View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Clemson University: TigerPrints

Clemson University
TigerPrints

All Theses | Theses

12-2007

THE EFFECTIVENESS OF MUSICAL STIMULI IN SPATIAL AUDIO DISPLAYS: AN EMPIRICAL INVESTIGATION OF THE EFFECTS OF VOLUME AND SPATIAL PROCESSING DETAIL

Matthew Crisler
Clemson University, crisler@clemson.edu

Follow this and additional works at: https://tigerprints.clemson.edu/all_theses

Part of the Psychology Commons

Recommended Citation
Crisler, Matthew, "THE EFFECTIVENESS OF MUSICAL STIMULI IN SPATIAL AUDIO DISPLAYS: AN EMPIRICAL INVESTIGATION OF THE EFFECTS OF VOLUME AND SPATIAL PROCESSING DETAIL" (2007). *All Theses*. 293.
https://tigerprints.clemson.edu/all_theses/293

This Thesis is brought to you for free and open access by the Theses at TigerPrints. It has been accepted for inclusion in All Theses by an authorized administrator of TigerPrints. For more information, please contact kokeefe@clemson.edu.

THE EFFECTIVENESS OF MUSICAL STIMULI IN SPATIAL AUDIO DISPLAYS: AN EMPIRICAL INVESTIGATION OF THE EFFECTS OF VOLUME AND SPATIAL PROCESSING DETAIL

A Thesis
Presented to
The Graduate School of
Clemson University

In Partial Fulfillment
Of the requirements for the Degree
Master of Science
Applied Psychology

---

by
Matthew Crisler
December 2007

Accepted by
Dr. Richard Tyrrell, Committee Chair
Dr. Chris Pagano
Dr. Rich Pak

ABSTRACT

Audio displays have potential to convey spatial information to users without taxing their visual resources, but have been shown to annoy some users. Musical stimuli have the potential to reduce user annoyance, but their potential to be localized spatially is untested. These experiments tested how well musical stimuli can be localized at different volumes and when using different spatial processing techniques to manipulate the spatial information.

The two experiments presented participants with brief musical stimuli simulating spatial locations between -40° and 40° from the saggital plane and asked participants to report the perceived direction of the sound. In Experiment 1, two spatial processing techniques were compared, and it was determined that a simple processing technique involving only manipulating the relative volume of two speakers is as effective as a more resource-intensive processing technique that incorporates multiple spatial cues. Experiment 2 manipulated the overall volume from 55 dBA to 65 dBA and showed that, throughout this range, there are no significant differences in spatial location ability.

TABLE OF CONTENTS

Page

TITLE PAGE i

ABSTRACT ii

LIST OF TABLES v

LIST OF FIGURES vi

CHAPTER

I. INTRODUCTION 1

II. EXPERIMENT 1 METHODS 11

Participants 11
Apparatus 12
Experimental Design 13
Procedure 14

III. EXPERIMENT 2 METHODS 17

Participants 17
Apparatus 18
Experimental Design 18
Procedure 18

IV. RESULTS 21

Data Analysis – Experiments 1 and 2 21
Experiment 1 - ±45° Speaker Configuration 25
Experiment 1 - ±90° Speaker Configuration 29
Experiment 2 - ±45° Speaker Configuration 31
Experiment 2 - ±90° Speaker Configuration 35

V. DISCUSSION 37

VI. CONCLUSION 42

Table of Contents (Continued)

Page

APPENDICES 44

A: Pilot Research Details 45
B. Individual Participant Graphs 60

REFERENCES 85

LIST OF TABLES

Table Page

1. Intensity differences used for IID processing 15

A.1 Mean perceived location at each Spatial Processing Location for individual speaker configurations 52

A.2 Mean differences of adjacent spatial processing pairs..... 53

LIST OF FIGURES

Figure Page

1. Protractor image used by participants to report perceived locations 13

2. Perceived location of sounds processed using both processing methods as a function of processed location. 23

3. Perceived location of sounds at each volume level as a function of processed location. 24

4. Mean perceived location as a function of processed location for each processing method 26

5. Perceived location of the sound source as a function of the processed location for the HRTF processing technique in the ±45° speaker configuration 27

6. Perceived location of the sound source as a function of the processed location for the IID processing technique in the ±45° speaker configuration 28

7. Perceived location of the sound source as a function of the processed location for the HRTF processing technique in the ±90° speaker configuration 30

8. Perceived location of the sound source as a function of the processed location for the IID processing technique in the ±90° speaker configuration 31

9. Mean perceived location as a function of processed location for the two musical clips (jazz and rock) 32

10. Perceived angle as a function of processed location for all participants and all volume levels in the ±45° speaker configuration 34

11. Perceived location of the sound source as a function of the processed location for all volume levels in the ±90° speaker configuration 36

A.1 Protractor used for reporting angular location of sound source 47

List of Figures (Continued)

Page

A.2 Speaker Location X Spatial Processing Interaction 50

A.3 Pilot Test data showing perceived angle of spatially processed sound in different speaker configurations 52

A.4 Scatterplot of Perceived Angle vs Spatially Processed Angle with Regression Lines 54

A.5 Typical individual participant regression result for participants that discriminated locations as well as direction 56

A.6 Typical individual participant regression result for participants that discriminated locations as well as direction (left vs. right) for speaker configurations, but only direction for headphone condition..... 57

A.7 Typical individual participant regression result for participants that discriminated most directions (left vs. right), but had trouble discriminating exact locations..... 58

B.1 Perceived location of sounds processed using both processing methods as a function of processed location 60-71

B.2 Perceived location of sounds at each volume level as a function of processed location..... 72-83

INTRODUCTION

According to the National Highway Traffic Safety Administration, 49,559 vehicles were involved in fatal crashes on US roads in 2004. Of these vehicles, 6624, or 13 percent, occurred while maneuvering around a curve. Unfortunately, roadside signs warning of curves and other dangers are fairly ineffective and generally result in less than a 50% accurate response rate (Fisher, 1992). This confirms a problem identified by Drory and Shinar who showed that only 5 to 10 percent of drivers could accurately recall a warning sign only 200 meters after passing it (as cited in Neurater, 2005). The implementation of a more salient curve awareness system could potentially enhance drivers' situation awareness in these situations (Neurater, 2005).

The present research sought to enhance our knowledge of the effectiveness of spatial audio displays with complex musical stimuli. This could enhance the design of a spatial audio based curve awareness display based on the possibility of adding location information into the vehicle's existing primary audio system used for the vehicle's entertainment system. The current research is based on earlier research showing the effectiveness of an auditory modality for curve awareness systems (Neurater, 2005) as well as pilot research supporting the effectiveness of spatial audio displays using musical stimuli (see appendix A).

Neurater tested auditory, visual, and haptic-based curve awareness systems using a high fidelity driving simulator. Systems were compared based on their effectiveness in terms of altering throttle reaction time, brake reaction time, and curve entrance speed. In addition, subjective measures assessed feelings about the urgency, annoyance,

appropriateness, interference, and desirability of the systems. The results of the study suggest that a verbal warning consisting of the words, "Curve ahead, Reduce Speed to 20 mph", would be the most effective system due to decreased curve entrance speeds and reaction times as well as high subjective ratings of urgency, appropriateness and desirability (Neurater, 2005). However, the subjective results from this study contradict previous research by Lerner, Decker, Steinberg, and Huey (1996) which showed that a digitized vocal stimulus resulted in higher ratings for annoyance than a rapidly beeping tone. Thus, vocal warnings like the one suggested by Neurater might, in long-term real world use, become unacceptably annoying to drivers. A more subtle modality that can be expected to result in lower ratings for annoyance should be designed and tested in order to produce a system that drivers would be less likely to disable.

Driving is an inherently spatial task that requires drivers to control a vehicle through a dynamic and unpredictable environment. Therefore, a more dynamic curve awareness system that has a more natural spatial mapping would seem appropriate; however, the visual system is already relied upon heavily by the driver for the purpose of collision avoidance and vehicle guidance (Norman, 2002). Therefore, consistent with a multiple resource model of attention, a dynamic visual display would seem to be a less appropriate option to achieve the goal of enhancing the situation awareness of the driver as compared to other modalities that are utilized less by the driving task (Wickens and Hollands, 2000). The human perceptual system is fairly effective at localizing sounds within space; suggesting a possible auditory option to warn drivers of upcoming curves. A system using spatial audio to enhance curve awareness could alert drivers to an

upcoming curve, as well as its severity relative to the vehicle's current speed and distance from the curve. A dynamic spatial audio curve awareness system could help to meet many of Norman's (2005) seven design principles including simplifying the structure of the task, making road features visible, getting the mapping correct (a spatial display for a spatial task), and designing for error. Many of the advantages that can be expected by using a spatial system involve the concept of stimulus-response compatibility (Kornblum & Lee, 1995). Many of Neurater's (2005) displays had poor stimulus response compatibility (vocal, tone, and auditory icon) and would therefore require additional cognitive processing in order to interpret and react to the display. In addition, the visual conditions alone tended to not be salient. This was particularly surprising due to the laboratory context in which the drivers were using the system for the first time where novelty would be expected to result in an overstated salience (Neurater, 2005).

In order to produce a curve awareness display that is effective and salient while not being overly annoying to drivers, I propose that a spatial audio display using a musical (or vocal) stimulus should be designed to portray curve information in a dynamic manner. It is important that such a system be designed to be consistent with the methods used by humans to perceive the location of sounds within space, and designers must also understand how effective and accurate these methods are for complex stimuli such as music. The human auditory perceptual system has been studied fairly extensively, and the mechanisms used to localize sounds within space are well documented; however, from a practical standpoint, there are few data that address how effectively humans perceive spatial audio with complex auditory stimuli such as music.

The human perceptual system uses two main cues to localize sounds within space. These cues are interaural intensity difference (IID) and interaural time difference (ITD). These cues are used by the brain to identify the position of a sound within space based on the difference in sound intensity at each ear and the difference in arrival time of the sound.

The human perceptual system can detect interaural time differences as small as 10 μs, which is enough to localize sound within 1 degree. The range of interaural time delay for sounds from 0 degrees to 90 degrees is approximately 0-640 μs (Wolfe, 2005). Although the exact mechanism used by humans to process and localize sounds using ITD is not known, a model was proposed by Jeffress (1948). This model proposes that the nerves within the ear send signals to the brain via a series of delay lines that serve to create a spatial map of the lateral plane. Neurons that detect coincidence (two similar signals arriving at the same time) fire at different rates based on the level of coincidence. These neurons have been found within the lateral superior olive and the inferior colliculus; however, the mechanism of the delay lines has not been isolated, and therefore this model has yet to be fully supported by neurological research (Behrend et al, 2002; Hartmann, 1999; Goldberg and Brown, 1969).

Interaural intensity difference is generally effective as a spatial cue only in the higher frequency ranges (above 4000 Hz). This is due to the fact that lower frequency sound waves have longer wavelengths relative to the head, and therefore are able to bend around the head with relatively little attenuation. However, the higher frequency sounds with shorter wavelengths are blocked by the head resulting in a significant reduction in

perceived sound level at the far ear. At 5,000 and 6,000 Hz, IIDs range from 0 dB directly in front of the perceiver to approximately 20 dB at 90 degrees to the side (Wolfe, 2005).

The IID and ITD cues for sound localization have been tested for accuracy across the audible frequency spectrum, and it has been found that they are highly accurate for most frequencies except those between 2000 and 4000 Hz. This finding has resulted in a duplex theory of sound localization that asserts that the two spatial cues, IID and ITD, are used to localize sounds of different frequencies. Low frequencies are localized using the ITD cue, and high frequencies are localized using the IID cue. Sounds with intermediate frequencies (2000-4000 Hz), between the optimum ranges of these two cues, are harder to locate within space. Fortunately most real world sounds consist of a wide range of frequencies, and can therefore be localized using one or both of the cues (Blake, 2006). This duplex theory is supported by neurological research suggesting that different areas of the brain (Medial Superior Olive and Later Superior Olive) appear to process IID and ITD information (McAlpine, 2005).

In order to reproduce all of the IID and ITD cues along with other subtle location cues that allow for localizing sound in the vertical plane and discerning the differences between sounds coming from in front of and behind an observer, careful measurements have been taken of humans and models of human ears and heads to produce a set of head related transfer functions that can be used to spatialize a musical source to be played using speakers. This is accomplished by playing different stimuli at different locations around the head model and recording the resulting sound produced within the model ear.

This process results in an impulse response for each location that can be used by digital signal processors (DSPs) to reproduce the effect of the head, ears, pinna, etc. for any sound input presented to the DSPs. The data resulting from this measurement process are referred to as a head-related transfer function (HRTF; Gardner and Martin, 1994).

Though the use of a HRTF to produce spatialized sound can be effective, it is also a fairly complex process that requires significant signal processing. A simpler system using IID manipulation alone has successfully been used by researchers of spatial audio displays that can be effective and less complicated. A spatial audio display using a musical source was designed as an attitude indicator for an aircraft using the pan dimension (resulting in a change in IID) to portray roll information and another dimension, emphasis on low and high frequencies, to portray pitch information. Although the study was conducted with limited statistical power due to a small sample size, the attitude indicator was effective in both the roll and pitch dimensions for 2 of the 3 pilots tested (Simpson, 2005). This suggests that simplified spatial audio systems involving only IID manipulations can effectively portray dynamic information using musical sources. In addition, research has shown that the smoothing of HRTFs by reducing the spectral detail does not result in significant decreases in localization ability. This suggests that humans have the ability to use fairly impoverished spatial signals to locate sounds (Kulkarni and Colburn, 1998).

The effectiveness of these cues and the human ability to perceive them for locating sounds within space has been extensively researched for pure tones and other simple stimuli; however, for practical applications including a system to enhance curve

awareness in automobiles, more complex stimuli could be quite effective in order to enhance the acceptance of the system (Begault and Wenzel, 1993; Wenzel, Arruda, Kistler and Wightman, 1993). In addition, a number of systems have been successfully designed and implemented using spatial audio to enhance performance in real world and laboratory tasks (Bolia, 2004; Simpson, 2005; Begault, Wenzel and Shrum, 1996; Holland and Morse, 2001; Grohn, Lokki and Takala, 2004; Bolia and D'Angelo, 1999). A system designed to enhance curve awareness using musical stimuli could be quite effective without resulting in a significant annoyance to the driver or vehicle passengers. In order to implement a system of this nature, it is important to understand how well potential drivers can localize musical stimuli, and how accurately the system must process the sound to produce the IID and ITD cues.

The existing literature on spatial audio is very limited in terms of the use of musical sources. Most previous research has focused on using pure tones (sine waves), noise bursts, and other auditory stimuli that could result in significant annoyance if used continuously in a driving task. As with all spatial audio, the design of a spatial audio display using musical stimuli could range from simple to extremely complex and require vastly different amounts of audio processing depending on how complex a model is used to produce the spatial locations based on combinations of IID, ITD, and HRTF. In order to minimize the complexity of the system and understand the advantages and limitations of music as a source for spatial audio displays, some fundamental questions must be answered before completing the design of the display.

In order to design an effective display, one must know how effective the spatial audio system is at producing sounds that can be discriminated by users based on their spatial location. Pilot research (see appendix A for details) has established that signals processed to be separated by 10 degrees can be effectively discriminated under a number of speaker configurations through a range of -30 degrees to +30 degrees. This is consistent with perceptual research on absolute judgments suggesting that errors will begin to occur in absolute perceptual judgments when the number of stimulus levels exceeds five to six (Wickens & Hollands, 2000). A full HRTF model that takes advantage of IID as well as ITD location cues was used in this experiment. This study also showed that there were few instances (<5%) where a signal on the left was perceived as being on the right or vice versa which is important for the design of a curve awareness system. This type of error would result in the driver perceiving a curve in the wrong direction; whereas, under or over-stating the magnitude of the spatial deviation, though still an error, is likely less of a problem in this context as it would not be expected to result in a grossly improper response (expecting a turn in the wrong direction). The granularity of a display that could be produced based on the pilot study is expected to be effective for a curve awareness display; however, further investigation is necessary to determine if a simpler spatial audio system could be used effectively as well as to determine if there are other factors, such as volume level, that affect the sensitivity of the display system. Investigating the effects of volume level may also help to explain some of the results of pilot research indicating that it is nearly impossible to process audio such

that it will be perceived as further away from the saggital plane than the location of the speakers.

I am reporting two experiments that explored both questions of model complexity as well as volume level. In both experiments, participants were presented with musical clips processed to produce a simulated location of -40 to +40 degrees in 10 degree intervals. These clips were presented using different processing methodologies (HRTF vs IID only) in the first experiment, and at different volume levels in the second experiment. The data provide valuable insight to the level of processing that will be required in order to produce an effective curve awareness display as well as appropriate volume ranges for a spatial audio display system using musical stimuli. In addition to these design insights for this specific system, the information could be used for designing any spatial display with musical stimuli.

Based on the results of Simpson (2005), I hypothesized that experiment 1 would show that the IID-only processing method can produce an effective curve awareness display, but I expected that the full HRTF processing method would result in more accurate judgments of spatial location Experiment 2 tested the hypothesis that the low volume levels used in the pilot study resulted in the inability to perceive sounds as being located more peripheral than the location of the speakers (in the 45° speaker configuration) as well as potentially causing the nonlinear results that emerged from the pilot study. Due to the fact that the stimuli were presented at 55 dBA and the audio processing for IID further attenuated the signal from one of the speakers, it is possible that the attenuated speaker was operating at a volume so low that it could not be

accurately perceived for the purpose of decoding ITD and therefore spatial location. By presenting similar stimuli at higher volume levels, it will be possible to determine if the spatial cues are decoded more accurately at different volume levels. This information will allow systems to be designed such that minimum sound levels are achieved for proper localization of complex stimuli.

This report describes the methods of these two experiments in sequence. However, because there were data analysis issues that apply to both experiments, the results of the two experiments are presented together.

EXPERIMENT 1

METHODS

Participants

Participants were selected from the Clemson University Psychology Department subject pool. Participants were undergraduates currently enrolled in a psychology class, and were awarded class credit for their participation in the experiment. Participants were screened for any self-reported aural pathology that would prevent them from successfully completing the experiment. In addition, participants were given a hearing test using an Earscan audiometer that measured pure-tone hearing loss at 500, 1000, 2000, 3000, 4000, 6000, and 8000 Hz. Any participants with an average hearing loss (average of hearing loss at each frequency measured) greater than 20 dB in either ear were excluded from analysis. In addition, any participants with a hearing loss greater than 40 dB at any one frequency were excluded from the analysis. According to the device manufacturer, under ideal testing conditions, hearing loss between 0 and 20 dB falls within normal limits, and losses between 25 and 40 dB represent slight to mild hearing loss. Participant 22 was excluded from analysis due to a self-reported hearing loss in the left ear that was confirmed by the audiometer testing. All other participants met the screening criteria for hearing.

A total of 40 participants completed the experiment. Of those 40, 36 participants met the auditory screening criteria and did not make any obvious errors in following the experiment's instructions. Participant 1's data were lost due to a computer error, and

participant 31 and 45 failed to follow the instructions by clicking on or near the head of the experimental apparatus instead of on the curved portion of the protractor. These participants were replaced to maintain the balanced latin squares design. Data analysis was completed on the remaining 36 participants (13 male and 23 female). Age ranged from 18 to 29 years (M=20.47 years).

Apparatus

Spatial audio stimuli were processed using Sony Media Software's Sound Forge and associated audio processing subsystems. The MIT KEMAR Head Related Transfer Function was used for all full spatial processing stimuli.

All stimuli were presented to participants by an IBM X41 Tablet PC running Windows XP Tablet PC edition and the E-Prime experiment operating system version 1.1 Service Pack 3. E-Prime was used to randomize the order of presentation of the different spatially processed locations as well as counterbalance the experimental conditions (Processing type, Volume Level, and Speaker Location). Stimuli were presented using Altec Lansing BX2 computer speakers mounted at 45° and 90° from the participant's saggital plane. The speakers were mounted at approximately ear height and approximately 27 inches from the center of the participant's head while seated in an anechoic chamber. Participants' heads were positioned in a chinrest throughout the experiment in order to avoid head motions that may affect the spatial audio processing. Speaker locations were masked by an opaque black cloth that covered approximately 200° of the participant's forward field of view. A system to automatically switch the speaker configuration eliminated the clicking sound during speaker changes that may

have identified speaker locations in the pilot study. Participants used the Tablet PC stylus to report the perceived spatial location by clicking on the perceived location on a protractor displayed on the screen with the image of a head at the focal point (see Figure 1).

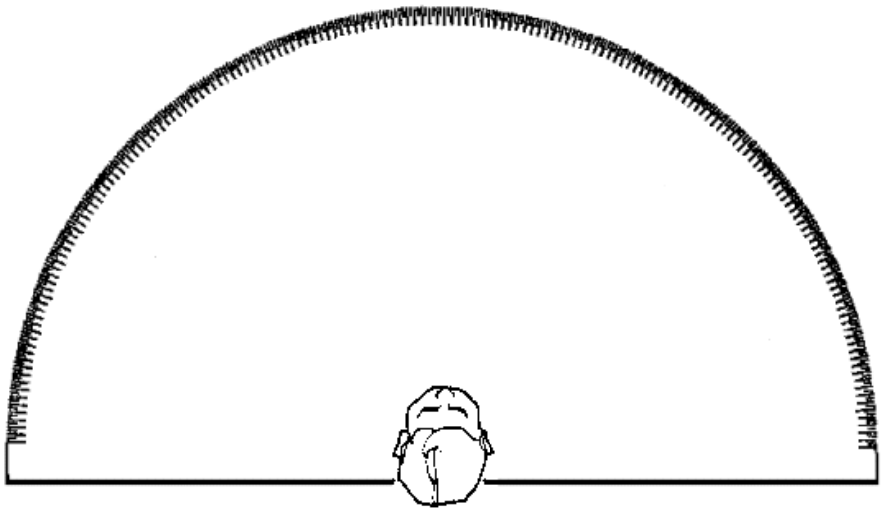


Figure 1: Protractor image used by participants to report perceived locations

Experimental Design

A 9X2X2 (Location X Processing Method X Clip) within-subjects design was used for the spatial processing experiment. There were 9 levels of the spatial location variable (-40° to 40° in 10° intervals), 2 levels of spatial processing (full HRTF and intensity differences only), and 2 speaker locations (45° and 90° from the participant's saggital plane). All participants were exposed to all levels of the independent variables. Speaker location and spatial processing were counterbalanced using a balanced latin squares design. Within each combination of speaker location and spatial processing technique, each of the nine spatial locations was presented using a new random order. Two different

sound clips were presented for each simulated angle resulting in two data points for each simulated angular location. These stimuli were the same as the two musical stimuli used in the pilot study (Pop Rock and Jazz - see Appendix A). Dependent variables include perceived angle as reported on the tablet PC and response time from stimulus presentation to reporting of perceived angle.

Procedure

After providing their informed consent, participants either completed the hearing test, or entered the apparatus and completed the experimental procedure. Those participants that completed the experiment first completed the hearing test immediately following the experiment. In general, two participants completed the experiment at one time, with one participant completing the hearing test while experimental data was collected using the experimental apparatus from the other participant.

Each participant was presented with the same musical clips at each spatial location (-40° to 40° in 10° steps) processed using 2 different methods. The first method was the same as was used in the pilot study (see Appendix A), utilizing the complete MIT KEMAR HRTF (using both IID and ITD information). The second method manipulated only IID and assumed that the stimuli consist of a constant frequency of 5000 Hz. 5000 Hz is in the highly effective range for localization by IID, and should result in noticeable attenuation of the speaker further from the simulated sound location. This resulted in stimuli processed such that the speaker further from the simulated sound location is attenuated by the amount shown in Table 1.

Table 1: Intensity differences used for IID processing (Wolfe, 2006)

Location	0°	±10°	±20°	±30°	±40°
IID	0 dB	4 dB	7.5 dB	11.5 dB	13.5 dB

Participants were seated in the experimental apparatus and the height of the chin rest was adjusted for comfort. Basic instructions and a general description of the procedure were provided by the experimenter. This included the use of the tablet PC stylus. Participants were then asked to read the detailed instructions that were available on the tablet PC and to use the stylus to change pages within the instructions. After reading the instructions, 18 practice trials were completed (using each speaker location 9 times). Nine of the practice trials were full HRTF processing and 9 were IID only, and the practice trials were presented in a new random order for each participant. All stimuli were presented at approximately 65 dBA throughout the practice and experimental trials. After the practice trials, the participant was given an opportunity to ask questions before beginning the experiment, and then the participant was asked to click the screen to begin the experiment. At this point, a musical stimulus was presented via the speaker configuration and processing method as proscribed by the balanced latin squares design. The participant responded using the stylus to click on the protractor at a point that corresponded to the apparent source of the sound. Response time (the time interval from the initial onset of the stimulus to the participant's stylus click) and perceived location were stored by the E-Prime software. After each trial, the participant clicked the screen to advance to the next trial. After being presented all of the spatial locations at a given combination of speaker and processing method, the speaker location and/or processing

method was changed according to the counterbalanced design, and the procedure was repeated until all combinations of location, processing method, and speaker location had been presented to the participant. Participants were then debriefed and given an opportunity to ask any other questions before being allowed to leave. In total, the experiment required approximately 20 minutes from each participant.

EXPERIMENT 2

METHODS

Participants

Participants were selected from the Clemson University Psychology Department subject pool. Participants were undergraduates currently enrolled in a psychology class, and were awarded class credit for their participation in the experiment. Participants were screened for any self-reported aural pathology that would prevent them from successfully completing the experiment. In addition, participants were given a hearing test using an Earscan audiometer that measured pure-tone hearing loss at 500, 1000, 2000, 3000, 4000, 6000, and 8000 Hz. Any participants with an average hearing loss (average of hearing loss at each frequency measured) greater than 20 dB in either ear were excluded from analysis. In addition, any participants with a hearing loss greater than 40 dB at any one frequency were excluded from the analysis. According to the device manufacturer, under ideal testing conditions, hearing loss between 0 and 20 dB falls within normal limits, and losses between 25 and 40 dB represent slight to mild hearing loss. All participants met the screening criteria for hearing. Participant 5 gave inconsistent responses during the hearing test and was excluded from analysis and replaced to maintain the balanced latin squares design.

A total of 38 participants completed the experiment. Of these, 36 participants made no obvious errors in following the experimental procedures and were used for the analysis. Participant 11 was replaced in the analysis due to repeated clicking on (or near)

the head instead of on the curved portion of the protractor image. Data analysis was completed on the remaining 36 participants (13 male and 23 female). Age ranged from 18 to 27 years (M=19.97 years).

Apparatus

The same apparatus was used as in experiment 1.

Experimental Design

A 9X3X2 (Location X Volume Level X Clip) within-subjects design was used for the volume level experiment. There were 9 levels of the spatial location variable (-40° to 40° in 10° intervals), 3 levels of the volume variable (55 dBA, 60 dBA, and 65 dBA), and 2 speaker locations (45° and 90° from the participant's saggital plane). All participants were exposed to all levels of the independent variables. Speaker location and volume level were counterbalanced using a balanced latin squares design, and the nine spatial locations were presented in a new random order within each of the six combinations of speaker location and volume level. Two different sound clips were presented for each spatial location resulting in two data points for each of the 54 combinations of the independent variables. These stimuli were identical to the two musical stimuli used in the pilot study. Dependent variables included perceived angle as reported on the tablet PC and response time from stimulus presentation to reporting of perceived angle.

Procedure

After providing their informed consent, participants either completed the hearing test, or entered the experimental apparatus and completed the procedure. Those

participants that completed the experiment first completed the hearing test immediately following the experiment. In general, two participants completed the experiment at one time, with one participant completing the hearing test while experimental data was collected using the experimental apparatus from the other participant.

Participants were seated in the experimental apparatus, and the height of the chin rest was adjusted for comfort. Basic instructions and a general description of the procedure were provided by the experimenter. This included the use of the tablet PC stylus. Participants were then asked to read the detailed instructions that were available on the tablet PC and to use the stylus to change pages within the instructions. After reading the instructions, 18 practice trials were completed (using each speaker location 9 times). All of the trials were full HRTF processing presented at a new set of volume levels that were randomly chosen from 55, 60, and 65 dBA for each participant. After the practice trials, the participant was given an opportunity to ask questions before beginning the experiment, and then the participant was asked to click the screen to begin the experiment. At this point, a musical stimulus was presented via the speaker configuration and volume level appropriate for the balanced latin squares design. The participant responded using the stylus to click on the protractor at a point that corresponded to the apparent source of the sound. Response time (the time interval from the initial onset of the stimulus to the participant's stylus click) and perceived location were stored by the E-Prime software. After each trial, the participant clicked the screen to advance to the next trial. . After being presented all of the spatial locations at a given combination of speaker location and volume level, the speaker location and/or volume

level was changed according to the counterbalanced design, and the procedure was repeated until all combinations of location, volume, and speaker location were presented to the participant. Participants were then debriefed, given an opportunity to ask questions, and released. In total, the experiment required approximately 20 minutes from each participant.

RESULTS

Data Analysis – Experiments 1 and 2

The perceived location and response time data from both experiments were analyzed using within-subjects ANOVAs with $\alpha=0.05$, and Greenhouse-Geisser degrees of freedom adjustments.

Prior to performing the inferential analyses, however, the data from both experiments were examined to determine the extent to which the participants provided valid responses. This section describes a problem that emerged in the ±90° speaker configuration in both experiments, and summarizes the manner in which this problem was addressed. In both experiments, a number of participants were identified as having had trouble with front/back confusions in the ±90° speaker configuration. This was identified based on stylus responses behind the protractor on the input display; this type of response never occurred in the ±45° speaker configuration. Participants whose data showed responses behind the area of the protractor were excluded from the analysis for the ±90° speaker configuration. Participants 3, 4, 5, 7, 8, 9, 13, 19, 21, and 23 were excluded from the Experiment 1. Participants 7, 13, 17, 19, 26, and 36 were excluded from Experiment 2. It is important to note that these front/back confusions would be expected to cause significant problems and confusion for users of any spatial display system. Therefore, the finding that front/back confusions occurred so frequently in the ±90° speaker configuration confirms and underscores the results of the pilot study

suggesting that a display consisting of speakers mounted at ±90 degree angles should be avoided.

Figures 2 and 3 show the valid and invalid responses of representative individual participants (See Appendix B for all participants). Invalid responses (i.e., responses that were beneath the head in the protractor image) are shown at a value of 100°; valid data points range between -90° and 90°. Note that as discussed above, all data were valid for the ±45° speaker configuration; however, 10 participants gave invalid responses in the ±90° speaker configuration of Experiment 1, and six participants gave invalid responses in the ±90° speaker configuration of Experiment 2.

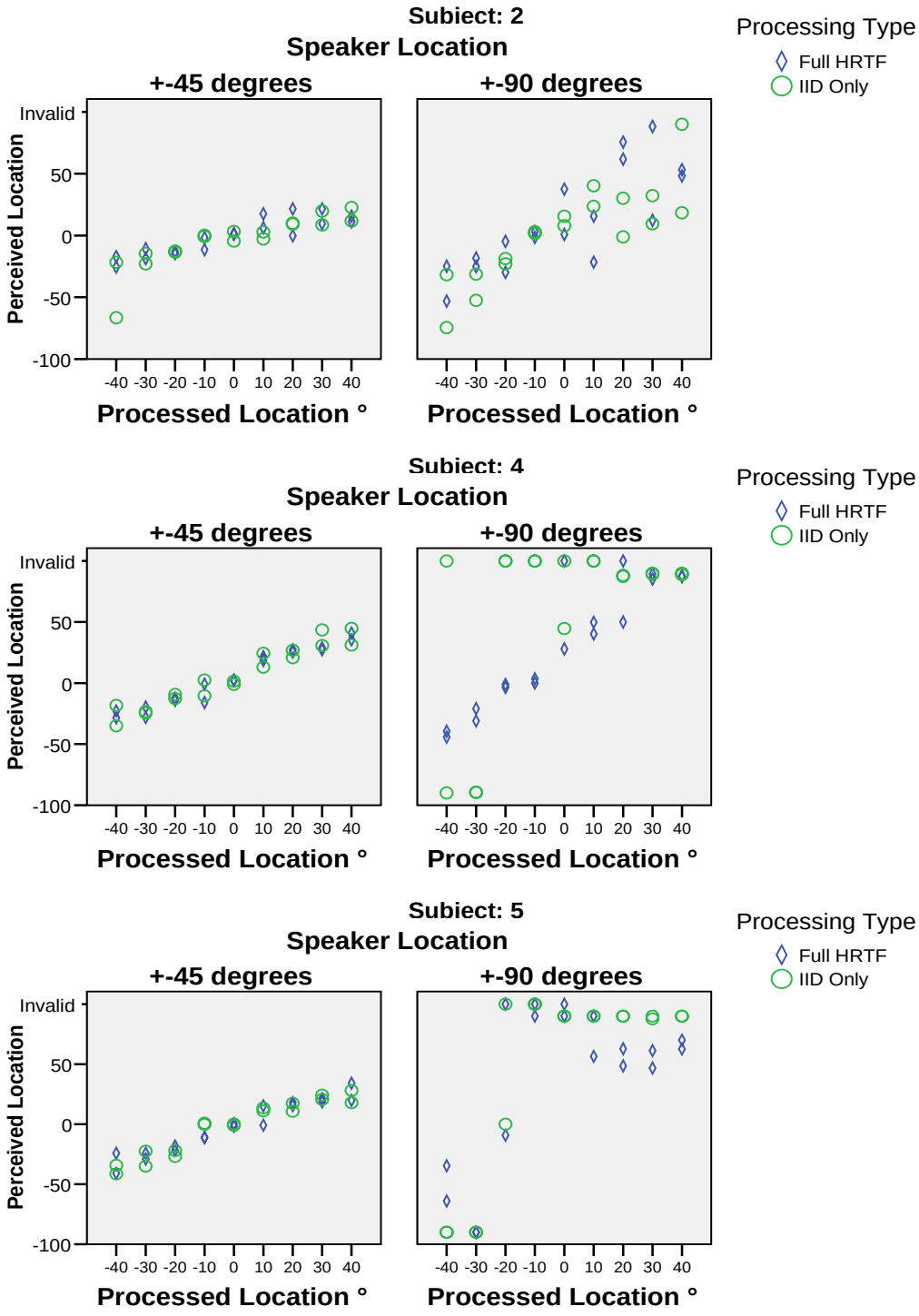


Figure 2: Perceived location of sounds processed using both processing methods as a function of processed location. Data are presented from individual participants. Invalid data are marked as a +100° response. Note the prevalence of invalid responses in ±90° speaker configuration. The units of all axes are degrees.

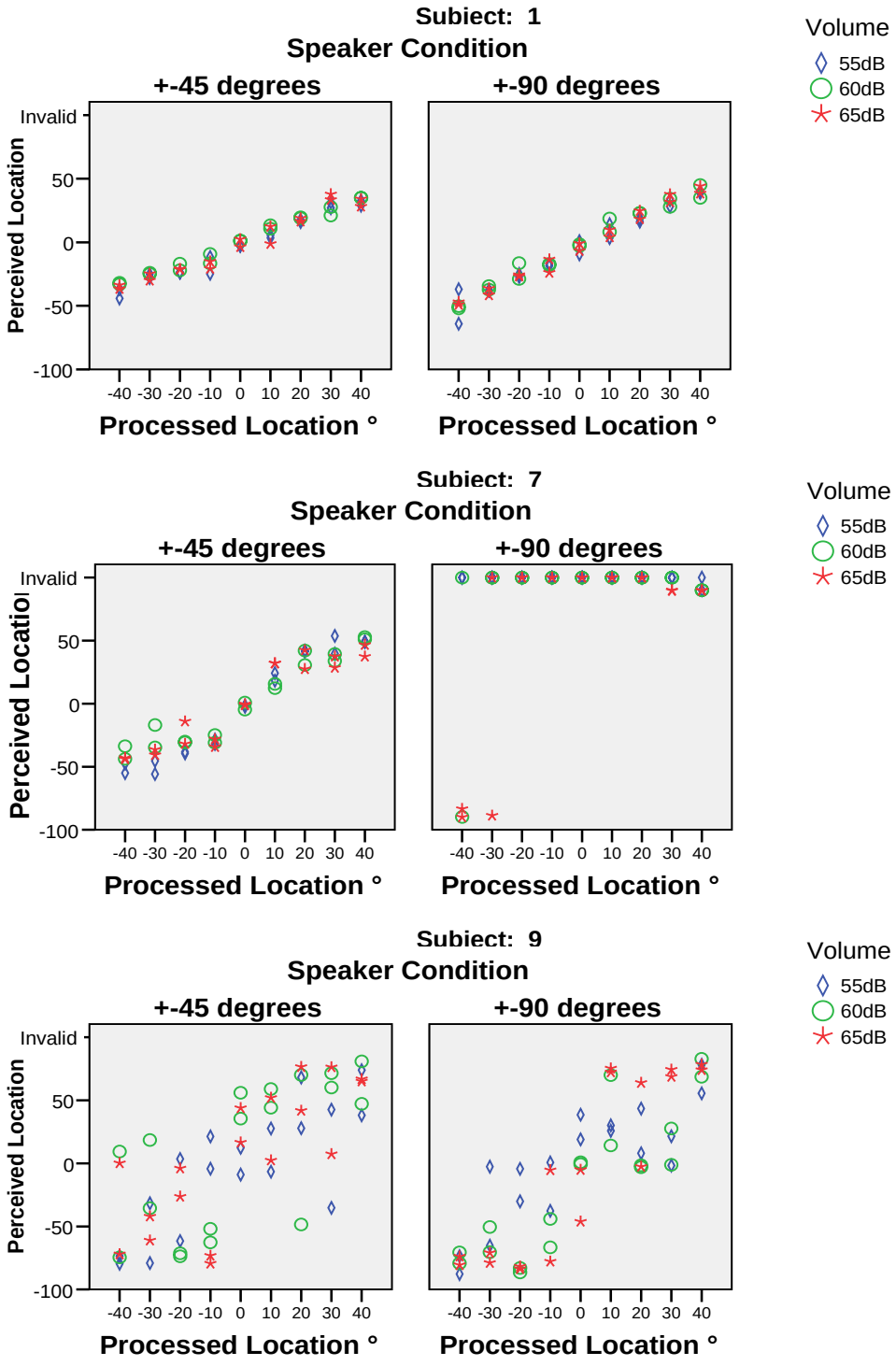


Figure 3: Perceived location of sounds at each volume level as a function of processed location. Data are presented from individual participants. Invalid data are marked as a +100° response. Note the prevalence of invalid responses in ±90° speaker configuration. The units of all axes are degrees.

Due to these exclusions and the counterbalanced within-subjects design, the data from the two speaker configurations were analyzed separately in both experiments. In both experiments, the ANOVA for the ±45° speaker configuration analysis included all participants and remained fully counterbalanced. The analysis of the ±90° speaker configuration, however, excluded the participants with invalid responses (listed previously) and was therefore not counterbalanced. Due to the lack of counterbalancing, together with the results from the pilot study (see Appendix A) that suggest order effects will dominate the response time data without proper counterbalancing, the response time data were not analyzed for the ±90° speaker configuration. In addition, interpretation of the results of the data from the ±90° speaker configuration must be handled cautiously given the frequency with which participants provided invalid responses.

Experiment 1 - ±45º Speaker Configuration

As expected, there was a significant effect of processed location ($F(1.621, 56.749)=338.628$, $p<0.001$). However, this was qualified by a location X processing method interaction ($F(5.311, 185.876)=6.276$, $p<0.001$) (see Figure 4). There were no significant differences between the perceived location of sounds processed using the full HRTF method as compared to those processed using only the IID processing ($F(1, 35)=0.839$, $p=0.366$).

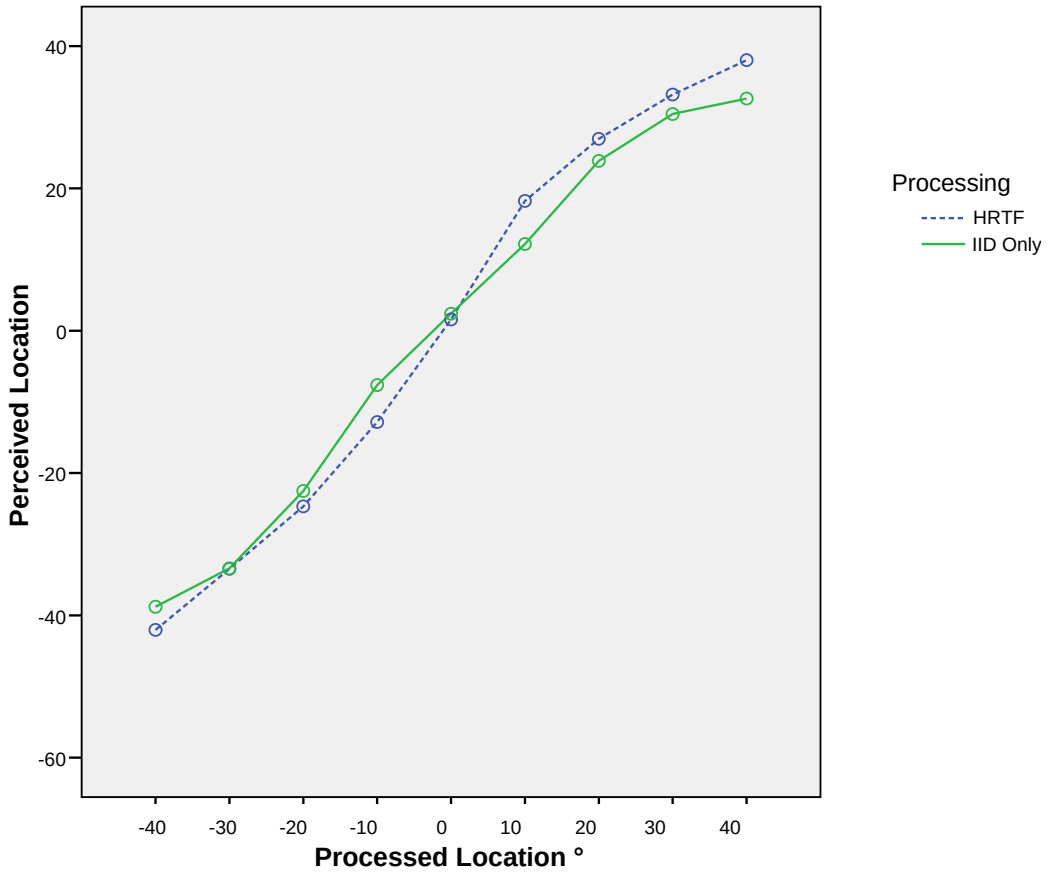


Figure 4: Mean perceived location as a function of processed location for each processing method (full Head Related Transfer Function and Interaural Intensity Difference only).

Linear regression analysis was performed in order to test the simple effects of the location X processing interaction, and the results showed that the perceived location corresponded well with the processed location across all subjects and conditions. Separate regressions were completed for the two processing methods. For the HRTF method, the regression coefficient of 1.091 with an intercept of 0.554 shows that on average, perception of the location of sounds processed using the HRTF method was nearly veridical (Slope=1, Intercept=0). For the IID only processing method, the regression coefficient of 0.983 with an intercept of -0.087 also shows near-veridical

perception of location. The regression analysis showed that the spatial processing manipulation accounted for 83% and 81% of the variance of the perceived location data for the HRTF and IID processing methods respectively. Figure 5 and Figure 6 below show scatterplots of the perceived location data for the HRTF and IID only processing methods respectively for the ±45° speaker configuration.

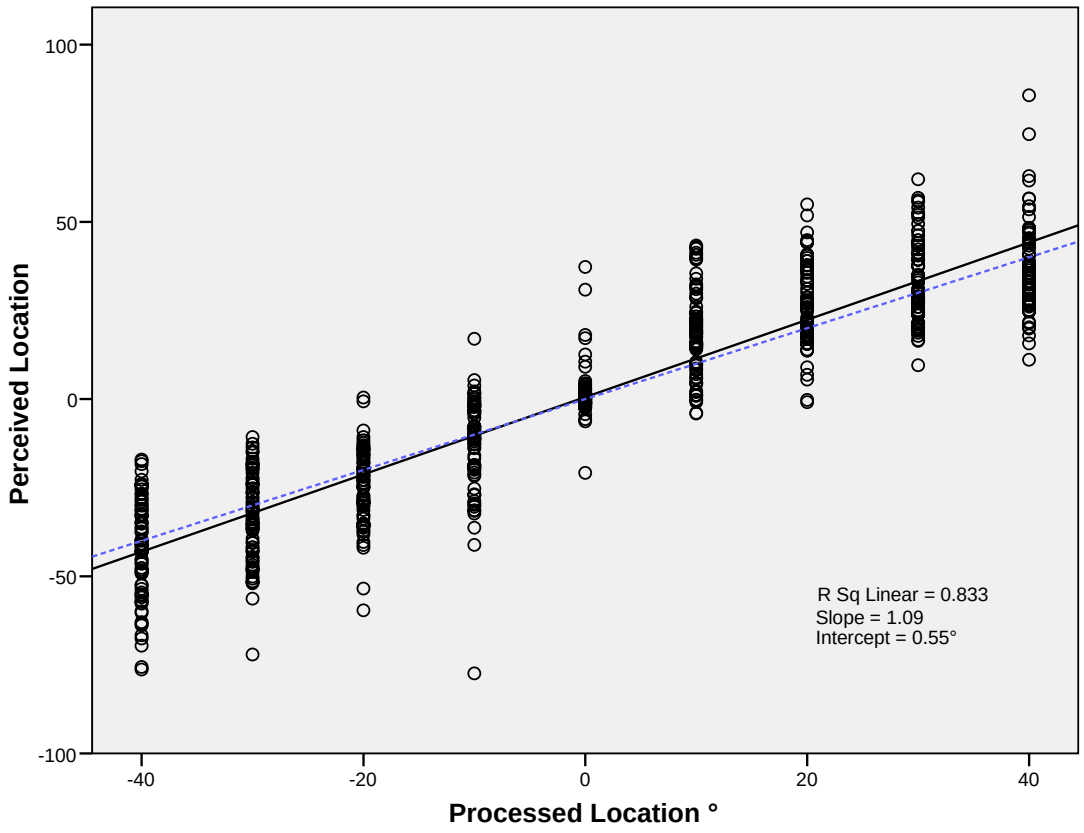


Figure 5: Perceived location of the sound source as a function of the processed location for the HRTF processing technique in the ±45° speaker configuration. The black line represents a regression line from the experimental data, and the blue line represents veridical perception of location.

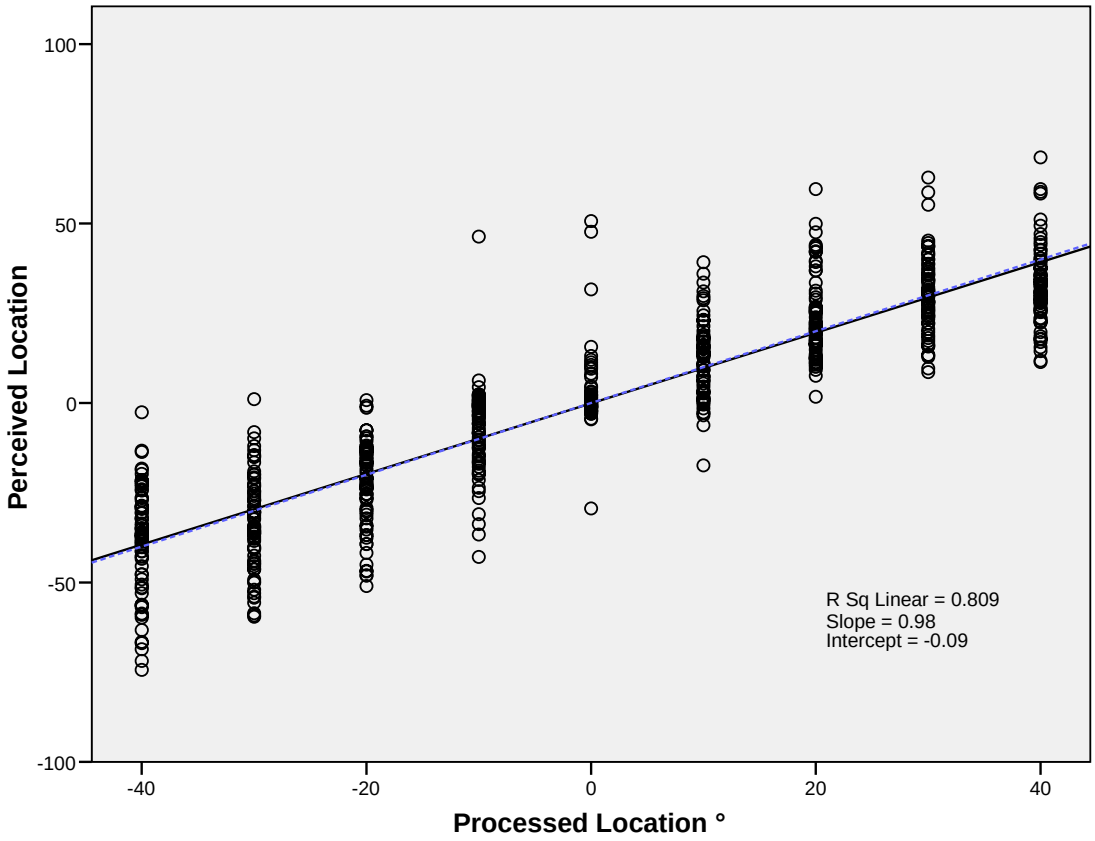


Figure 6: Perceived location of the sound source as a function of the processed location for the IID processing technique in the ±45° speaker configuration. The black line represents a regression line from the experimental data, and the blue line represents veridical perception of location.

The response time data did not show the expected faster response times for the full HRTF processing method $F(1, 35)=0.335$, $p=0.566$. Although this analysis is qualified by the significant Clip X Processing Method interaction ($F(1, 35)=4.503$, $p=0.041$), this interaction only shows a smaller effect size ($\eta^2=0.03$ for the jazz clip and $\eta^2<0.001$ for the rock clip) for processing for the rock clip as compared to the jazz clip, but though the effect size differs, the simple effect of processing method on response

times is not significant for either clip alone ($F(1, 35)=1.088$, $p>0.05$ for the jazz clip and $F(1,35)=0.016$, $p>0.05$ for the rock clip).

Experiment 1 - ±90° Speaker configuration

The same 40 participants from the ±45° configuration completed the experiment with speakers at ±90°. Of these, 36 met the screening criteria for inclusion in the experiment; however, 10 of those participants gave invalid data for this speaker configuration. Of those, eight gave invalid data almost solely in the IID only processing condition. An ANOVA was completed for the response angles for the 26 participants that gave no invalid responses. As expected, the main effect of processing technique was significant ($F(1, 25)=22.802$, $P<0.001$) such that the perceived locations were further to the right than they were processed to be by approximately 6º when using the HRTF processing method. As expected, there was also a significant main effect of location ($F(2.49, 62.36)=213.469$, $p<0.001$) as well as a processing X location interaction ($F(2.971, 74.295)=3.882$, $p=0.013$). As a result of the significant interaction, simple effects tests examined the effects of location within each of the processing techniques.

To examine the simple effects of location within each processing method, a regression analysis was completed for each processing method. The regression coefficient for the HRTF processing method was 1.787 with an intercept of 13.561°. The regression coefficient for the IID only processing method was 2.143 with an intercept of 7.297°. The processed location explained 72.5 % of the variance with the HRTF method and 77.9% with the IID only method; both slopes were significantly different from zero; ($t(466)=35.014$, $p<0.001$ and $t(466)=40.529$, $p<0.001$ for HRTF and IID processing

methods respectively). However, as discussed above, this analysis excludes a disproportionate amount of data from participants that gave invalid data for the IID processing method while giving valid data for the HRTF processing method. Figure 7 and Figure 8 below show the regression lines produced from the data in the HRTF and IID only processing methods respectively.

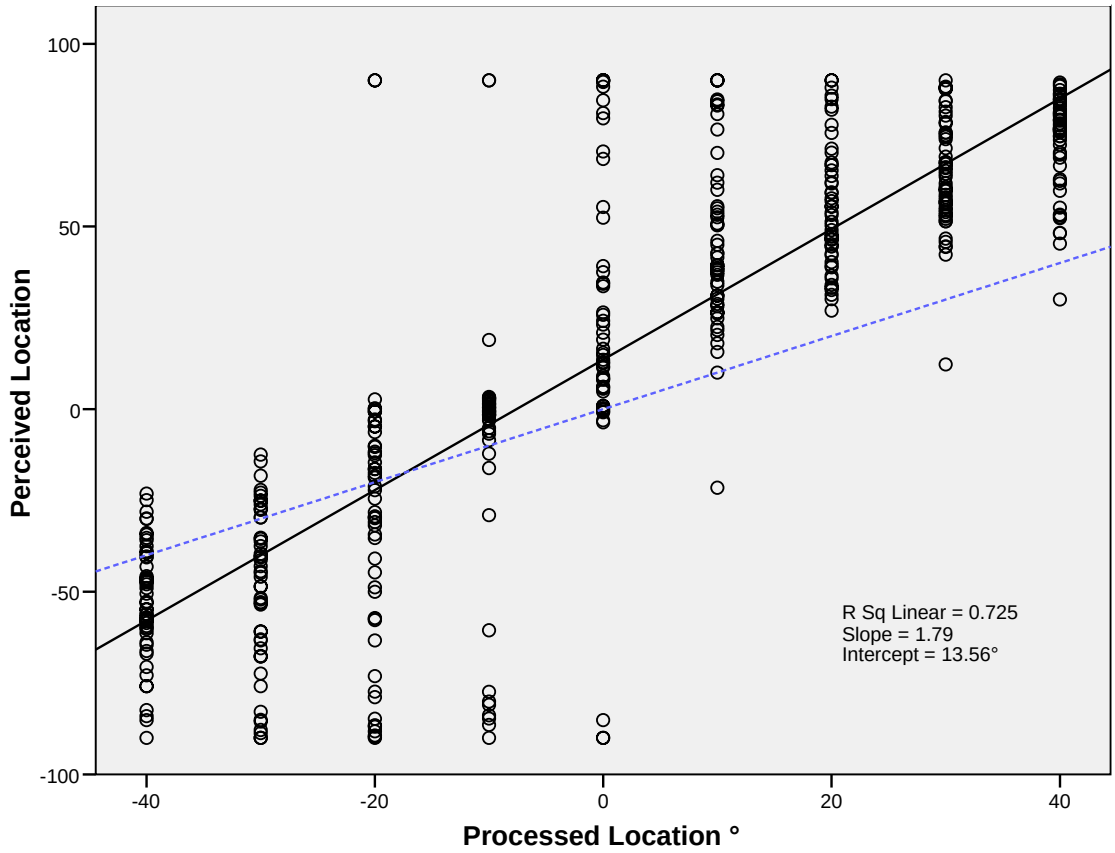


Figure 7: Perceived location of the sound source as a function of the processed location for the HRTF processing technique in the ±90° speaker configuration. The black line represents a regression line from the experimental data, and the dashed blue line represents veridical perception of location.

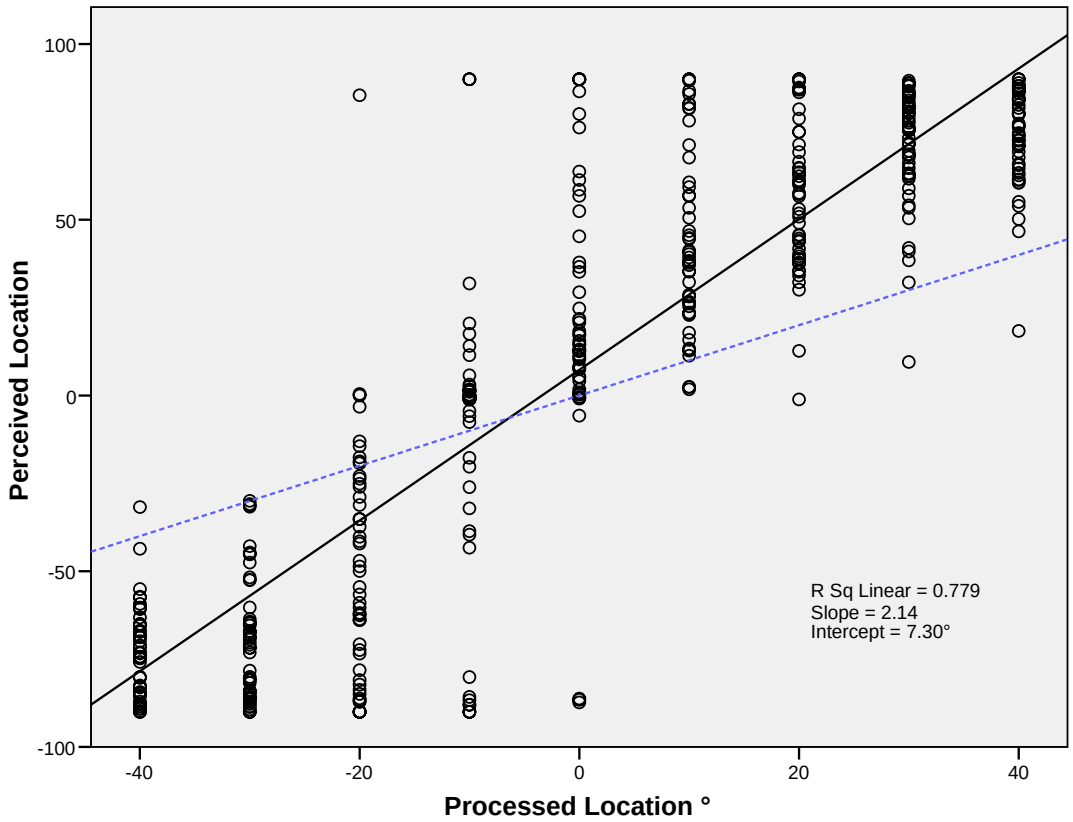


Figure 8: Perceived location of the sound source as a function of the processed location for the IID processing technique in the ±90° speaker configuration. The black line represents a regression line from the experimental data, and the blue line represents veridical perception of location.

Experiment 2 - ±45° Speaker configuration

There was no main effect of volume on localization responses, $F(2, 70)=0.054$, $p>0.05$. Although these data are inconsistent with the predicted increase in accuracy for louder stimuli, this result confirms that spatialized audio stimuli can be localized throughout the volume range from 55 dBA to 65 dBA. There was a significant effect of processed location, $F(1.856, 64.961)=536.183$, $p<0.001$. Post-hoc paired comparisons (LSD with $\alpha=0.05$) showed significant differences in perceived location at each of the

nine different locations as compared to each adjacent location. In addition, there was a significant clip X location interaction (see Figure 9), $F(3.905, 136.666)=3.715$, $p=0.007$. Separate regression analyses were conducted with processed location predicting perceived location for the rock and jazz clips. For the rock clip, the regression coefficient was 1.11 ($t(970)=71.5$, $p<0.001$) with an intercept of 1.175°. For the jazz clip, the regression coefficient was 1.103 ($t(970)=68.564$, $p<0.001$)with an intercept of 0.951.

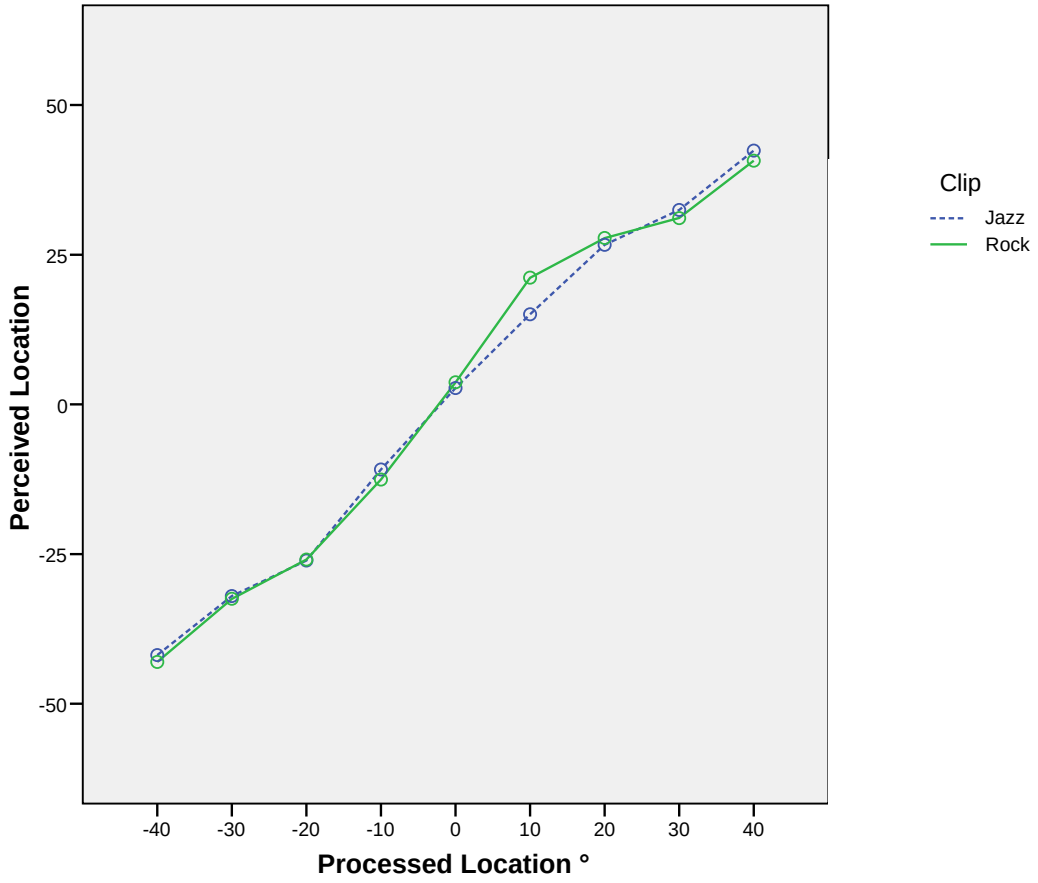


Figure 9: Mean perceived location as a function of processed location for the two musical clips (jazz and rock)

Linear regression analysis showed that the perceived locations corresponded quite well with the processed locations. Given the ANOVA results suggesting that there were no significant differences in perceived location between the volume levels, a single regression analysis was completed on the combined data from the two volume levels. The regression coefficient of 1.107 with an intercept of 1.063 shows that on average, the localization of sounds at all volume levels was quite accurate. The regression analysis showed that the volume manipulation accounted for 83.5% of the variance of the data. Figure 10 below shows a scatter plot of all data for the ±45° speaker configuration.

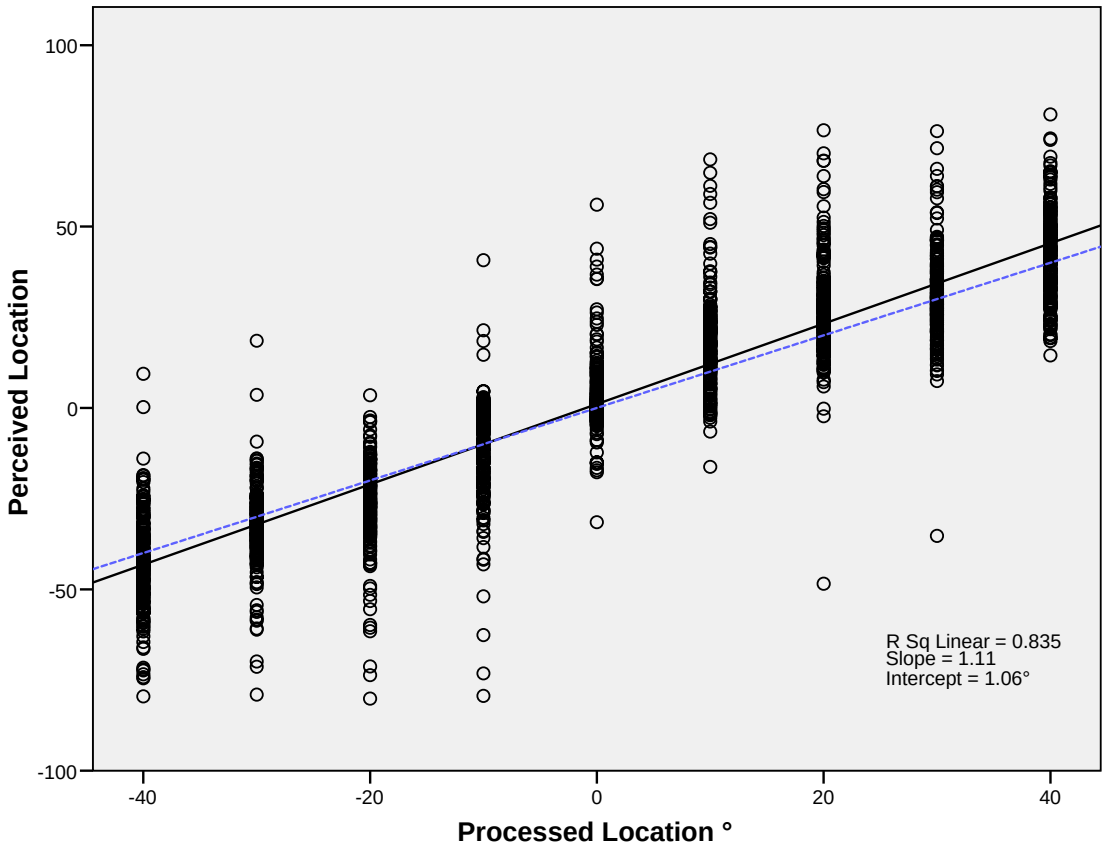


Figure 10: Perceived angle as a function of processed location for all participants and all volume levels in the ±45° speaker configuration. The black line represents a regression line from the experimental data, and the blue line represents a theoretical veridical perception.

As expected, the response time data did not show any significant differences for the volume manipulation, $F(2, 70)=0.551$, $p=0.579$; however, the response time for the rock clip (M=2879 ms) was significantly higher than that of the jazz clip (M=2716 ms), $F(1, 35)=6.334$, $p=0.017$. There were no significant interactions between the independent variables for the response time dependent variable ($p>0.05$).

Experiment 2 - ±90° Speaker configuration

The same 38 participants from the ±45° configuration completed the experiment with speakers at ±90°. Of these, 36 met the screening criteria for inclusion in the experiment; however, six of those participants gave invalid data for this speaker configuration. No trend was identified that would suggest that the volume level manipulation resulted in differences in the number of invalid responses.

An ANOVA was completed for the response angles for the 30 participants that gave valid data for all trials. The results showed a main effect of location ($F(1.984, 57.531)=371.208$, $p<0.001$). The data showed no effect of volume ($F(1.812, 52.555)=0.158$, $p=0.854$) or clip ($F(1, 29)=0.359$, $p=0.554$. In addition, the data showed no significant interaction effects ($p>0.05$).

Post-hoc follow up tests for the significant location effect showed significant differences between all pairs of adjacent spatial locations (eg -40° and -30°); however, the mean response at all locations was biased towards the right. This suggests that the ±90° configuration was not as close to veridical perception as was the ±45° configuration. A linear regression analysis produced a regression coefficient of 1.539 with an intercept of 9.463°. This is similar to the results from the HRTF processing method in experiment 1 which used the same stimuli and volume level (regression coefficient of 1.787 with an intercept of 13.561°). Figure 11 below shows all valid data from the ±90° speaker configuration along with the regression line fit to the data.

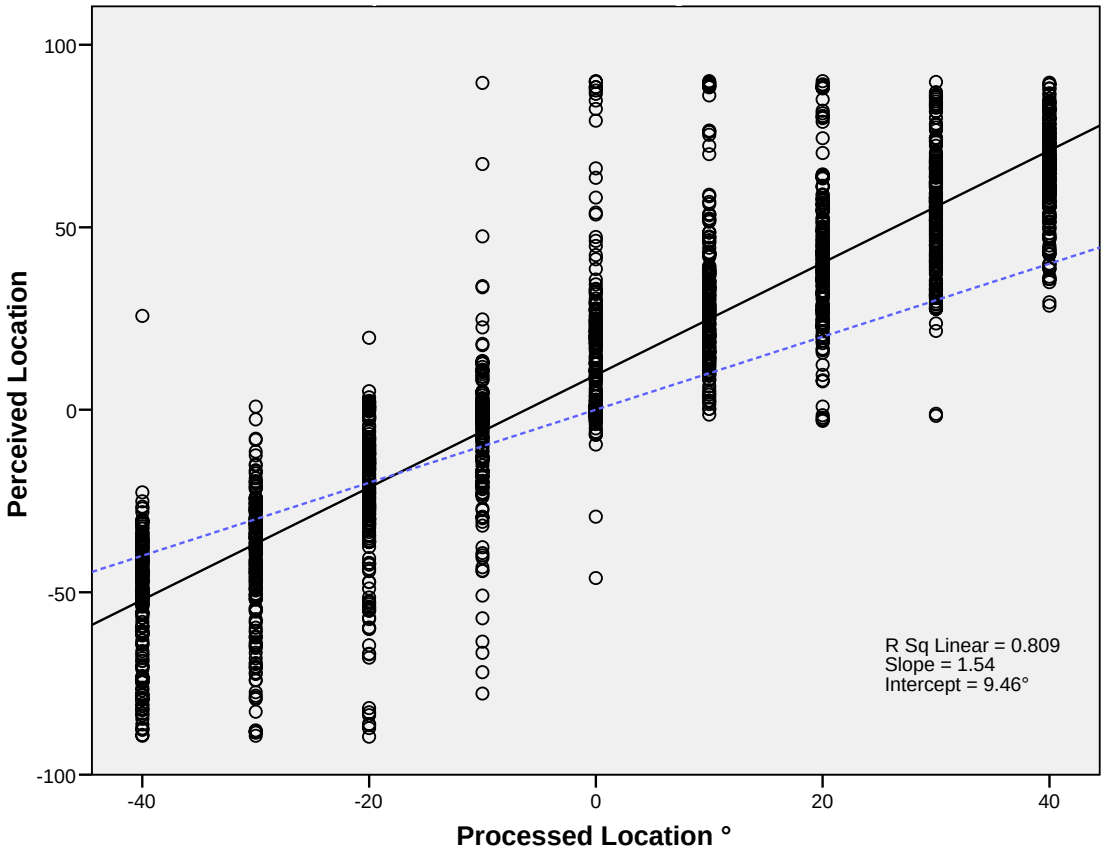


Figure 11: Perceived location of the sound source as a function of the processed location for all volume levels in the ±90° speaker configuration. The black line represents a regression line from the experimental data, and the blue line represents veridical perception of location.

DISCUSSION

This study was designed to show that a spatial audio display utilizing musical stimuli could be designed and used to portray directional information in a system such as a curve awareness display. Two experiments tested the effectiveness of localization using two different processing methods (i.e., Experiment 1, which utilized either a complete head-related transfer function or only interaural intensity differences) as well as a range of volume levels (i.e., Experiment 2, which used volumes ranging from 55 dBA to 65 dBA) in order to guide the design of a display that is not overly complex, but remains effective in the design context. Multiple speaker configurations (located at ±45º and ±90º from the saggital plane of the participant) were investigated in order to identify appropriate speaker configurations for such a display. In both experiments, participants listened to a short audio clip and then reported the direction from which the clip appeared to originate by clicking on the perceived location of the musical clip on a protractor displayed on a tablet PC screen.

It is important to note that the most important conclusion to be made concerning the data from the ±90° speaker configuration of both experiments is that it resulted in numerous invalid responses that were most likely caused by front/back confusions. As noted above, ten participants in Experiment 1 (28%) gave at least one invalid response at some point during the trials using the ±90° configuration. Given the fact that none of the participants gave invalid responses for the ±45° speaker configuration, it can be assumed that the ±90° speaker configuration resulted in an unexpected perception of location (most likely behind the participant since invalid responses occurred behind the head in

the experimental apparatus). The increased number of invalid responses seen in the IID only condition of Experiment 1 as compared to the HRTF condition suggests that the major advantage of the more powerful processing methodology for the ±90° speaker configuration is a reduction in front/back confusions. Experiment 2 showed a similar pattern, with six participants (17%) giving invalid responses at some point during the ±90º speaker configuration.

The finding from Experiment 1 that the two processing techniques did not produce meaningful differences in sound localization is promising, as it supports the use of simpler spatial displays that rely on only the IID cue to location; however, it must be noted that this experiment did not use individualized HRTF processing, and therefore did not take full advantage of the possibilities of using the HRTF processing method. Still, however, while it is possible to use semi-individualized HRTF processing by testing an individual and selecting and/or tweaking the closest matching HRTF from a database (Tan and Gan, 1998), the use of individualized HRTF's seems unnecessary and overly complex for a display that would be effective for enhancing curve awareness in a driving task. A system making use of IID only processing instead of utilizing a complete HRTF would be less complicated and less expensive to produce, implement, and adjust for the variability in speaker configurations in different vehicle models. However, in more demanding localization tasks, as well as tasks that make use of speaker configurations located on the axis of the perceiver's ears (such as headphones), the use of individualized HRTF's should be investigated in order to minimize or eliminate front/back confusions.

It had been expected that the processing methods would influence response times, since there are multiple cues that are in agreement in the case of the HRTF display, whereas the interaural time difference cue will be at odds with the intensity difference cue when using the IID-only processing method. The fact that no processing-related response time effect was found in either experiment suggests that there would not be a significant performance decrement when utilizing a display that makes use of only IID directional information as compared to a display using a full HRTF. Although this is promising for simplistic display designs, it must be noted that this experiment did not use HRTFs that were customized for each individual participant, and that there was large variability in the response time data that could make it difficult to identify a relatively small response time effect (M=3187.4 ms and SD=2041.4 ms for HRTF processing and M=3297.5 ms, SD=2353.1 for IID only processing). Future experiments should refine the data collection procedure such that the participant is required to start each trial with the stylus located at the focal point of the protractor in order to force the participant to move the stylus the same distance for each trial. This may minimize the noise in the response time data associated with different distances moved while responding.

As was expected from the pilot study, the results of this experiment suggest that the ±45° speaker configuration was more effective at influencing perceived location. The ±45º condition produced regression coefficients representing near veridical perception (ranging from 0.98 to 1.11) with intercepts near zero degrees (ranging from 0.09º to 1.06º) ; whereas the ±90º speaker condition produced regression coefficients ranging from 1.54 to 2.14 and intercepts ranging from 7.3º to 13.56º. In addition, the ±90°

speaker configuration resulted in a large number of invalid responses that probably represent perceived locations far from those desired based on the audio processing. After examining the data suggesting that the ±90° speaker configuration produced perceptions that were biased to the right (in absolute terms as well as relative to the ±45° speaker configuration), the experimental apparatus was checked to confirm the location of the participant relative to the speakers. It was found that at some point in the experiment, the table had shifted approximately 1 inch to the right of dead center. This would result in an imposed interaural time difference of approximately 15 μS (right leading) as well as a negligible interaural intensity difference due to the difference in path loss. Both of these changes would be expected to produce a rightward shift in perceived location as was seen in this experiment. This highlights another advantage of the ±45° speaker configuration the ±45° speaker configuration which is more robust to small changes in the location of the listener (from left to right) as a leftward or rightward shift in listener location in the ±90° configuration changes the distance to each speaker more dramatically than occurs in the ±45° configuration.

The results from Experiment 2 indicate that the effectiveness of spatial audio displays with musical stimuli is consistent across a range of volume levels. No significant differences either in perceived location or response times were identified at the three volume levels tested (55, 60, and 65 dBA). The lack of a response time effect suggests that there is no tradeoff for allowing display users to control the volume of the display even if it is important for users to respond quickly to changes in the display output. Extremely high or extremely low volume levels might result in distraction or

ambiguity and therefore increased response times that were not identified in the volume range tested, but these volume levels are not expected to be practical for a curve awareness display due to the fact that extremely high volumes could result in hearing loss and extremely low volumes would be barely audible in a vehicle due to vehicular noise. The effectiveness of localization throughout a range of volume levels provides additional support for the potential to implement an effective curve awareness display that makes use of the auditory stimuli (music) that many drivers already enjoy in their vehicles. Although the significant response time effect for musical clip seen in Experiment 2 (jazz vs rock) is somewhat troubling in that it suggests some music may be more effective than others, the difference in response time is only 163 ms (approximately 6% of total response time), and this effect was not observed in Experiment 1 using the same musical clips (the trend was smaller, but in the same direction). This suggests that future research should investigate more audio clips to establish any potential advantages for one type of stimuli over another.

CONCLUSION

The results of these experiments confirm that a simple spatial audio display with musical stimuli can be a feasible option for a curve awareness display as well as other displays that manipulate complex auditory stimuli to present directional information to users without taxing their visual resources. While such displays deserve further attention from researchers, the results of the present study suggest that a simple processing method using only interaural intensity difference processing methods at a volume range of 55-65 dBA could be successfully implemented. If possible, a speaker configuration with speakers mounted at ±45º is preferable for such displays as this configuration produced near veridical perception of location. Speakers mounted on the axis of users' ears should be avoided as this configuration tends to produce front/back confusions. If such a configuration must be used, further investigation of individualized HRTFs would be warranted.

One key example of a spatial audio display using musical stimuli would be presenting drivers with an auditory display that specifies the direction and intensity of an upcoming curve. The results of these experiments indicate that the source of the auditory information need not be simple (predictable) tones, but could instead be the music that the driver had already chosen through his or her in-vehicle entertainment system. Although the use of a full HRTF to process the spatial audio cues could be implemented, the present data suggest that a design utilizing only the interaural intensity difference cue may be effective, simpler, and more cost effective.

Further testing of spatial audio with musical stimuli could still enhance our understanding of the capabilities and limitations of such systems. However, for the purpose of designing an in-vehicle curve awareness display, these experiments have shown that a ±45° speaker configuration using IID location cues should provide adequate localization ability. Substantial research on these displays would be required in order to establish the effectiveness and desirability of such a display as compared to other options for enhancing curve awareness. This research should focus on testing steering performance around curves as well as increasing driver awareness of other roadside hazards. In addition, establishing appropriate onset times/distances and angular deviations for curves of varying radii would be necessary.

Although this research suggests potential applications for simple spatial audio displays with musical stimuli, more advanced display systems that would require the processing of cues to appear at different elevations or behind the participant would likely require the use of individualized HRTFs. Further testing should also examine methods of choosing from pre-measured HRTFs in order to enhance the ability to locate sounds throughout space. The use of individualized or matched HRTF measurements could potentially allow for the use of headphones which was determined to be relatively ineffective in the pilot study using non-individualized HRTF processing. The ability to localize effectively in headphones and process sounds through a full 360° range as well as potentially including elevation information would open up a number of options for more portable applications of music based spatial displays.

APPENDICES

APPENDIX A: PILOT RESEARCH DETAILS

INTRODUCTION

In order to assess the feasibility of spatial audio displays using musical stimuli in environments such as motor vehicles, a study was designed to test how effectively humans could discriminate between different simulated spatial locations produced using different speaker configurations similar to those that may be found in motor vehicles. This study shows that the concept of a spatial audio display with musical stimuli can be expected to effectively portray information in this context. It also provides evidence to support that the device will operate consistently across a number of different musical stimulus styles.

METHODS

Participants

23 Participants were selected from the Clemson University Psychology Department subject pool. Participants were university underclassmen enrolled in an introductory psychology course. Participants received class credit for their participation

Apparatus

Spatially processed audio clips were presented to participants using an IBM X41 Tablet PC running Windows XP Tablet PC Edition and the E-Prime experiment operating system version 1.1 (SP3). Four different stimuli were presented at 9 different spatially processed locations in the horizontal plane corresponding to -40°, -30°, -20°, -10°, 0°,

10°, 20°, 30°, and 40°. The four stimuli conditions included 2 musical stimuli (1 Pop/Rock and 1 Jazz Clip), 1 talk radio clip (www.redbarradio.com) and white noise (20 Hz-20 kHz). The stimuli were processed using Sony Media Software Sound Forge and the Wave Arts Panorama DirectX plugin. All 3 musical and talk radio clips were mixed (left + right channel) to produce a monaural stimulus prior to spatial processing. The white noise clip was generated by Sound Forge as a monaural stimulus. The MIT KEMAR head related transfer function was used for all conditions, and each clip was normalized after spatial processing to ensure consistent volume levels.

All audio clips were presented to each participant at approximately 55 dB (A weighted) in an anechoic chamber using 3 different speaker configurations (headphones, speakers at ±45 degree angles facing the participant, and speakers at ±90 degree angles facing directly at the participants ears). Sony MDR7506 professional headphones were used for the headphone conditions, and Altec Lansing BX2 computer speakers were used for the headphone conditions. The speakers were mounted at positions 27 inches from the center of a chinrest that was used to minimize participant head movements and maintain proper head positioning within the experimental apparatus. Speaker locations were masked using a large black cloth that obscured approximately 180° of the participant's field of view in order to avoid participants reporting that the sound came from one speaker or another instead of truly reporting the perceived location of the sound.

EPrime experiment operating system software (Version 1.1 SP3) was used to collect all data including perceived angle and response time. The software was

programmed such that participants were presented with the auditory stimulus and a picture of a protractor (see Figure A.1). Participants were instructed to click (using the tablet stylus) on the perceived angular location of the musical clip relative to the head superimposed on the protractor. The position of the click on the screen was recorded and converted to an angle.

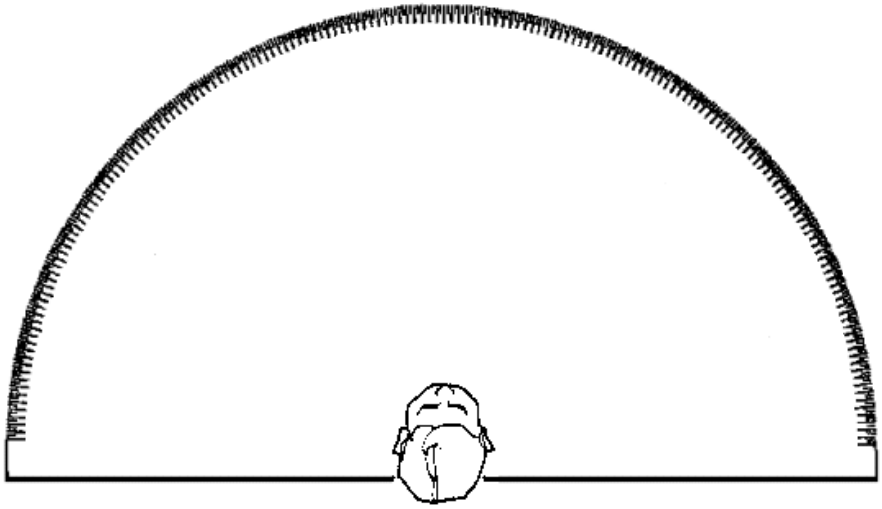


Figure A.1: Protractor used for reporting angular location of sound source

Design

The experimental design was a 9 X 4 X 3 within subjects factorial design. The independent variables included spatial processing, clip style, and speaker location. Each participant was exposed to all 108 combinations of levels of the independent variables and reported the perceived location of each sound as the dependent variable. There were 9 different spatial locations, ranging from -40° to +40° in 10° intervals (including 0°) from the saggital plane. Four different sound clips were used including 2 musical clips (soft rock and jazz), a talk radio clip, and a white noise clip. Each clip was 10 seconds

long. In addition to the factorial design analyzed using a Within Subjects ANOVA, a regression analysis was completed for each speaker configuration. This resulted in a regression coefficient that estimates perceived location based on processed location. Veridical perception of location would result in a regression coefficient of 1. In addition, regression analysis was performed on each individual participant's data to show that correct spatial location perception was not just a product of averaging across participants as well as to highlight patterns of individual differences seen in the data.

Procedure

After giving informed consent to participate in the experiment, participants were given instructions on how to complete the experiment using the experimental apparatus and positioned within the apparatus. Participants were given instructions to listen to each clip and report the spatial location of the sound on the protractor shown on the tablet PC's screen by clicking at that location on the graphic. Participants were required to click the screen using the stylus to proceed through the instructions, and the experimenter corrected any gross misuse of the stylus at this time.

After four practice trials (one of each clip at the same spatially processed location) and an opportunity to ask the experimenter any questions about the experimental apparatus, clips were presented in randomized order within each speaker configuration. All clips for a single speaker configuration were presented in random sequence, and then the next configuration was used until all three configurations were completed. After the software presented all clips at all locations for a given speaker configuration, the experimenter switched the speaker configurations manually resulting in

a slight buzzing sound as the speakers were removed and reconnected to the computer. This may have revealed the speaker locations in the ±45° and ±90° speaker configurations.

RESULTS

The response time data showed a significant main effect for speaker location order of presentation (most likely due to a learning or fatigue effect), F(5, 1781)=61.39, MSe=4416886.3, p<0.0005. Due to the order effect explaining more variance than processing angle and speaker location (partial eta^2 for order=0.15, partial eta^2 for processing angle=0.00, and partial eta^2 for speaker location =0.00), further analysis of the reaction time data would be misleading. In order to reduce the influence of this learning effect, further experiments should be carefully counterbalanced and allow more training time with the apparatus if useful response time data are to be collected.

Greenhouse-Geisser degrees of freedom adjusted F values were used for all ANOVA analyses. A 9 X 4 X 3 (Spatial Processing X Clip Type X Speaker Location) within subjects ANOVA on the data from perceived angles reported by participants identified significant main effects for Spatial Processing, F(3.47, 76.25)=642.03, MSe=1859.60, *p*=0.000, and Speaker Location, F(1.37, 30.11)=11.154, MSe=1559.83, *p*=0.001, as well as a Spatial Processing X Speaker Configuration interaction, F(4.78, 105.24)=25, MSe=2471.31, *p*=0.000.

The data showed no significant main effect for clip type (Rock/Pop, Jazz, Talk Radio, and White Noise), *F*(3,66)=0.09, MSe=599.63, *p*=0.965. The Spatial ProcessingXClip interaction was also non-significant, *F*(8.09, 177.96)=0.86,

MSe=1557.20, p=0.551. There was no significant 3 way interaction between Spatial Processing, Speaker Location, and Clip Type, F(9.23, 203.01)=0.65, MSe=2444.90, p=0.759.

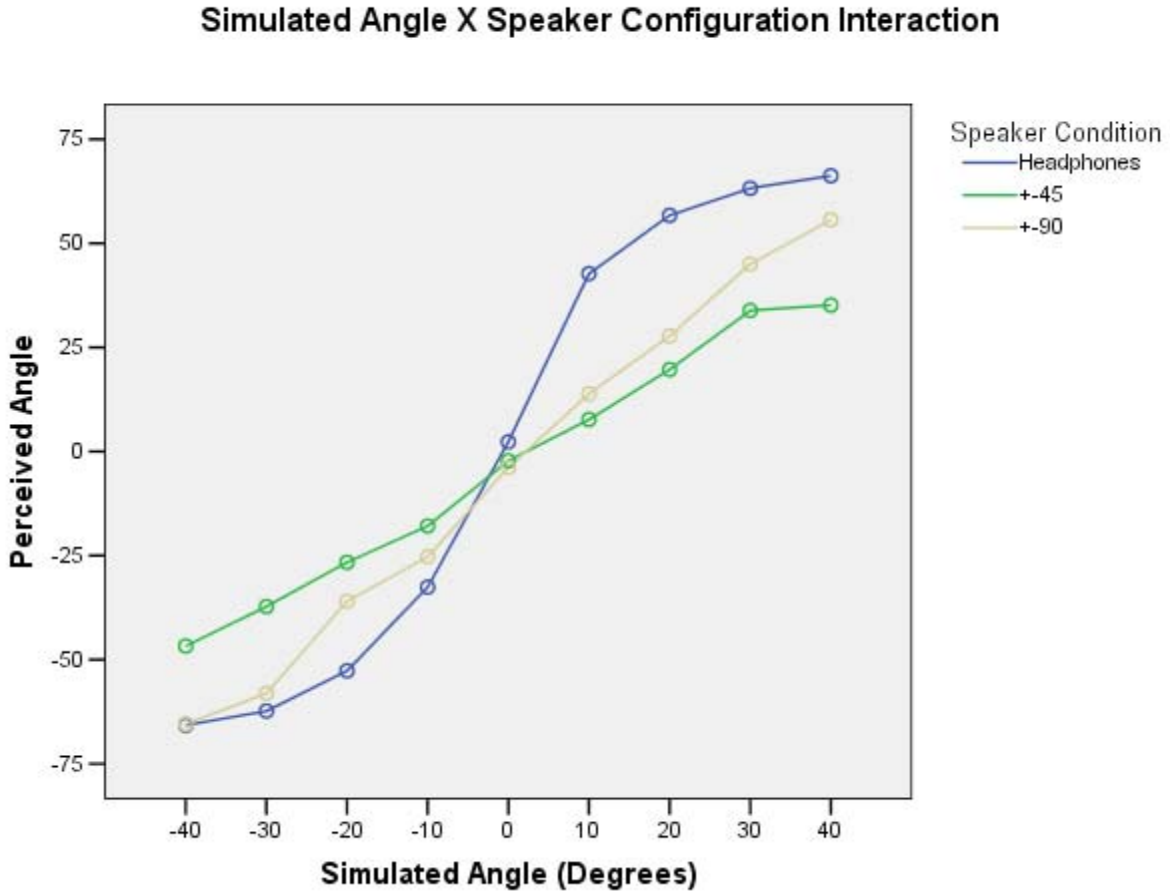


Figure A.2: Speaker Location X Spatial Processing Interaction. Notice that the perceived angles in the ±90° and Headphone conditions are larger at lower processed angles as compared to the ±45° speaker configuration.

The significant Spatial Processing X Speaker Location interaction (see Figure A.2) shows that using speakers located on the axis of the ears (especially headphones)

results in larger perceived deviations at lower processed deviations as compared to speakers at ±45° from the saggital plane (ex: -32.6° perceived corresponding to processing of -10 degrees using headphones as compared to -17.9° perceived corresponding to -10° using speakers located at ±45°). See table A.1 and figure A.3 for mean perceived locations for each spatially processed location. These larger perceived deviations, corresponding to changes in processing from -20° to 20°, did not seem to limit the ability to discriminate between processing levels as compared to the ±45° speaker configuration. In all speaker configurations, LSD post-hoc comparisons ($p<0.05$) show that at least 7 different spatially processed locations result in significantly different perceived locations as compared to the adjacent spatially processed location (see table A.2). In general, the -40° and -30° as well as the 30° and 40° processing conditions are either marginally significant or non-significant, resulting in only 7 instead of 9 locations that are perceived at significantly different locations. The paired comparisons for adjacent pairs of processed angle in the -30°-30° range are all significant at the $p<0.01$ level except in the ±90° speaker configuration where the -20° and -10° was marginally significant, $p=0.076$. As seen in figure A.3 and A.4, the variability in the ±90° speaker configuration was larger than that of the headphone and ±45° conditions. The ±90° condition resulted in a larger number of direction errors (perceiving a location to the left when processing dictates a direction to the right or vice versa), which could cause major problems in spatial displays attempting to portray directional information.

Headphones		±45° Speakers		±90° Speakers	
Processed Location	Perceived Location	Processed Location	Perceived Location	Processed Location	Perceived Location
-40°	-65.7°	-40°	-46.7°	-40°	-65.4°
-30°	-62.3°	-30°	-37.2°	-30°	-58.1°
-20°	-52.7°	-20°	-26.6°	-20°	-36.0°
-10°	-32.6°	-10°	-17.9°	-10°	-25.3°
0°	2.3°	0°	-2.2°	0°	-3.8°
10°	42.7°	10°	7.7°	10°	13.8°
20°	56.7°	20°	19.7°	20°	27.7°
30°	63.2°	30°	33.9°	30°	45.0°
40°	66.2°	40°	35.2°	40°	55.6°

Table A.1: Mean perceived location at each Spatial Processing Location for individual speaker configurations

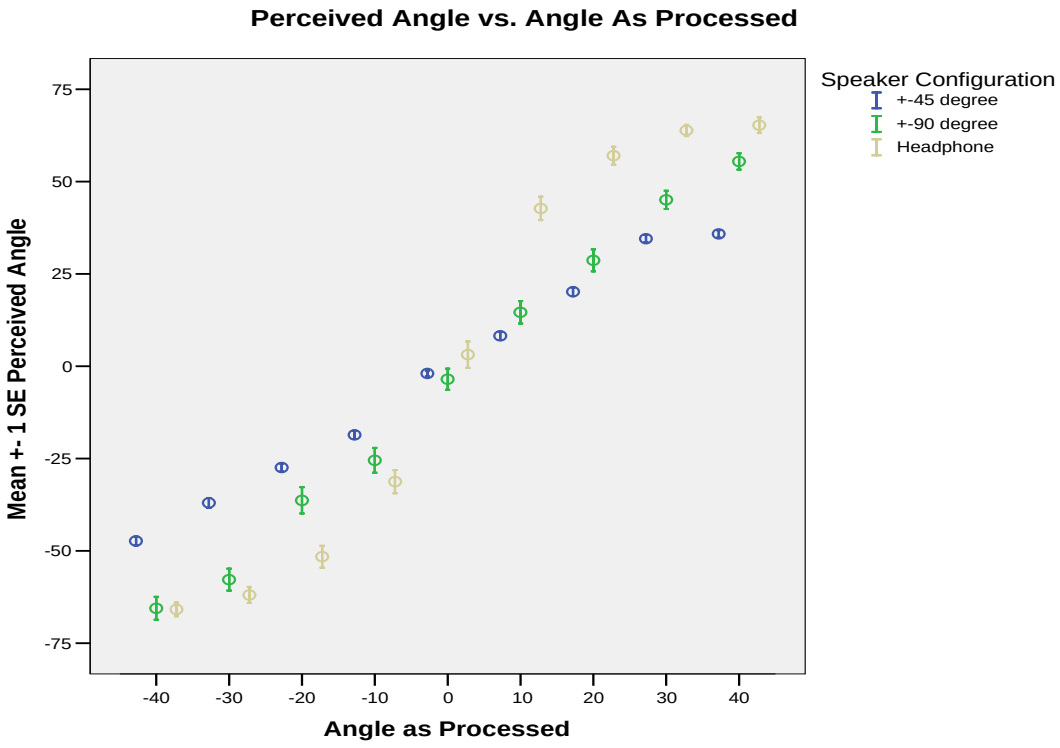


Figure A.3: Pilot Test data showing perceived angle of spatially processed sound in different speaker configurations

Table A.2: Mean Differences of Adjacent Spatial Processing Pairs			
Spatial Processing Pair	**Headphone**	**±45°**	**±90°**
-40/-30	-3.4	-9.480*	-7.344
-30/-20	-9.671*	-10.614*	-22.122*
-20/-10	-20.117*	-8.739*	-10.697**
-10/0	-34.862*	-15.653*	-21.440*
0/10	-40.417*	-9.915*	-17.675*
10/20	-13.993*	-12.005*	-13.896*
20/30	-6.508*	-14.185*	-17.262*
30/40	-3.015	-1.286	-10.647*

* Significant at $p < 0.05$, ** $p=0.076$

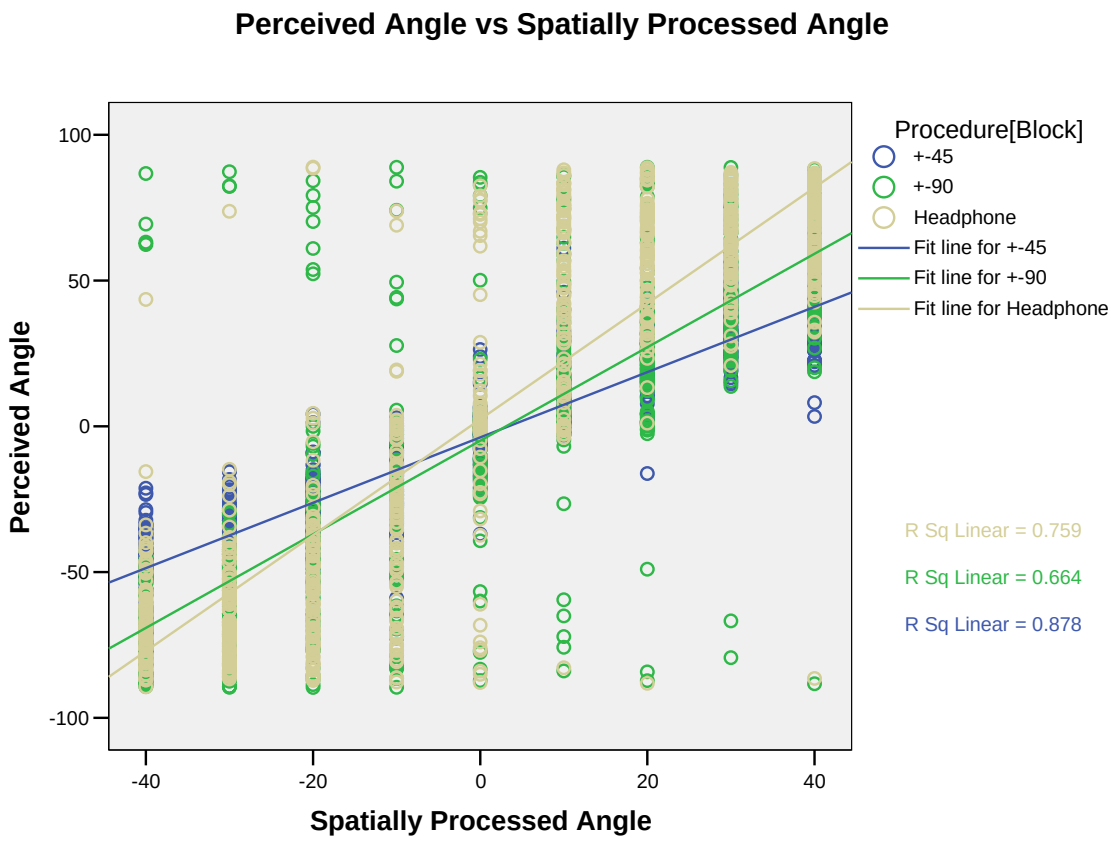


Figure A.4: Scatterplot of Perceived Angle vs Spatially Processed Angle with Regression Lines

A linear regression analysis was completed for each speaker configuration with spatially processed angle as the independent variable, and regression coefficients ranged from 1.12 for the ±45° speaker configuration to 1.61 for the ±90° condition to 1.99 for the headphone condition. This shows that the 45° angle speaker configuration produced the closest results to a regression coefficient of 1 corresponding to veridical perception of spatial locations (1 degree increase in processing resulting in 1 degree increase in perceived angle); however, this could be due to the fact that processing conditions ranged

across only 40 degrees of deviation from the center, and volume levels were relatively low which may have limited the effectiveness of spatial processing at large deviations and resulted in a perception of the location coming from the location of the higher volume speaker in the large deviation conditions (due to IID manipulation, one speaker is attenuated significantly at large deviations).

Separate linear regressions and scatter plots were also completed for each participant's data. This analysis highlights patterns of individual differences in perceiving spatial locations. The analysis revealed two fairly distinct patterns where there was a group of people who could perceive the location of the sounds very accurately (see figure A.5), and another group that could generally perceive that a sound was either left or right of center, but were not nearly as accurate in determining exact locations (see figure A.6). In general, the ±45° speaker configuration produced the most accurate results and least variability, and in many cases, the two speaker configurations were much more effective than the headphone configuration. Figures A.5, A.6,and A.7 show typical scatterplots with regression lines representing cases from each of these groups.

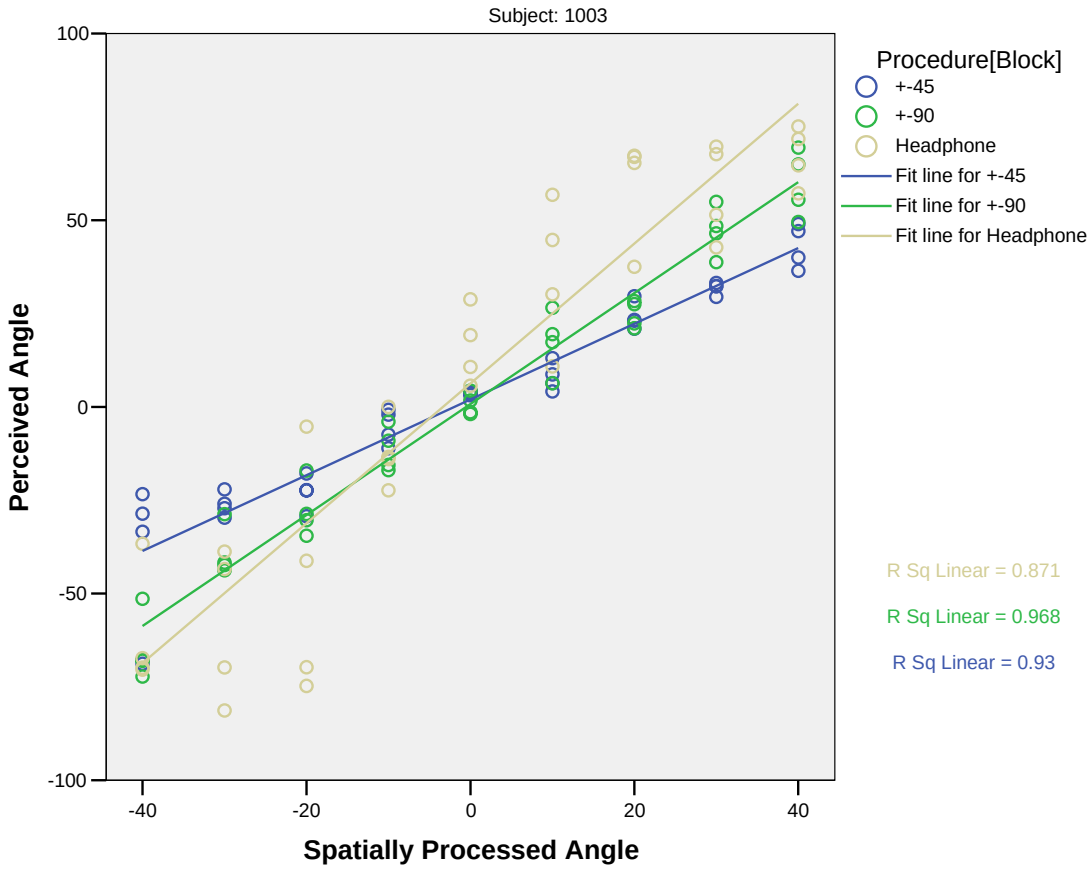


Figure A.5: Typical individual participant regression result for participants that discriminated locations as well as direction

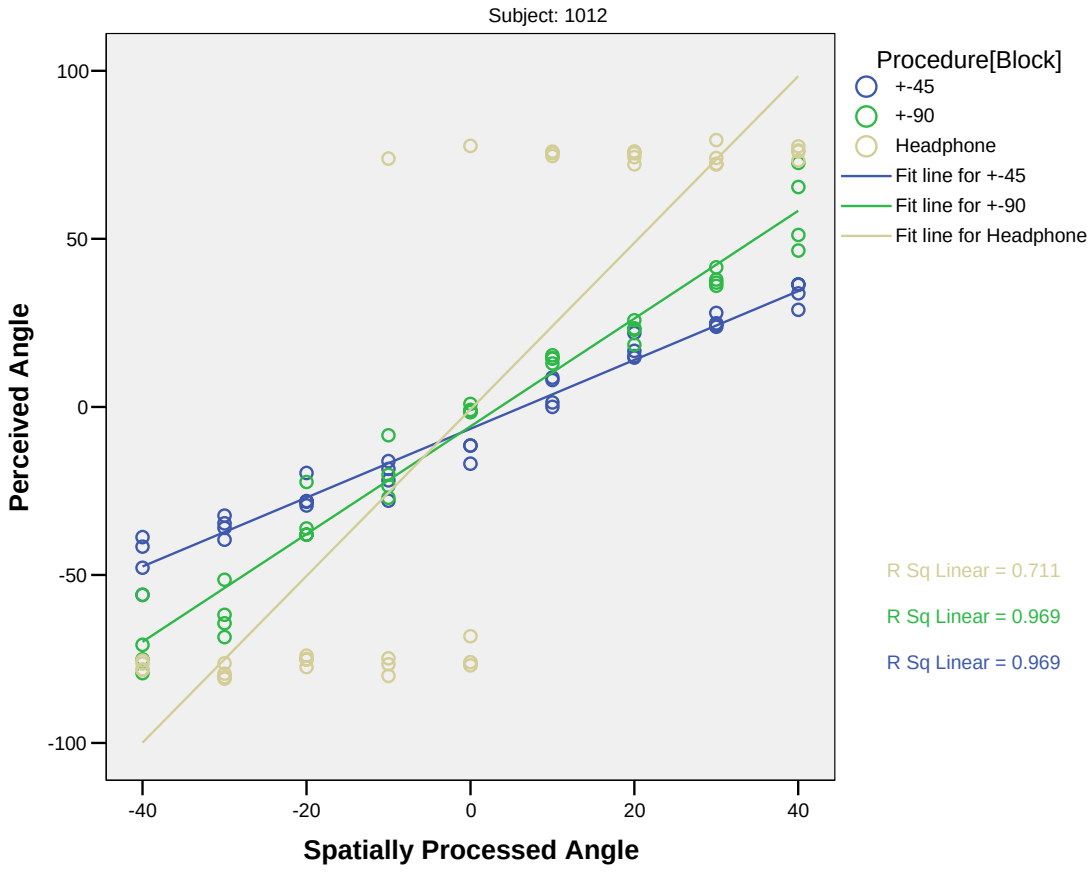


Figure A.6: Typical individual participant regression result for participants that discriminated locations as well as direction (left vs. right) for speaker configurations, but only direction for headphone condition.

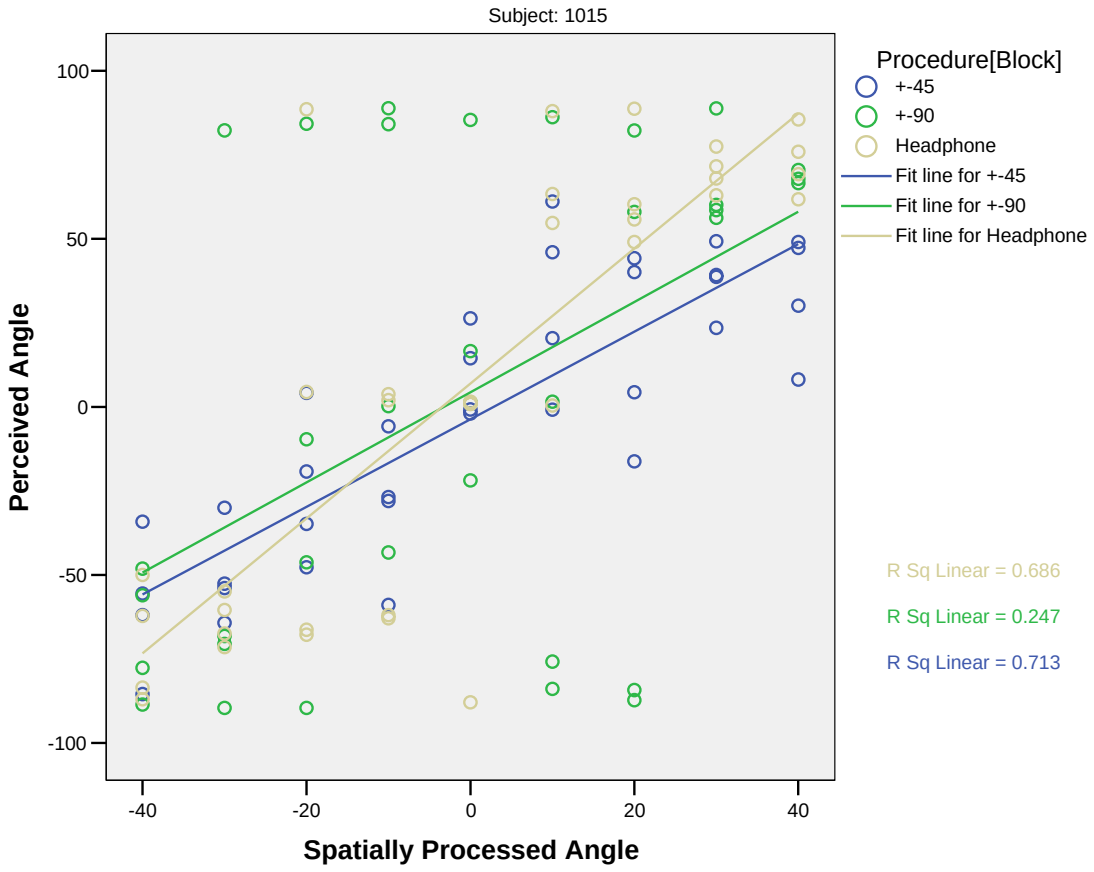


Figure A.7: Typical individual participant regression result for participants that discriminated most directions (left vs. right), but had trouble discriminating exact locations

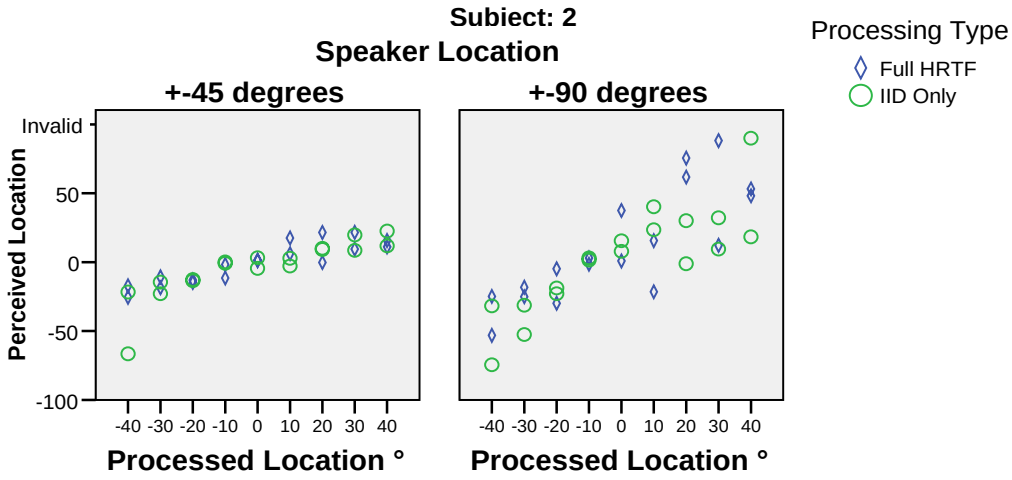
DISCUSSION

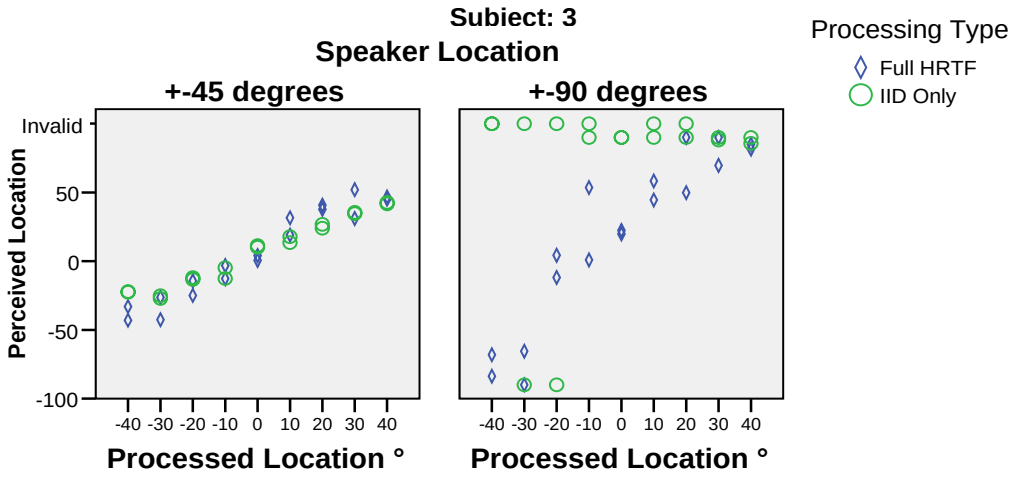
This experiment shows that the concept of using spatial audio with musical stimuli can be used to successfully display information. In general, the ±45° speaker configuration was most effective at producing accurate spatial perceptions, but all 3 speaker configurations resulted in perceptions of location that could, at a minimum, be used in a display to show an event to the left versus an event to the right. This suggests that such a spatial audio display could potentially be effective in a curve awareness

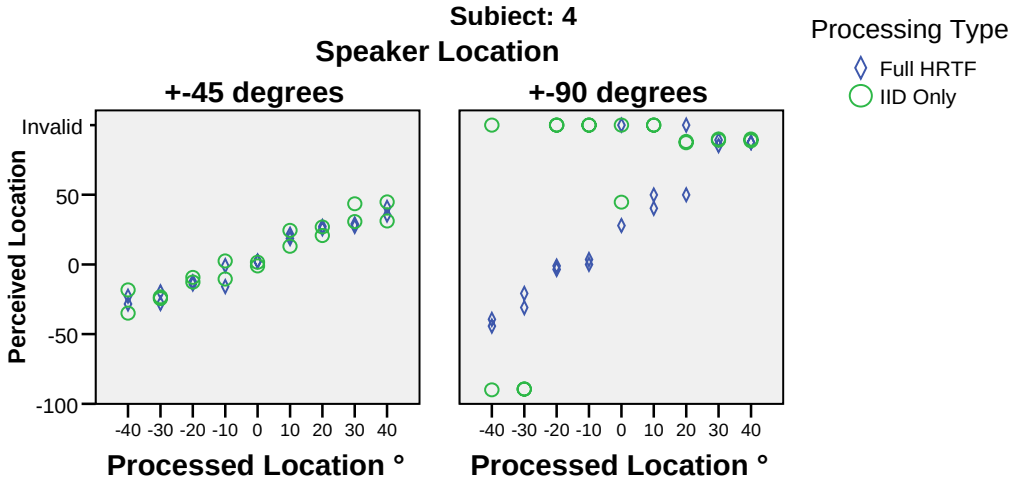
context, but careful speaker placement will allow the display to include more information as to the characteristics of an upcoming curve.

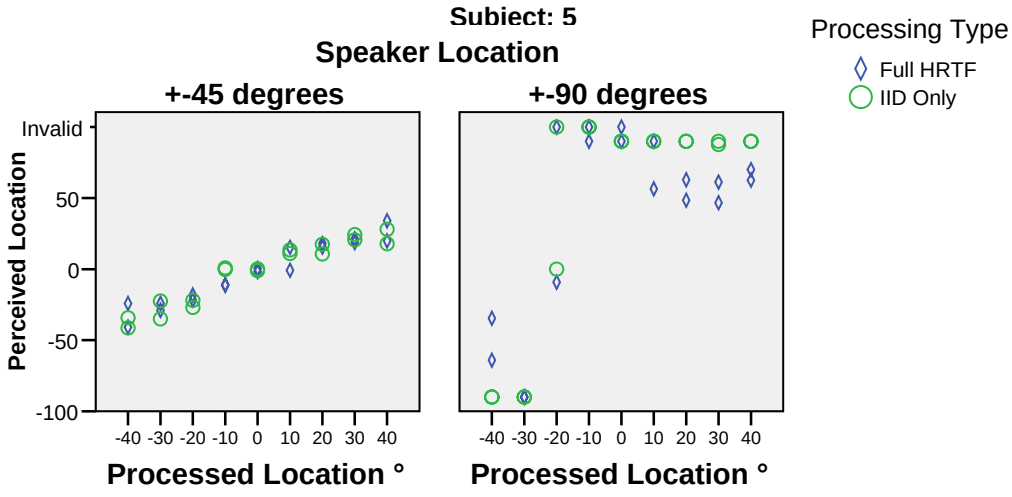
This experiment produced a number of interesting results that should be further investigated. First, there was a tendency to over-estimate the angle of sounds processed to be close to the saggital plane. This resulted in very large perceived deviations at small processed locations and ever decreasing deviations as processed angle increased. Further experiments should investigate why this occurred. In addition, a number of methodological flaws (such as speaker clicks revealing speaker locations during speaker configuration changes) can be corrected in future experiments in order to produce response time data that is not dominated by order and learning effects.

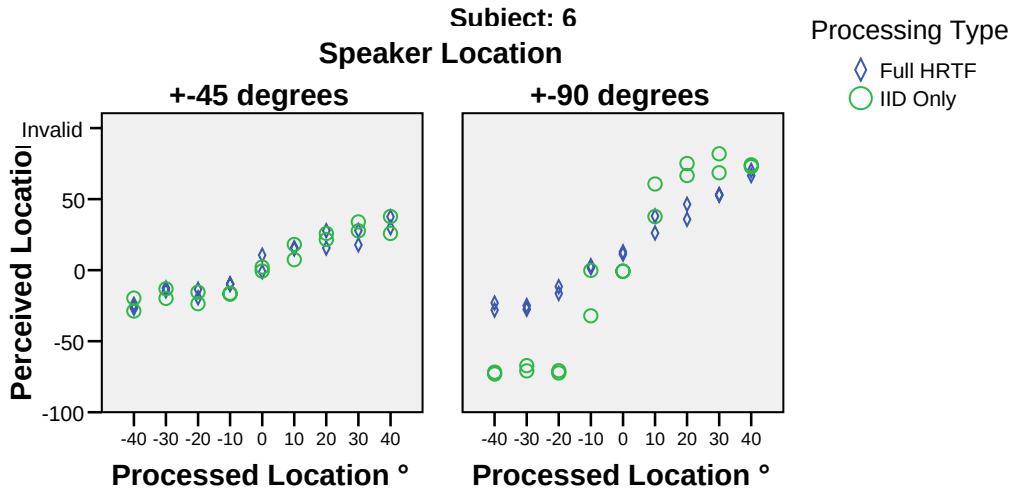
APPENDIX B: INDIVIDUAL PARTICIPANT GRAPHS

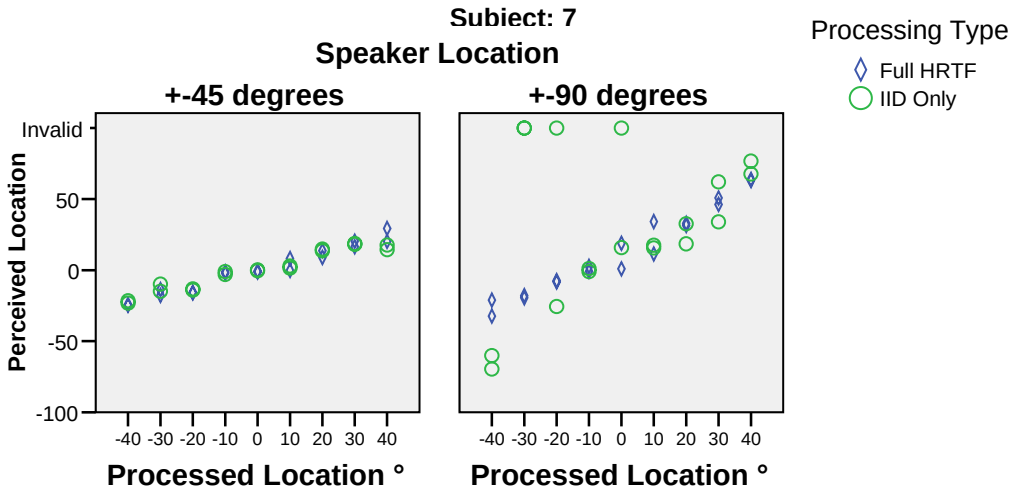


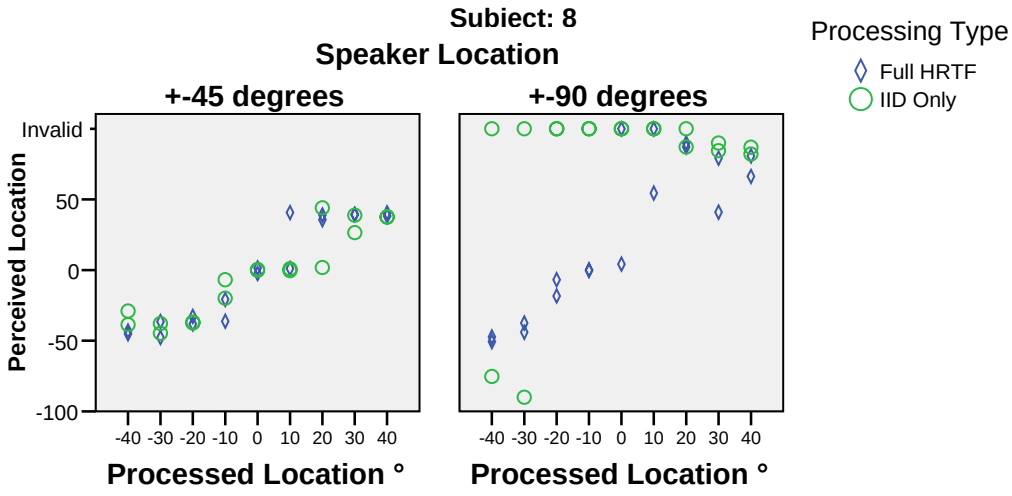


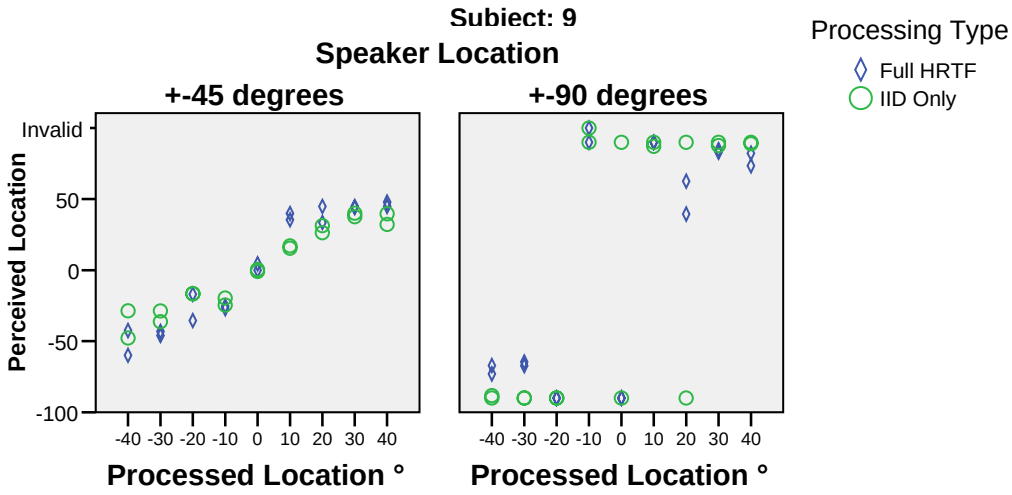


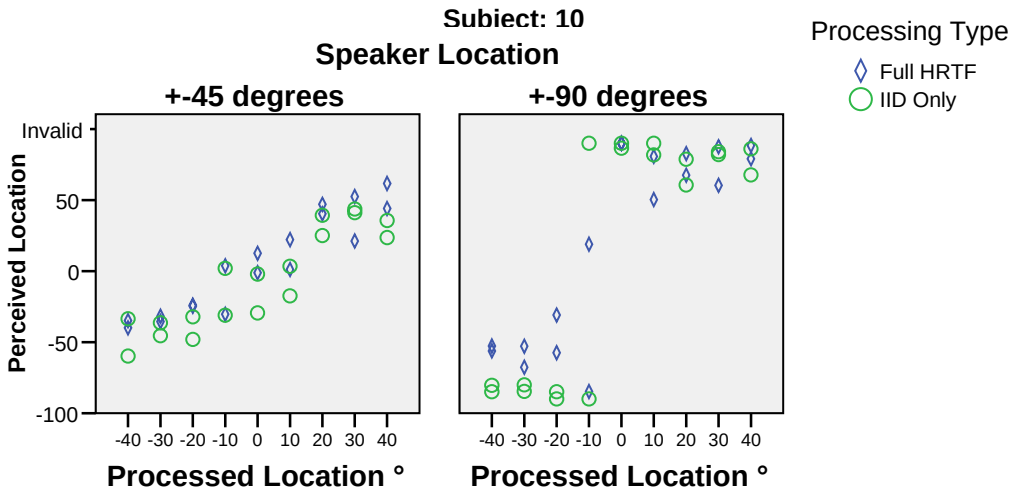

Subject: 5
Speaker Location
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Processed Location °
Processing Type
Full HRTF
IID Only

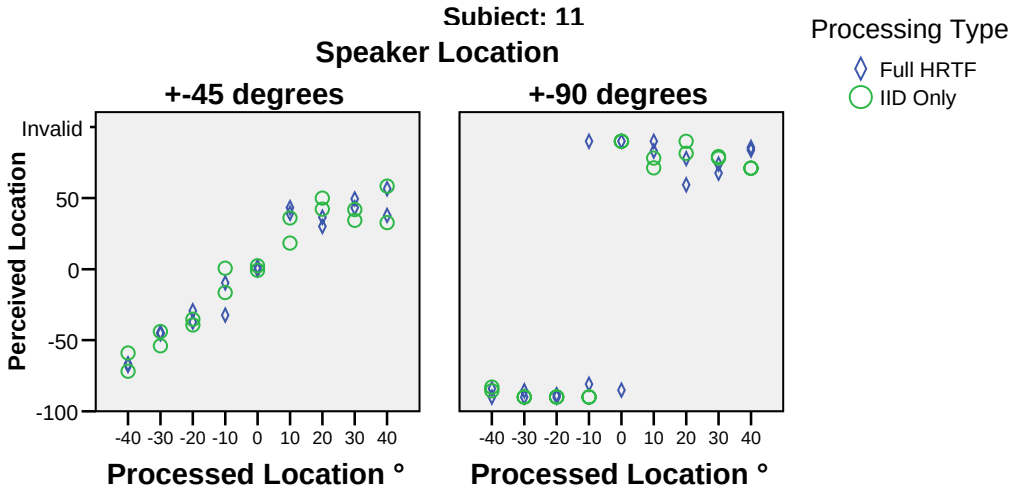

Subject: 6
Speaker Location
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Processed Location °
Processing Type
Full HRTF
IID Only

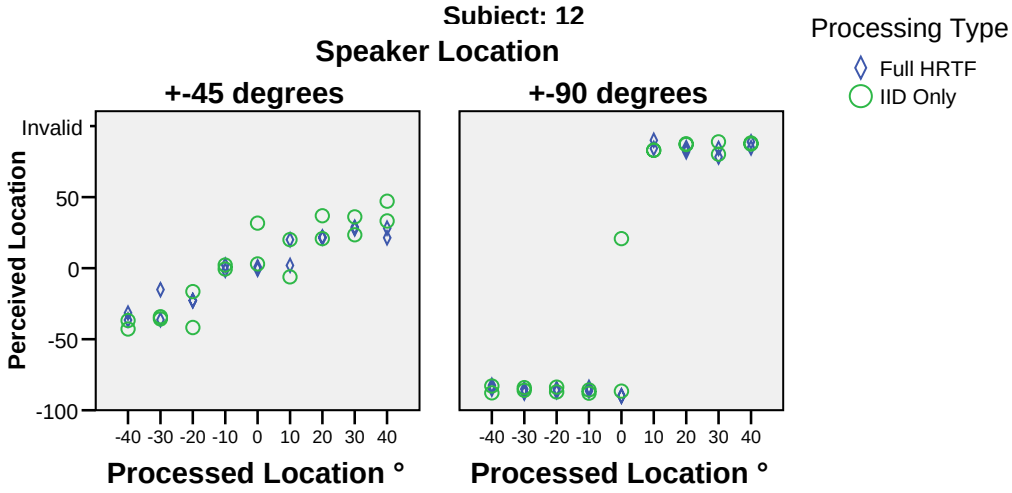

Subject: 7
Speaker Location
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Processed Location °
Processing Type
Full HRTF
IID Only

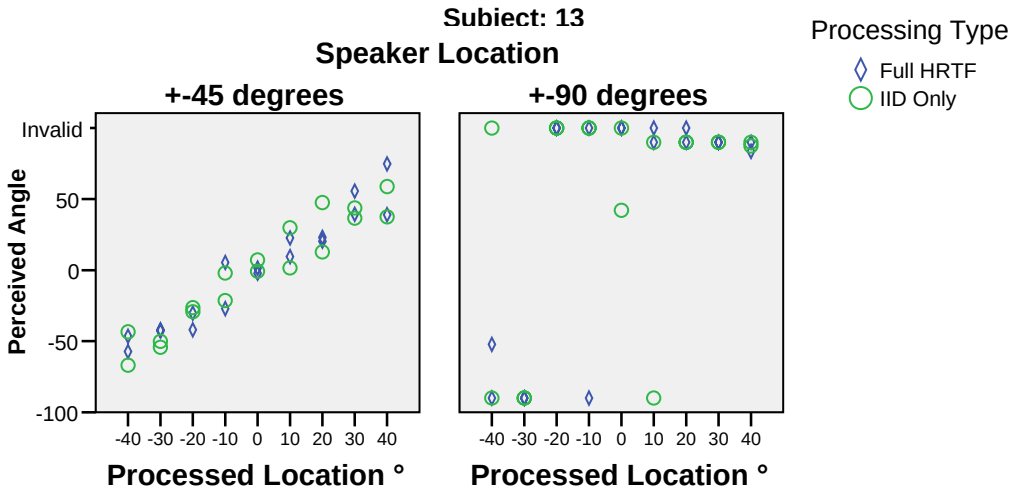

Subject: 8
Speaker Location
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Processed Location °
Processing Type
Full HRTF
IID Only

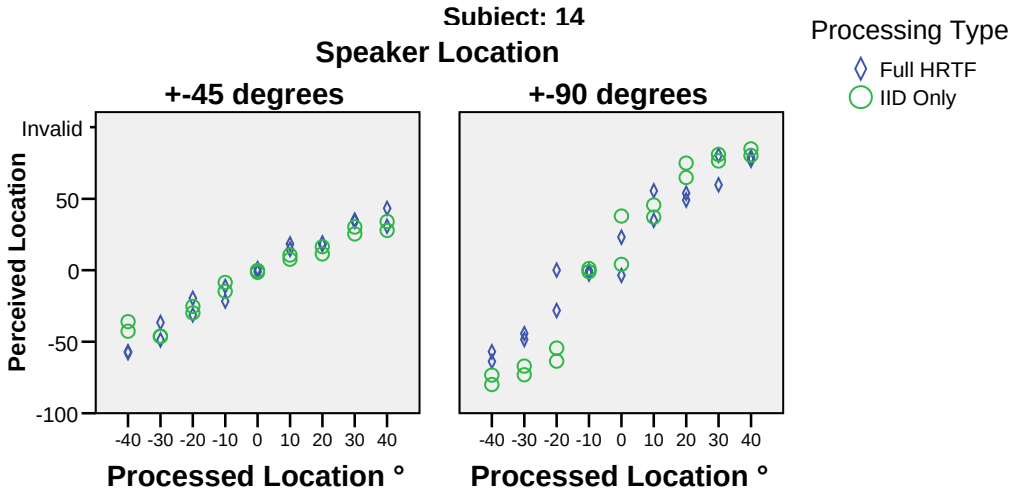

Subject: 9
Speaker Location
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Processed Location °
Processing Type
Full HRTF
IID Only

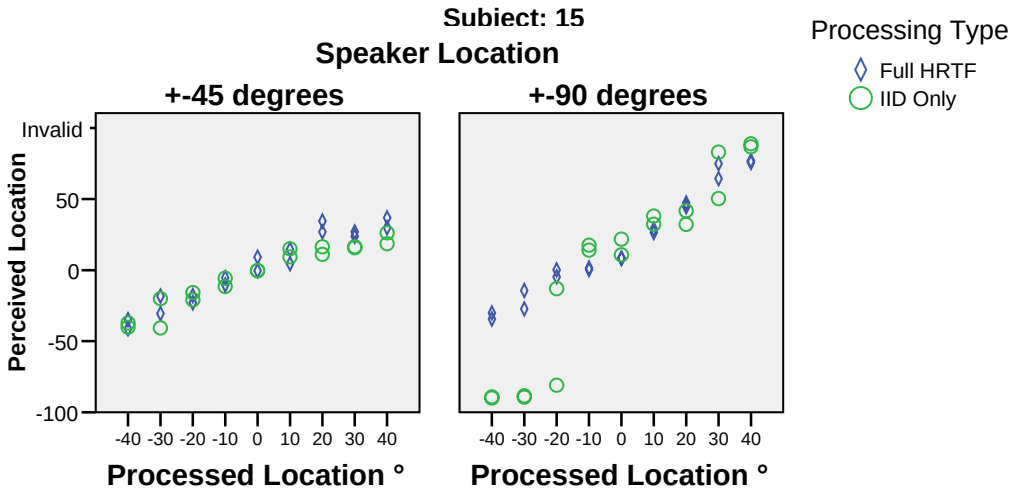

Subject: 10
Speaker Location
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Processed Location °
Processing Type
Full HRTF
IID Only

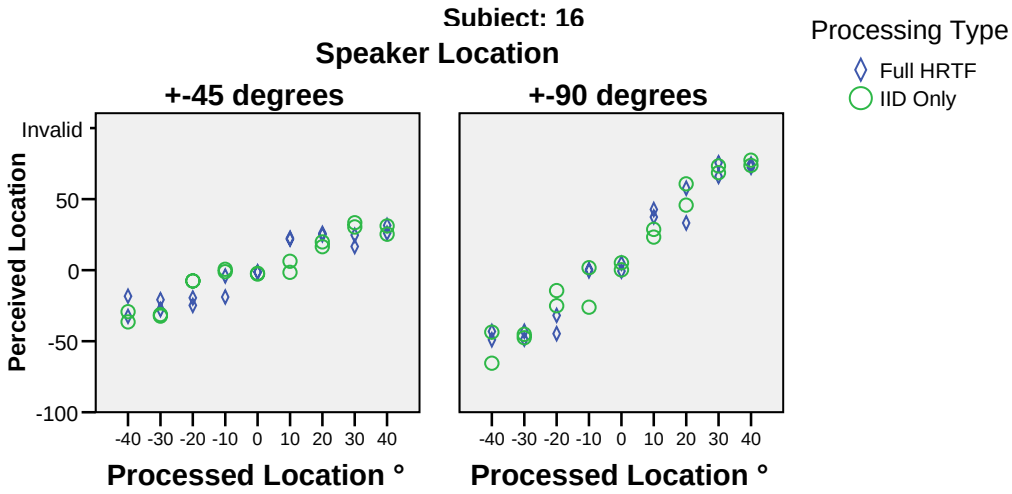

Subject: 11
Speaker Location
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Processed Location °
Processing Type
Full HRTF
IID Only

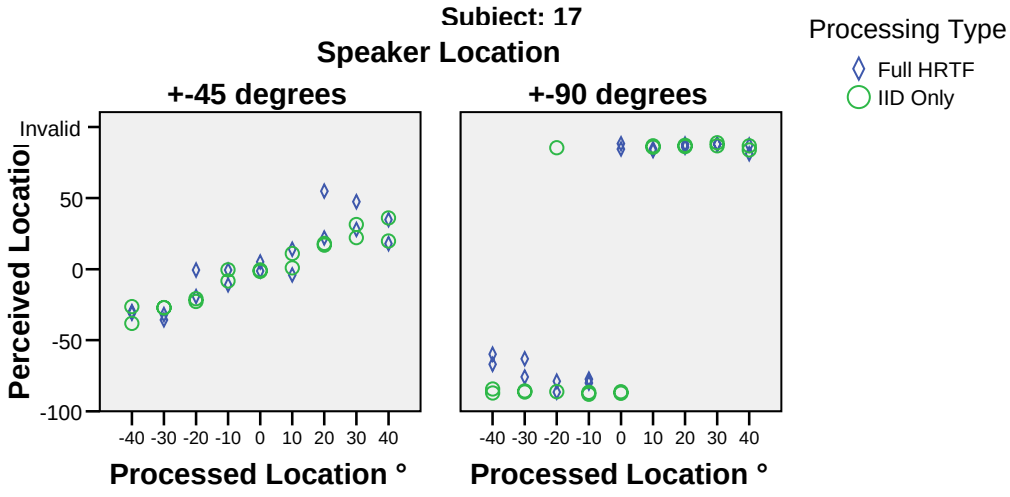

Subject: 12
Speaker Location
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Processed Location °
Processing Type
Full HRTF
IID Only

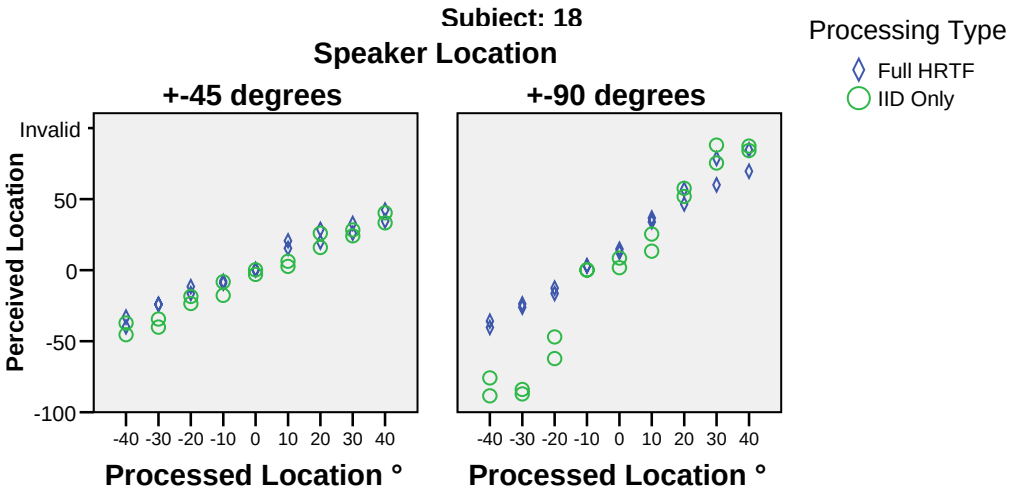

Subject: 13
Speaker Location
+-45 degrees
+-90 degrees
Perceived Angle
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Processed Location °
Processing Type
Full HRTF
IID Only

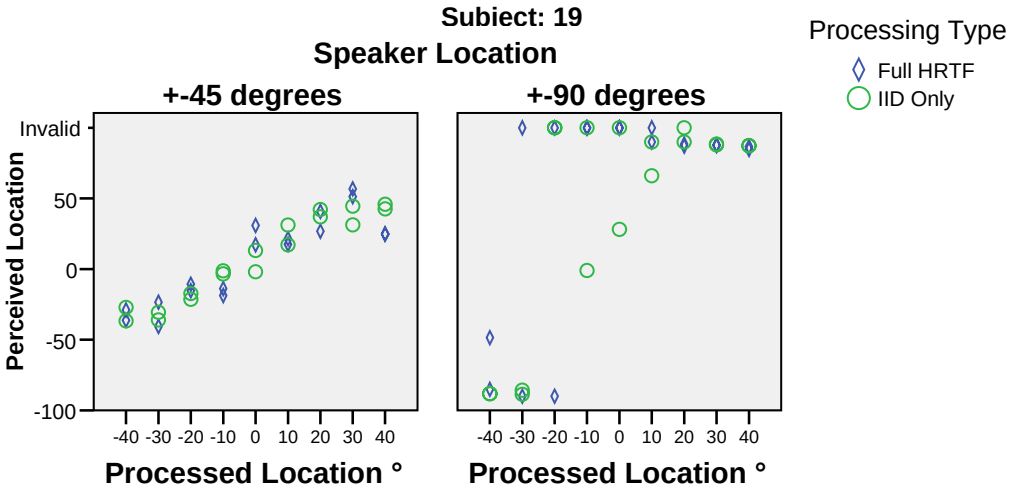

Subject: 14
Speaker Location
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Processed Location °
Processing Type
Full HRTF
IID Only

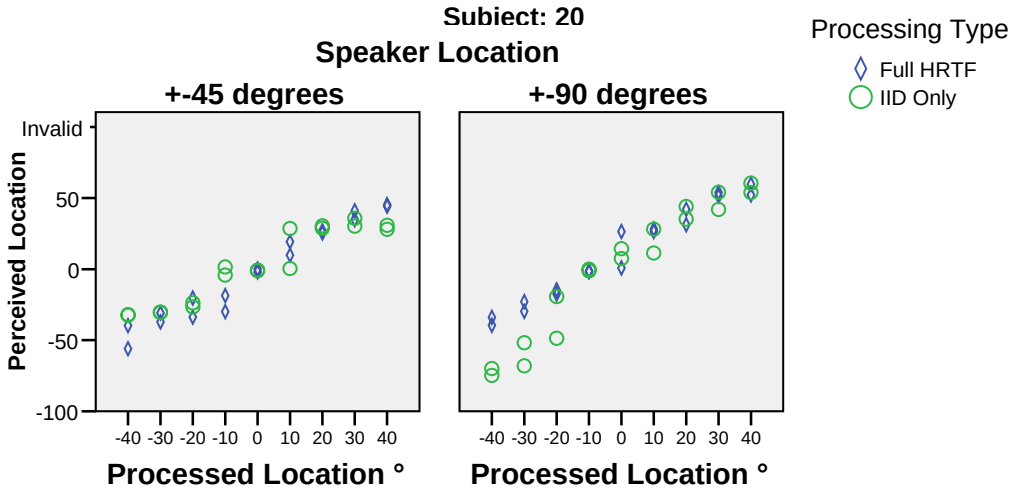

Subject: 15
Speaker Location
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Processed Location °
Processing Type
Full HRTF
IID Only

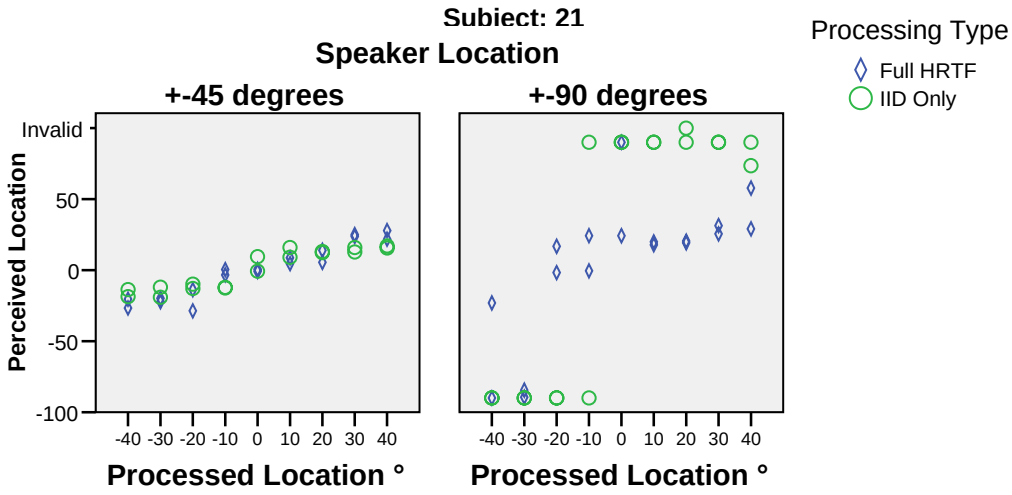

Subject: 16
Speaker Location
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Processed Location °
Processing Type
Full HRTF
IID Only

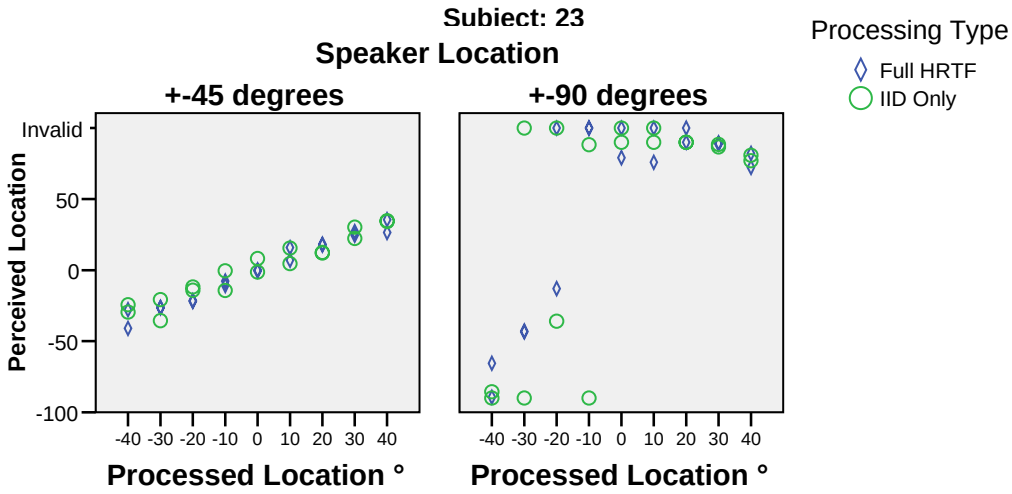

Subject: 17
Speaker Location
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Processed Location °
Processing Type
Full HRTF
IID Only

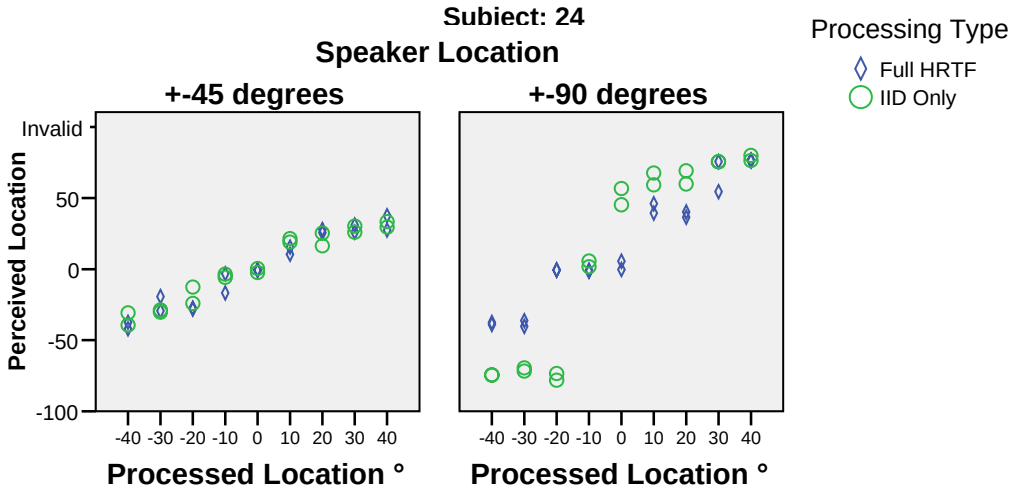

Subject: 18
Speaker Location
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Processed Location °
Processing Type
Full HRTF
IID Only

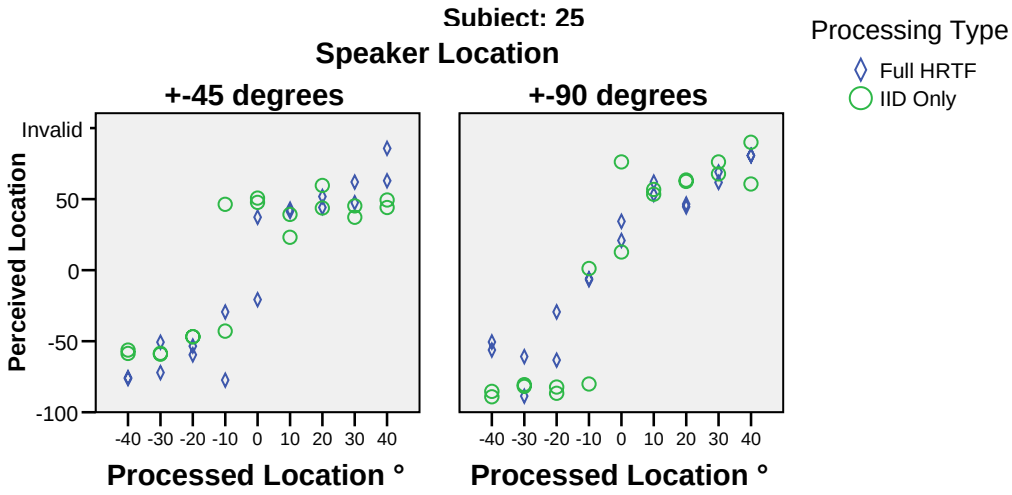

Subject: 19
Speaker Location
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Processed Location °
Processing Type
Full HRTF
IID Only

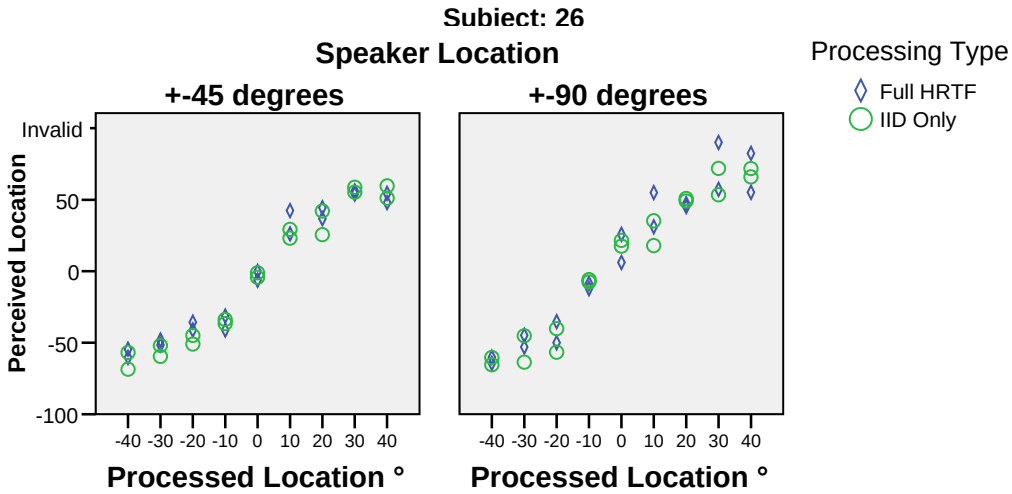

Subject: 20
Speaker Location
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Processed Location °
Processing Type
Full HRTF
IID Only

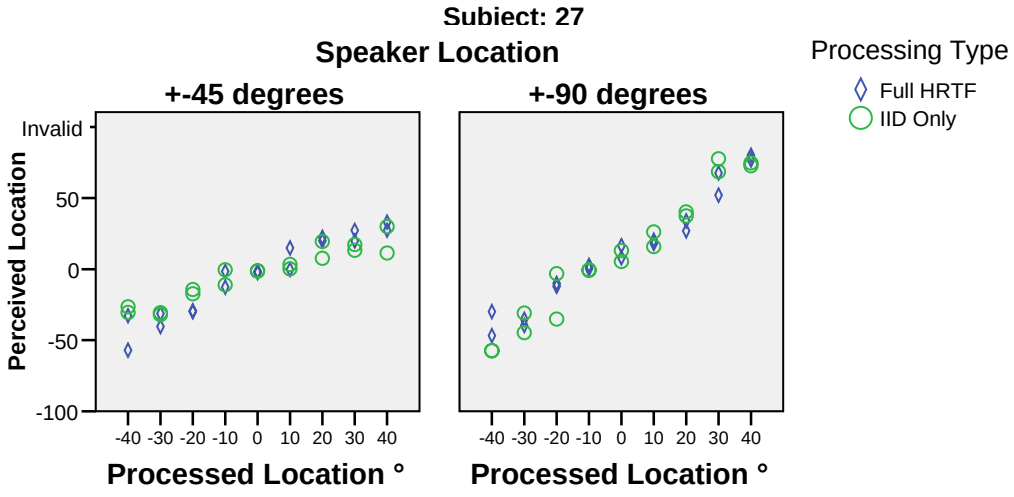

Subject: 21
Speaker Location
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Processed Location °
Processing Type
Full HRTF
IID Only

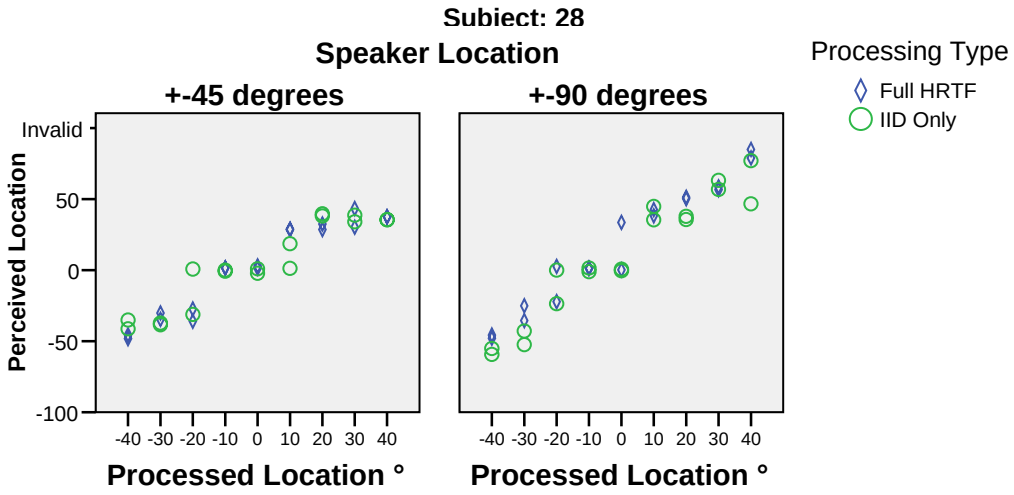

Subject: 23
Speaker Location
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Processed Location °
Processing Type
Full HRTF
IID Only

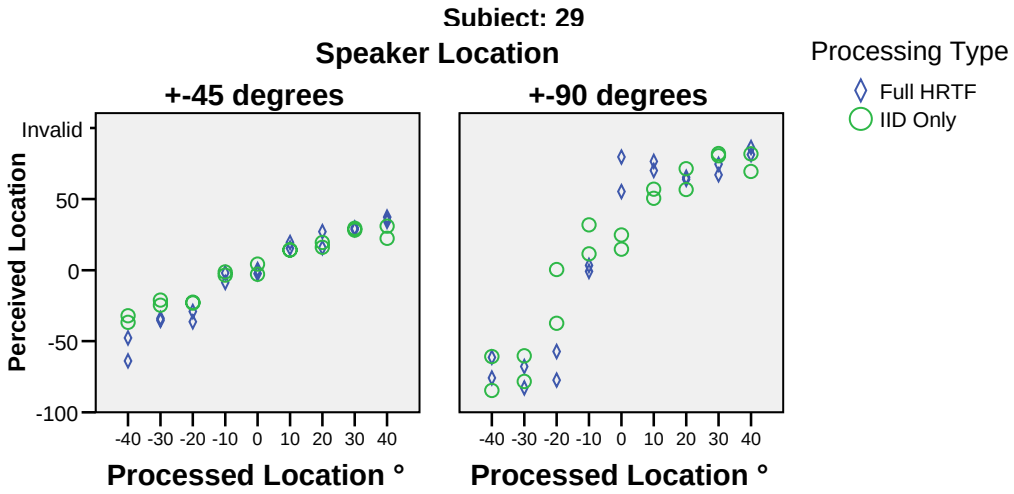

Subject: 24
Speaker Location
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Processed Location °
Processing Type
Full HRTF
IID Only

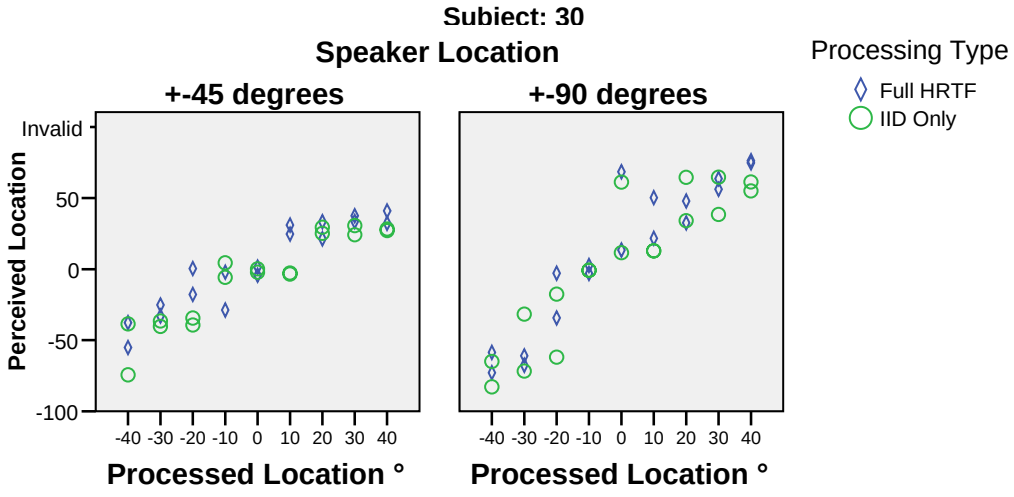

Subject: 25
Speaker Location
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Processed Location °
Processing Type
Full HRTF
IID Only

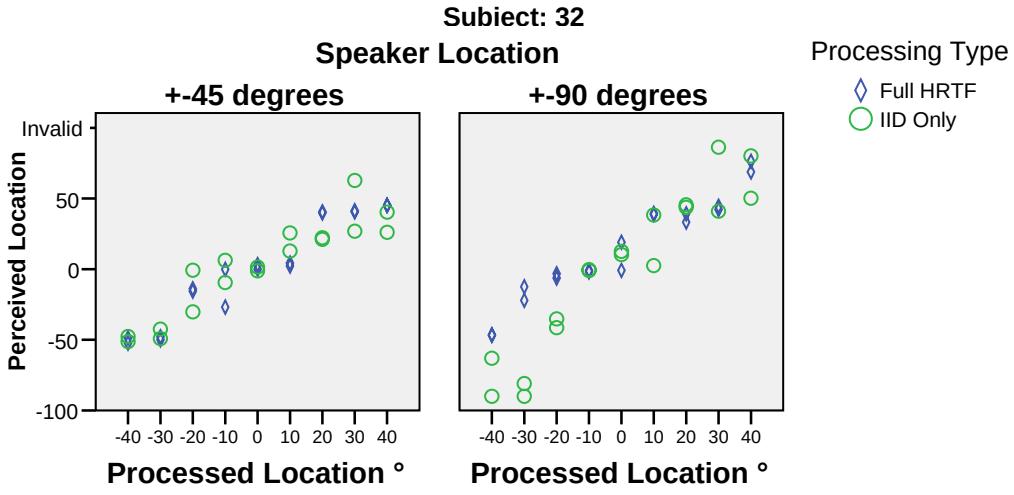

Subject: 26
Speaker Location
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Processed Location °
Processing Type
Full HRTF
IID Only

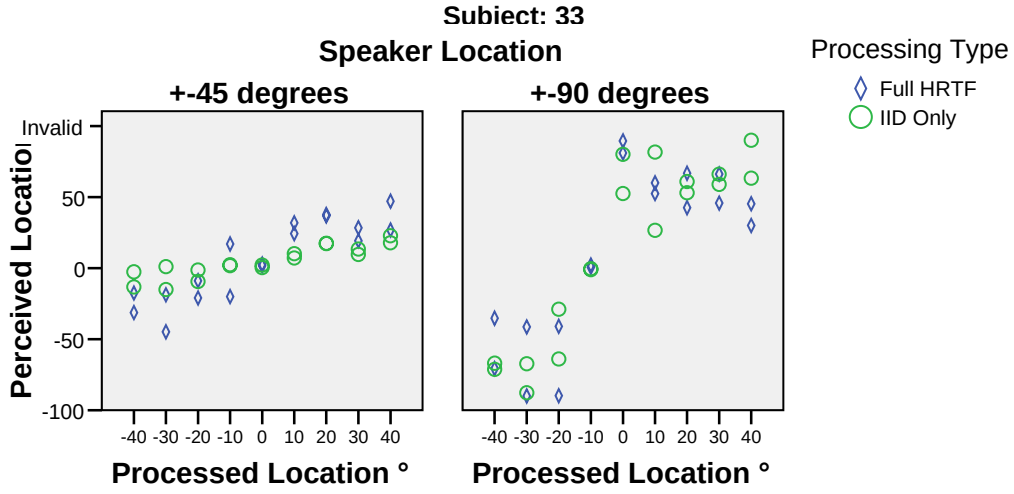

Subject: 27
Speaker Location
+-45 degrees
+-90 degrees
Processing Type
Full HRTF
IID Only
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °

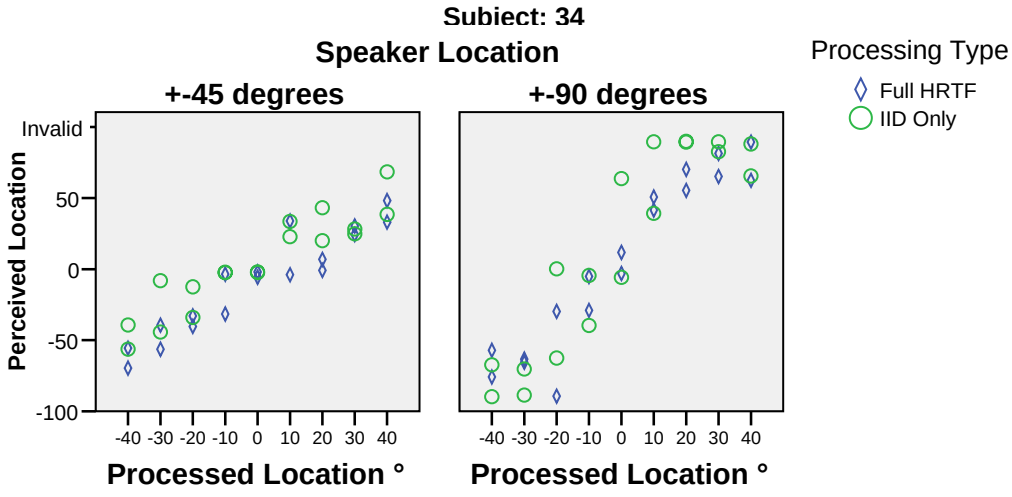

Subject: 28
Speaker Location
+-45 degrees
+-90 degrees
Processing Type
Full HRTF
IID Only
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °

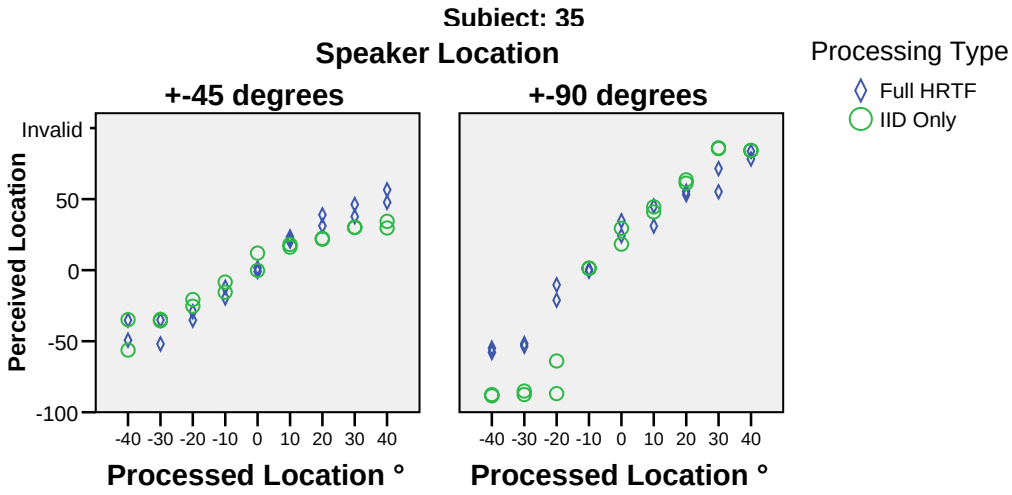

Subject: 29
Speaker Location
+-45 degrees
+-90 degrees
Processing Type
Full HRTF
IID Only
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °

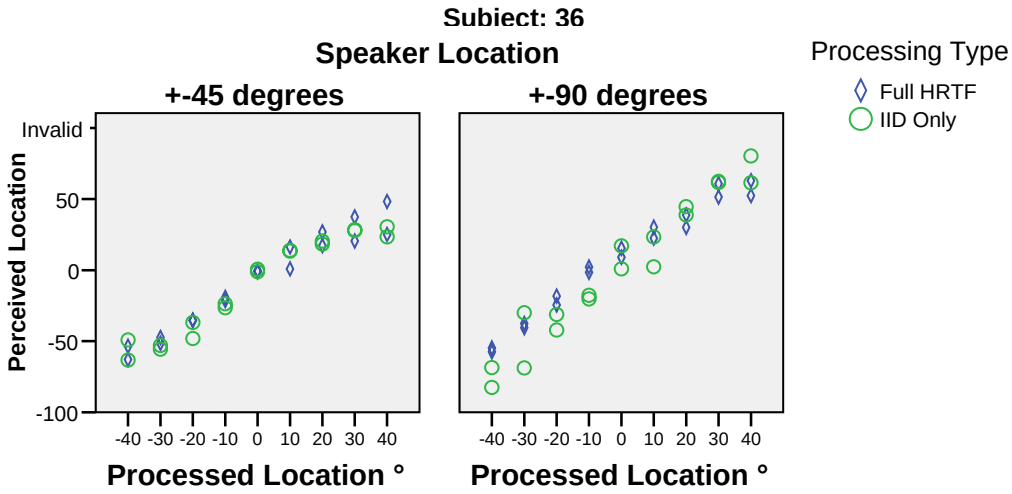

Subject: 30
Speaker Location
+-45 degrees
+-90 degrees
Processing Type
Full HRTF
IID Only
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °


Subject: 32
Speaker Location
+-45 degrees
+-90 degrees
Processing Type
Full HRTF
IID Only
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °


Subject: 33
Speaker Location
+-45 degrees
+-90 degrees
Processing Type
Full HRTF
IID Only
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °


Subject: 34
Speaker Location
+-45 degrees
+-90 degrees
Processing Type
Full HRTF
IID Only
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °


Subject: 35
Speaker Location
+-45 degrees
+-90 degrees
Processing Type
Full HRTF
IID Only
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °


Subject: 36
Speaker Location
+-45 degrees
+-90 degrees
Processing Type
Full HRTF
IID Only
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °

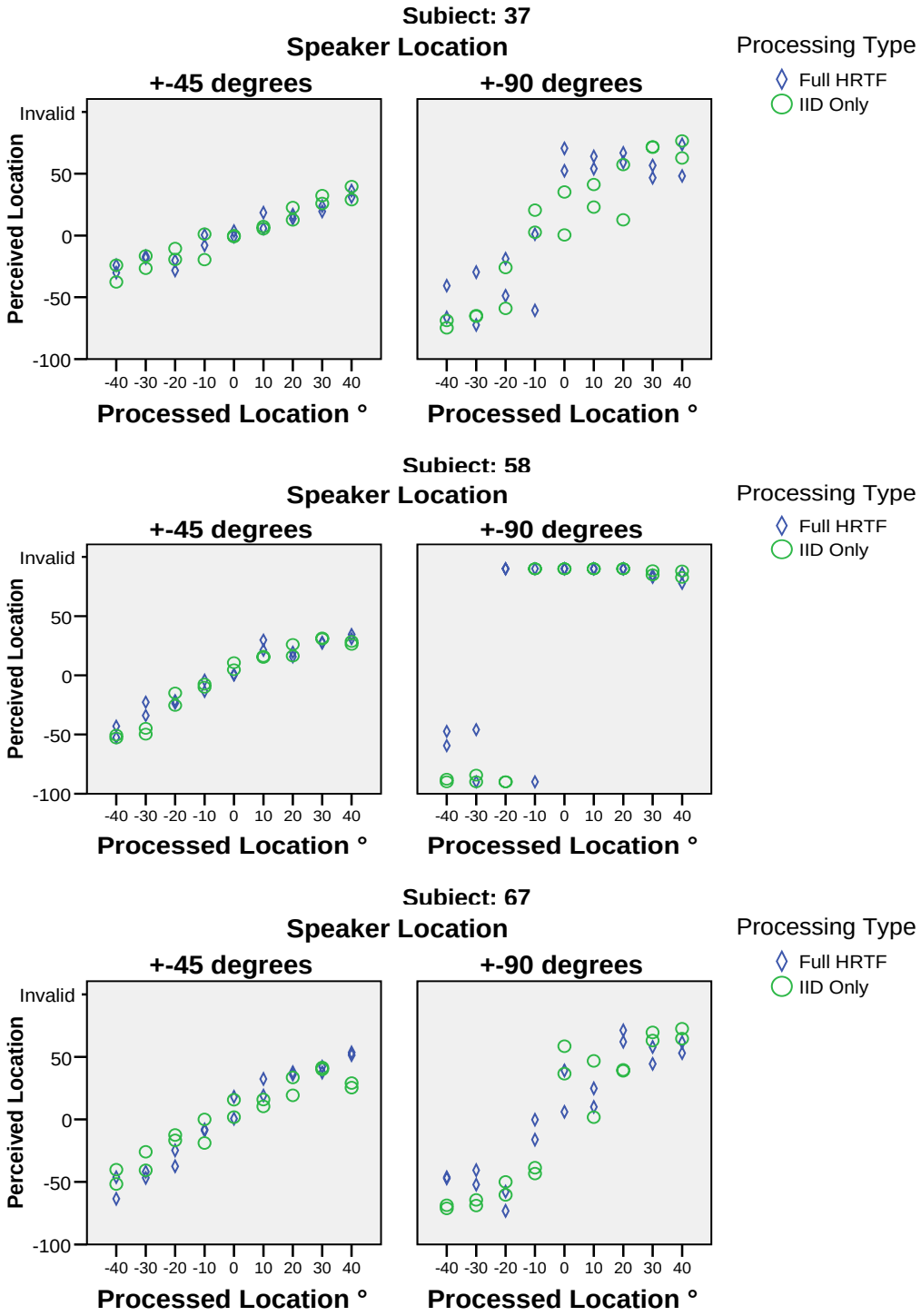
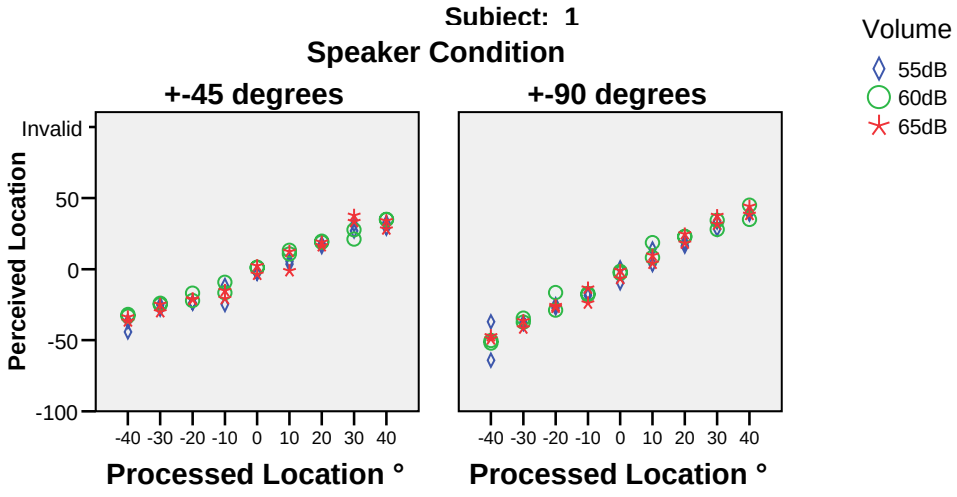
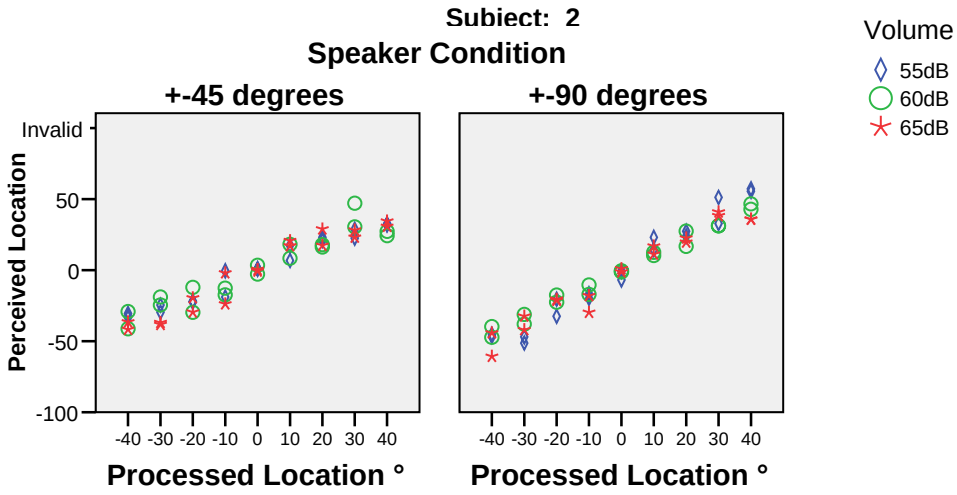
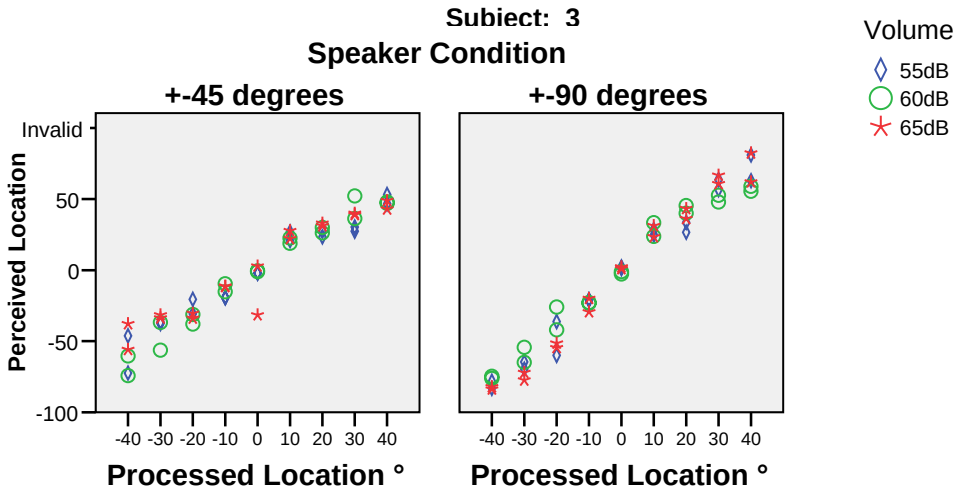
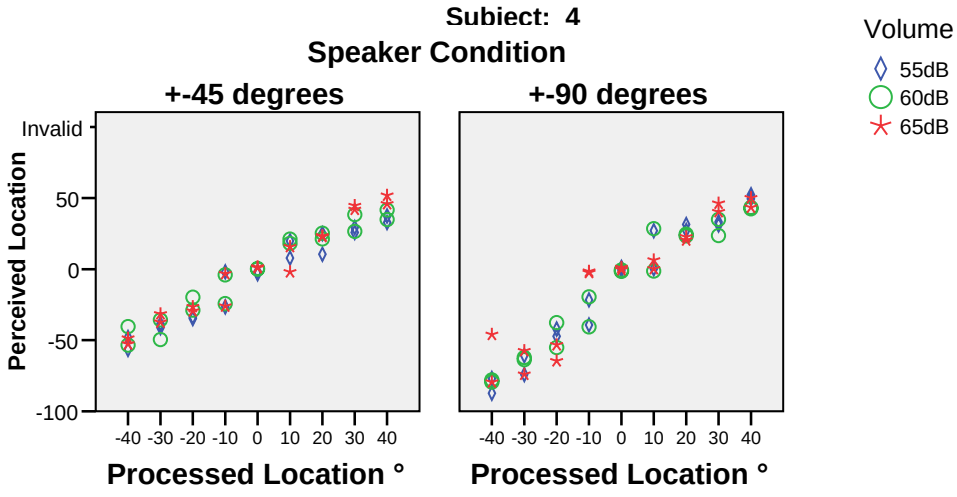


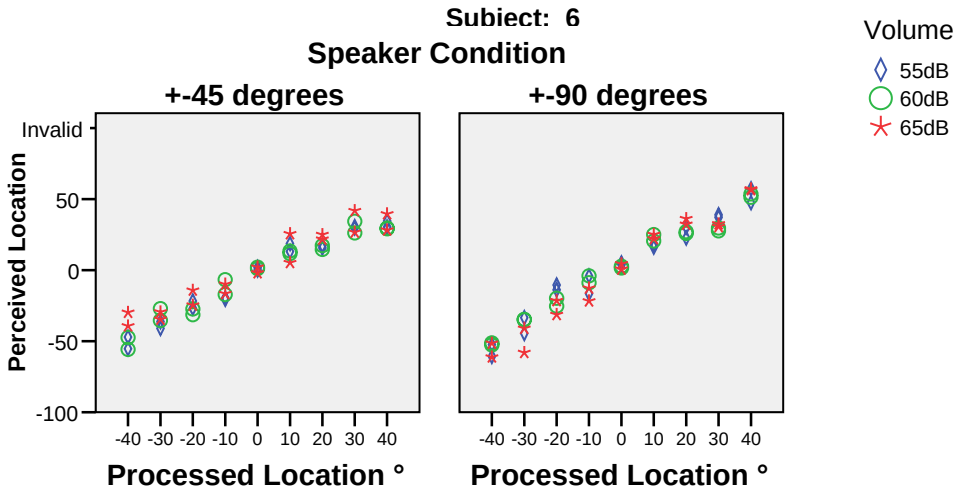
Figure B.1: Perceived location of sounds processed using both processing methods as a function of processed location. Data are presented from individual participants. Invalid data are marked as a +100° response. Note the prevalence of invalid responses in ±90° speaker configuration. The units of all axes are degrees.

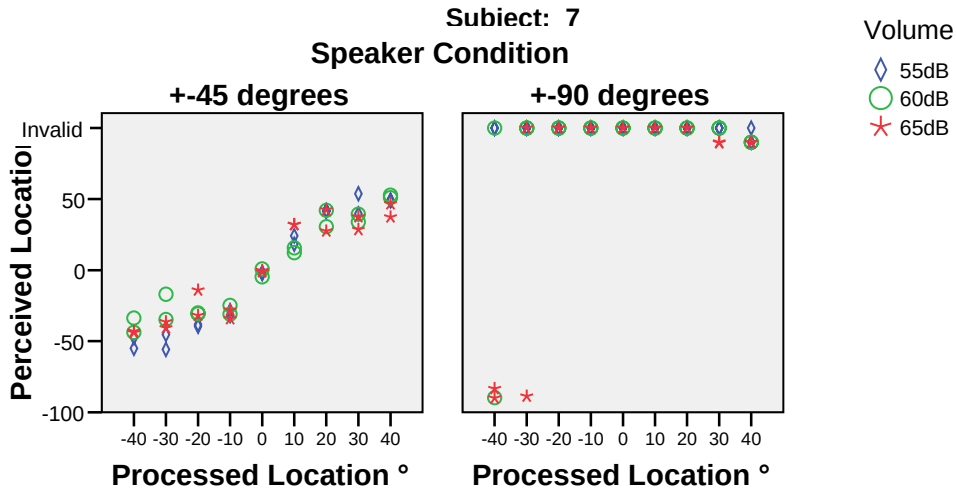

Subject: 1
Speaker Condition
+-45 degrees
+-90 degrees
Invalid
50
0
-50
-100
Perceived Location
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

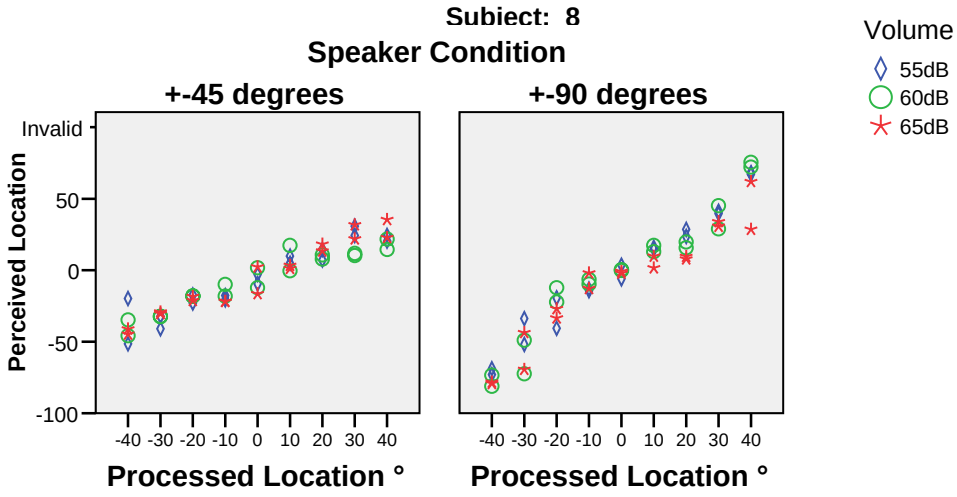

Subject: 2
Speaker Condition
+-45 degrees
+-90 degrees
Invalid
50
0
-50
-100
Perceived Location
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

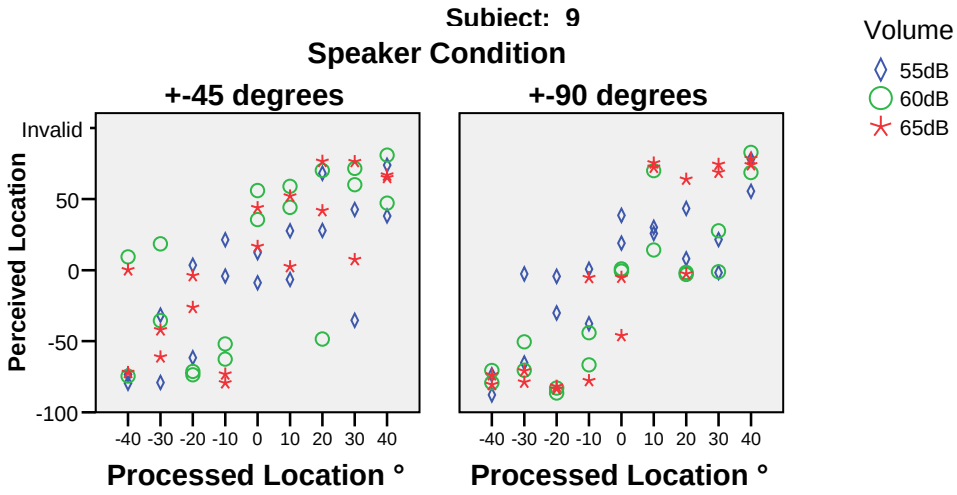

Subject: 3
Speaker Condition
+-45 degrees
+-90 degrees
Invalid
50
0
-50
-100
Perceived Location
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

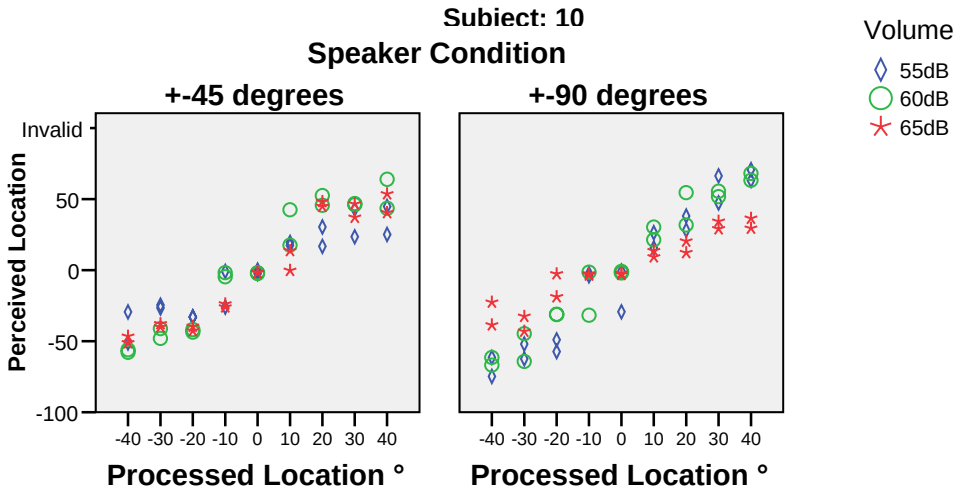

Subject: 4
Speaker Condition
+-45 degrees
+-90 degrees
Invalid
50
0
-50
-100
Perceived Location
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

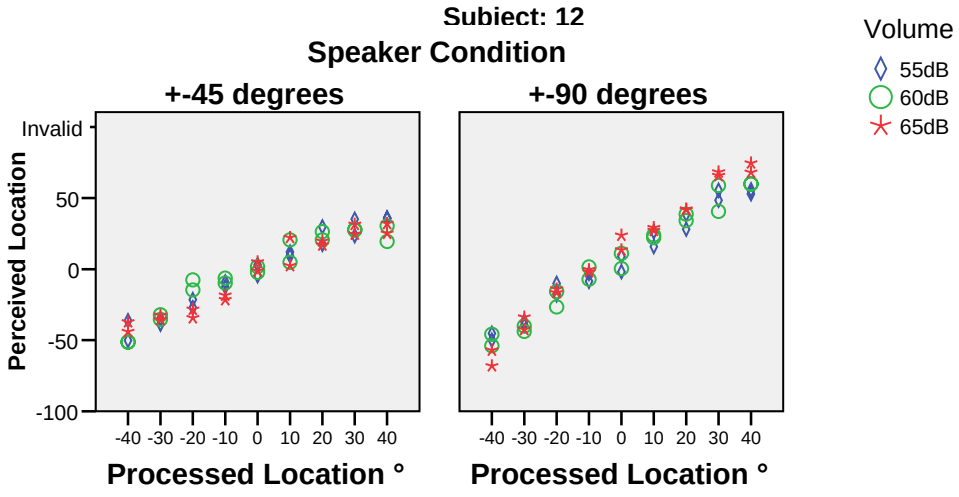

Subject: 6
Speaker Condition
+-45 degrees
+-90 degrees
Invalid
50
0
-50
-100
Perceived Location
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

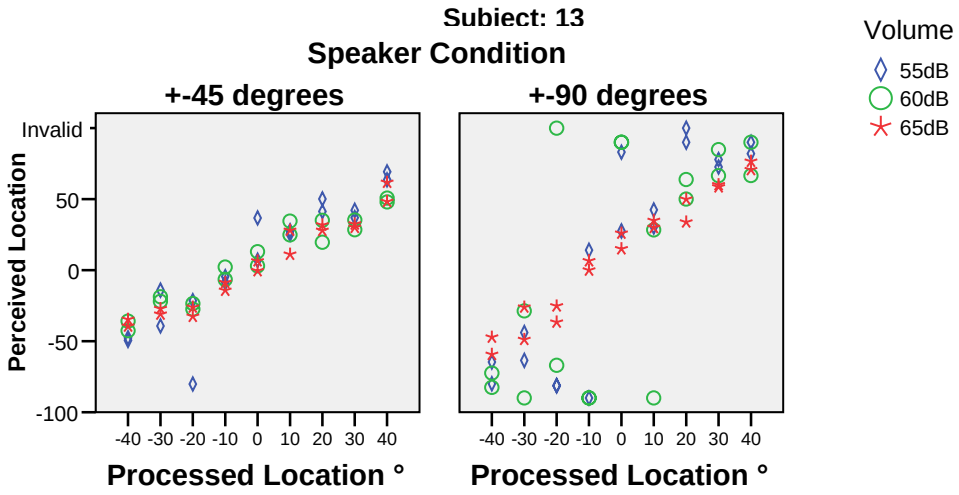

Subject: 7
Speaker Condition
+-45 degrees
+-90 degrees
Invalid
50
0
-50
-100
Perceived Location
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

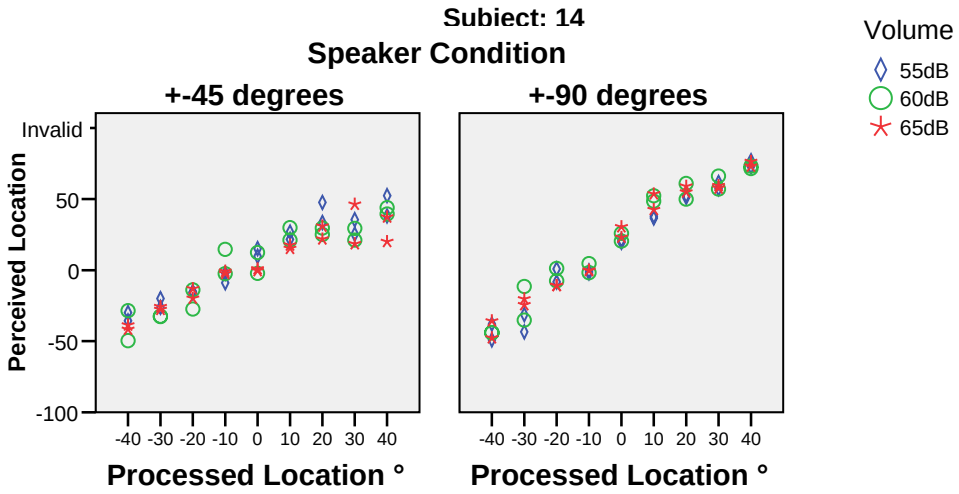

Subject: 8
Speaker Condition
+-45 degrees
+-90 degrees
Invalid
50
0
-50
-100
Perceived Location
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

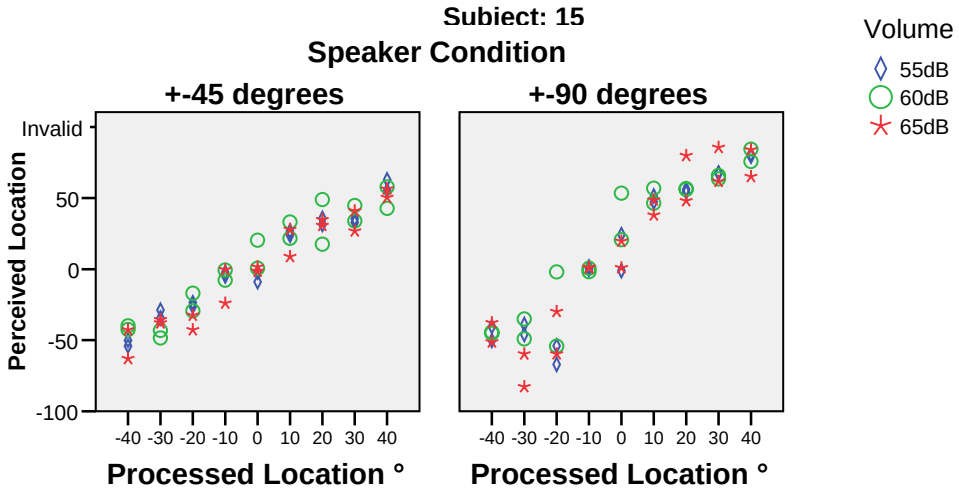

Subject: 9
Speaker Condition
+-45 degrees
+-90 degrees
Invalid
50
0
-50
-100
Perceived Location
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

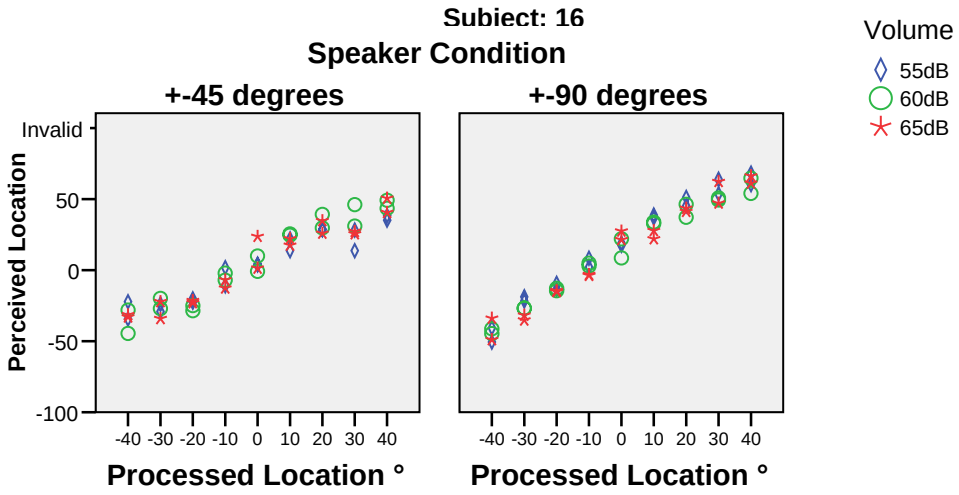

Subject: 10
Speaker Condition
+-45 degrees
+-90 degrees
Invalid
50
0
-50
-100
Perceived Location
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

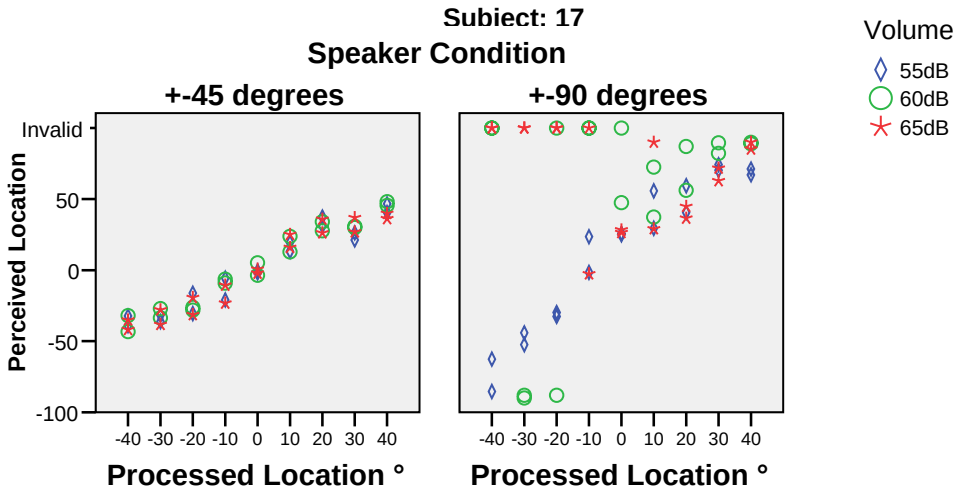

Subject: 12
Speaker Condition
+-45 degrees
+-90 degrees
Invalid
50
0
-50
-100
Perceived Location
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

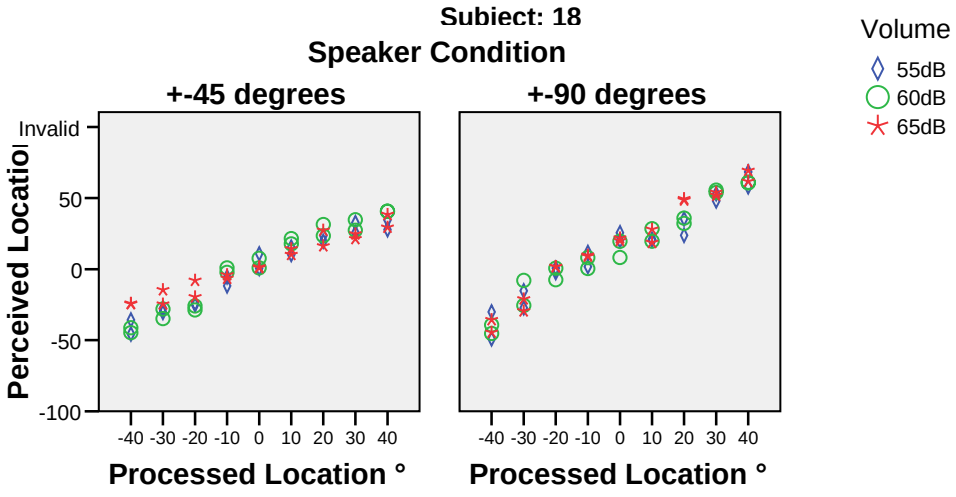

Subject: 13
Speaker Condition
+-45 degrees
+-90 degrees
Invalid
50
0
-50
-100
Perceived Location
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

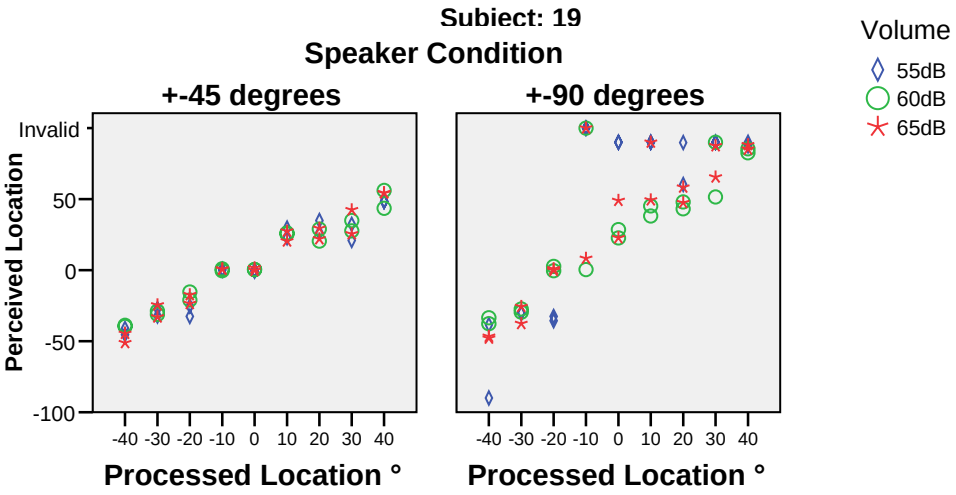

Subject: 14
Speaker Condition
+-45 degrees
+-90 degrees
Invalid
50
0
-50
-100
Perceived Location
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

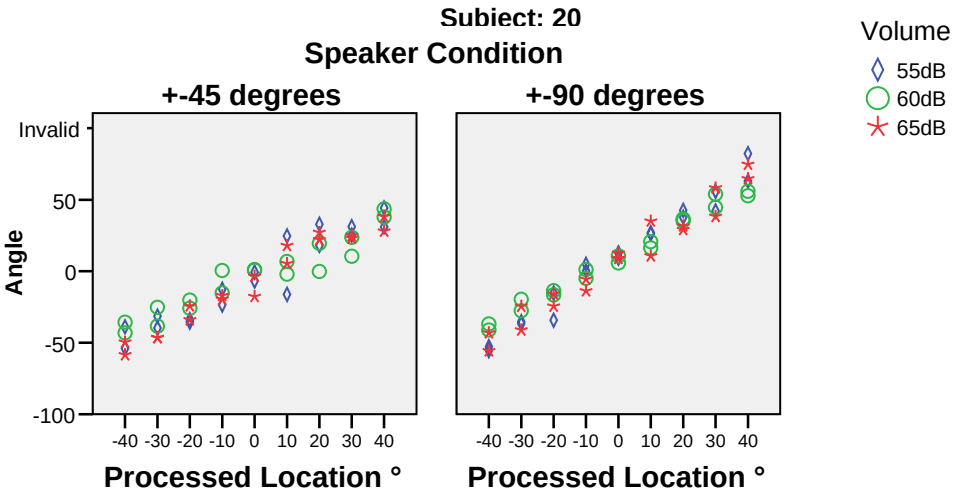

Subject: 15
Speaker Condition
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

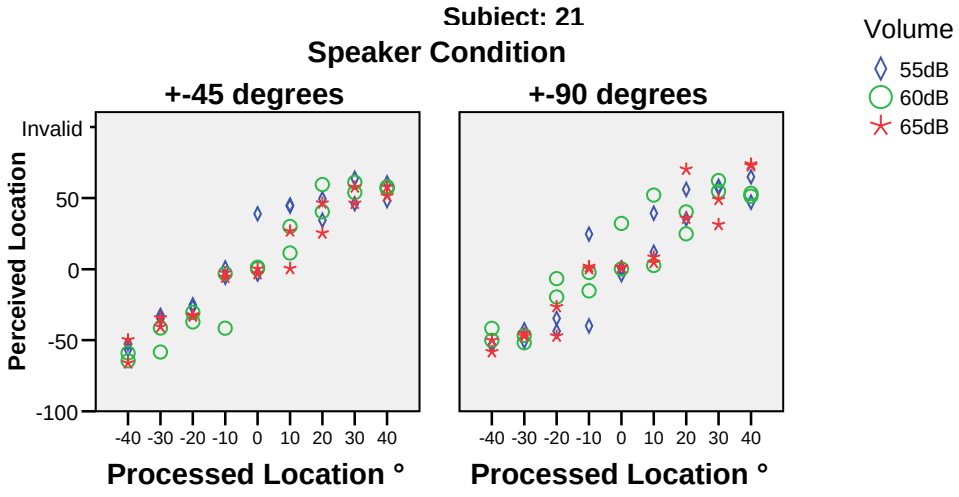

Subject: 16
Speaker Condition
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

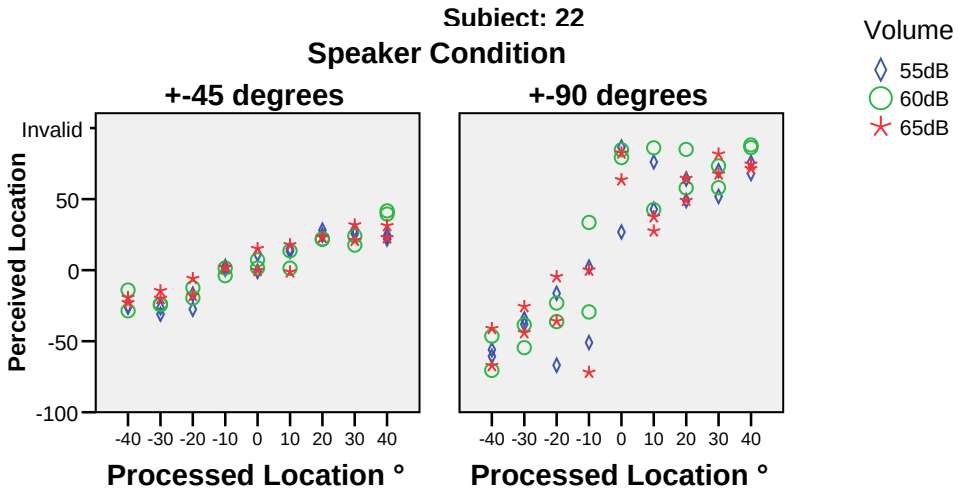

Subject: 17
Speaker Condition
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

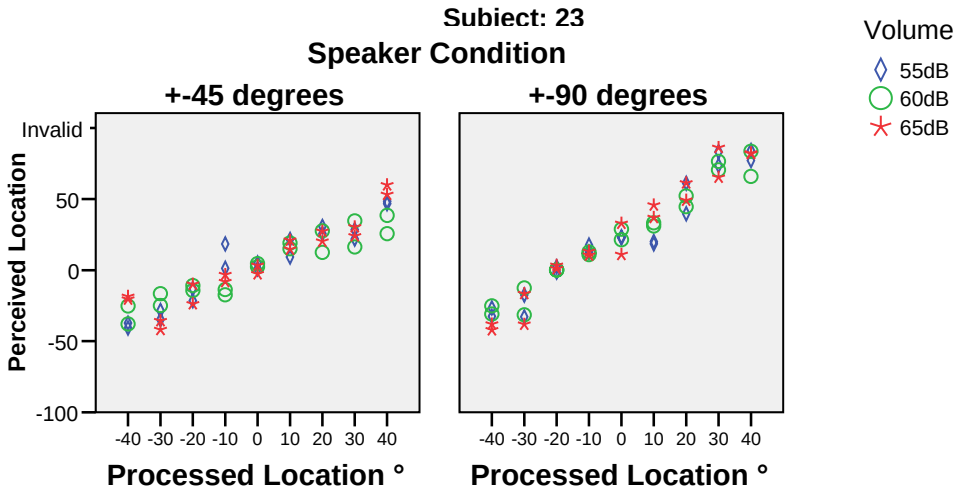

Subject: 18
Speaker Condition
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

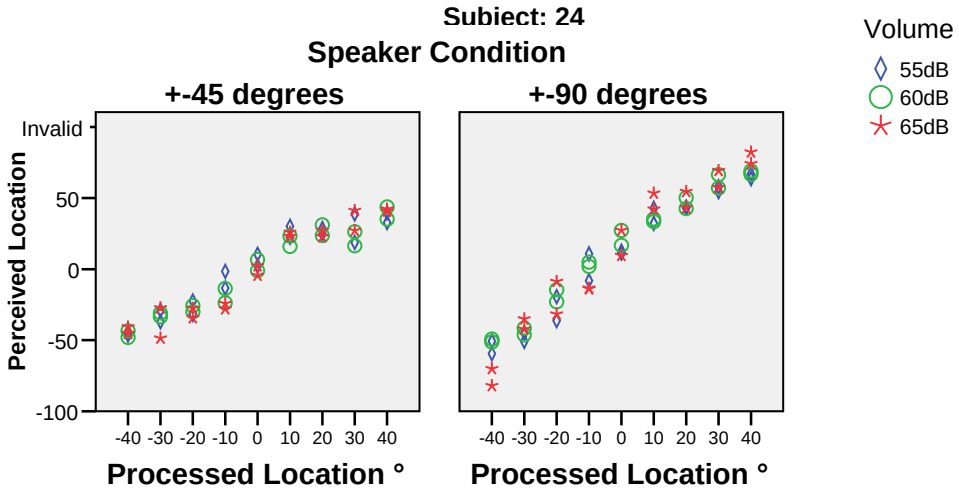

Subject: 19
Speaker Condition
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

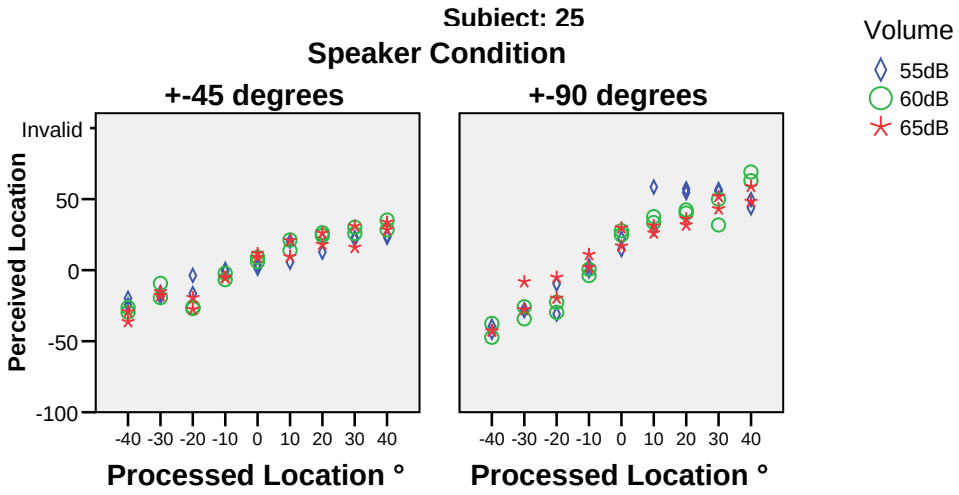

Subject: 20
Speaker Condition
+-45 degrees
+-90 degrees
Angle
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

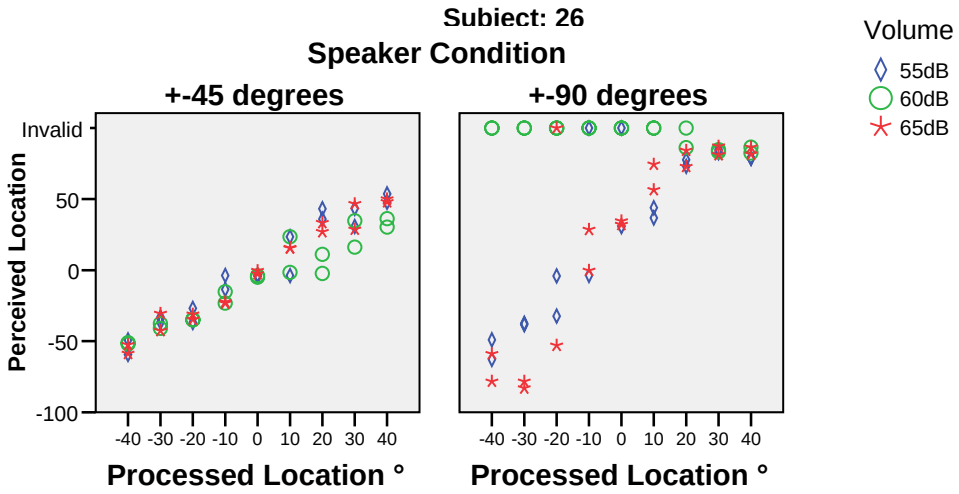

Subject: 21
Speaker Condition
+-45 degrees
+-90 degrees
Invalid
50
0
-50
-100
Perceived Location
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

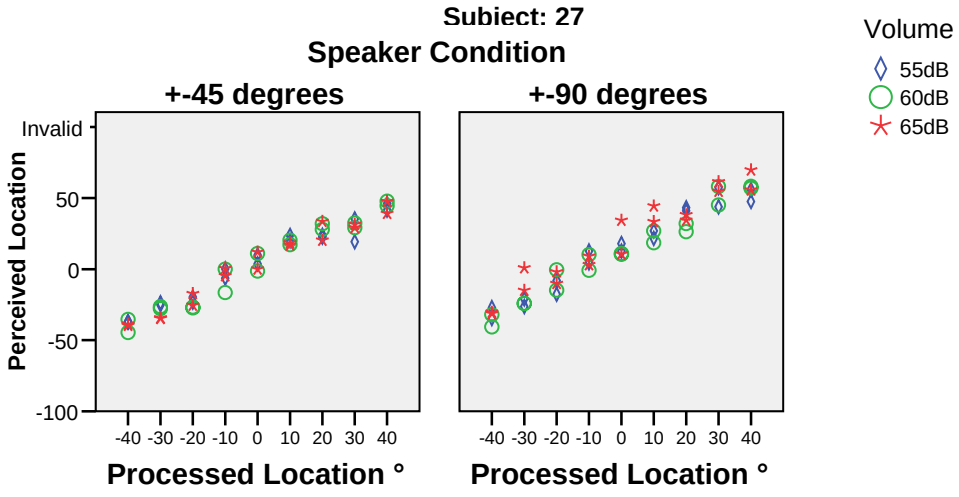

Subject: 22
Speaker Condition
+-45 degrees
+-90 degrees
Invalid
50
0
-50
-100
Perceived Location
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

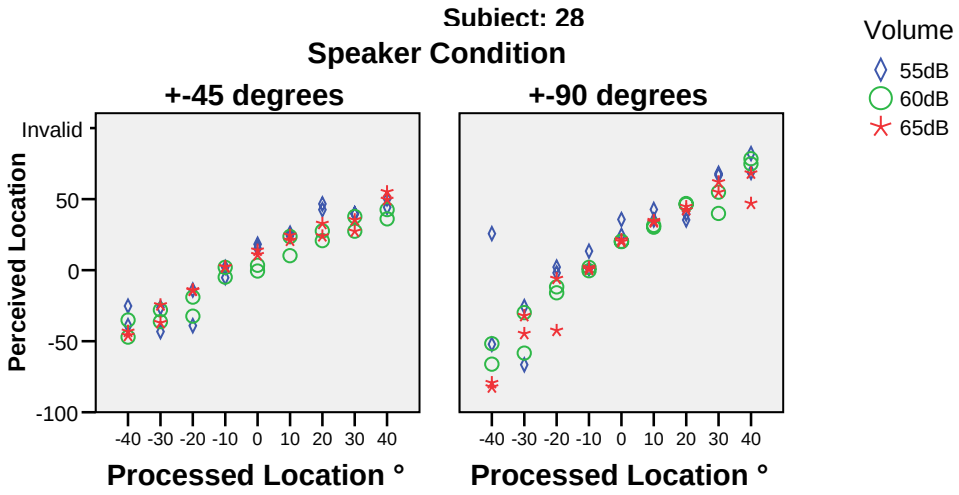

Subject: 23
Speaker Condition
+-45 degrees
+-90 degrees
Invalid
50
0
-50
-100
Perceived Location
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

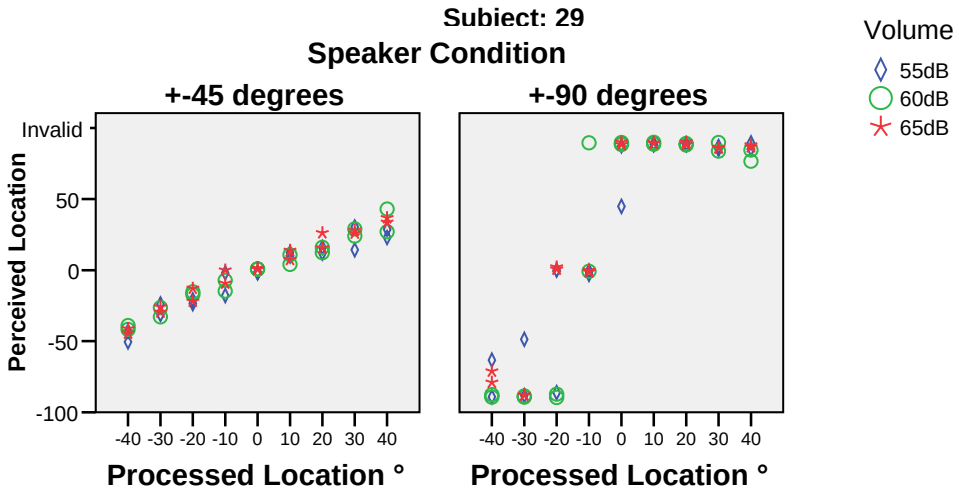

Subject: 24
Speaker Condition
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

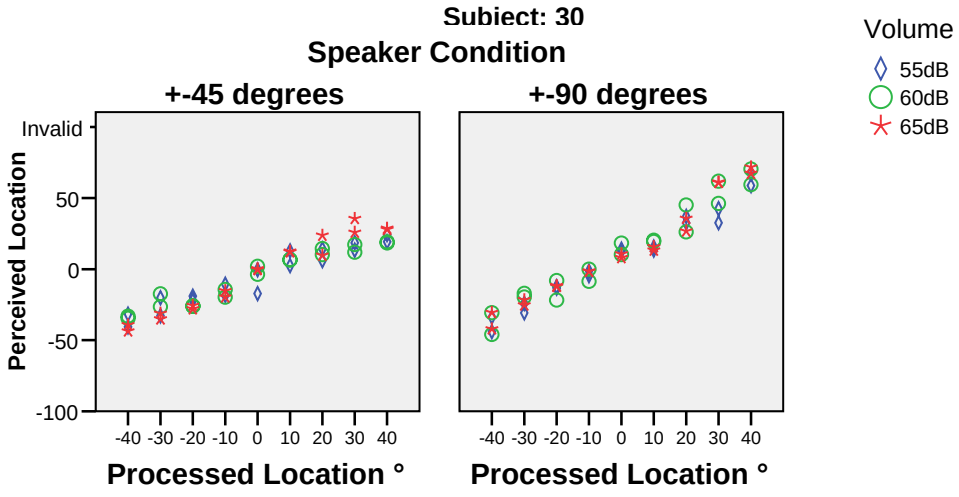

Subject: 25
Speaker Condition
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

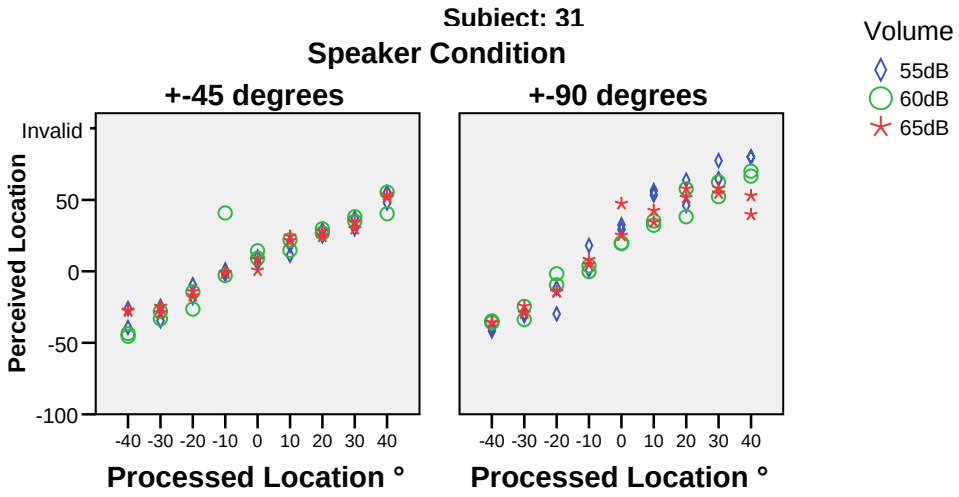

Subject: 26
Speaker Condition
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

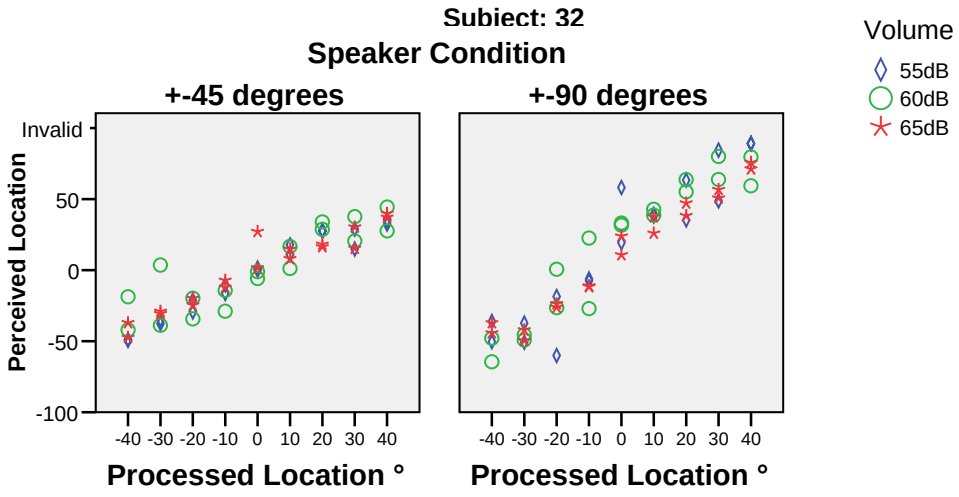

Subject: 27
Speaker Condition
+-45 degrees
+-90 degrees
Invalid
50
0
-50
-100
Perceived Location
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

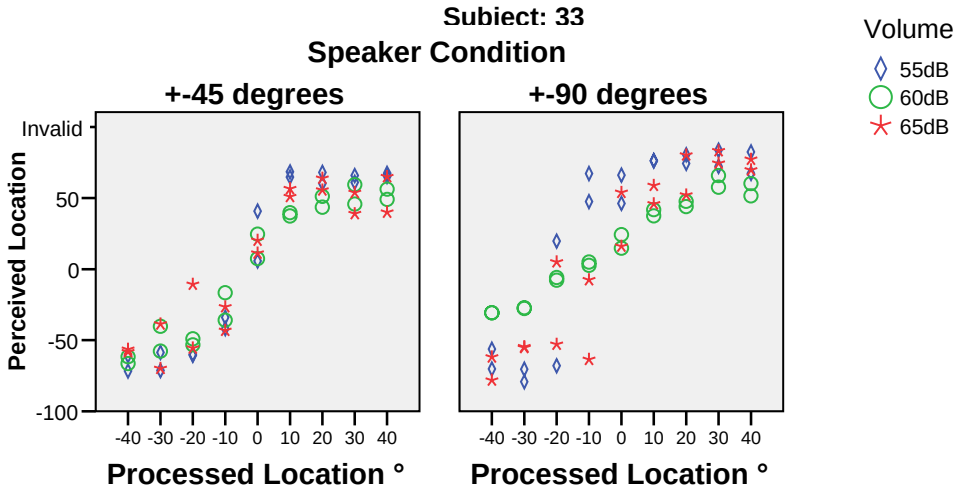

Subject: 28
Speaker Condition
+-45 degrees
+-90 degrees
Invalid
50
0
-50
-100
Perceived Location
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

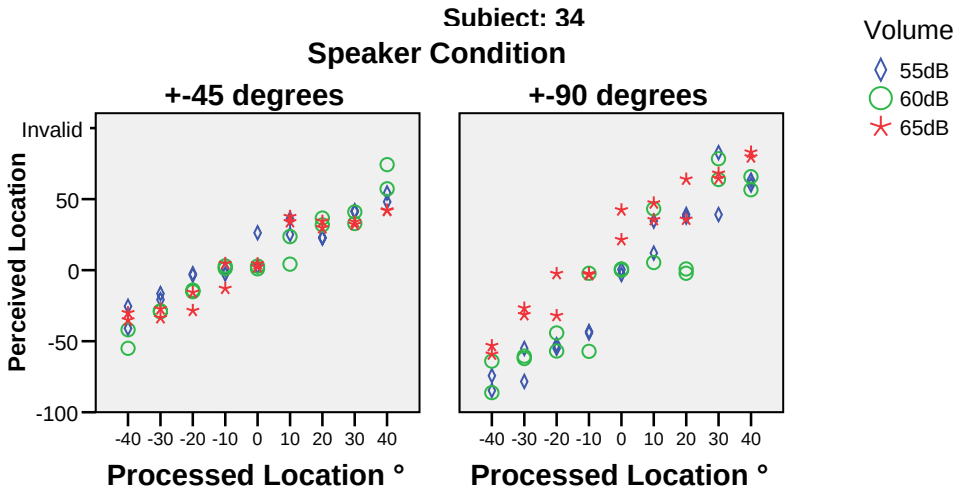

Subject: 29
Speaker Condition
+-45 degrees
+-90 degrees
Invalid
50
0
-50
-100
Perceived Location
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

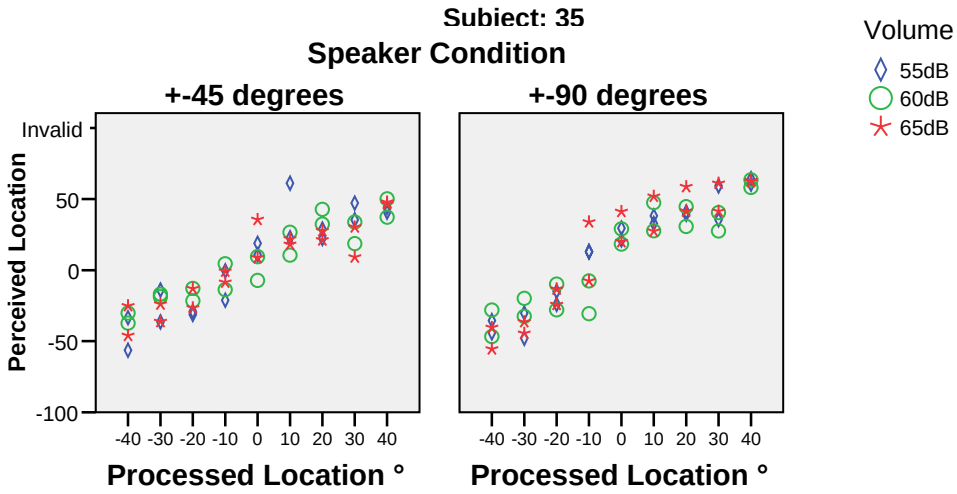

Subject: 30
Speaker Condition
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB


Subject: 31
Speaker Condition
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB


Subject: 32
Speaker Condition
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB


Subject: 33
Speaker Condition
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB


Subject: 34
Speaker Condition
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB


Subject: 35
Speaker Condition
+-45 degrees
+-90 degrees
Perceived Location
Invalid
50
0
-50
-100
-40 -30 -20 -10 0 10 20 30 40
Processed Location °
Volume
55dB
60dB
65dB

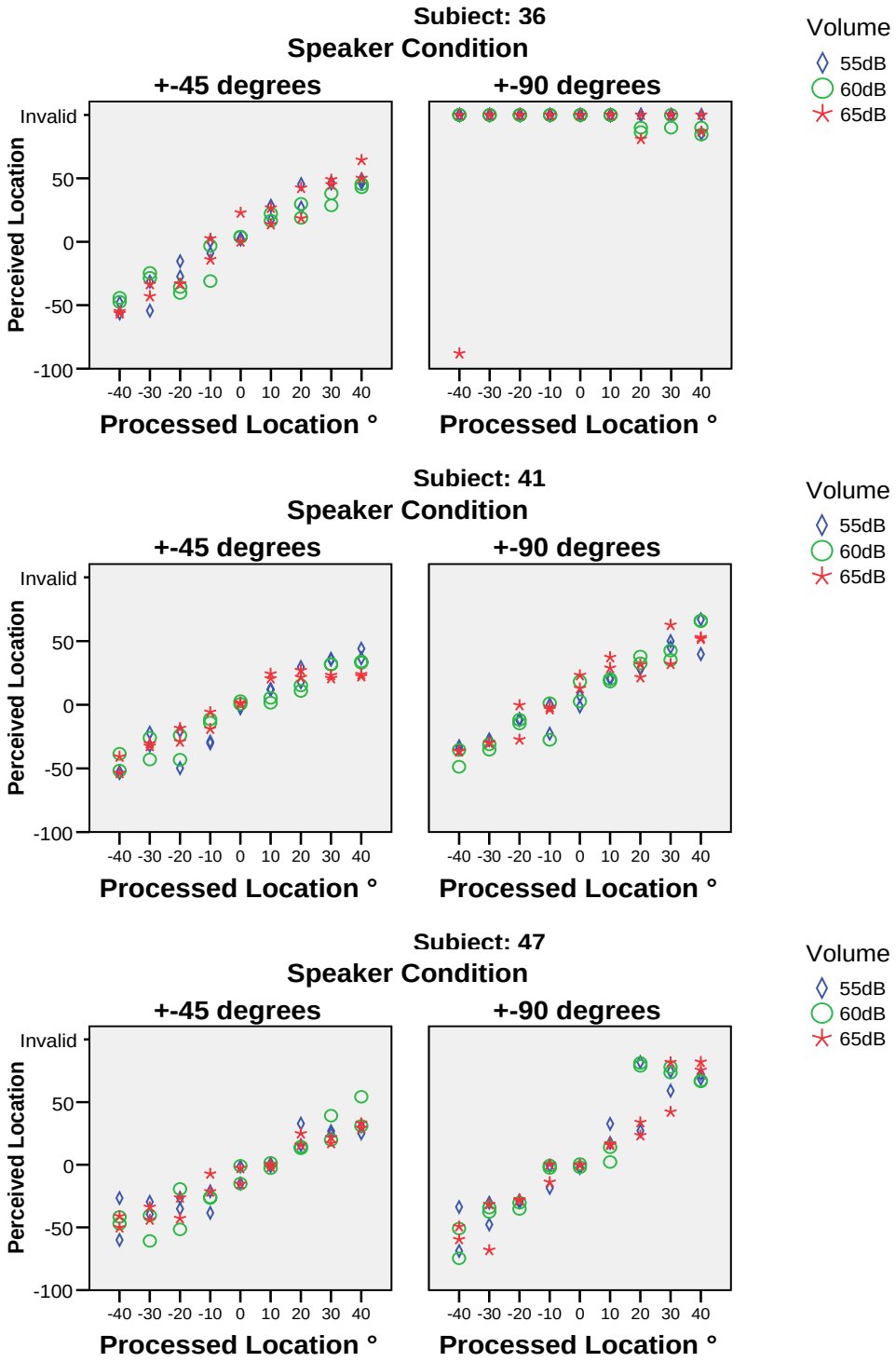


Figure B.2: Perceived location of sounds at each volume level as a function of processed location. Data are presented from individual participants. Invalid data are marked as a +100° response. Note the prevalence of invalid responses in ±90° speaker configuration. The units of all axes are degrees.

REFERENCES

Behrend, O., Brand, A., Kapfer, C., and Grothe, B. (2002). Auditory Response Properties in the Superior Paraolivary Nucleus of the Gerbil. *J. Neurophysiol,* 87, 2915-2928.

Begault, D., & Wenzel, E. (1993). Headphone Localization of Speech. *Human Factors, 35* (2), 361-376.

Begault, D. R., Wenzel, E., Shrum, R., & Miller, J. (1996). A Virtual Audio Guidance and Alert System for Commercial Aircraft Operations. *Proceedings of ICAD 96: International Conference on Auditory Display.* Palo Alto.

Blake, R., Sekuler, R. (2006). *Perception (fifth ed).* New York: McGraw-Hill.

Bolia, R. (2004). Spatial Audio Displays for Military Aviation. *International Journal of Aviation Psychology, 14* (3), 233-238.

Bolia, R. S., & D'Angelo, W. R. (1999). Aurally Aided Visual Search in Three-Dimensional Space. *Human Factors, 41* (4), 664-669.

Fisher, J. (1992). Testing the Effect of Road Traffic Signs' Informational Value on Driver Behavior. *Human Factors, 34* (2), 231-237.

Gardner, B., & Martin, K. (1994). *HRTF Measurements of a KEMAR Dummy-Head Microphone.* Massachusetts Institute of Technology, Perceptual Computing. Cambridge: MIT Media Lab.

Goldberg, J. M., & Brown, P. B. (1969). Response of Binaural Neurons of Dog Superior Olivary Complex to Dichotic Tonal Stimuli: Some Physiological Mechanisms of Sound Localization. *J. Neurophysiol, 32* (4), 613-636.

Grohn, M., Lokki, T., & Takala, T. (2004). An Orientation Experiment Using Auditory Artificial Horizon. *Proceedings of ICAD 04 - Tenth Meeting of the International Conference on Auditory Display.* Sydney, Australia.

Hartmann, W. M. (1999). How we localize sound. *Physics Today*, 24.

Holland, S., & Morse, D. (2001). Audio GPS: spatial audio in a minimal attention interface. *Proceedings of Human Computer Interaction with Mobile Devices*, (pp. 28-33).

Jeffress, L. A. (1948). A place theory of sound localization. *J. Comp. Physiol. Psychol.*, 35-39.

Kornblum, S. & Lee, Ju-Wei. (1995). Stimulus-Response compatability with relevant and irrelevant stimulus dimensions that do and do not overlap with the response. *Journal of Experimental Psychology: Human Perception and Performance, 21* (4), 855-875.

Kulkarni, A., & Colburn, H. S. (1998). Role of Spectral Detail in Sound Source Localization. *Nature, 396*, 747-749.

Lerner, N.D., Dekker, D.K., Steinberg, G.V., and Huey, R.W. (1996). *Inappropriate Alarm Rates and Driver Annoyance* (DOT HS 808 533). Silver Spring, Maryland: Office of Crash Avoidance Research (NHTSA).

McAlpine, D. (2005). Creating A Sense of Auditory Space. *J. Physiol, 566* (1), 21-28.

Neurater, M. (2005). Multimodal Warnings: Curve-Warning Design. *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting* (pp. 1945-1949). Orlando: Humand Factors and Ergonomics Society.

Norman, D. (2002). *The Design of Everyday Things.* New York: Basic Books.

Simpson, B.D., Brungart, D.S., Dallman, R.C., Joffrion, J., Presnar, M.D., Gikley, R.H. (2005). Spatial Audio As A Navigation Aid And Attitude Indicator. *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting* (pp. 1602-1606). Orlando: Human Factors and Ergonomics Society.

Tan, C. and Gan, W. (1998). User-defined spectral manipulation of HRTF for improved localisation in 3D sound systems. *Electronics Letters*, 34 (25), 2387-2389.

US Department of Transportation - NHTSA. (n.d.). *Fatality Analysis Reporting System Web Based Encyclopedia.* Retrieved October 10, 2006 from http://www-fars.nhtsa.dot.gov/: http://www-nrd.nhtsa.dot.gov/pdf/nrd-30/NCSA/TSFAnn/TSF2004.pdf.

Wenzel, E., Arruda, M., Kistler, D., & Wightman, F. (1993). Localization using nonindividualized head-related transfer functions. *The Journal of the Acoustical Society of America, 94* (1), 111-123.

Wickens, C., & Hollands, J. (2000). *Engineering Psychology and Human Performance, 3rd edn.* Upper Saddle River, NJ: Prentice Hall.

Wolfe, J. K. (2006). *Sensation&Perception.* Sunderland, MA: Sinauer Associates, Inc.