this study show a similar trend where the techniques related to an omnidirectional mid component display significant preference towards M-S. For all four ratings, the techniques related to a bidirectional mid component display marginally significant or non-significant preference for X/Y. The techniques related to a hypercardioid mid component are the only techniques that display preference for M-S in at least one condition as well as preference for X/Y in any other condition. For all rated conditions, the relationship between the five technique pairs is highly linear in nature, with $R2$ values approaching 1. As a result, adjacent techniques often do not display a significant difference for each rating, however significance often occurs when the techniques are two or more variations apart. The technique most often utilized for M-S, with a cardioid mid component, displays moderately significant preference for M-S over the equivalent X/Y technique. Non-significant differences between subjects for all rated conditions provides evidence that the listeners reliably indicated their preference based on perceived differences between each of the M-S techniques and their equivalent X/Y technique.

5. DISCUSSION

Testing of the AKG C426-B microphone verified the polar and frequency response specifications published in [36]. Stereophonic technique testing reveals significant inter- channel correlation, as well as similar increasing trends in both M-S and X/Y techniques as the mid component of the M-S configuration decreases in directionality. Averaged frequency response measurements indicate increased high-frequency attenuation for each X/Y configuration. Goniometer measurements show decreased stereophonic width for X/Y configurations, with differences between M-S and X/Y configurations increasing as the mid component of the M-S configuration decreases in directionality.

ABX testing performed during this study provides ample evidence to support the claim that listeners can identify differences between M-S and equivalent X/Y techniques. The high degree of success found in this test indicates that significant perceptual differences exist between M-S and equivalent X/Y techniques.

Through the use of preference testing, this study has shown that listeners often prefer the spatial attributes of M-S techniques. Of the five technique pairs tested, three displayed significant preference for M-S when rating width. One technique, based on the hyper- cardioid mid component, displayed marginally significant preference for M-S. The final technique, based on the bidirectional mid component, displayed a non-significant preference for X/Y. The ratings for depth show a similar pattern, with the techniques based on the hypercardioid mid component displaying a non-significant preference for X/Y.

Preference for timbral characteristics similarly favors M-S in most cases. Three of the five technique pairs display significant preference for M-S when rated for both timbral balance and sound source definition. The techniques based on a hypercardioid mid component and those based on a bidirectional mid component display non-significant to marginally significant preference for X/Y techniques.

Trendlines for all four attributes show a highly linear relationship. As the “B” value of the M-S mid component increases, mean ratings decrease. The trendlines may indicate that the M-S to X/Y transformation process is less successful as the directionality of the mid component decreases. This is likely the result of the wider included angle forcing the X/Y microphones further off-axis from the center of the sound source.

Combining the ABX and preference testing results, evidence suggests that the mathematical equivalency of the M-S to X/Y transformation process is not supported by perception. Additionally, preference testing shows that M-S techniques are preferred over equivalent X/Y techniques for the majority of tested techniques. Those techniques that do not show a significant preference for M-S show no preference, or in some cases only marginal preference for X/Y.

6. CONCLUSIONS

Evidence gathered from this study rejects the equivalency of the M-S to X/Y transformation model when applied to practical usage scenarios. Frequency response, as shown in Figures 13 through 17, can account for the timbral differences as suggested by [5]. Significant differences in spatial ratings, however, indicate the influence of additional factors. The preference for M-S recording techniques found in this study justifies the increased use of these techniques for stereophonic recording scenarios.

Additional benefits of the M-S recording technique, including monophonic compatibility and post-production variability, should not be overlooked by engineers when selecting a stereophonic microphone technique. The results of this study show that these benefits are not counterweighted by impaired sonic attributes. On the contrary, M-S recording techniques produce superior quality recordings than other coincident methods.

Averaged frequency response measurements indicate that the off-axis coloration of directional microphones causes noticeable high-frequency attenuation when recording with X/Y microphone techniques. This is particularly significant as the test signals were produced from the center of the sound image. As a result, the use of X/Y techniques to capture soloist or small ensemble performances may not be ideal. Inter-channel correlation and goniometer output, however, indicate a narrower stereophonic image for X/Y techniques.

This may indicate that X/Y techniques do not provide a realistic sense of stereophonic width. Taking these conclusions together, X/Y techniques are less capable than M-S techniques of effectively capturing audio across all portions of the stereophonic sound field.

The results of this study are relevant to the development of new microphones, particularly those designed for various modeling platforms. Currently, at least one major modeling microphone manufacturer is producing a dual-capsule microphone, similar to the AKG C426-B used in this study. The advantage to this type of microphone is that each capsule provides two separate outputs, one for each diaphragm. By varying the level and polarity of these outputs an engineer can alter the polar pattern of the capsule remotely, or even in post-production. The conclusions of this study imply that using such a microphone in the standard M-S configuration would provide the engineer with unlimited control of the polar response and included angle of the resulting stereophonic recording. Using such a microphone in a more traditional X/Y configuration would limit the engineer to control over the polar response only, with no ability to alter the included angle of the stereophonic image.

With the results of this study lending evidence towards increased use of M-S techniques, further research could be conducted to discover new ways to implement M-S into existing arrays. Similarly, new combinations of microphones that rely on the principles of M-S could provide additional advantages over current stereophonic microphone techniques.

6.1 Further Research

6.1.1 Limiting Factors

The physical testing for this experiment largely took place in an anechoic environment. Unfortunately, this space was designed for additional uses that limit the effectiveness of the absorptive and diffusive properties of the room. In order to conduct more reliable measurements of the microphones and techniques used in this study, a more suitable environment for audio research would be required. Due to the requirement for a multi- pattern microphone, this research relied heavily on the dual-diaphragm microphone design first developed by [44]. As noted by [45], these types of microphones respond to proximity much differently than single-diaphragm microphones. In particular, dual-diaphragm microphones have been shown to exhibit uneven polar response, especially toward the rear of the capsule. When using such a microphone for M-S techniques, the asymmetry present in the bidirectional polar pattern can have a detrimental effect on the stereophonic image of the produced recording. The use of a dual-diaphragm microphone was necessary to directly compare equivalent techniques, however, the imperfections inherent to the design should not be overlooked.

While the results of ABX testing were more conclusive than the preference testing, the limited sample size of both tests could be improved. Additional subjects could increase the reliability of the preference testing and further solidify the conclusions of this research. The preference tests also could benefit from further consideration with reference to the ratings schemes as well as the definitions of the rated attributes. Providing more detailed definitions, as well as additional attributes, could more accurately describe the differences detected between M-S and X/Y techniques.

6.1.2 Additional Research

Expanded physical testing of the microphone and patterns used in this study could be a useful way to explain the preference for M-S over X/Y. Direct comparisons could be made between the polar responses of M-S techniques and their equivalent X/Y technique.

Capturing impulse responses in a reverberant environment could provide insight into the ability of each technique to effectively capture the recording space. Direct-to-Reverberant Energy Ratio measurements could be used to compare equivalent techniques. Additionally, measurements of clarity (C80) could be used to support or dispute the preference ratings.

Several microphone arrays exist that rely on two or more coincident or near-coincident microphones. The Faulkner Array, for example, attempts to correct the emphasis that a Blumlein array can place on the ambient field [46]. By using a stereophonic microphone with adjustable polar patterns, an engineer could maintain the benefits of M-S while maintaining control over the direct-to-reverberant ratio of the resulting recording.

Additional research could compare this ratio for various M-S techniques. Another microphone array that relies on coincident microphones is the Stereo Technique for Augmented Ambience Gradient, or STAAG. This technique, developed by [47], uses two pairs of near-coincident microphones placed so that the left and right microphones form their own coincident pair. Since the purpose of this technique is to provide an adjustable polar pattern, the use of multi-pattern M-S techniques would seem appropriate. Further research using M-S microphone techniques in place of X/Y techniques could focus on improving the results found with arrays such as STAAG.

Lastly, this research has shown that altering the Mid component of a M-S configuration is an effective way to alter the included angle and polar pattern of the resulting stereophonic image. Altering the ratio of level between the Mid and Side components has a similar effect. Additional research could be used to develop algorithms that provide an engineer with the appropriate Mid component polar pattern and level ratio for the desired stereophonic image. This would allow audio engineers to specify the desired angle and polar pattern of the stereophonic image while software determines the necessary polar and level adjustments to create the desired output. This type of software would be particularly useful with the multi-pattern modeling microphones mentioned in Section 6.

BIBLIOGRAPHY

Citations

[1] T. D. Rossing, The Science of Sound. *Addison-Wesley Publishing Company, Inc.*, 1982.
[2] J. L. Butler, "Microphone Selection," in *86th Convention of the Audio Engineering Society*, 1989.
[3] G. Davis and R. Jones, *The Sound Reinforcement Handbook*. Milwaukee, Wisconsin: Hal Leonard Corporation, 1989.
[4] R. D. Streicher and W. L. Dooley, "Basic Stereo Microphone Perspectives - A Review," *J. Audio Eng. Soc.*, vol. 33, no. 7/8, pp. 548–556, 1985.
[5] W. L. Dooley and R. D. Streicher, "M-S Stereo: A powerful technique for working in stereo," *J. Audio Eng. Soc.*, vol. 30, no. 10, pp. 707–718, 1982.
[6] MATLAB, Version 9.4.0.813654 (R2018a). Natick, Massachusetts: *The MathWorks Inc.*, 2018.
[7] B. Barltett, Stereo Microphone Techniques. Boston: *Focal Press*, 1991.
[8] D. M. Huber and R. E. Runstein, Modern Recording Techniques, 7th Edition. *Focal Press*, 2010.
[9] B. Owsinski, The Recording Engineer's Handbook, Fourth Edition. Burbank, CA: *Bobby Owsinski Media Group*, 2017.
[10] B. F. Hertz, "100 Years with Stereo: The Beginning," *J. Audio Eng. Soc.*, vol. 29, no. 5, pp. 368–372, 1981.
[11] A. C. Keller, "Early Hi-Fi and Stereo Recording at Bell Laboratories (1931-1932)," *J. Audio Eng. Soc.*, vol. 29, no. 4, pp. 274–280, 1981.
[12] F. Rumsey and R. Auld, "Stokowski and the Evolution of Recording," *J. Audio Eng. Soc.*, vol. 60, no. 7/8, pp. 627–630, 2012.
[13] A. D. Blumlein, "British Patent Specification 394,325," *J. Audio Eng. Soc.*, vol. 6, no. 2, pp. 91–98, 1958.
[14] D. Griesinger, "New perspectives on coincident and semi coincident microphone arrays," in *82nd Convention of the Audio Engineering Society*, 1987.
[15] S. P. Lipshitz, "Stereo Microphone Techniques: Are the Purists Wrong?," in *78th Convention of the Audio Engineering Society*, 1985.
[16] J. C. Steinberg and W. B. Snow, "Auditory Perspective - Physical Factors," *Electr. Eng.*, vol. 53, no. 1, pp. 12–17, 1934.
[17] H. A. M. Clark, G. F. Dutton, and P. B. Vanderlyn, "The 'Stereosonic' Recording and Reproducing System: A Two-Channel System for Domestic Tape Records," *J. Audio Eng. Soc.*, vol. 6, no. 2, pp. 102–117, 1958.
[18] Lord Rayleigh, "XII. On our perception of sound direction," London, Edinburgh, Dublin *Philos. Mag. J. Sci.*, vol. 13, no. 74, pp. 214–232, 1907.
[19] F. L. Wightman and D. J. Kistler, "The dominant role of low-frequency interaural time differences in sound localization," *J. Acoust. Soc. Am.*, vol. 91, no. 3, pp. 1648–1661, 1992.
[20] H. Wallach, E. B. Newman, and M. R. Rosenzweig, "The Precedence Effect in Sound Localization," *J. Audio Eng. Soc.*, vol. 21, no. 10, pp. 817–826, 1973.
[21] B. Bernfeld, "Attempts for better understanding of the directional stereophonic listening mechanism," in *44th Convention of the Audio Engineering Society*, 1973.
[22] M. Williams, "Unified theory of microphone systems for stereophonic sound recording," in *82nd Convention of the Audio Engineering Society*, 1987.

[23] M. A. Gerzon and J. M. Woram, "Blumlein Stereo Microphone Technique and Author's Reply," *J. Audio Eng. Soc.*, vol. 24, no. 1, p. 36,38, 1976.
[24] M. Hibbing, "XY and MS Microphone Techniques in Comparison," *J. Audio Eng. Soc.*, vol. 37, no. 10, pp. 823–831, 1989.
[25] J. Braasch, "A binaural model to predict position and extension of spatial images created with standard sound recording techniques," in *119th Convention of the Audio Engineering Society*, 2005.
[26] M. Plewa and G. Pyda, "Localization curves in stereo microphone techniques - comparison of calculations and listening tests results," in *128th Convention of the Audio Engineering Society*, 2010.
[27] H. Riekehof-Boehmer and H. Wittek, "Prediction of perceived width of stereo microphone setups," in *130th Convention of the Audio Engineering Society*, 2011.
[28] F. E. Toole, "Listening Tests-Turning Opinion into Fact," *J. Audio Eng. Soc.*, vol. 30, no. 6, pp. 431–445, 1982.
[29] T. Letowski, "Sound Quality Assessment: Concepts and Criteria," in *87th Convention of the Audio Engineering Society*, 1989.
[30] F. Rumsey, "Spatial Quality Evaluation for Reproduced Sound: Terminology, Meaning, and a Scene-Based Paradigm," *J. Audio Eng. Soc.*, vol. 50, no. 9, pp. 651–666, 2002.
[31] S. Le Bagousse, C. Colomes, and M. Paquier, "State of the Art on Subjective Assessment of Spatial Sound Quality," in Audio Engineering Society *38th International Conference: Sound Quality Evaluation*, 2010.
[32] G. A. Soulodre and M. C. Lavoie, "Subjective Evaluation of Large and Small Impairments in Audio Codecs," in Audio Engineering Society *17th International Conference: High-Quality Audio Coding*, 1999, pp. 329–336.
[33] EBU Tech 3286-E, "Assessment methods for the subjective evaluation of the quality of sound programme material – Music," *European Broadcasting Union*. Geneva, 1997.
[34] C. Ceoen, "Comparative Stereophonic Listening Tests," *J. Audio Eng. Soc.*, vol. 20, no. 1, pp. 19–27, 1972.
[35] W. Lim, "An Objective Comparison of Stereo Recording Techniques Through the Use of Subjective Listener Preference Ratings," in *135th Convention of the Audio Engineering Society*, 2013.
[36] AKG Acoustics GmbH, C426-B Stereo Condenser Microphone Cutsheet. *Harman International*.
[37] AP2700, Version 3.30 (build 118). Beaverton, Oregon: *Audio Precision, Inc.*, 2007.
[38] G. Martin, "The Significance of Interchannel Correlation, Phase and Amplitude Differences on Multichannel Microphone Techniques," in *113th Convention of the Audio Engineering Society*, 2002.
[39] E. Tarr, Hack Audio: An Introduction to Computer Programming and Digital Signal Processing in MATLAB, 1st ed. Waltham, MA: *Focal Press*, 2018.
[40] Maurizio Pagnutti Sextet, "All the Gin is Gone." Amerio Stefano Edizioni Musicali, 2010.
[41] ITU-R BS.1284-1, "General methods for the subjective assessment of sound quality," *International Telecommunications Union*. Geneva, 2003.
[42] Room EQ Wizard, Version 5.19. John Mulcahy, 2018.
[43] ITU-R BS.1116-3, "Methods for the subjective assessment of small impairments in audio systems," *International Telecommunication Union*. Geneva, 2002.
[44] H. J. Von Braunmuehl and W. Weber, "Condenser microphone," *U.S. Patent No. 2179361*, 1939.

[45] G. Torio and J. Segota, "Unique Directional Properties of Dual-Diaphragm Microphones," in *109th Convention of the Audio Engineering Society*, 2000.
[46] T. Faulkner, "A Phased Array," *Hi-Fi News & Records Review*, pp. 44–46, 1981.
[47] D. J. Tagg, "A Microphone Technique for Improved Stereo Image, Spatial Realism, and Mixing Flexibility: STAGG (Stereo Technique for Augmented Ambience Gradient)," in *133rd Convention of the Audio Engineering Society*, 2012.

Resources

J. Berg and F. Rumsey, "Identification of Quality Attributes of Spatial Audio by Repertory Grid Technique," *J. Audio Eng. Soc.*, vol. 54, no. 5, pp. 365-379, 2006.

B. Bernfeld and B. Smith, "Computer-Aided Model of Stereophonic Systems," in *56th Convention of the Audio Engineering Society*, 1978.

J. Blauert, Spatial Hearing, Cambridge, MA: *MIT Press*, 1997.

G. F. Dutton, "The Assessment of Two-Channel Stereophonic Reproduction Performance in Studio Monitor Rooms, Living Rooms and Small Theatres," *J. Audio Eng. Soc.*, vol. 10, no. 2., pp. 98-105, 1962.

M. A. Gerzon, "The Design of Precisely Coincident Microphone Arrays for Stereo and Surround Sound," in *50th Convention of the Audio Engineering Society*, 1975.

D. Griesinger, "Spaciousness and Localization in Listening Rooms and Their Effect on the Recording Technique," *J. Audio Eng. Soc.*, vol. 34, no. 4, pp. 255-268, 1986.

B. F. Hertz, "100 Years with Stereo – The Beginning," in *68th Convention of the Audio Engineering Society*, 1981.

S. Julstrom, "An Intuitive View of Coincident Stereo Microphones," in *89th Convention of the Audio Engineering Society*, 1990.

H. Lauridsen, F. Schlegel, "C," *Gravesaner Blätter*, vol. 5, pp. 28-50, 1956.

H. F. Olson, "A History of High Quality Studio Microphones," in *55th Convention of the Audio Engineering Society*, 1976.

I. Pulkki, "Microphone techniques and directional quality of sound reproduction," in *112th Convention of the Audio Engineering Society*, 2002.

APPENDIX

A: MATLAB® Code

A1. Generate polar patterns for the M-S and equivalent X/Y configurations.

```
% MSEquiv.m

% Initialize theta
theta = linspace(0,2*pi);

% Determine Mid to Side Ratio
MidGain = 0.5;
SideGain = 1 - MidGain;

% Determine Mid Microphone Type
% 0 = Figure-8, 0.25 = Hypercardioid, 0.5 = Cardioid, 0.75 = Wide Cardioid,
1 = Omnidirectional
MidPattern = 0.5;

% Calculate M and S Plots
M = MidGain * (MidPattern + (1 - MidPattern) * cos(theta-(pi/2)));
S = SideGain * sin(theta-(pi/2));

% Convert M-S to X/Y
Left = M + S;
Right = M - S;

% Calculate Included Angle of Equivalent X/Y pattern
y = SideGain;
x = MidGain * (1 - MidPattern);
angle = 2 * atan2d(y,x);

% Plot M-S
subplot(2,1,1);
polarplot(theta,abs(M)); hold on;
polarplot(theta,abs(S)); hold off;

% Plot X/Y Equivalence and Included Angle
subplot(2,1,2);
angleTitle = ['Included Angle: ', num2str(angle)];
polarplot(theta,abs(Left)); hold on; title(angleTitle);
polarplot(theta,abs(Right)); hold off;
```

A2. Plot averaged frequency response and inter-channel correlation coefficient for each M-S and equivalent X/Y.

```
% corrfreq.m
clear all, close all, clc;

% Import Audio Files
[sweep,Fs] = audioread('Sweep.wav');
[ms1,Fs1] = audioread('MS1_Sweep.wav'); ms2 = audioread('MS2_Sweep.wav');
ms3 = audioread('MS3_Sweep.wav'); ms4 = audioread('MS4_Sweep.wav');
ms5 = audioread('MS5_Sweep.wav'); xy1 = audioread('XY1_Sweep.wav');
xy2 = audioread('XY2_Sweep.wav'); xy3 = audioread('XY3_Sweep.wav');
xy4 = audioread('XY4_Sweep.wav'); xy5 = audioread('XY5_Sweep.wav');

%% Interchannel Correlation
% Calculate Correlation
ms1C = rcorr(ms1); xy1C = rcorr(xy1); C1 = ms1C - xy1C; ms2C = rcorr(ms2);
xy2C = rcorr(xy2); C2 = ms2C - xy2C; ms3C = rcorr(ms3); xy3C = rcorr(xy3);
C3 = ms3C - xy3C; ms4C = rcorr(ms4); xy4C = rcorr(xy4); C4 = ms4C - xy4C;
ms5C = rcorr(ms5); xy5C = rcorr(xy5); C5 = ms5C - xy5C;

% Concatenate Coefficients
msCorr = [ms5C ; ms4C ; ms3C ; ms2C ; ms1C]; xyCorr = [xy5C ; xy4C ; xy3C ;
xy2C ; xy1C];

% Plot Correlation Coefficients plot(msCorr); hold on; plot(xyCorr); hold off;
xticks([1 2 3 4 5]); ylim([0.5 1]); yticks([0.5 0.6 0.7 0.8 0.9 1]);
xticklabels({'Bidirectional (B=0)','hypercardioid (B=0.25)','Cardioid (B=0.5)','Wide
Cardioid (B=0.75)','Omnidirectional (B=1)'}); legend('MS','XY');
xlabel('M-S Mid Component'); ylabel('Correlation Coefficient'); figure;

%% Averaged Frequency Response
% Frequency Analysis Function
[ms1Avg, ms1f] = freqAvg(ms1,Fs); [ms2Avg, ms2f] = freqAvg(ms2,Fs); [ms3Avg, ms3f] =
 freqAvg(ms3,Fs); [ms4Avg, ms4f] = freqAvg(ms4,Fs); [ms5Avg, ms5f] = freqAvg(ms5,Fs);
 [xy1Avg, xy1f] = freqAvg(xy1,Fs); [xy2Avg, xy2f] = freqAvg(xy2,Fs); [xy3Avg, xy3f] =
 freqAvg(xy3,Fs); [xy4Avg, xy4f] = freqAvg(xy4,Fs); [xy5Avg, xy5f] = freqAvg(xy5,Fs);


% Plot using Logarithmic Frequency Scale
semilogx(ms1f, smooth(ms1Avg,2000)); hold on; semilogx(xy1f, smooth(xy1Avg,2000));
hold off; axis([20 20000 -20 70]); legend('MS','XY');
xlabel('Frequency (Hz)'); xticks([31.25, 63.5, 125, 250, 500, 1000, 2000, 4000, 8000,
16000]); ylabel('Amplitude (dB)'); figure;

semilogx(ms2f, smooth(ms2Avg,2000)); hold on; semilogx(xy2f, smooth(xy2Avg,2000));
```

```
hold off; axis([20 20000 -20 70]); legend('MS','XY');
xlabel('Frequency (Hz)'); xticks([31.25, 63.5, 125, 250, 500, 1000, 2000, 4000, 8000,
16000]); ylabel('Amplitude (dB)'); figure;

semilogx(ms3f, smooth(ms3Avg,2000)); hold on; semilogx(xy3f, smooth(xy3Avg,2000));
hold off; axis([20 20000 -20 70]); legend('MS','XY');
xlabel('Frequency (Hz)'); xticks([31.25, 63.5, 125, 250, 500, 1000, 2000, 4000, 8000,
16000]); ylabel('Amplitude (dB)'); figure;

semilogx(ms4f, smooth(ms4Avg,2000)); hold on; semilogx(xy4f, smooth(xy4Avg,2000));
hold off; axis([20 20000 -20 70]); legend('MS','XY');
xlabel('Frequency (Hz)'); xticks([31.25, 63.5, 125, 250, 500, 1000, 2000, 4000, 8000,
16000]); ylabel('Amplitude (dB)'); figure;

semilogx(ms5f, smooth(ms5Avg,2000)); hold on; semilogx(xy5f, smooth(xy5Avg,2000));
hold off; axis([20 20000 -20 70]); legend('MS','XY');
xlabel('Frequency (Hz)'); xticks([31.25, 63.5, 125, 250, 500, 1000, 2000, 4000, 8000,
16000]); ylabel('Amplitude (dB)');
```

A2.1. "freqAvg" function called by the corrfreq.m script.

```
% Averaged Frequency Response Function
function [y, f] = freqAvg(x,Fs)

% Initialize Variables
xL = x(:,1); xR = x(:,2);
N = length(x);

% Normalize Signal
maxL = abs(max(xL));
maxR = abs(max(xR));

if maxL > maxR
xL = xL/maxL;
xR = xR/maxL;
else
xL = xL/maxR;
xR = xR/maxR;
end

% Perform FFT
xLFFT = fft(xL);
xRFFT = fft(xR);

% Average L/R FFT
M = length(xLFFT);
for n = 1:M
```

```
xFFT(n) = (xLFFT(n) + xLFFT(n)) / 2;
end

% Generate Frequency Bins
k = 0:N-1;
f = k*Fs/N;

f = f(1:floor(N/2));

% Convert Linear Amplitude to Decibels
xAmp = abs(xFFT);
y = 20 * log10(xAmp(1:floor(N/2))); end
```

A2.2. "rcorr" function called by the corrfreq.m script.

```
% Correlation Coefficient Function
function [r] = rcorr(x)

xLeft = x(:,1); xRight = x(:,2);
leftMean = mean(xLeft); rightMean = mean(xRight);

for n = 1:length(x)

sLeft(n) = (leftMean - xLeft(n)).^2;
sRight(n) = (rightMean - xRight(n)).^2;

s(n) = (leftMean - xLeft(n)) * (rightMean - xRight(n));

end

sL = sqrt(mean(sLeft));
sR = sqrt(mean(sRight));

sMean = mean(s);

r = sMean / (sL * sR);

end
```

B. REM Studio D Measurements

In [43], several properties for the shape and dimensions of the listening room are given. For monophonic or stereophonic reproduction room should be between 20 and 60 square meters, symmetrical along the vertical plane and either rectangular or trapezoidal in shape. To ensure uniform distribution of low frequencies, the room dimensions should fulfill the following equation from [43]:

$$1.1 * \frac{Width}{Height} \leq \frac{Length}{Height} \leq 4.5 * \frac{Width}{Height} - 4$$

$$where\ \frac{Length}{Height} < 3\ and\ \frac{Width}{Height} < 3 \tag{5}$$

The dimensions of REM Studio D are below the recommendation at 17.54 square meters, however, [43] notes that smaller rooms will limit the number of participants. Since this research only tested a single listener at a time, these dimensions were deemed adequate. The room is symmetrical along the vertical plane and is trapezoidal in shape.

The dimensions of the room are as follows:

- Mean Width: 3.77 meters
  - Front width: 3.2 meters
  - Rear width: 4.34 meters
- Length: 4.65 meters
- Height: 2.6 meters

These values, when used to calculate the room dimension recommendation shows that REM Studio D does meet the requirements for room dimensions as shown below:

$$1.1 * \frac{3.77}{2.6} \leq \frac{4.65}{2.6} \leq 4.5 * \frac{3.77}{2.6} - 4$$

$$or$$

$$1.56 \leq 1.79 \leq 2.53 \tag{6}$$

The acoustical properties of the room are also codified by [43] in relation to reverberation time, frequency response, early reflections, operational room noise, and background noise. Reverberation time recommendations are frequency dependent and are shown in Figure 28. REM Studio D meets these tolerances for all frequencies above 125 Hz. As noted by the recommendation, reverberation times can be difficult to measure at low frequencies. REM Studio D was measured to be below the recommended reverberation time at frequencies below 125 Hz. Measurements of reverberation time can be seen in Figure 29.

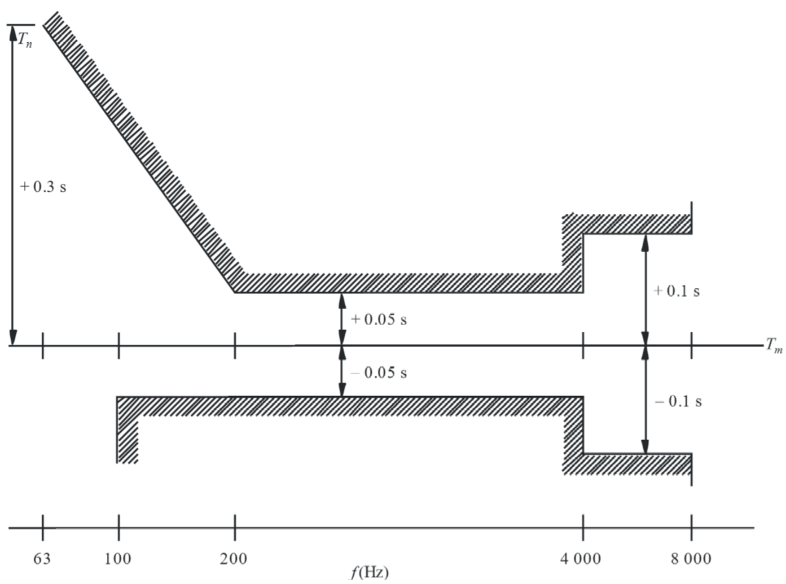


Figure 28. Tolerance for reverberation time, relative to the average value, Tm [43].

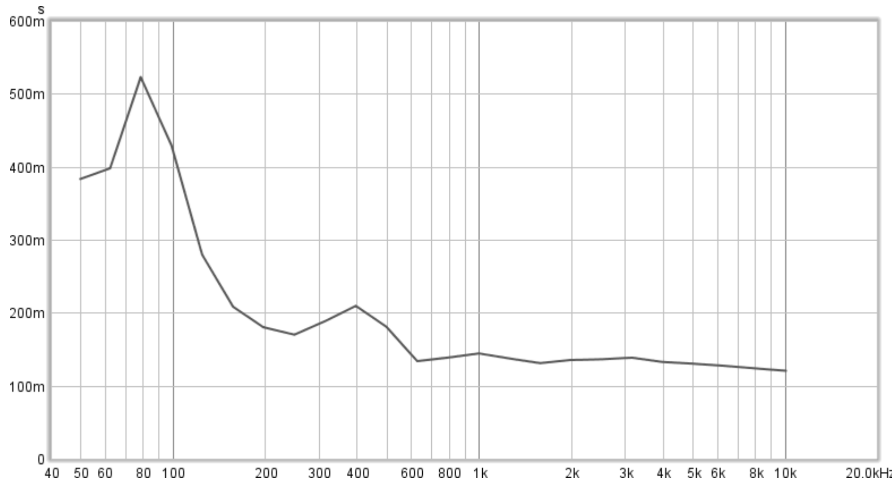


Figure 29. Reverberation time measurement (1/3 octave bands) for REM Studio D.

Frequency response, per the recommendation, should fall within a 4-decibel band when measured in one third octave bands from 40 Hz to 16 kHz. The frequency response of the REM Studio D loudspeakers is shown in Figure 30.

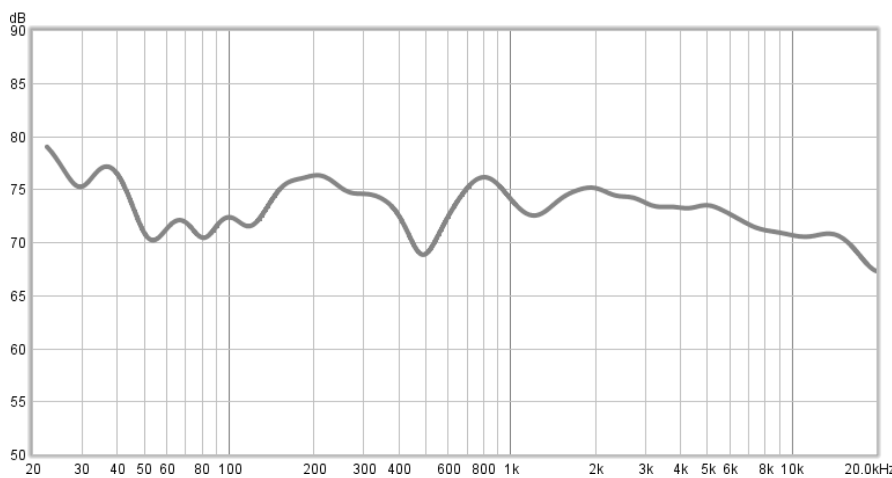


Figure 30. Frequency response measurement (1/3 octave smoothing) for REM Studio D.

Early reflections are defined by [43] as those which reach the listening area during a time interval up to 15 ms after the direct sound. The recommendation is for these reflections to be attenuated by at least 10 dB in the range from 1 kHz to 8 kHz. Measurements of early reflection attenuation in REM Studio D are shown in Table 8.

Table 8. Early reflection (15 ms) attenuation for REM Studio D.

Frequency	1 kHz	1.25 kHz	1.6 kHz	2 kHz	2.5 kHz
Early Reflection Attenuation	-12.9 dB	-14.8 dB	-14.5 dB	-15.9 dB	-13.9 dB
Frequency	3.15 kHz	4 kHz	5 kHz	6.3 kHz	8 kHz
Early Reflection Attenuation	-14.7 dB	-14 dB	-15.1 dB	-14.2 dB	-14.1 dB

Operational room noise is measured in relation to the average sound pressure level from 50 Hz to 16 kHz and is defined in [43] with the chart shown in Figure 31. As shown in the frequency response measurement in Figure 30, REM Studio D displays an operational room response within the recommended limits.

The final measurement conducted to ensure that REM Studio D falls within the recommend specifications of [43] is background noise. The background noise of a listening room should fall below the NR10 noise rating curve shown in Figure 32. Measurements for the background noise of REM Studio D are shown in Figure 33. The measurements show that the background noise falls well below the NR10 rating.

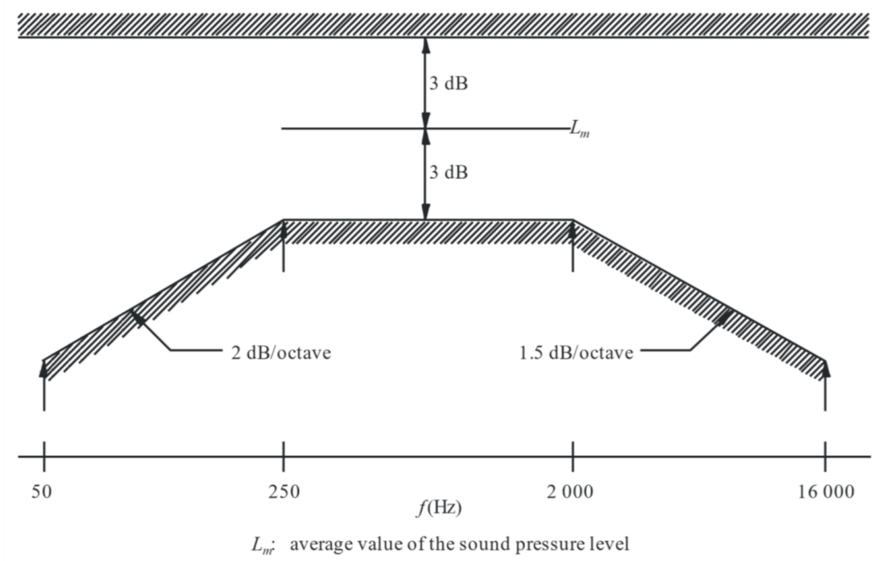


Figure 31. Tolerance for operational room response curve [43].

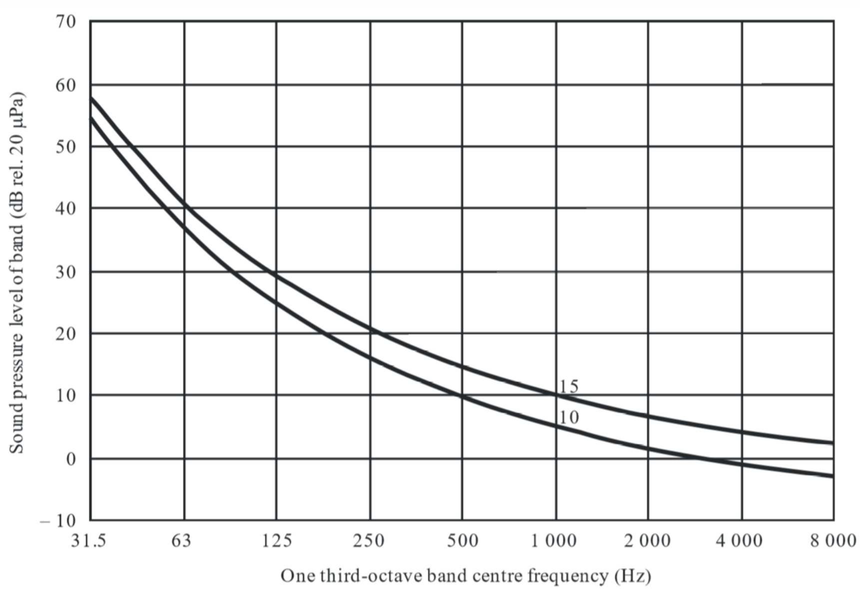


Figure 32. One-third-octave band background noise level limits and noise rating curves [43].

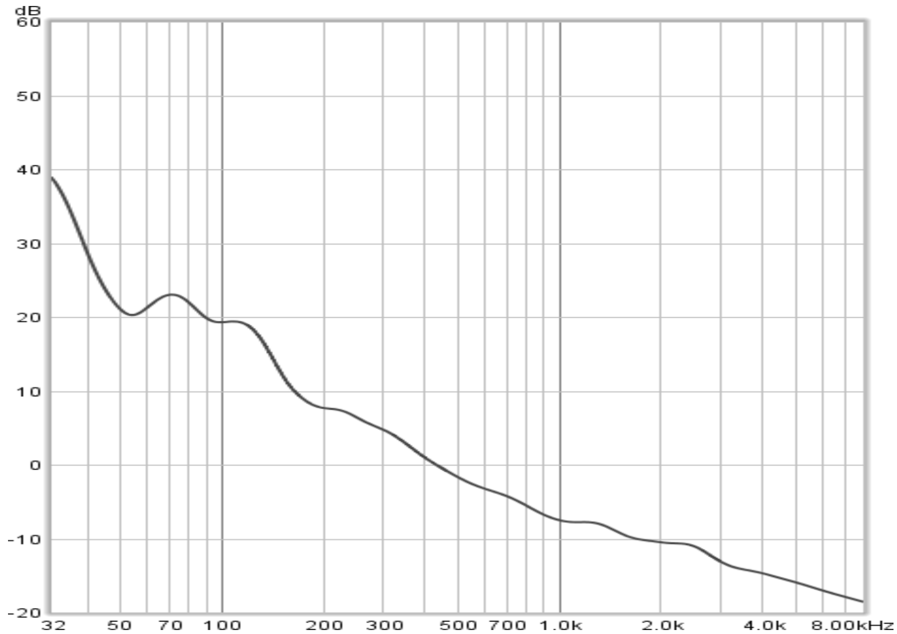


Figure 33. One-third-octave band measurements for background noise.

C. Listening Test Data

Table 9. ANOVA data for width preference.

Source of Variation	*d.f.*	*F*	*p*-value
Subject	12	1.1273	.3614
Technique	4	10.5609	< .0001
Within Groups	48		

Table 10. *Post-hoc* analysis (*Fisher* LSD) of technique for width preference.

Contrast	Difference	Test Statistic	*p*-value
Omni vs Wide	-0.5077	1.9936	.0466
Omni vs Cardioid	-0.7077	2.779	.0056
Omni vs Hyper	-1.2231	4.8028	< .0001
Omni vs Bi	-1.4769	5.7997	< .0001
Wide vs Cardioid	-0.2	0.7854	.4325
Wide vsHyper	-0.7154	2.8092	.0051
Wide vs Bi	-0.9692	3.806	.0002
Cardioid vs Hyper	-0.5154	2.0238	.0434
Cardioid vs Bi	-0.7692	3.0207	.0026
Hyper vs Bi	-0.2538	0.9968	.3192

Table 11. ANOVA data for depth preference.

Source of Variation	*d.f.*	*F*	*p*-value
Subject	12	1.546	.1408
Technique	4	9.5858	< .0001
Within Groups	48		

Table 12. *Post-hoc* analysis (*Fisher* LSD) of technique for depth preference.

Contrast	Difference	Test Statistic	*p*-value
Omni vs Wide	-0.2769	1.1003	.2716
Omni vs Cardioid	-0.5462	2.1701	.0304
Omni vs Hyper	-1.0846	4.3096	< .0001
Omni vs Bi	-1.3231	5.2571	< .0001
Wide vs Cardioid	-0.2692	1.0698	.2851
Wide vsHyper	-0.8077	3.2093	.0014
Wide vs Bi	-1.0462	4.1568	< .0001
Cardioid vs Hyper	-0.5385	2.1395	.0328
Cardioid vs Bi	-0.7769	3.087	.0021
Hyper vs Bi	-0.2385	0.9475	.3437

Table 13. Uncorrected ANOVA data for timbral balance preference.

Source of Variation	*d.f.*	*F*	*p*-value
Subject	12	1.9883	.0465
Technique	4	22.291	< .0001
Within Groups	48		

Table 14. Uncorrected *post-hoc* analysis (*Fisher* LSD) of technique for timbral balance preference.

Contrast	Difference	Test Statistic	*p*-value
Omni vs Wide	-0.2769	1.1003	.2716
Omni vs Cardioid	-0.5462	2.1701	.0304
Omni vs Hyper	-1.0846	4.3096	< .0001
Omni vs Bi	-1.3231	5.2571	< .0001
Wide vs Cardioid	-0.2692	1.0698	.2851
Wide vsHyper	-0.8077	3.2093	.0014
Wide vs Bi	-1.0462	4.1568	< .0001
Cardioid vs Hyper	-0.5385	2.1395	.0328
Cardioid vs Bi	-0.7769	3.087	.0021
Hyper vs Bi	-0.2385	0.9475	.3437

Table 15. Uncorrected *post-hoc* analysis (*Fisher* LSD) of subject for timbral balance preference.

Contrast	Difference	Test Statistic	*p*-value
1 vs 2	1.1	2.985	.0029
1 vs 3	0.28	0.7598	.4476
1 vs 4	0.34	0.9226	.3566
1 vs 5	0.84	2.2794	.0230
1 vs 6	0.72	1.9538	.0512
1 vs 7	-0.02	0.0543	.9567
1 vs 8	0.1	0.2714	.7862
1 vs 9	-0.02	0.0543	.9567
1 vs 10	0.78	2.1166	.0347
1 vs 11	0.18	0.4884	.6254
1 vs 12	0.52	1.4111	.1587
1 vs 13	0.32	0.8684	.3855
2 vs 3	-0.82	2.2252	.0264
2 vs 4	-0.76	2.0623	.0396
2 vs 5	-0.26	0.7055	.4807
2 vs 6	-0.38	1.0312	.3029
2 vs 7	-1.12	3.0392	.0025
2 vs 8	-1.0	2.7136	.0068
2 vs 9	-1.12	3.0392	.0025
2 vs 10	-0.32	0.8684	.3855
2 vs 11	-0.92	2.4965	.0128
2 vs 12	-0.58	1.5739	.1160
2 vs 13	-0.78	2.1166	.0347
3 vs 4	0.06	0.1628	.8707
3 vs 5	0.56	1.5196	.1291
3 vs 6	0.44	1.194	.2329
3 vs 7	-0.3	0.8141	.4159
3 vs 8	-0.18	0.4884	.6254
3 vs 9	-0.3	0.8141	.4159
3 vs 10	0.5	1.3568	.1753
3 vs 11	-0.1	0.2714	.7862
3 vs 12	0.24	0.6513	.5151
3 vs 13	0.04	0.1085	.9136
4 vs 5	0.5	1.3568	.1753
4 vs 6	0.38	1.0312	.3029
4 vs 7	-0.36	0.9769	.3290
4 vs 8	-0.24	0.6513	.5151
4 vs 9	-0.36	0.9769	.3290
4 vs 10	0.44	1.194	.2329
4 vs 11	-0.16	0.4342	.6643
4 vs 12	0.18	0.4884	.6254
4 vs 13	-0.02	0.0543	.9567
5 vs 6	-0.12	0.3256	.7448
5 vs 7	-0.86	2.3337	.0199

Contrast	Difference	Test Statistic	*p*-value
5 vs 8	-0.74	2.0081	.0451
5 vs 9	-0.86	2.3337	.0199
5 vs 10	-0.06	0.1628	.8707
5 vs 11	-0.66	1.791	.0738
5 vs 12	-0.32	0.8684	.3855
5 vs 13	-0.52	1.4111	.1587
6 vs 7	-0.74	2.0081	.0451
6 vs 8	-0.62	1.6824	.093
6 vs 9	-0.74	2.0081	.0451
6 vs 10	0.06	0.1628	.8707
6 vs 11	-0.54	1.4653	.1433
6 vs 12	-0.2	0.5427	.5875
6 vs 13	-0.4	1.0854	.2781
7 vs 8	0.12	0.3256	.7448
7 vs 9	0.0	0.0	1.0
7 vs 10	0.8	2.1709	.0303
7 vs 11	0.2	0.5427	.5875
7 vs 12	0.54	1.4653	.1433
7 vs 13	0.34	0.9226	.3566
8 vs 9	-0.12	0.3256	.7448
8 vs 10	0.68	1.8452	.0655
8 vs 11	0.08	0.2171	.8282
8 vs 12	0.42	1.1397	.2548
8 vs 13	0.22	0.597	.5507
9 vs 10	0.8	2.1709	.0303
9 vs 11	0.2	0.5427	.5875
9 vs 12	0.54	1.4653	.1433
9 vs 13	0.34	0.9226	.3566
10 vs 11	-0.6	1.6282	.104
10 vs 12	-0.26	0.7055	.4807
10 vs 13	-0.46	1.2483	.2124
11 vs 12	0.34	0.9226	.3566
11 vs 13	0.14	0.3799	.7041

Table 16. Corrected ANOVA data for timbral balance preference.

Source of Variation	*d.f.*	*F*	*p*-value
Subject	11	1.4711	.177
Technique	4	20.9347	$< .0001$
Within Groups	44		

Table 17. Corrected *post-hoc* analysis (*Fisher* LSD) of technique for timbral balance preference.

Contrast	Difference	Test Statistic	p-value
Omni vs Wide	-0.4583	1.9414	.0527
Omni vs Cardioid	-0.9333	3.9535	.0001
Omni vs Hyper	-1.5	6.3538	< .0001
Omni vs Bi	-1.8917	8.0128	< .0001
Wide vs Cardioid	-0.475	2.012	.0447
Wide vsHyper	-1.0417	4.4124	< .0001
Wide vs Bi	-1.4333	6.0714	< .0001
Cardioid vs Hyper	-0.5667	2.4003	.0167
Cardioid vs Bi	-0.9583	4.0594	.0001
Hyper vs Bi	-0.3917	1.659	.0976

ACKNOWLEDGMENTS

First and foremost, I must thank my wife, Megan McCracken, for her unwavering support throughout the pursuit of this degree. Without the guidance of Wesley Bulla this research would simply not have been possible. Thank you to the Starbucks Madison crew, and my manager Melissa Antocicco, for covering my shifts when I needed to escape to Nashville to conduct research. To SSG Jonathan Pusztai for pushing me to become a better audio engineer, but also for always being a sounding board not only for my mixes, but also for my thoughts on life in general. Lastly, a huge shout-out to the members of the 101st Airborne Division Band (Air Assault). Your support of my various projects, and your willingness to leave your comfort zones and step into the studio was invaluable to my growth as an audio engineer.

AUTHOR BIOGRAPHY

Brent T. Hauer is currently finishing his Master of Science degree in Audio Engineering Technology at Belmont University. He has previously earned a Bachelor of Music degree in Trombone Performance at Lawrence University in 2010. Brent's research interests include auditory perception of stereophonic microphone techniques and monaural compatibility of stereophonic recording techniques. Prior to attending Belmont University, Brent served as a musician and audio engineer for the 2d Marine Division Band and the 101st Airborne Division Band. He has also worked as a production manager for the Aspen Music Festival and School. Brent aspires to a career in higher education as an audio engineer and professor.